المجلة العراقية للعلوم الاحصائية (18) 2010
ص ص [155–176]

تقدير نماذج فضاء الحالة باستخدام أسلوب انحدار الحرف مع التطبيق

هيام عبد المجيد حياوي*

الملخص:

تم في هذا البحث توفيق عدة نماذج لفضاء الحالة وتم اختيار أفضل نموذج باستخدام المعايير الإحصائية ومن ثم قدرت معلماته باستخدام طريقة انحدار الحرف بعد الكشف عن وجود مشكلة التعدد الخطي بين المتغيرات التوضيحية للبيانات التي تم استخدامها والتي تمثل متغير الاستجابة وهو الحرارة وعدة متغيرات توضيحية كانت السطوع والرياح والإشعاع الشمسي والتبخر.

Estimation of State Space Models by using Ridge Regression Technique with Application

Abstract

In this paper many state space models are fitted. The statistical criteria used to choose the best estimated model whose parameters estimated by the Ridge regression method. This method can be used when the predicator variables are multicollinear.The data used in this paper represented by the temperature, which is the response variable and the predictor variables effulgence, wined, sunshine and evaporation.

---

* مدرس/قسم الاحصاء والمعلوماتية/ كلية علوم الحاسوب والرياضيات

تاريخ التسلم :2010/3/22 ـــــــــــــــــ تاريخ القبول : 4/21/2010

1- المقدمة

يعد تحليل الانحدار من أوسع الطرائق الإحصائية استخداماً في مختلف العلوم اذ يوضح العلاقة بين المتغيرات على هيئة معادلة يستدل من تقدير معلماتها علـى أهمية هذه العلاقة وقوتها واتجاهها كما يبين تقدير الاستجابة والتنبؤ بما يفيد كثيـرا في التخطيط واتخاذ القرارات الرصينة حولها.

تواجه الباحث مشكلات عديدة اغلبها عدم توفر فروض التحليل عند استخدام طريقة المربعات الصغرى الاعتيادية ومنها مشكلة تعدد العلاقة الخطية التي تـؤثر على نتائج التقديرات والاختبارات وتظهر هذه المشكلة نتيجة وجود ارتبـاط بـين المتغيرات التوضيحية التي تؤدي إلى إعطاء تقديرات ضعيفة لا يمكـن أن يعـول عليها كما أن تباينات هذه المقدرات تكون ضخمة وغير مقبولة فضلا عن ذلك عند اختبار فرضيات حول بعض المعلمات التي تعاني من كبر عامل تـضخم التبـاين سوف يكون القرار المتخذ غير صحيح فقد تقبل الفرضية في حين أنهـا يجـب أن ترفض كما أن فترات الثقة حول تلك المعلمات تكون واسعة وغير منـسجمة مـع الواقع.

هدف البحث :-

يهدف البحث إلى تقدير معلمات أفضل نموذج لفضاء الحالة باستخدام أسـلوب انحدار الحرف وذلك بعد الكشف عن وجود مشكلة تعـدد العلاقـة الخطيـة بـين المتغيرات التوضيحية وتطبيق ذلك على بيانات حقيقية والتـي تمثـل بالمعـدلات الشهرية لدرجات الحرارة والعوامل المؤثرة فيها في محافظة نينوى للفترة (1985-2000) وهي المدخلات المتمثلة بالإشعاع الشمسي الذي يقاس بـ (واط .يـوم / م^2)، سرعة الرياح التي تقاس (م / ساعة)، سطوع الشمس الذي يقاس (ساعة) ، التبخر الذي يقاس بـ (ملمتر) فضلا عن المخرجات وهي درجة الحرارة.

2- مشكلة تعدد العلاقات الخطية

إن نموذج الانحدار الخطي المتعدد $\underline{Y} = X\underline{\beta} + \underline{U}$ حيث أن Y يمثل متجه مشاهدات متغير الاستجابة ذا سعة $(n \times 1)$ و X مصفوفة المتغيرات التوضيحية ذات سعة $(n \times m)$ و β متجه المعلمات المجهولة ذات سعة $((m+1) \times 1)$ و U متجه الأخطاء العشوائية ذا سعة $(n \times 1)$ حيث أن $E(\underline{U}) = \underline{0}$ و $Var(\underline{U}) = \sigma^2 I_n$ ، وقد بني نموذج الانحدار على مجموعة من الفروض منها عدم وجود علاقة خطية تامة أو شبه تامة بين المتغيرات المستقلة أي غياب مشكلة تعدد العلاقة الخطية (Multicollinearity) وبسبب ما ذكر في مقدمة البحث عن خطورة هذه المشكلة فقد اهتمت الدراسات والبحوث اهتماماً كبيراً بمشكلة تعدد العلاقات الخطية بين المتغيرات المستقلة وما تسببه من تشويه للمقدرات التي تقدر بطريقة المربعات الصغرى الاعتيادية ، ويشير المصطلح إلى وجود علاقة خطية تامة أو غير تامة بين اثنين أو أكثر من المتغيرات المستقلة في نموذج الانحدار مما يؤدي إلى خرق فرض من فروض التحليل . (النعيمي،2005)

ويمكن تعريف التداخل الخطي رياضياً بالصورة الآتية : تسمى المتجهات $\underline{X}_1, \underline{X}_2, ..., \underline{X}_m$ مرتبطة خطياً إذا وجد (m) من القيم (الثوابت) مثل $C_1, C_2, ..., C_m$ ليست كلها أصفاراً (أي موجبة وسالبة وصفر) وتحقق العلاقة الآتية: (سعيد، 1996)

$$\sum_{j=1}^{m} C_j \underline{X}_j = \underline{0} \qquad ...(1)$$

إذا تحققت العلاقة (1) فيقال أن مجموعة المتجهات غير مستقلة خطياً ، أو انه إذا كان $C_1 = C_2 = ... = C_m$ أيضاً تحقق المعادلة (1) . أما إذا كان $C_j \neq 0$ فانه يمكن الحل بالنسبة إلى:-

$$\underline{X}_j = \frac{1}{C_1}[C_1X_1 + C_2X_2 + \ldots + C_mX_m] \qquad \ldots(2)$$

من المعادلة (2) يلاحظ بأنه إذا كان هناك (m) من المتجهات مرتبطة خطياً فيمكن التعبير عن احدها بأنها تركيب خطي من المتجهات الأخرى.

أنواع تعدد العلاقات الخطية

يوجد نوعان لتعدد العلاقات الخطية بين المتغيرات التوضيحية هما: (النعيمي،2005)

1. تعدد العلاقات الخطية التامة Exact Multicollinearity

تكون فيه العلاقة الخطية التامة موجودة عندما يتحقق الشرط المتمثل بالعلاقة (3) $C_1, C_2, \ldots, C_m$ ثوابت احدها على الأقل لا يساوي صفراً وعند المقدرات بطريقة المربعات الصغرى الاعتيادية تتطلب إيجاد معكوس المصفوفة $(X'X)$ ويصعب ذلك لكون $|X'X| = 0$ عندما تكون رتبة المصفوفة $(X'X)$ اقل من m ، أي عند وجود علاقة خطية تامة بين المتغيرات التوضيحية لا يمكن تقدير معلمات النموذج بطريقة المربعات الصغرى الاعتيادية.

