مقارنة بين أسلوب بيز وكرنل لتقدير الانحدار اللامعلمي بالطاقة الكهربائية في العراق
(دراسة تطبيقية)

أ.د/ البيومي عوض عوض طاقية
أستاذ الإحصاء التطبيقي
ووكيل كلية التجارة لشؤون التعليم والطلاب
كلية التجارة- جامعة المنصورة

د/ هشام محمد رجب المنجي
مدرس الإحصاء التطبيقي
كلية التجارة- جامعة المنصورة

عادل احمد صالح جاسم

المستخلص :

يتمثل الهدف من هذا البحث في اختيار نموذج إحصائي يعطي أفضل وأدق التقديرات الشهرية لإنتاج واستهلاك الطاقة الكهربائية في العراق للفترة من يناير عام ٢٠٠٧ وحتى ديسمبر عام ٢٠١٣ في القطاع المنزلي من خلال المقارنة بين أسلوب بيز وكرنل لتقدير الانحدار اللامعلمي . وترجع أهمية هذا البحث إلى أهمية الطاقة الكهربائية واهتمام الدول والحكومات بهذه الطاقة الحيوية لاعتبارها من المصادر الرئيسية للطاقة التي يعتمد عليها في مختلف جوانب الحياة.

وقد تبين وفقا للبيانات المستخدمة أن أفضل النماذج كان النموذج المقدر باستخدام بيز ، وأن أهم المتغيرات التي تؤثر معنوياً في الكمية المستهلكة في القطاع المنزلي هي عدد المشتركين والرطوبة النسبية وفي القطاع الصناعي هي عدد المشتركين فقط

ABSTRACT

The objective of this research is to select of a statistical model which gives the best and most accurate monthly production and consumption of electric power in Iraq for the period from January 2007 estimates to December 2013 in the domestic sector through comparison between Bayes and kernel method to estimate Allamwalima decline. The importance of this research return to the importance of electrical energy and the attention of State and Government in this vital energy to the mind of the main sources of energy upon which the various aspects of life.

The according to the data used to show that the best models were estimated using the bayes model, and that the most important variables that significantly affect the amount consumed in the domestic sector are the number of subscribers and the relative humidity, in the industry sector is the number of subscribers only.

المقدمة :

يتطلب التقدم العلمي في كافة المجالات والاختصاصات في عصرنا الحالي تطوراً إن يكون هنالك تطوراً في علم الإحصاء ، تعتبر أساليب الانحدار اللامعلمي من الأدوات

Introduction

الإحصائية المهمة لوضع الخطط والبرامج المرتبطة بالدراسة ظاهرة معينة على فترة من الزمن وتعتمد اغلب الدول في بناء خططها وبرامجها على الأسس

والمعـــايير العلميـــة الدقيقـــة للوصـول إلـى تحلـيلات دقيقـة تسهم في تطوير مجـالات الحيـاة المختلفـة ،ولعلـم الإحصـاء دور مهم وكبير في بنـاء هذه الخطط والبرامج الموضوعة من خـلال دراسة الظـواهر والتنبؤ بهـا فـي المستقبل .

اهـتم البحـث بـالتطبيق علـى البيانـات الخاصـة باسـتهلاك القطـاع المنزلـي مـن الطاقـة الكهربائيـة ،حيـث يسـتهلك هـذا القطـاع النسـبة الكبـرى مـن الكهربـاء ، مـع الأخـذ فـي الاعتبـار المتغيـرات التـي تـؤثر معنويـاً علـى الكميـة المسـتهلكة من الكهرباء (عدد المشتركين ، درجة الحرارة درجة الرطوبة النسبية).

ونقارن هذا البحث بين أسلوب بيـز وكرنـل لتقـدير الانحـدار اللامعلمـي والمفاضـلة بيـنهم واختيـار الأسـلوب الافضـل والأدق .[٣]

<u>مشكلة البحث</u> :

أصـبح انقطـاع التيـار الكهربائي هي الحالة السائدة في كافـة مـدن العـراق والانقطـاع لفترات طويلـة تصل أحيانـا إلـى اكثـر مـن ١٢ساعة يومياً، ممـا يعـود إلـى مشـاكل اقتصـادية واجتماعيـة عديـدة فضـلاً عـن الانعكاسات النفسية.

ويتـأثر الطلـب علـى الطاقـة الكهربائيـة بنـوع القطاعـات المسـتخدمة وتختلـف الكميـة المسـتهلكة مـن القطاعـات بـاختلاف سـاعات العمـل يومياً وتبعاً للظروف المناخية ، ففـي قطـاع الصـناعة مـثلاً يـنخفض الطلـب علـى الطاقـة الكهربائيـة فـي الليـل إلـى أدنـى مسـتواه، وفـي الإنـارة العامـة (التـي تتبـع قطـاع الخـدمات العامـة) تـنخفض الكميـة المستهلكة من الكهربـاء نهـاراً، وترتفع الكمية المستهلكة ليلاً .

وتتـأثر كميـة الطاقـة المسـتهلكة بعدة متغيرات تفسيرية مستقلة مثـل متوسـط درجـة الحـرارة ومتوسط درجة الرطوبة النسبية وعـدد المشـتركين المسـتهلكين للطاقة .

