Warith Scientific Journal

مقارنة بعض طرائق تقدير إنموذج الانحدار اللامعلمي(مع التطبيق)
Comparison of some methods for estimating a nonparametric regression model (with application)

علياء محمد جعفر غضبان
Aliaa Mohammed jaafar ghadban
alaia.jaafar2201m@coadec.uobaghdad.edu.iq
جامعة بغداد/كلية الادارة والاقتصاد

أ.د صباح منفي رضا
Sabah Manfi Rada
drsabah@coadec.uobaghdad.edu.iq
جامعة بغداد/كلية الادارة والاقتصاد

المستخلص:

يهدف هذا البحث إلى تقدير نموذج انحدار لامعلمي باستخدام ثلاث طرائق لامعلمية هي: طريقة ناداريا واتسن(NW) ، طريقة الجار الأقرب(KNN) ، والطريقة الشرائحية (SS) وتم اعتماد أسلوب المحاكاة لتوليد بيانات تماثل البيانات الحقيقية الخاصة بنهر الفرات، حيث اعتُبر متغير TDS كمتغير تابع، والمتغيرات Cl و SO_4وTH كمتغيرات مستقلة تم رصدها من 22 محطة موزعة على ثماني محافظات. وأظهرت نتائج المقارنة باستخدام معيار MSE تفوق طريقة SS من حيث الدقة والكفاءة، مما جعلها تُعتمد في الجانب التطبيقي للدراسة.

الكلمات المفتاحية: الانحدار اللامعلمي، طريقة NW، طريقة KNN، طريقة SS، معيار MSE.

Abstract :

This study aims to estimate a nonparametric regression model using three nonparametric methods: Nadaraya-Watson (NW), K-Nearest Neighbor (KNN), and Spline Smoothing (SS). A simulation approach was used to generate data reflecting real measurements from the Euphrates River, where TDS was the dependent variable and Cl, SO_4, and TH were independent variables from 22 monitoring stations across eight provinces. Results based on MSE showed that the SS method outperformed the others in accuracy and efficiency, and was thus adopted in the applied part of the study.

Keywords: Nonparametric Regression, NW, KNN, SS, MSE.

1- المقدمة

يُعد تحليل الانحدار أحد الفروع المهمة في علم الإحصاء، ويُعنى بدراسة وتقدير العلاقات بين المتغير التابع والمتغيرات المستقلة، بهدف تفسير تلك العلاقات أو التنبؤ بقيم مستقبلية. وتُعتبر النماذج الإحصائية أدوات جوهرية لفهم واستيعاب العلاقات الكامنة بين المتغيرات، خاصة في ظل التطور الكبير في مجالات تحليل البيانات، حيث برزت النماذج اللامعلمية كأدوات أكثر مرونة وأهمية مقارنة بالنماذج المعلمية التي تعتمد على افتراض شكل معين للدالة.

تواجه العديد من الدراسات ولا سيما الاقتصادية والبيئية منها تحديات كبيرة عند التعامل مع بيانات معقدة البنية كبيانات جودة المياه والتي تتصف بالاعتمادية المكانية والزمانية، إذ تكون المشاهدات مترابطة وغير مستقلة. ويؤدي تجاهل هذا الاعتماد إلى تقديرات مضللة ويؤثر سلبًا على دقة النماذج الإحصائية التقليدية. لذا أصبح من الضروري اللجوء إلى طرائق تقدير لا تفترض شكلاً معينًا للعلاقة بين المتغيرات، أي الاعتماد على النماذج اللامعلمية، لما توفره من مرونة عالية في تحليل البيانات غير الخطية والمعقدة دون الحاجة إلى فرضيات مسبقة.

ومن هنا، تنبع أهمية هذا البحث من كونه يسهم في تطوير أساليب التقدير الإحصائي من خلال اعتماد نماذج لامعلمية قادرة على التعامل بكفاءة مع البيانات المعتمدة مكانيًا وزمانيًا. وتبرز الأهمية التطبيقية في استخدام بيانات حقيقية تتعلق بنوعية المياه في نهر الفرات، أحد أهم الموارد المائية في العراق، مما يمنح النماذج المقترحة بعدًا عمليًا يمكن أن يُسهم في دعم قرارات إدارة المياه وحماية البيئة. كما تسهم المقارنة بين ثلاث طرائق لامعلمية مختلفة في تقديم تصور علمي واضح حول كفاءة كل طريقة، وتحديد الأنسب منها لتقدير دالة الانحدار اللامعلمي في سياقات بيئية مماثلة.