2. تعدد العلاقات الخطية غير التامة (الجزئية) Non Exact Multicollinearity

تحدث هذه الحالة عندما تكون بعض المتغيرات التوضيحية دالة في التركيبة نفسها للمتغيرات الأخرى مع قيم عشوائية (ω_i) فإذا تحققت المعادلة الآتية:

$$a_1x_1 + a_2x_2 + a_3x_3 + \ldots + a_mx_m + \omega_i = 0 \qquad \ldots(3)$$

$$i = 1,2,\ldots,m$$

إذ أن $a_1, a_2, \ldots, a_m$ ثوابت احدها لا يساوي صفرا ،كذلك $(\omega_i \neq 0)$، في هذه الحالة يكون محدد المصفوفة $(X'X)$ صغيرا ، وعند التقدير بطريقة المربعات الصغرى الاعتيادية سوف يكون التقدير غير دقيق وغير ممثل لواقع التجربة

المدروسة ،والتي يترتب عليه أن يكون تباين المعلمات المقدرة كبيرا جداً وذلك لان :

$$V(\hat{\underline{\beta}}_{OLS}) = S^2 (X'X)^{-1}$$
$$= \frac{adj(X'X)}{|X'X|} S^2 \qquad ...(4)$$

ومن ثم ينتج خطأ بأن بعض المتغيرات التوضيحية غير مهمة،إذ يظهر اختبار (t) عدم معنوية معلمات تلك المتغيرات في حين أنها في الواقع معنوية ولكن بناء النموذج يعجز عن إظهار اثر كل منها منفصلا نظراً لارتباط هذه المتغيرات ببعضها البعض.

إن تعدد العلاقة الخطية ليس من الضروري ان تكون تامة لكي تؤثر في تحليل الانحدار جدياً، كذلك فإن تعدد العلاقة الخطية غير التامة تكون مسببة لمشاكل اكبر من تعدد العلاقة الخطية التامة وذلك بسبب أن المشاكل تبقى مولدة ولكن في معظم الأحيان يكون الباحث غير مدرك لوجود مثل هذه المشكلة.

تشخيص التداخلات الخطية

توجد أساليب عديدة للكشف عن التداخلات الخطية منها: (سعيد، 1996)

1. **الارتباط البسيط بين المتغيرات التوضيحية**

يعد الارتباط البسيط من ابسط مقاييس الكشف عن التداخلات الخطية. فإذا كان $r(x_i, x_j) = \pm 1$ دل ذلك على وجود علاقة خطية بين x_i, x_j وبما أن أعمدة المصفوفة X هي متغيرات ذوات شكل ثابت بطول واحد (unity length and scaling) فإن $(X'X)$ هي مصفوفة الارتباط بين المتغيرات التوضيحية،اقترح Gunst and Mason (1980) فحص الارتباطات بين x_i, x_j التي يكون فيها $r(x_i, x_j) > 0.7$ ذلك من خلال رسم شكل الانتشار بين x_i و x_j فإذا أمكننا التعبير عن الشكل الانتشاري بخط مستقيم معنى

ذلك هناك علاقة خطية بين x_i, x_j ،وإذا كان معامل الارتباط بـين x_i, x_j سالباً معنى ذلك هناك علاقة خطية تأخذ الشكل الآتي:

$$X_i \approx -X_j \quad \text{or} \quad X_i + X_j = 0$$

وإذا كان الارتباط البسيط بين x_i^*, x_j^* موجباً فإن العلاقـة الخطيـة تأخـذ الشكل الآتي:

$$X_i \approx X_j \quad \text{or} \quad X_i - X_j = 0$$

وفي حالة ارتباط أكثر من متغيرين بعلاقة خطية فإنه ليس من الضروري ان يكون $r(x_i, x_j)$ قريبا من الواحد أو حتى كبير ولهذا السبب فان مقياس الارتباط وحده غير كاف في الكشف عن التداخلات الخطية.

2. **مقياس عامل تضخم التباين**

اقتـرح هـذا المقيـاس (1967) Glaube, Ferrar وسـماه (1970) Minquardt بـ (VIF) ،ويعتمد هذا المقيـاس علـى فحـص عناصر القطر الرئيس للمصفوفة $(X'X)^{-1}$ حيث أن :

$$VIF = a_{ij} - (1 - R_j^2)^{-1} \qquad \text{...(5)}$$

$$j = 1,2,...,m$$

حيث أن m هو عدد المتغيرات التوضيحية، R^2 هو معامل التحديد لانحدار X_j على بقية المتغيرات التوضيحية.

توضح المعادلة (5) عند وجود تداخل خطي بين المتغيرات التوضيحية فإن R_j^2 ستكون قيمته قريبة أو مساوية للواحـد وبـذلك تكـون a_{jj} كبيـرة الحجم،أما إذا كان X_j مستقلاً عن بقية المتغيرات التوضيحية الأخرى فإن $R_j^2 = 0$ وبـذلك فـإن $a_{jj} = 1$.وقـد ذكـر بعـض البـاحثين إذا كـان $a_{jj} = VIF \geq 10$ فإن هذا المقدار يكون سببا كافياً لإهمال المتغير x_j من التحليل أو استخدام طريقة أخرى كبديل عن المربعـات الـصغرى فـي التقدير.

اقترح Gunst and Mason (1980) بأنه إذا كان $a_{jj} \geq 4$ فـإن ذلك دليل كاف على كون المتغير X_j يعاني من تضخم في تباين معامله وانه مرتبط ببقية المتغيرات التوضيحية.(سعيد، 1996)

معالجة مشكلة تعدد العلاقة الخطية

هناك العديد من الأساليب التي يتم فيها معالجة التعدد الخطي بين المتغيرات منها:

1. التحويل المعياري للمتغيرات التوضيحية.
2. تضخيم البيانات.
3. حذف بعض المتغيرات.
4. استخدام طرائق التقدير التي تعتمد على المعلومات الكمية الأولية ومنها:

أ) طريقة المربعات الصغرى المقيدة.

ب) طريقة دمج السلاسل الزمنية مع البيانات المقطعية.

ت) استخدام الطرائق المتحيزة في التقدير وأهمها:

1- طريقة المكونات الرئيسة Principle Component Analysis

2- طريقة انحدار الحرف Ridge Regression Method

وسيتم الاعتماد في هذا البحث الاعتماد على طريقة انحدار الحرف في تقدير معلمات النماذج .