مما سبق تتمثل مشكلة البحث فـي عمليـة اختيـار نمـوذج إحصـائي للوصـول إلـى أفضـل نمـوذج للكميـة المنتجـة والمستهلكة من الكهرباء ، حيث تعاني الدول من نقص الكهربـاء بشكل مستمر خلال الزمن .

<u>هدف البحث</u> :

يهدف البحث إلى استخدام الأساليب التقدير والخاصة بالانحدار اللامعلمي بهدف إيجاد أفضلها ملائمة للنماذج المتنوعة وتوزيع الخطأ العشوائي ، ودلك من خلال استخدام أسلوبي بيز وكرنل لتقدير الانحدار اللامعلمي بالطاقة الكهربائية في العراق، والمفاضلة بين النماذج واختيار الأسلوب المناسب للبيانات .

أهمية البحث:

تتمثل أهمية هذا البحث في أهمية الطاقة الكهربائية ، واهتمام الدول والحكومات بهذه الطاقة الحيوية التي لا غنى عنها في حياتنا العاصرة ،لاعتبارها من المصادر الأساسية للطاقة التي تعتمد عليها مختلف جوانب الحياة ومن ثم نشأت فكرة هذه الدراسة حول الطاقة الكهربائية، لذا يهدف البحث إلى اختيار نموذج إحصائي يعطي أفضل وأدق التقديرات الممكنة لانتاج واستهلاك الطاقة الكهربائية في محافظات العراق من خلال استخدام أسلوب بيز وكرنل لتقدير الانحدار اللامعلمي.

الاساليب الإحصائية المستخدمة :

مفهوم نظرية بيز Concept of bayesian theorem

نشأت مفهوم نظرية بيز في أواسط القرن الثامن عشر من قبل الكاهن (thom as bays) وسميت باسمة وعلية ظهر رأيان مختلفان ليشكلا مدرستين أو اتجاهين متضادين لتقويم نظرية بيز، فيما يوجد معارض مؤيد، ليستمر النقاش بين أصحاب هاتين المدرستين .

تفترض الطرق التقليدية في التقدير أن للمعلمة بأنها ثابتة وليست متغيرة ، وتم الاعتماد على بيانات العينة المُشاهدة فقط.[٧]

بينما تعتمد طريقة بيز في التقدير على المعلومات المسبقة حول المعلمات المطلوب تقديرها على اعتبار هذه المعلمات متغيرات عشوائية وليست كميات ثابتة ويضاف على تلك المعلومات المسبقة معلومات العينة المشاهدة. وعلى هذا الأساس فان مقدر

بيز يعتمد على دالة الكثافة الاحتمالية اللاحقة (Posterior dinesty function) ويرمز لها (p.d.f) [٩] الناتجة من دمج دالة الكثافة الاحتمالية السابقة للمعلمات مع دالة الخسارة المستخدمة . هذه التقديرات يمكن الحصول عليها من خلال تقليل [(risk function)٥]، ويرمز لها

$$L = L(\hat{\theta}, \theta)$$

أسلوب بايز في عملية التقدير: Bayes's Approach in the Estimation

يعتمد أسلوب بيز في التقدير بشكل عام على توظيف معلومات مسبقة (prior Information) حول المعلمات المجهولة $\theta = [\theta_1, \theta_2 \ldots \theta_p]$ المطلوب تقديرها على اعتبار أن هذه المعلمات متغيرات عشوائية، وليست كميات ثابتة يمكن وصفها على شكل توزيع احتمالي يعرف بدالة الكثافة الاحتمالية السابقة (Proir p.d.f) ، وهذه المعلومات معروفة من بيانات وتجارب سابقة أو من النظرية التي تحكم تلك الظواهر. [٥]

ومن الجدير بالذكر أن هناك عدة أنواع من دول الخسارة، والنوع الأكثر شيوعاً واستخداماً هو دالة الخسارة التربيعية (Quadratic Loss Function) [١٠] والتي تفترض وجود متجه بدالة خسارة مربع الخطأ الموزون النسبي (Weighted Squared Error Loss Function) والتي يعبر عنها بالشكل الآتي: [٨]

$$L(\hat{\theta}, \theta) = (\hat{\theta} - \theta)' C (\hat{\theta} - \theta)$$

أما دالة التوزيع الحالي لمشاهدات العينة الحالية قيد البحث يكون فيها قيمة المتغيرات العشوائية (y) لهذه المشاهدات دالة توزيعية تعتمد على θ، ويرمز لها P(y/θ) تسمى دالة الامكان (Likelihood Function)ويعتمد تقدير بيز لهذه المعلمات على دالة الكثافة الاحتمالية اللاحقة (Posterior p.d.f) والتي يمكن الحصول عليها من خلال دمج دالة الكثافة الاحتمالية الأولية للمعلمات مع دالة

الإمكـــان للمشـــاهدات ،حيـــث تعرف دالـة الاحتماليـة اللاحقـة للمعلمـــة (θ) بالدالـــة الشـــرطية لمجـــال المعلمـــات (θ) بوجـــود معلمات العينة الحالية .
لتوضيح أسلوب بيز في التقدير توضـــيحاً عامـــاً بعـــد أن يـــتم

أسلوب كرنل :