أما مشكلة البحث فتتمثل في الحاجة إلى إيجاد نماذج تقدير مناسبة للتعامل مع بيانات تعتمد مكانياً وزمانياً، تعجز الطرائق المعلمية التقليدية عن معالجتها بدقة. وفي ضوء هذه المشكلة، جاء هذا البحث ليقدم بدائل تقديرية لامعلمية مرنة ودقيقة.

وعليه، فإن هدف البحث يتمثل في تقدير إنموذج الانحدار اللامعلمي للبيانات المعتمدة مكانياً وزمانياً باستخدام بعض طرائق التقدير اللامعلمية، وهي: طريقة ناداريا واتسن(NW) ، طريقة الجار الأقرب(KNN) ، والطريقة الشرائحية(SS) ، والمقارنة فيما بينها بالاعتماد على بيانات محاكاة. وقد طُبقت هذه الطرائق تجريبياً لتحديد أكثرها دقة وكفاءة، ثم تم اعتماد الطريقة الأفضل في الجانب التطبيقي باستخدام البيانات الحقيقية الخاصة بجودة مياه نهر الفرات، والتي تتأثر بالعوامل الجغرافية والبيئية المحيطة بها على امتداد الزمن.

2- الجانب النظري:

2-1 نماذج الانحدار اللامعلمي Non-parametic Regression models [8][7][2]

ان نماذج الانحدار اللامعلمي هي نماذج محببة ومرغوب بها من قبل الباحثين لانها اقل صرامة من نماذج الانحدار المعلمي وذلك بان الباحث يهتم بالوصف الشامل للعلاقة ولا يبحث عن التفاصيل الدقيقة وغالباً ماتكون البيانات الحقيقة لا تتصف بالمثالية مما جعل نماذج الانحدار اللامعلمي اكثر مرونة وقابلية في التعامل مع هذه البيانات ويمكن كتابة الصيغة العامة بالشكل الاتي :

$$Yi=m(Xi)+\varepsilon i \qquad ...(1)$$

حيث ان :

Y_i:هو قيمة متغير الاستجابة .

$m(X_i)$: تمثل دالة الانحدار المراد تقديرها .

X_i :تمثل قيمة المشاهدات للمتغير التفسيري (التوضيحي).

ε_i: قيمة الخطأ العشوائي الذي يتوزع توزيعاً طبيعياً بمتوسط يساوي صفر وتباين σ^2.

نلاحظ الانموذج اعلاه انه لا يشمل على معالم وبذلك فأنه لا يحتاج الى تقدير للمعالم وانما يحتاجه هو تقدير دالة الانحدار $m(X_i)$ او تقدير منحنى الانحدار عند النقاط (X1,…,Xn).

2-2 اختيار دالة النواة Kernel:[8][9]

عند أختيار دالة Kernel يجب ان تتوفر فيها الشروط التالية

1 - دالة ذات قيم حقيقية غير سالبة (موجبة) $k(.) > 0$

2- دالة مستمرة .

3- k هي دالة متماثل

وفي ادناه اهم دوال Kernel التي تستعمل في تقدير نماذج الانحدار اللامعلمي :

Function Name	Mathematical Formula	Interval
Gaussian	$k(u) = \frac{1}{\sqrt{2\pi}} e^{\frac{-u^2}{2}}$	$I[-\infty, \infty]$
Epanechnikov	$\frac{3}{4}(1-u^2)$	[-1,1]

2-3 طريقة ناداريا واتسن Nadaraya-watson kernel method (NW) [3][11][13][14] وهومن ابسط مقدرات الانحدار اللامعلمي واكثرها استخداماً وقد اقترح هذا النوع من المقدرات Nadaraya وWatson (1964) لتقدير الانحدار اللامعلمي وتكتب الصيغة كالاتي:

$$\widehat{m}_{NW}(X) = \frac{\sum_{i=1}^{n} k\left(\frac{X-X_i}{h}\right)}{\sum_{i=1}^{n} k\left(\frac{X-X_i}{h}\right)} Y_i \qquad ...(2)$$

حيث ان :

$k(.)$: هي دالة النواة kernel function .

h : معلمة التمهيد (smoothing parametic) h>0 .