3- انحدار الحرف

تعد طريقة انحدار الحرف طريقة معدلة لطريقة المربعات الصغرى الاعتيادية في حالة وجود تداخل خطي بين المتغيرات التوضيحية،وتعتبر من الطرائق المتحيزة اذ تعتمد على إضافة المصفوفة (kIp) إلى المصفوفة $(X'X)$ قبل إيجاد المعكوس لها وكالآتي: (Hoerl and Kemard(1970a))

$$\hat{\beta} = \left(X'X + kIp\right)^{-1} X'Y \qquad ..(6)$$

حيث أن : (kIp) مصفوفة قطرية ذات بعد $(p \times p)$

ان طريقة انحدار الحرف مبنية على مبدأ يجعل مجموع مربعات البواقي اقل ما يمكن وفق قيد متجه المعلمات $\underline{\beta}$ اذ يجعل بعد هذا المتجه عن نقطة الأصل مساوية لكمية ثابتة مساوية لـ m^2 وفي الوقت نفسه إن هذا المقدر يمثل مقدر الإمكان الأعظم المقيد وفق القيد $\beta'\beta \le m^2$ وباستخدام نظرية Kuhn – Tucker في الامثلية المقيدة فإن معادلة لاكرانج تكون كالآتي:

$$L(\beta, C) = -(Y - X\beta)'(Y - X\beta) = K(m^2 - \beta'\beta) \qquad ...(7)$$

يتم الحصول على القيم الحرجة من حل المعادلة (7) أعلاه والتي تجعل $L(\beta, C)$ اكبر ما يمكن ويمكن أن يكون المقدر للمتجه β هو نفس ما تم تعريفه في المعادلة (6) أعلاه.(Davidov,2006)

1. معلمة الحرف: يمكن أن يعرف انحدار الحرف على انه دالة لمعلمة الحرف k و k عدد موجب صغير يضاف إلى القطر الرئيس للمصفوفة $(X'X)$ بعد ضربه بالمصفوفة الأحادية (Ip) ، $0 < k < 1,$ إذ يعمل هذا العدد على تقليل تباين المعلمات المقدرة، ولكن حساب إدخال بعض التحيز على مقدرات تلك المعلمات ،الا ان متوسط مربعات الخطأ mean square error لمقدرات انحدار الحرف اقل من متوسط مربعات الخطأ لطريقة المربعات الصغرى الاعتيادية.
2. تقدير معلمة الحرف:أن اختيار معلمة الحرف k يعود إلى خبرة الإحصائي وعندما تكون فترة زيادة قيمة k زيادة معقولة وصحيحة وغير سريعة،إذ يكون تحسن في المقدرات وهذا التحسن غالباً ما يؤخذ من الواقع الذي تكون فيه المقدرات أكثر استقراراً أو ملاحظة التحسن في التنبؤ
3. وكلما تكبر قيمة k ضمن الفترة $(0,1)$ فإن التباين يقل والمعلمات تصبح أكثر استقراراً.

توجد عدة طرائق لاختيار معلمة الحرف منها الطريقة التكرارية التي تعد من الطرائق المهمة والناجحة لاختيار معلمة الحرف ،إذ أنها تتمتع بصغر متوسط مربعات الخطأ مقارنة بالطرائق الأخرى.(القصيمي،2000)

تعتمد الطريقة التكرارية في تقدير معلمة الحرف على قيمة عامل الحرف الابتدائية وعادة في البداية يتم اختيارها مساوية للصفر k وتعويضها في معادلة انحدار الحرف لإيجاد عدة قيم لمقدرات الحرف،ومن ثم ملاحظة استقرار هذه المقدرات من خلال اخذ الفرق بين المقدرات اللاحقة والسابقة لهذه المقدرات بعد افتراض خطأ معين يمكن السماح به ويتم اختيار احد المقدرات ويعد هذا المقدر هو الأفضل وقيمة k المستخدمة هي القيمة المثلى. (القصيمي،الطالب،الشبخون،2007) (الصفاوي،يحيى،2009).

4- نماذج فضاء الحالة

إن نظم السيطرة الحركية Dynamic Control تتصف بعدم إمكانية تحليل تركيبها وتحسين مستوى أدائها وذلك لعدم التعرف على الحركية الداخلية والخارجية للنظام، اذ إن تحديد المتغيرات الداخلية لنظام السيطرة وإمكانية التحكم بها وملاحظتها يجعل من الممكن الوصول إلى تطوير المفاهيم التي تعمل عليها ومن ثم الحصول على نظام سيطرة مثلى.

تتميز نماذج فضاء الحالة من نظم السيطرة الحركية بأنها توفر إمكانية التعرف على الحركية الداخلية والخارجية للنظام وهي ميزة أساسية في وصف نظام السيطرة، اذ يمكن القول بان تمثيل فضاء الحالة هو أسلوب رياضي يستخدم مع النظم الديناميكية الاعتيادية والمعقدة، تبرز أهمية أسلوب فضاء الحالة عند وجود نظم بمدخلات ومخرجات متعددة ،كما تتميز بأنها تكون في موازاة المتغيرات المعروفة أو التي يمكن قياسها،إذ أن المتغيرات داخل العملية لا يمكن قياسها ولهذا السبب في تسمية نموذج فضاء الحالة بالنموذج الداخلي، وللتفاصيل انظر (البدراني،2002).

يتكون نموذج فضاء الحالة عادةً من معادلتين الأولى تصف متجه المدخلات عند الزمن $(t+1)$ ويرمز له X_{t+1} بدلالة X_t وكذلك المدخلات U_t وتسمى هذه المعادلة بمعادلة الحالةState Equation في حين تصف الثانية المخرجات Y_t بدلالة كل من المدخلات X_t والمدخلات U_t وتسمى هذه المعادلة بمعادلة المشاهدة

Observation Equation ويمكن توضيح هاتين المعادلتين في حالة وجود مدخل ومخرج واحد(SISO (Single Input Single Output كالآتي:-

$$X_{t+1} = AX_t + BU_t + He_t \qquad \dots (8a)$$
$$Y_t = CX_t + DU_t + e_t \qquad \dots (8b)$$

إذ أن:- A : تمثل الديناميكية المستقلة للنظام ببعد (n*n).

B : يمثل تأثير أفعال السيطرة ببعد (n*n).

$\underline{C}$: يمثل الإسقاط على المتغيرات المشاهدة ببعد (1*n).

D : يمثل قيمة حقيقية. H : تمثل متجه عمودي يوضح خصائص البواقي في النظام ذي بعد (1*n).

وتؤدي هاتان المعادلتان دوراً مهماً في دراسة النظم الديناميكية اذ يعبر عن المدخلات والمخرجات من خلال معادلة فرقية عند وجود نظام محدد في الزمن المتقطع وتكون هناك عادةً اضطرابات غير مسيطر عليها تعامل كمتغيرات عشوائية أو تشويش يؤثر في المخرجات (Nelles,2001).