إن تمهيد Kernel يعطينـا طريقـة مبسـطة لإيجـاد هيكـل مجموعـات مـن البيانـات دون الحاجـة إلـى نمـوذج معلمـي، وبـــالنظر لوجـــود معوقـــات لاستخدام نموذج معلمي لتقدير الدالـــة m حيـث m دالـــة فـــي المتغيـرات التفسـيرية فالتقييـد للدالـة(m) فيمـا يخـص نمـوذج الأخير يعني إن تكون m بعض الأحيان صارمة، مثلا أن تكون دوريـــة (Periodic)، قطـــع مكـــــافئ (Parabolic)، أو رتيبـــة (Monotone) وكـــل واحـدة مـن الخصـائص أعـلاه تكون قيوداً تعيق جعل التقدير يلائم دالـة الانحدار الصحيحة ، كمـــا إن اســـتخدام النمـــوذج المعلمـي سـيؤدي إلـى خطـورة الوصـول إلـى استنتاجـات غيـر صحيحـة فـي تحليـل الانحدار، والجدير بالإشارة إلى إن تقدير الانحـدار اللامعلمـي باسـتخدام أســـلوب Kernel يصـــف أن الحصـــول علـــى دالـــة الكثافـــة الاحتماليـــة اللاحقـــة لمتجـــه معلمات (θ) يتم تحديد ما يعرف بدالـــة الخســـارة (Loss Function) [٥] التي يرمز لها عادة $L = L(\hat{\theta}, \theta)$

يكون لدينا عـائلات اكثر مرونـة مـن المنحنيـات للاختيـار فيمـا بينهـا، لكـن زيـادة المرونـة لـه كلفته إذ يقود إلى عدة تساؤلات منهـــا مـــا هـــي الخصـــائص الإحصـــائية لتقـــدير انحـــدار Kernel، ومـا هـو تـأثير الشـكل (Shape) والقيـــــــــــاس (Scaling) لدالـــة Kernel علـى التقـدير، وكيـف تسـتخدم ممهدات Kernel لعمل فترة ثقة وليس مجرد تقدير نقطي، ومـا هـو تـأثير عـدم الاسـتقلالية فـي البيانـات علـى تقـدير انحـدار Kernel .

امـــا أســـلوب Kernel المبسـط هـو لتمثيـل سلسلـة الـوزن، فيوصـف دالـة الـوزن $w_{ni}(x)$ بواسـطة دالـة الكثافـة مع معلمة قياس التي تعدل حجم وشـكل الأوزان القريبـة مـن X، والشائع هو الإشارة لشكل الدالـة بـ K (Kernel) والتـي تختـار لتكـون دالـة كثافـة احتماليـة

تكاملها يساوي واحداً، علماً أن الدالة هي حقيقية محددة مستمرة ومتماثلة حول الصفر.وتعرف سلسلة الوزن لتقديرات Kernel ذات البعد الواحد [٧] علما أن شكل الوزن تسمى دالة Kernel وعرض الحزمة h (Bandwidth) أو تسمى (Window Width)، مع العلم إن أوزان (Kernel) $[w_{hi}(x)]$ تعتمد على قيم المشاهدات X خلال تقدير الكثافة.

الانحدار اللامعلمي : Non –parametric regression

تطورت نظرية وأساليب التقدير خلال السنوات الأخيرة ،حيث إن التقدير المعلمي البحث في تقدير منحني الانحدار لا يتوافق مع الحاجة إلى المرونة في تحليل البيانات ،كما إن تطور أجهزة الحاسبه من الناجيتين المادية والبر مجيه (Hardwer -sotwear) أدى إلى تطوير الجانب النظري لحساب المقدرات المعلمية ، وقد أصبح من الممكن تقدير الانحدار اللامعلمي في علوم الحاسبات والهندسة البيولوجية والإحصاء والرياضيات ، ويهدف منحني الانحدار لتوضيح العلاقة العامة بين المتغير التفسيري x والمتغير الاستجابة y ولتقدير دالة m .فان نموذج الانحدار يفترض بأن دالة m معلوم ما عدا بعض المعالم النهائية . [٦ ويكون الاختلاف بين طريقتي الانحدار المعلمي والانحدار اللامعلمي في درجة دالة m الحاصل عليها والبيانات ، حيث يتم في نموذج الانحدار اللامعلمي يتم اختيار مدى دالة مناسب يلائم الدالة m ، وهذا الاختبار يتعلق بخصائص النموذج والذي يفترض أن تكون في دالة الانحدار ، ويستفاد من البيانات للحصول على معلومات عن الدالة m، وعندما تكون المعلومات لدينا معلومات قليلة عن دالة الانحدار وأن حجم البيانات للدالة m تكون من البيانات الأكثر استخداما ، فان تقنيات الانحدار اللامعلمي تكون اكثر ملائمة لمسائل الاستدلال حيث المعرفة المتوفرة عن الدالة m تكون ذات طبيعة مفيدة ، ومن نتائج تحليل الانحدار اللامعلمي هو منحني يطابق مجموعة البيانات .