5- معايير اختيار أفضل نموذج

من المعايير الإحصائية التي سيتم تناولها وذلك لاعتمادها للوصول إلى أفضل نموذج

1. دالة الخسارة Loss Function

تعد المتطلب الأساسي والأفضل في اختيار الأنموذج مـن خـلال ملاحظـة سلوك هذه الدالة مع زيادة رتبة النموذج ، اذ تقل قيمتها بزيادة رتبة النمـوذج وان التناقص في قيمة هذه الدالة يتوقف عند نقطة معينة مما يعنـي أن زيـادة رتبـة النموذج تصبح عديمة الفائدة و تعد هذه الميزة من أهم المزايا فـي هـذه الدالـة (Nelles, 2001).

ويمكن حساب هذه الدالة رياضياً بالشكل الآتي :-

$$V = \frac{1}{2}\sum_{t=1}^{N} e_t^2 \qquad \dots (9)$$

إذ إن :-

V : تمثل دالة الكلفة، N : تمثل عدد المشاهدات ، e_t: تمثل البواقي.

2. معيار اكاكي لخطأ التنبؤ النهائي

Akaik's Final Prediction Error Criterion

ويعد من المعايير المهمة في تحديد النموذج المناسب،فقد عرفَ من قبل العـالم Akaike عام 1969 ويرمز له اختصاراً FPE ويمثل مقياس خطأ التنبؤ النهائي ويعرف على انه تباين خطأ التنبؤ لفتـرة قادمـة ويحـسب بالـشكل الآتـي :- (Ljung,2004)

$$FPE = \frac{1+\frac{m}{N}}{1-\frac{m}{N}} V \qquad \dots (10)$$

إذ إن :-

FPE :تٌعبر خطأ التنبؤ النهائي، m : تٌعبر عدد المعلمات في النموذج .

3.معيـار اكـاكي للمعلومـات Akaike's Information Criterion

عرف هذا المعيار أيضا من قبـل العـالم Akaike عـام (1973-1974) ويرمز له اختصارا AIC اذ قدم من خلاله معلومات لاختيار الرتبـة الملائمـة لنموذج ARIMA من بين عدة نماذج بحيث تقابل الرتبة المناسبة اقل قيمة لمعيار AIC وتمثل الرتبة الأكثر ملاءمة، وهذا المعيار يساوي ضعف عـدد المعلمـات مطروحاً منه ضعف لوغاريتم دالة الترجيح للمشاهدات،ويعبر عنه رياضياً كالآتي: (Wei ,1990) و(Nelles, 2001)

$$AIC = 2\,p - 2\ln L \qquad \dots (11)$$

إذ إن :-

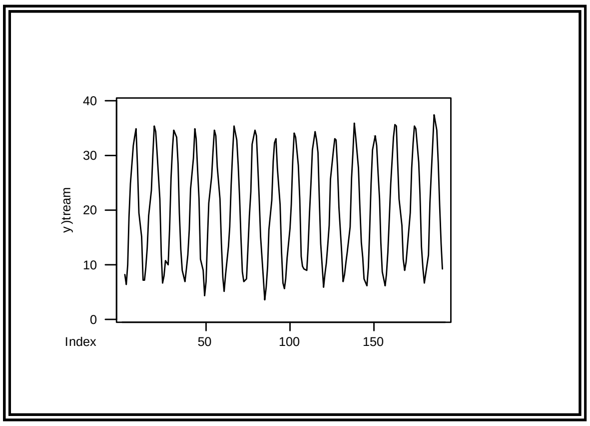
AIC : يٌعبر عن معيار اكاكي للمعلومات، L : تٌعبر عن دالة الترجيح.

علما أن معظم البرمجيات الجاهزة مثل Matlab تستخدم دالة الكلفـة فـي حساب قيمة هذا المعيار و يمكن التعبير عنه كالآتي:-(Ljung, 2004)

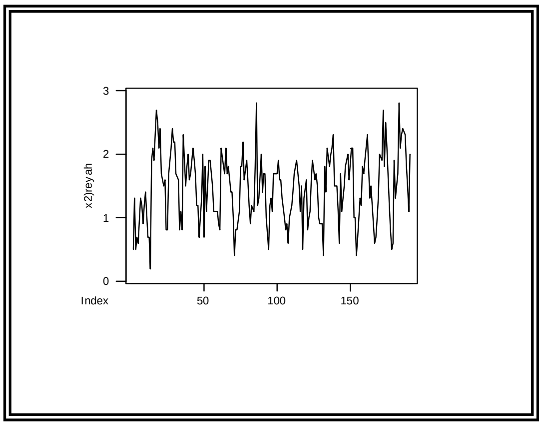
$$AIC = \log\left(V\left(1 + 2\frac{m}{N}\right)\right) \qquad \dots (12)$$

6- الجانب التطبيقي

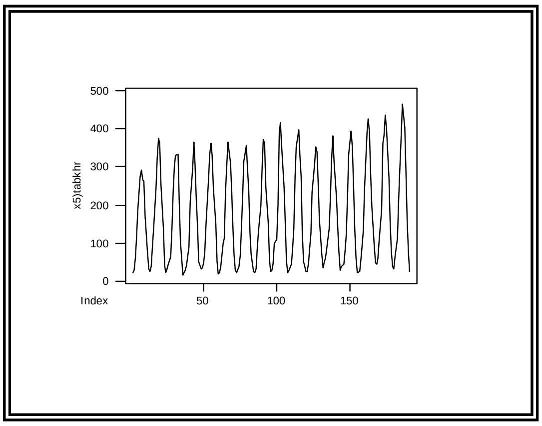
يستند البحث الى بيانات حقيقية تتمثل بالمعدلات الشهرية لدرجات الحـرارة والعوامل المؤثرة فيها في محافظة نينوى للفترة (1985-2000) وهي عبارة عـن المدخلات المتمثلة بالإشعاع الشمسي الذي يقاس بـ (واط .يـوم / م^2)، سـرعة الرياح التي تقاس (م / ساعة)، سطوع الشمس الذي يقاس (ساعة) ، التبخر الذي يقاس بـ (ملمتر) فضلا عن المخرجات وهي عبارة عن درجة الحرارة،وكما مبين في الملحق ويمكن تمثيل هذه البيانات بالأشكال الآتية :