ومن المعروف أن تقنيتي الانحدار اللامعلمي

والانحدار المعلمي تمثلان طريقتين مختلفتين لمسائل تحليل الانحدار ، فالأساليب المعلمية تتطلب حد أدنى من المعلومات عن الدالة m تجعله مقيدا في البيانات الخاصة بدالة الانحدار . [٦]

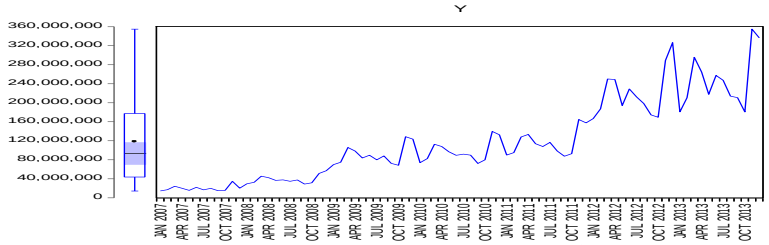
علما أن الانحدار اللامعلمي التي تعمل على المجرب لتزويد معلومات عن الدالة m والحصول على البيانات ، التي تفسر الشكل الفعلي لمنحني الانحدار، ويقدم الانحدار اللامعلمي أدوات مرنه لتحليل علاقات الانحدار ويكون شكل لهذه النماذج كما يلي .

$$Yi = m(xi) + ei$$

<u>الجانب التطبيقي</u> :

ان القراءات المسجلة زمنيا قد تأخذ نمطا معينا قد يكون تصاعديا أو تنازليا أو ثابتا، لذا فان التنبؤ المستقبلي يتطلب أخذ هذا النمط بعين الاعتبار، كما أن استخدام عدد مناسب من القراءات يعطى نموذج مناسب يكون أكثر دقة خاصة عند استخدام أسلوب السلسلة الزمنية حيث يتم التنبؤ بالظاهرة مستقبلاً من البيانات المسجلة سابقا دون الحاجة للحصول على بيانات للمتغيرات الكثيرة التي تعتمد عليها الظاهرة، ولصعوبة تمثيل وحصر هذه المتغيرات خاصة وأن هذه النماذج تعطى وزنا أكثر للقراءات القريبة، ويقل هذا الوزن كلما ابتعدنا عن القراءات المسجلة حاليا وهذا أكثر منطقية. [١]

ويوضح شكل (١) المنحنى التاريخي لسلسلة كمية الاستهلاك الشهري من الكهرباء بالعراق (Y) خلال الفترة من يناير ٢٠٠٧ حتى ديسمبر ٢٠١٣ ومنه يتضح إن هناك اتجاه عام تزايد خلال الفترة حسب المتاح من بيانات لكل سلسلة، مع عدم وجود قيم متطرفة أو شاذة بالسلسلة.



شكل (١)النحنى التاريخي للبيانات

نمودج الانحدار المقترح لتفسير العلاقة :

فى ضوء ما سبق سيتم اختبار علاقة (تأثير)المؤثرات المستقلة الأربع هي عدد المشتركين، درجة الحرارة، الرطوبة النسبية، كمية المطر $(X_4، X_3، X_2، X_1)$ على التابع (Y) لاختيار أفضل نموذج يفسر العلاقة، عند تطبيق أسلوب الانحدار التدريجي على بيانات الدراسة والمتضمنة ٤ متغيرات ومن النتائج الواردة بالجدول التالي وعند قيمة (α=٠.٠٥) كمعيار لدخول المتغيرات في نموذج الانحدار التدريجي، و(α=0.10) كمعيار لخروج المتغيرات من النموذج أظهرت النتائج تكوين نموذجين وقد تبين أيضا أن هذه النماذج كانت ذات دلالة معنوية.وفيما يل يجدول ملخص تحليل الانحدار لتحديد العوامل المستقله على التابع .

جدول (١) ملخص نتائج تحليل الانحدار المتعدد لتحديد تأثير العوامل المستقلة علي التابع

المتغيرات المستقلة	المعالم غير المعيارية		المعالم المعيارية	اختبار ت	المعنوية	معامل التحديد %	معامل تضخم التباين	مؤشر الشرط 7.25
	التقدير	الخطأ المعياري						
الثابت	-70983835.493	11067521.839		-6.414	.000			
X_2	1560892.479	60755.080	.995	25.692	.000	.856	1.133	
X_3	1797869.821	342948.843	.203	5.242	.000	.043	1.133	

يتضح من الجدول (٣-١)أن المتغيرات المعنوية تساهم بنحو 89% من التباين في التابع، ولا تساهم باقي العوامل في تفسير التباين في التابع، وترجع عدم مساهمه العوامل الأخرى نظرا لان معاملات انحدار هذه المتغيرات المستقلة علي المتغير التابع ليست ذات دلالة إحصائية عند مستوي ٥%.