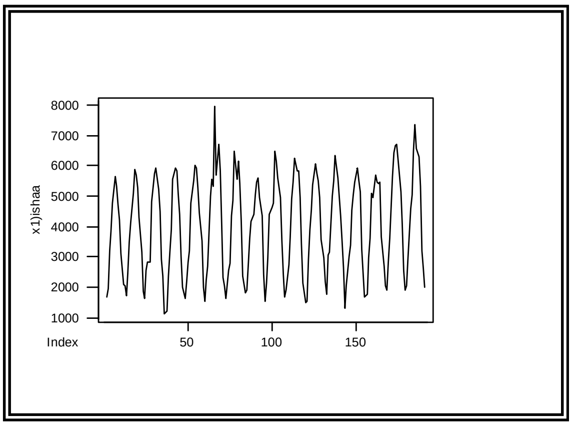
الشكل رقم (1) : الرسم الزمني لمخرجات الحرارة y_t



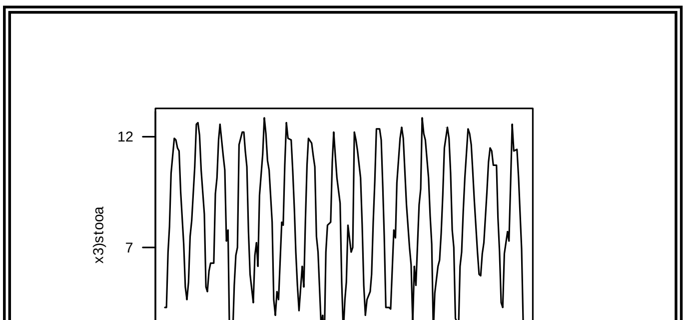
الشكل رقم (2) : الرسم الزمني لمدخلات الرياح x_{2t}



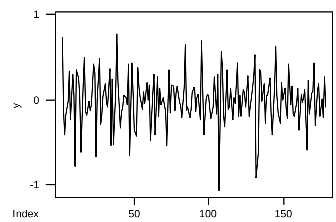
الشكل رقم (3) : الرسم الزمني لمدخلات التبخر x_{4t}



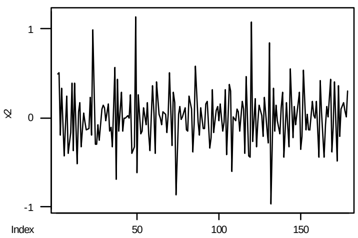
الشكل رقم (4) : الرسم الزمني لمدخلات الإشعاع x_{1t}



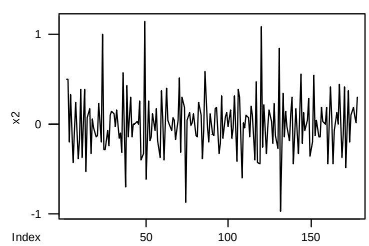
نلاحظ من هذه الأشكال أن سلسلة المدخلات والمخرجات غير مستقرة في الوسـط والتباين ولأجل جعلها مستقرة تم استخدام احد التحويلات وهي اخذ الجذر التربيعي لثبوت التباين والفرق الموسمي لإزالة اثر الموسم من هذه السلسلة،ثم تم اخذ الفرق الأول لجعلها مستقرة في الوسط وكما مبين في الأشكال الآتية:



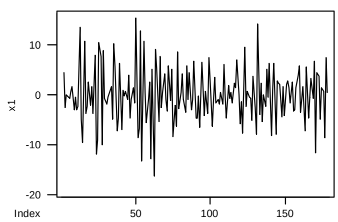
الشكل (6) : الرسم الزمني بعد إزالة الأثر الموسمي واستقرار مخرجات الحرارة



الشكل (7) : الرسم الزمني بعد إزالة الأثر الموسمي واستقرار لمدخل الرياح

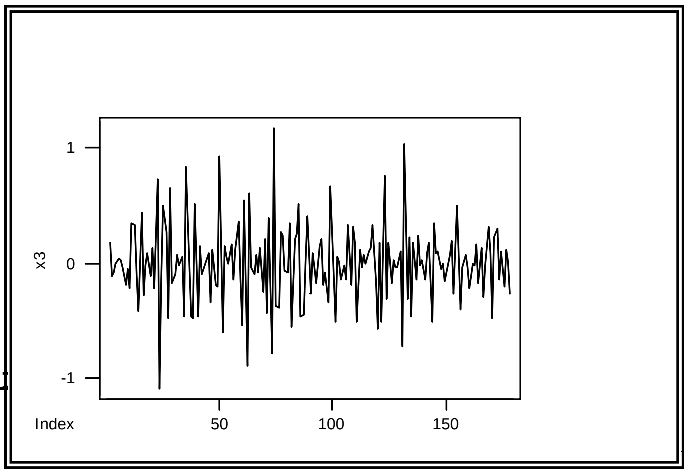


الشكل (8) : الرسم الزمني بعد إزالة الأثر الموسمي واستقرار لمدخل التبخر



الشكل (9) : الرسم الزمني بعد إزالة الأثر الموسمي واستقرار لمدخل الإشعاع

تم توفيق ... ستخدام البرنـامج
الجاهز B



الشكل (10) : يمثل الرسم الزمني بعد إزالة الأثر الموسمي

الجدول(1): نماذج فضاء الحالة بالرتب (1-10) مع المعايير

المعايير / الرتب	AIC	Loss fun.	FPE
1	-2.7275	0.06092	0.0653
2	-2.6674	0.06029	0.0694
3	-2.6805	0.05544	0.06858
4	-2.2995	0.07563	0.10049
5	-3.0618	0.03288	0.04697
6	-2.5689	0.05016	0.07711
7	-1.2597	0.1731	0.2867
8	-2.6260	0.04114	0.07351
9	-1.5479	0.112	0.217
10	-0.8399	0.213	0.4456

وتم إيجاد قيم المعايير الإحصائية ومنها معيار اكاكي AIC ،معيار دالة الخسارة Loss Function ، فضلا عن معيار خطأ التنبؤ النهائي FPE واخترنـا نمـوذج فضاء الحالة من الرتبة الخامسة ليكون أفضل نموذج وذلك لأنه يقابل اقل قيم لهـذه

المعايير كما تم أيضا تقدير معلمات هذا النموذج باستخدام إحدى طرائق التقدير في الانحدار وهي طريقة المربعات الصغرى وكانت صيغته كالآتي:

State space model :

$$A=\begin{bmatrix} -0.48633 & 0.46964 & 0.48097 & 0.25821 & 0.30226 \\ -0.62875 & -0.3991 & -0.34703 & -0.081675 & -0.082651 \\ 0.20777 & 0.10041 & 0.11191 & -0.0042911 & -1.1727 \\ 0.25077 & 0.67577 & -0.18866 & -0.41521 & -0.7686 \\ -0.056559 & 0.039169 & -0.013694 & 0.74186 & -0.26141 \end{bmatrix}$$

$$B=\begin{bmatrix} 0.0028249 & 0.079402 & 0.068167 & -0.066512 \\ -0.0037625 & 0.051898 & 0.37275 & -0.050124 \\ 0.0010686 & -0.16341 & -0.90413 & 0.12487 \\ 0.0043885 & 0.11034 & 0.2813 & -0.05409 \\ -0.0025527 & 0.052015 & 0.20093 & -0.026866 \end{bmatrix}$$

$$H = \begin{bmatrix} 0.02384 \\ -0.39616 \\ 1.1947 \\ -0.15996 \\ -0.32353 \\ \end{bmatrix} , \qquad X(0) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$C = \begin{bmatrix} -1.0056 & 1.4557 & -0.15762 & -1.2118 & -0.56404 \end{bmatrix}$$

وكانت قيمة D= Constant وتم تمثيل هذا النموذج من خلال صيغة احد نمـاذج الصندوق الأسود وهو ARMAX باستخدام النظام الجاهز الـ MATLAB وتـم الحصول على المعادلة الآتية المتمثلة بالنموذج ARMAX وكالآتي:

$$\begin{aligned} y_t = &-1.45\ y_{t-1} - 1.521\ y_{t-2} - 0.6169\ y_{t-3} - 0.04336\ y_{t-4} + 0.3787\ y_{t-5} - 0.01236\ u_{1,t-1} - 0.01706\ u_{1,t-2} \\ &- 0.01334\ u_{1,t-3} - 0.0006575\ u_{1,t-4} + 0.003769\ u_{1,t-5} - 0.1416\ u_{2,t-1} - 0.2253\ u_{2,t-2} - 0.05502\ u_{2,t-3} \\ &+ 0.132\ u_{2,t-4} + 0.1569\ u_{2,t-5} + 0.1623\ u_{3,t-1} + 0.2561\ u_{3,t-2} + 0.5577\ u_{3,t-3} + 0.5447\ u_{3,t-4} + 0.4863\ u_{3,t-5} \\ &+ 0.05494\ u_{4,t-1} + 0.1212\ u_{4,t-2} + 0.02225\ u_{4,t-3} - 0.05795\ u_{4,t-4} - 0.1092\ u_{4,t-5} + a_t + 1.038\ a_{t-1} \\ &+ 0.5173\ a_{t-2} - 0.7052\ a_{t-3} - 0.7709\ a_{t-4} - 0.7337\ a_{t-5} \end{aligned} \quad \ldots (13)$$

وبعد فحص البيانات تبين وجـود علاقـة خطيـة بـين المتغيـرات (المـدخلات والمخرجات) وذلك من خلال استخدام أساليب الكشف عن تعدد العلاقـة الخطيـة المبينة في الجانب النظري ومنها أسلوب الارتباط البسيط اذ وجـد أن مـصفوفة الارتباط كانت بالشكل الآتي:

$$R = \begin{bmatrix} 1.00000 & 0.88609 & 0.34162 & 0.94002 & 0.96552 \\ 0.88609 & 1.00000 & 0.49878 & 0.92472 & 0.91152 \\ 0.34162 & 0.49878 & 1.00000 & 0.34420 & 0.45028 \\ 0.94002 & 0.92472 & 0.34420 & 1.00000 & 0.93539 \\ 0.96552 & 0.91152 & 0.45028 & 0.93539 & 1.00000 \end{bmatrix}$$

نلاحظ من المصفوفة أعلاه أن قيمة الارتباط بين المتغيرات يتجاوز 0.7 أي أن هناك ارتباطا يبلغ 0.88609 و 0.94002 و 0.96552 في الصف الأول أي العلاقة بين المتغيرات التوضيحية وهي السطوع والرياح والإشعاع الشمسي والتبخر مما يدل على وجود علاقة خطية بين المتغيرات،وبما أن هذا المقياس غير كاف لمعرفة تعدد العلاقة الخطية بين المتغيرات كما تم توضيحه في الجانب النظري عليه فقد تم الاعتماد على مقياس آخر.

وباستخدام مقياس عامل تضخم التباين كانت النتائج كالآتي:

الجدول (2) قيم معامل التضخم في تباينات المعاملات

المتغيرات	VIF
الإشعاع	9.6
الرياح	1.6
السطوع	13.0
التبخر	9.7

يلاحظ من الجدول المذكور انفا أن VIF للإشعاع والسطوع والتبخر كانت عالية وهذا دليل على أن تلك المتغيرات هي المسببة لتعدد العلاقة ،فضلا عن ذلك فقد ذكر (1980) Gunst and mason انه إذا كانت قيمة معامل التضخم VIF اكبر أو تساوي 4 فإن ذلك دليل كاف على كون المتغير يعاني من تضخم في تباين معامله وانه مرتبط ببقية المتغيرات التوضيحية.

ونظرا لوجود العلاقة الخطية بين المتغيرات حسب ما تم بيانه في مقاييس اختبار وجود تعدد العلاقة الخطية بين المتغيرات التوضيحية فقد تم استخدام أسلوب انحدار الحرف لتقدير قيم المعلمات للنماذج كبديل عن طريقة المربعات الصغرى وحددت قيمة k (عامل الحرف)التي تقع بين الصفر والواحد بالطريقة التكرارية وتم التوصل الى قيمة k المثلى وهي k=0.70 من خلال استخدام البرمجية الجاهزة MINITAB وعليه تم توفيق النموذج الآتي:

$$y=4.38803-0.00019x_1-1.46362x_2+0.99101x_3+0.05996x_4 \qquad \ldots (14)$$

وباستخدام معايير المعلومات تمت مقارنة النماذج التي تم الحصول عليها باستخدام أسلوب انحدار الحرف وطريقة المربعات الصغرى وكانت النتائج كالآتي:

الجدول (4) : يوضح معايير المقارنة

المعايير / النموذج	MSE	AIC	FPE	Loss Fun.
طريقة المربعات الصغرى	56529	2516.7	0.6587	0.5891
طريقة انحدار الحرف	4.959	860.3	0.04697	0.03288

نلاحظ من الجدول أعلاه أن النموذج المقدر باستخدام انحدار الحرف يقابل اقل قيم لمعيار متوسط مربعات الخطأ فضلا عن معيار اكاكي ومعيار خطأ التنبؤ النهـائي ودالة الخسارة مما يدل على تميز طريقة انحدار الحرف في تقدير معلمات أفـضل نموذج يمثل البيانات وذلك لأنه يقابل اقل قيم للمعايير الإحصائية أعلاه.

7- الاستنتاجات

خاص البحث إلى بعض الاستنتاجات نذكر منها:

1. لاحظنا في الجانب العملي وجود علاقة خطية بين المـدخلات والمتمثلـة بالرياح والتبخر والسطوع والإشعاع وبين المخرجـات المتمثلـة بدرجـة الحرارة وتم الكشف عنها باستخدام أساليب الكشف عـن تعـدد العلاقـة الخطية.
2. إن تقدير معلمات نموذج فضاء الحالة باستخدام أسلوب انحـدار الحـرف أعطى نتائج أفضل عن نموذج فضاء الحالة المقـدر بطريقـة المربعـات الصغرى وتبين ذلك من خلال مقارنة النمـوذجين المقـدرين باسـتخدام المعايير الإحصائية فقد أعطى نموذج انحدار الحرف اقـل قـيم لمعـايير المعلومات وكما موضح في الجدول 4.