بناء نموذج الانحدار اللامعلمي بطريقة كرنل:

يمكـن إيجـاد منحنـى الانحـدار اللامعلمـي مـن خـلال المعادلـة الأساسية:

$$y_i = f(x'_i) + \varepsilon_i$$

$$= f\left(x_{i1}، x_{i2}، \dots، x_{ik}\right) + \varepsilon_i \quad \dots (4-3)$$

مــن الناحيــة النظريــة يمكن لهذه الصيغة إيجاد منحنى الانحــدار اللامعلمــي متعــدد الحدود لكن من الناحية العملية هناك صعوبات كثيرة للحصول على المنحنى اللامعلمي متعدد الحدود، واهم هذه الصعوبات:

عندما نحدد النقطة x'_0 باعتبارها بؤرة ونقطة الالتقاء للقيم والبيانات مع المحور y ثم مـــــــــــا يليهـــــــــــا

$$x'_0 = (x_{01}، x_{02}، \dots، x_{0k})$$

عند دلك يمكن تطبيق المعادلة التقليدية التالية :

$$D\left(x_i، x_0\right) = \sqrt{\sum_{j=1}^{k} (z_{ij} - z_{0j})^2} \dots (2-3)$$

ويمثل z_i : المؤشرات المعيارية للبيانات في نموذج الانحدار.

وسنجد في الصيغة:

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}$$

ان $\bar{x}_j$ تدل على: الانحراف المعياري وسط المتغيرات .

عند حساب المعادلة :

$$w_i = W\left[\frac{D(x_i، x_0)}{h}\right]$$

بحيث تصبح قيمة $W(.)$: هي قيمة الوزن الخاص بالبيانات في المعادلة .

ويمثل الرمز h : حجم اتساع النطاق او القطاع في منحنى الانحدار .

عند تمثيل القيم والنقاط على منحنى y بالنسبة للقيم والنقاط على المنحنى x سيأخذ ذلك التمثيل شكل المنحنى الخطي للمعادلة:

$$y_i = \alpha + b_1(x_{i1} - x_{01}) + b_2(x_{i2} - x_{02}) + \cdots + b_k(x_{ik} - x_{0k}) + e_i \quad \ldots (6-3)$$

حيث: $\hat{y}_0 = \alpha$

الخصائص الإحصائية لنموذج الانحدار اللامعلمى كرنل:

مقدر خطي :

$$\hat{\mu}_{y|x_1,\ldots,x_p} = \frac{1}{n}\sum_{i=1}^{n} L_i(x) y_i$$

حيث:

$$L_i(x) = \frac{\prod_{j=1}^{p} \frac{1}{h_j} k\left(\frac{x_j - x_{ij}}{h_j}\right)}{n^{-1} \sum_{i=1}^{n} \prod_{j=1}^{p} \frac{1}{h_j} k\left(\frac{x_j - x_{ij}}{h_j}\right)}$$

مقدر متحيز:

$$\mathrm{Bias}\left(\hat{\mu}_{y|\underline{x}}\right) \approx \frac{1}{2}\sigma_K^2 \mathrm{tr}\left\{H^T H_\mu H\right\} + \frac{\sigma_K^2}{f(\underline{x})}\left\{\nabla_\mu^T H H^T \nabla_x\right\}$$

حيث: H : مصفوفة اتساع الحيز ، وهي مصفوفة قطرية.

H_μ :مصفوفة المشتقات الجزئية الثانية للدالة $\hat{\mu}_{y|x_1,\ldots,x_p}$.

∇: التفاضل أو المشتقة الأولى.

التباين :

$$\mathrm{var}\left(\hat{\mu}_{y|\underline{x}}\right) = \frac{1}{n.\det(H)} \cdot \frac{\sigma^2(x)}{f(\underline{x})} \int k^2(u)du$$

ويلاحظ وجود علاقة عكسية بين تباين $\hat{\mu}_{y|\underline{x}}$ واتساع الحيز H فمع القيم الصغيرة لاتساع الحيز يزداد التباين ومع القيم الكبيرة لاتساع الحيز يقل

التباين ، بالتالي تظهر مشكلة Bias-variance Trade-off نظرا لأن القيم الصغيرة لاتساع الحيز تؤدي إلى تقليل التحيز و في نفس الوقت زيادة التباين، وعلى العكس استخدام قيم كبيرة لاتساع الحيز تؤدي إلى تقليل التباين لكنها تؤدي إلى زيادة التحيز.

ولتقدير قيمة التابع المتوقعة يتم أتباع الأتي:

أولا : تحديد دالة النواة المستخدمة في التقدير:

$$|u| \leq 1 \ldots (3-3)$$

وهنا يتم استخدام دالة Epanechnikov لأنها دالة كثافة احتمال متماثلة حيث:

$$k(u) = \frac{3}{4}(1-u^2)$$

ثانيا: تحديد قيم اتساع الحيز h_j حيث j=1،2،3:

$$h_1 = \frac{1}{\sqrt[3]{n}}$$

$$h_2 = \frac{1}{\sqrt[5]{n}}$$

$$h_3 = \frac{1}{\sqrt[7]{n}}$$

ثالثا : التعويض:

تحديد قيم x_j التي يراد التقدير عندها. أي تحديد قيم من المشاهدات المقدرة.

التعويض للحصول على التابع المقدر وذلك بالتعويض في دالة الانحدار المقدرة التالية :

$$\hat{y} = \frac{\sum_{i=1}^{n}[\prod_{j=1}^{p}\frac{1}{h_j}k(\frac{x_j - x_{ij}}{h_j})]y_i}{\sum_{i=1}^{n}[\prod_{j=1}^{p}\frac{1}{h_j}k(\frac{x_j - x_{ij}}{h_j})]}$$

كالتالي:

التعويض عن n=13، Epanechnikovأي:
p=4 وبالتعويض عن بدالة

$$k_1(.)،k_2(.)،k_3(.) = k(u) = \frac{3}{4}(1-u^2)$$

التعويض بقيم اتساع الحيز التي تم الحصول عليها في ثانيا فيصبح تركيز الملوث كما يلي:

$$= \frac{\sum_{i=1}^{5}\left\{\left[\frac{3}{4(0.425)}\left[1-\left(\frac{x_1-x_{i1}}{0.425}\right)^2\right]\right]\left[\frac{3}{4(0.599)}\left[1-\left(\frac{x_2-x_{i2}}{0.599}\right)^2\right]\right]\left[\frac{3}{4(0.693)}\left[1-\left(\frac{x_3-x_{i3}}{0.693}\right)^2\right]\right]\right\}y_i}{\sum_{i=1}^{5}\left\{\left[\frac{3}{4(0.425)}\left[1-\left(\frac{x_1-x_{i1}}{0.425}\right)^2\right]\right]\left[\frac{3}{4(0.599)}\left[1-\left(\frac{x_2-x_{i2}}{0.599}\right)^2\right]\right]\left[\frac{3}{4(0.693)}\left[1-\left(\frac{x_3-x_{i3}}{0.693}\right)^2\right]\right]\right\}}$$

مرحلة الفحص التمهيدي للبيانات:

تعتبر مرحلة الفحص التمهيدي للبيانات Stage of Preliminary Examination of the Data مرحلة أساسية وهامة في بناء النموذج الملائم للبيانات، حيث تبدأ بالرسم البياني للعلاقة بين متغير الاستجابة والمتغيرات التفسيرية لتحديد مدى ملائمة طبيعية البيانات لتمثيلها باستخدام النماذج المعممة المضافة من أجل تحديد معاملات دوال التمهيد اللامعلمية في النموذج المقترح من خلال مجموعة من نقاط البيانات (المبعثرة)، وهناك العديد من الطرق التي يمكن استخدامها لتقدير دوال التمهيد اللامعلمية الموجودة ($S_1, S_2, ..., S_j$) بالنماذج المعممة المضافة وأهمها: [٢]

طريقة كرنل Kernel:

تعتمد الطريقة على حساب معاملات دوال التمهيد اللامعلمية باستخدام الوسط المرجح لجميع قيم متغير الاستجابة Y المتعلقة بالمتغير التفسيري X_i مستخدماً الصيغة التالية:

$$\hat{S} = \frac{1}{|N_i|}\sum_{j \in N_i} \hat{y}_i$$

where,

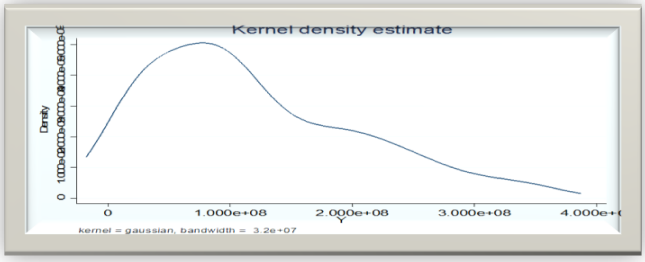
$$\sum_{j=1}^{n} w_{ij} = 1 ,$$

$$\hat{y}_i = \sum_{j=1}^{n} w_{ij}\, y_i .$$

وتعطى هذه الطريقة تقديرات لمعاملات دوال التمهيد اللامعلمية أقل تحيزاً من الطرق السابقة، حيث يفضل استخدام هذه الطريقة عندما تتعلق البيانات بأقل من ٣ متغيرات تفسيرية، حيث أنه بزيادة عدد المتغيرات التفسيرية التي تؤثر على متغير الاستجابة تكون التقديرات متحيزة.

تطبيق نموذج الانحدار اللامعلمى كرنل:

هنا نتعرض لدراسة بناء نموذج الانحدار اللامعلمى NPRM (كرنل)، وسيتم توفيق النموذج باستخدام الحزمة الإحصائية لبرنامج "STATA" وبالتطبيق تظهر نتائجه فيما يلي:



شكل (٢) تقدير دالة كثافة كرنل

ويعرض جدول (٢) مقاييس الجودة لنموذج الانحدار اللامعلمى كرنل وملائمته لتوفيق للبيانات باستخدام المعايير الإحصائية لقياس قدرة النموذج.

جدول (٢) المعايير الإحصائية لقدرة نموذج الانحدار اللامعلمى كرنل

الجذر ألتربيعي لخطأ التنبؤ (RMSE)	المتوسط النسبي لخطأ التنبؤ المطلق (MAPE)	متوسط القيمة المطلقة للأخطاء (MAE)	معامل ثيل	معامل التحديد
28018068	26.02	20887493	0.097	0.89

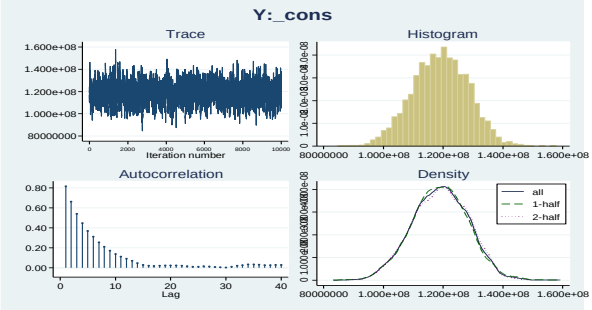
تطبيق نموذج الانحدار البيزى:

هنا نتعرض لدراسة بناء نموذج الانحدار البيزي، وسيتم توفيق النموذج باستخدام الحزمة الإحصـــــــــــــــــائية لبرنـــــامج "STATA" وبـالتطبيق تظهـر نتائجـه فيمـا يلي:

```
Burn-in ...
Simulation ...

Model summary
------------------------------------------------------------------------------
Likelihood :
  Y ~ normal({Y:_cons},7.41e+15

Prior :
  (Y:_cons} ~ 1 (flat)
------------------------------------------------------------------------------

Bayesian normal regression                      MCMC iterations  =
12,500
Random-walk Metropolis-Hastings sampling        Burn-in          =
2,500
                                                MCMC sample size =  10,000
                                                Number of obs    =  84
                                                Acceptance rate  =  .7496
Log marginal likelihood = -1636.4689                             Efficiency
=  .1016

------------------------------------------------------------------------------
                             |         Equal-tailed
           Y |      Mean   Std. Dev.     MCSE     Median  [95% Cred. Interval]
-------------+----------------------------------------------------------------
       cons |  1.19e+08   9327994    292588   1.19e+08   1.00e+08   1.36e+08
------------------------------------------------------------------------------
```

وتظهـر النتـائج بتبعيـة السلسلة للتوزيع الطبيعي بتباين مقدر 7.41e+15 وبمعدل دقة يبلـــغ 0.75 وكفـــاءة 0.10، وبمتوسط ثابت بلغ 1.19e+08 وبـانحراف 9327994 ووسـيط 1.19e+08 ، كمـا يظهر الشكل التالي معايير جودة النمـــــــــــــــــــــوذج

.



شكل (٨) معايير جودة نموذج الانحدار البيزى

ويعرض جدول (٤) مقاييس الجودة لنموذج البيزي وملائمته لتوفيق للبيانات باستخدام المعايير الإحصائية لقياس قدرة النموذج علي التنبؤ:

جدول (٤) المعايير الإحصائية لقدرة نموذج الانحدار البيزي التنبؤية

الجذر التربيعي لخطأ التنبؤ (RMSE)	المتوسط النسبي لخطأ التنبؤ المطلق (MAPE)	متوسط القيمة المطلقة للأخطاء (MAE)	معامل ثيل	معامل التحديد
8831139.4	5709612.4	5.391	0.030	0.99

مرحلة الفحوص التشخيصية (مقارنة وتقييم النماذج المختلفة):

بمجرد اختيار النماذج وتقدير معالمها وبعد التحقق من إمكانية عدم تحسينها وتطويرها وتبين ملائمتها لطبيعة البيانات، تتم مقارنة تقديرات النماذج المختلفة، نموذج الانحدار المعلمي، نموذج الانحدار اللامعلمي، والنماذج المعممة المضافة، ولتقييم قدرتها على التنبؤ خلال فترة الدراسة، وكذلك لمعرفة أيها أكثر ملاءمة لتتناسب مع مثل هذا النوع من البيانات فقد تم استخدام بعض المعايير الإحصائية المناسبة لقياس قدرة النموذج على التنبؤ، ومن أهم تلك المعايير.

١. معامل التحديد **(CD)**.
٢. جذر متوسط مربعات الخطأ **(RMSE)**.
٣. المتوسط النسبي للخطأ المطلق **(MAPE)**.
٤. معامل ثيل **(TC)**.
٥. متوسط القيمة المطلقة للخطأ (MAE).

ويوضح جدول (٥) التالي القيم المقدرة لتلك المعايير.

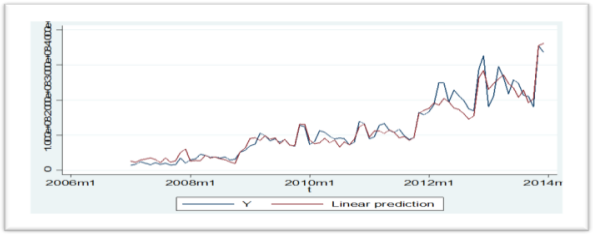
جدول (٥): أهم المعايير الإحصائية المستخدمة للمقارنة بين النماذج

النموذج	CD%	RMSE	MAPE	MAE	TC
الانحدار اللامعلمى كرنل	0.93	22523215.9	13160986.98	9.338	0.078
الانحدار البيزى	0.99	8831139.4	5709612.4	5.391	0.030

وباستعراض نتائج الجدول السابق يتضح ارتفاع قيمة معامل التحديد لنموذج الانحدار البيزى يليها نموذجي الانحدار اللامعلمى كرنل واريما، مما يرجح الانحدار البيزى لتمثيل البيانات، وكذلك انخفاض قيمة معامل ثيل للنموذج وهو بذلك يقترب من الصفر مما يؤكد دقة تلك النماذج في التنبؤ في مرحلتي التقدير والتنبؤ معاً، وأيضاً معياري الجذر التربيعى والنسبة RMSEلخطأ التنبؤ المطلقة لمتوسط الأخطاء MAPE يرجحا الانحدار البيزى و الانحدار اللامعلمى كرنل لاستخدامها في التنبؤ.