8- التوصيات

بالإمكان استخدام أسلوب انحدار الحرف في تقدير قيم معلمات نماذج الصندوق الأسود وهي نماذج ديناميكية أو حركية تتضمن مدخلات ومخرجات ومقارنة النتائج مع طرائق التقدير الأخرى كالمربعات الصغرى وطريقة الإمكان الأعظم .

9- المصادر

1. البدراني ،ظافر رمضان مطر ،(2002):"**دراسة في تشخيص النظم التصادفية مع إشارة خاصة إلى أسلوب فضاء الحالة والاستقرارية**"،أطروحة دكتوراه ،كلية علوم الحاسبات والرياضيات،جامعة الموصل،العراق.
2. سعيد،هيفاء عبد الجواد،(1996):"**طرق التعرف على تعدد العلاقة الخطية وكيفية معالجتها بطرائق التقدير المتحيزة**"،رسالة ماجستير،كلية الإدارة والاقتصاد،جامعة الموصل،العراق.
3. النعيمي ، أسوان محمد طيب، (2005):"**اختيار المتغيرات في انحدار الحرف**"،رسالة ماجستير،كلية علوم الحاسبات والرياضيات،جامعة الموصل،العراق.
4. القصيمي،عزة مصطفى عبد القادر،الطالب ،محاسن عبد ا ،الشبخون ،عبير أسامة،(2007)،:"**مقارنة طريقتي .H.S.L والطريقة التكرارية لتقدير معلمة الحرف**"،مجلة تنمية الرافدين،المجلد 29،العدد 86،ص.ص171-179.
5. القصيمي،عزة مصطفى عبد القادر ،(2000):"**استخدام أسلوب المحاكاة في مقارنة مقدرات انحدار الحرف**"، رسالة ماجستير،كلية علوم الحاسبات والرياضيات،جامعة الموصل،العراق.
6. الصفاوي،د.صفاء يونس،يحيى ،مزاحم محمد،(2009):"**استخدام انحدار الحرف والانتروبي العظمى العامة في تحليل التلوث لمعمل اسمنت كركوك**"،المؤتمر العلمي الثاني للرياضيات-الإحصاء والمعلوماتية للفترة 6-2009/12/7.

7. Davidov,O.,(2006):"**Constrained Estimation and the Theorem of KUHN-TUCKER**",Hindawi Publishing Corporation, Journal of Applied Mathematics and Decision Sciences ,Article ID 92970,PP.(1-13).
8. Horel, A.E & Kennard, R.W., (1970a):"**Ridge Regression: Biased Estimation for Nonorthogonal Problem** "; Techno metrics, No.1.12.O P.Cit.P.58.OP.Cit.PP.69-80.
9. Nelles, O., (2001):"**Nonlinear System Identification from Classical Approach to Neural Network and Fuzzy Models**", Springer Verlag Belin Heidelberg Germany.
10. Ljung, L., (2004):"**System Identification Toolbox for use with MATLAB**", 6.0 Math works Inc.
11. Wei, W.W.S., (1990):"**Time Series Analysis Univariate and Multivariate Methods**", Addison Wesley Publishing Company Inc.The advanced Book Program California U.S.A.

الملحق

البيانات التي تم استخدامها في البحث

ت	Y	X1	X2	X3	X4	ت	Y	X1	X2	X3	X4
1	8.2	1687	0.5	4.3	21.6	98	7.4	2951	1.7	5.4	44.2
2	6.4	1977	1.3	4.3	29.7	99	11.3	4392	1.7	8.0	100.8
3	10.0	3203	0.5	6.8	59.7	100	16.6	4627	1.7	6.8	109.5
4	18.6	3822	0.7	7.9	113.6	101	21.1	4754	1.9	7.0	197.3
5	24.9	4769	0.6	10.3	186.3	102	29.1	6449	1.6	12.2	386.6
6	31.8	5641	1.3	11.9	277.1	103	34.0	6138	1.6	11.8	418.2
7	33.3	5330	1.2	11.8	292.0	104	33.4	5606	1.3	11.3	353.6
8	34.9	4718	0.9	11.5	265.0	105	28.1	4943	1.0	10.1	246.5
9	28.1	4200	1.2	11.3	261.8	106	22.0	3542	0.8	8.0	147.7
10	19.4	3129	1.4	9.5	169.0	107	11.5	2411	0.9	5.3	52.5
11	15.2	2096	0.7	7.1	73.6	108	9.8	1666	0.6	3.9	24.0
12	7.3	2036	0.7	5.2	33.2	109	9.2	1899	1.0	4.6	28.6
13	7.3	1742	0.2	4.6	24.2	110	8.9	2724	1.2	5.0	46.4
14	9.6	2451	1.9	5.4	40.0	111	13.2	3666	1.4	5.8	83.3
15	13.0	3494	2.1	7.5	92.0	112	19.4	4881	1.7	8.1	137.0
16	18.9	4134	1.9	8.2	139.3	113	24.8	5446	1.8	10.0	276.7
17	23.5	5074	2.7	10.6	243.7	114	31.0	6237	1.9	12.3	352.0
18	30.1	5867	2.5	12.5	325.4	115	34.3	5836	1.5	12.3	397.0
19	35.5	5669	2.1	12.6	373.7	116	32.7	5836	1.1	11.8	329.9
20	34.4	5269	2.4	12.0	361.7	117	30.4	4932	1.5	9.7	272.8
21	30.8	4308	1.7	10.5	252.3	118	22.4	3415	0.5	7.5	124.1
22	22.1	3166	1.5	8.5	139.6	119	13.9	2129	1.3	4.3	51.2
23	11.7	1862	1.6	5.2	41.5	120	5.8	1480	1.6	4.3	25.3
24	6.6	1614	0.8	5.0	21.3	121	8.3	1553	0.8	4.2	25.2
25	8.0	2541	0.8	5.9	33.0	122	10.3	2928	1.0	6.0	46.9
26	10.7	2843	1.7	6.3	46.0	123	13.5	3907	1.1	7.8	87.2
27	10.0	2843	2.1	6.3	63.2	124	17.1	4528	1.6	7.4	124.4
28	16.7	4808	2.4	9.4	139.6	125	25.6	5344	1.9	9.9	233.9
29	25.9	5280	2.2	10.1	230.8	126	30.7	6029	1.6	11.9	307.9
30	31.3	5728	2.2	11.8	301.2	127	33.0	5713	1.7	12.4	352.8
31	34.6	5931	1.7	12.5	330.2	128	32.9	5483	1.5	11.9	339.2
32	33.3	5216	1.6	11.1	332.9	129	28.0	4926	1.0	10.4	250.4
33	28.4	4486	0.8	10.5	216.3	130	20.6	3568	0.9	8.9	159.1
34	19.6	2906	1.1	7.3	101.7	131	12.2	2967	0.9	7.0	68.9
35	12.7	2373	0.8	7.8	56.2	132	6.8	2192	0.4	6.2	34.2
36	9.1	1150	2.3	2.8	17.0	133	8.3	1756	1.8	3.6	51.6
37	6.8	1225	1.5	3.0	28.7	134	10.6	3061	1.4	6.1	59.9
38	9.3	2352	1.8	5.3	42.1	135	12.4	3149	2.1	5.3	87.7

39	11.7	3162	2.0	6.6	67.6	136	16.8	4980	1.8	8.9	137.2
40	16.5	3887	1.6	7.0	89.8	137	26.0	5499	2.0	9.6	220.6
41	23.9	5546	1.7	11.6	208.5	138	30.2	6312	2.1	12.8	316.3
42	29.5	5891	2.1	12.2	293.9	139	35.8	5941	2.3	12.1	380.4
43	34.9	5829	1.9	12.2	364.3	140	33.4	5604	1.5	11.8	311.6
44	33.2	5069	1.7	11.3	308.3	141	27.7	4353	1.5	10.1	218.3
45	27.4	4387	1.2	10.6	215.9	142	20.5	3523	1.1	8.4	138.9
46	22.1	3013	1.2	7.7	141.3	143	14.0	2664	0.6	7.1	76.5
47	10.9	2011	0.7	5.8	50.0	144	11.4	1302	1.7	3.2	29.9
48	9.0	1619	1.3	4.5	32.2	145	7.5	2067	1.1	4.9	39.9
ت	Y	X1	X2	X3	X4	ت	Y	X1	X2	X3	X4
49	4.3	2201	2.0	6.6	34.2	146	6.2	3021	1.5	6.1	44.8
50	6.8	2842	0.7	7.2	46.9	147	9.6	3375	1.8	6.4	78.0
51	14.1	3180	1.8	6.1	76.4	148	16.1	4517	1.9	7.6	121.0
52	21.3	4749	1.1	9.3	151.6	149	24.9	5025	2.0	9.3	224.4
53	26.1	5484	1.9	11.2	262.5	150	31.1	5473	1.6	11.5	332.3
54	30.4	5992	1.9	12.8	333.7	151	33.5	5898	2.1	12.4	392.9
55	34.7	5909	1.7	12.1	362.3	152	31.7	5501	2.1	11.9	352.3
56	33.6	5266	1.5	10.9	330.3	153	26.7	5134	1.0	10.3	242.8
57	28.0	4446	1.1	10.5	242.6	154	22.0	3227	1.0	7.8	137.9
58	22.1	3512	1.1	8.1	151.8	155	14.0	2514	0.4	7.0	59.7
59	13.8	1989	1.1	4.6	50.0	156	8.8	1672	0.7	3.8	23.7
60	7.6	1527	0.9	3.9	17.7	157	6.2	1794	1.3	3.5	24.6
61	5.1	2233	0.8	5.0	21.5	158	8.2	2982	1.2	6.1	53.7
62	8.2	2709	2.1	4.6	40.0	159	12.4	3602	1.8	6.8	94.2
63	13.4	4975	1.8	8.1	100.3	160	18.4	5075	1.7	8.6	136.2
64	16.8	5531	1.7	8.0	113.4	161	24.9	4968	1.9	10.1	241.8
65	24.7	5296	2.1	10.7	236.9	162	33.3	5679	2.3	12.3	391.5
66	30.6	7917	1.7	12.6	308.7	163	35.7	5473	1.8	12.1	426.3
67	35.3	5671	1.8	11.9	365.1	164	35.4	5399	1.3	11.6	390.8
68	32.9	6690	1.4	11.8	309.9	165	29.0	5433	1.5	10.5	284.8
69	28.3	5823	1.4	10.4	241.3	166	22.0	3641	1.2	9.1	197.1
70	22.0	4012	1.0	8.7	141.6	167	17.3	2712	0.6	6.8	92.9
71	15.2	2316	0.4	6.8	69.8	168	11.1	2069	0.7	5.8	49.4
72	8.6	2047	0.8	5.3	30.2	169	9.1	1915	1.0	5.7	46.1
73	6.8	1650	0.8	4.1	21.7	170	10.4	2812	1.3	6.7	61.7
74	7.5	2564	1.1	6.1	38.5	171	13.6	3598	2.0	7.2	107.7
75	13.1	2791	1.8	5.2	67.0	172	19.6	5652	1.9	9.5	185.4
76	19.0	4334	1.8	8.2	135.8	173	27.6	6429	2.7	10.8	361.9
77	23.3	4850	2.2	10.7	225.6	174	32.2	6661	1.8	11.5	380.6
78	32.0	6480	1.6	11.9	315.2	175	35.4	6711	2.5	11.3	434.5
79	34.5	5552	1.9	11.7	354.5	176	34.8	6175	2.1	10.7	397.0

80	33.6	6148	1.6	11.1	299.4	177	28.7	5116	1.2	10.7	272.0
81	28.4	5389	1.2	10.6	237.6	178	22.8	4055	0.8	8.3	161.6
82	22.3	4186	0.9	7.5	124.2	179	13.4	2574	0.5	6.8	76.8
83	15.0	2364	1.2	6.8	70.6	180	9.4	1915	0.6	4.5	38.3
84	7.7	1801	1.1	3.5	24.6	181	6.7	2053	1.9	4.3	31.9
85	3.7	1899	1.8	3.9	23.8	182	8.3	2897	1.3	6.7	65.4
86	5.8	2758	2.8	2.7	31.3	183	11.8	4572	1.7	7.7	110.9
87	9.8	3591	1.2	6.8	84.1	184	21.2	5054	2.8	7.3	216.5
88	16.4	4176	1.3	8.0	131.8	185	26.3	6482	2.1	9.8	294.2
89	21.9	4405	2.0	8.1	200.3	186	31.9	7335	2.3	12.5	375.6
90	29.0	5044	1.4	10.8	300.6	187	37.4	6562	2.4	11.3	464.1
91	32.2	5434	1.7	12.2	370.5	188	34.5	6261	2.3	11.4	403.3
92	33.2	5606	1.7	11.1	363.2	189	28.6	5307	1.8	10.1	275.3
93	27.7	4955	1.0	10.1	247.8	190	20.8	3192	1.5	8.4	158.3
94	21.0	4360	0.5	9.0	153.5	191	13.8	2618	1.1	6.9	72.8
95	12.4	2412	1.2	5.4	54.5	192	9.2	1992	2.0	3.5	25.1
96	6.6	1555	1.3	3.2	26.5						
97	5.6	2149	1.1	4.6	28.0						