ومن العرض السابق يتضح أن نموذج الانحدار البيزى يعد النموذج الأكفأ من بين النماذج التي تفسر العلاقة، وذلك وفقاً للمعايير السابقة والذي يرجحه لتمثيل البيانات ويؤكد دقة النموذج في التنبؤ في مرحلتي التقدير والتنبؤ

ويظهر الشكل التالي المقارنة بين تقديرات نموذج الانحدار البيزى خلال الفترة المتاحة.



شكل (٩) المقارنة بين الفعلي وتقديرات نموذج الانحدار البيزى خلال فترة الدراسة

ويمكن الاستفادة من النماذج المقدرة والتنبؤ بها لعام ٢٠١٤ كما يلي:

جدول (٦): التوقعات المستقبلية لكل من الانحدار البيزى والانحدار اللامعلمى كرنل

EK	EB	Month
152814000	239381530	Jan-14
175627000	265104517	Feb-14
241240000	340917603	Mar-14
218948000	318870894	Apr-14
190667000	273173281	May-14
205063000	307809029	Jun-14
215622000	297676578	Jul-14
170829000	270200475	Aug-14
177607000	261861598	Sep-14
173287000	240040146	Oct-14
264413000	393626059	Nov-14
250894000	387775094	Dec-14

نتائج البحث :

١. أثبتت نتائج الدراسة لنموذج الانحدار البيزي ، ارتفاع قيمة معامل التحديد، لذلك يعتبر

تمثيل البيانات الأفضل والأدق ،حيث بلغت نسبته (0.99) في حين بلغت نسبة الانحدار اللامعلمي كرنل (0.93) لذلك يعد أسلوب بيز أفضل في هذه الدراسة .

٢. توصل البحث إلى انه عند انخفاض قيمة معامل ثيل ، واقترابه من الصفر ، يكون الانحدار البيزي أكثر دقة في مرحلتي التقدير والتنبؤ معاً.

٣. بالنظر للكفاءة العالية التي أبداها أسلوب بيز للانحدار اللامعلمي تجاه القيم الشاذة ، حيث أعطى نتائج فعالة ، ومقاومة للقيم الشاذة.

٤. ارتفاع قيمة معامل التحديد لنموذج الانحدار البيزي ، ويليه النموذج اللامعلمي كرنل مما يؤكد أن تمثيل البيانات لانحدار بيز هو الأفضل.

توصيات البحث :

١- يوصي البحث العلمي الاهتمام بالمتغيرات التي تؤثر في كمية استهلاك الطاقة الكهربائية.

٢- الاعتماد والاستمرار في استخدام النماذج المتنوعة لحل المشكلة الاستهلاك، وكذلك المرونة في استخدام الأسلوب إحصائي الافضل.

٣- يوصي البحث باستخدام الانحدار اللامعلمي لاهميتة التطبيقية في دراسة لتمييز الأصوات، والهندسة في السطوح .

المراجع

١- **العباسي ،عبد الحميد محمد (٢٠١٠) ، التحليل الإحصائي المتقدم باستخدام spss** ، معهد الدراسات والبحوث الإحصائية –جامعة ألقاهره – مصر

٢- **العباسي ،عبد الحميد محمد (٢٠١١)، التحليل الحديث للسلسلة الزمنية باستخدام eviwes** ، معهد الدراسات والبحوث الإحصائية – جامعة القاهرة – مصر

٣- **القزاز، قتيبة نبيل نايف، (٢٠٠٧)، مقارنة أساليب بيز الحصين مع طرائق أخرى لتقدير معالم أنموذج الانحدار الخطي المتعدد في حالة**

البيانات غير التامة"؛ أطروحة دكتوراه فلسفة في الإحصاء، كلية الإدارة والاقتصاد، جامعة بغداد.

٤- خلود يوسف، (٢٠٠٤) مقارنة أساليب بيز مع طرائق أخرى لتقدير منحنى الانحدار اللامعلمي" أطروحة دكتوراه إحصاء / كلية الإدارة والاقتصاد / جامعة بغداد .

٥- مناف يوسف (٢٠٠٠) مقارنة مقدرات Kernel اللامعلمية لتقدير دوال الانحدار" رسالة ماجستير في الإحصاء / كلية الإدارة والاقتصاد / جامعة بغداد .

6-Cao.R(2002) "**Asltor introduction to nonparametric curve estimation**" Universidad decoruna spain .pp1-4.

7-Hoti, F., (2001), " ***Kernel Regression via Binned Data*****", Academic** Dissertation for the Degree of Licentiate of Philosophy, 1-55, http://www.rni.helsinki.fi/˜ fjh/lisuri.pdf.

8-Zellner, A, (2003);"Some a spects of the History of Bayesian Information proressing" H.G.B. Alexander Recerch foundation

9-Zellner, A., (1971) "**An Introduction to Bayesian Inference in** Econometrics" John Wiley and sons, Inc.

10- Mood, A.M., Greybill, F.A. & Boes, D.c, (1985);"Introduction to theory of statistics". Third edition Mc Graw . Hill, Inc.