وتكتب الصيغة اعلاه بالشكل :

$$\hat{m}_{NW}(X) = 1/n \ (W_{ni}(X_i)Y_i) \qquad ...(3)$$

حيث ان $W_{ni}(X_i)$ هي متسلسلة الاوزان (دالة الوزن) وتكون موجبة وتكتب :

$$W_{ni}(X_i) = \frac{k(\frac{X-X_i`}{h})}{\sum_{i=1}^{n} k(\frac{X-X_i`}{h})} \qquad ...(4)$$

4-2 طريقة الجار الاقرب K-Nearast Neighbor method (KNN) [5][6][12]

وهي من اشهر الطرق اللامعلمية والتي تستخدم لحل مشاكل الانحدار والتصنيف ويعد من الاساليب غير المعقدة . ويعرف بأنه المتوسط المرجح (الموزون) لقيم متغير الاستجابة (Y) المقابلة لقيم X والتي تكون ضمن حيز الجوار K الاكثر قربا" من النقطة X في المسافة الاقليدية والتي عندها نقوم بتقدير الدالة m .

وتكتب صيغة مقدر الجار الاقرب بالشكل التالي:

$$\hat{m}_{\mathrm{k}}(X) = \frac{1}{n}\sum_{i=1}^{n} W_{ni}(X)\, Y_i \qquad ... (5)$$

حيث ان :

$$W_{ni}(X) = \begin{cases} \frac{n}{k} \ ; & i \in gx \\ 0 & O/w \end{cases} \qquad ... (6)$$

gx :هي المجموعة الحاوية على الجيران K الاقرب للنقطة X .

K : يمثل عدد الجيران وتحدد حسب القرب من النقطة X .

و K هي معلمة التمهيد و وتعرف kn بالمسافة التقليدية بين النقاط وتكتب بالشكل التالي:

$$K = Kn = d(X_i, X_j) = \sqrt{\sum_{i=1}^{n} (X_{ik}, X_{jk})^2} \qquad , \qquad i,j = 1,2,\dots,n$$

5-2 الطريقة الشرائحية smothing spline [15] [10][4][1]

وهي من الطرائق الشائعة وكثيرة الاستخدام في مجالات عديدة مثل الرسومات الحاسوبية والمحاكاة والتصميم الهندسي وتقدير الدوال وغيرها من المجالات .

هذه الطريقة تستخدم عندما يقسم خط الانحدار الى عدد من قطع من الخطوط وتكون هذه القطع مفصولة بعقد knots تسمى عقد الشريحة spline knots وهي عبارة عن نقاط ربط توصل وتربط القطع ببعضها وتوجب الشرائح spline شرط قيود

الاستمرارية عند نقاط الربط (العقد) حيث ان خط الانحدار يغير اتجاهه من دون القفز عند العقد الشرائحية لان في الغالب تكون عدم الاستمرارية عند نقطة الاتصال ليس لها معنى وتتطلب تحديد عدد العقد ومواقعها وعموماً فأن عند العقد K يكون اقل من حجم العينة n ويجب ان يكون عدد هذه العقد عدداً صحيحاً وبذلك فأن الخيارات المتاحة والمناسبة لايجاده فتكون محدودة ويتم اعتماد عدد كبير من العقد عند كل نقطة لذلك يكون المنحنى كثير التذبذب ولعلاج هذه المشكلة يتم اضافة (penalty) قيد جزائي على الجزء الغير ممهد لذلك المنحنى وبذلك سوف نحصل على معيار المربعات الصغرى الجزائية penalized least squars :

$$\sum_{i=1}^{n}[Y_i - m(X_i)]^2 + \lambda \int_a^b \left[m^{\mathrm{k}}(X)\right]^2 dx \qquad \dots (7)$$

حيث ان:

$\sum_{i=1}^{n}[Y_i - m(X_i)]^2$ يمثل مجموع مربعات البواقي .

$\lambda \int_a^b \left[m^{\mathrm{k}}(X)\right]^2 dx$ يمثل جزاء الخشونه (Roughness penalty)

λ تمثل معلمة التمهيد (λ) والتي تحدد قيمة حد الجزاء .

m(Xi) هي الدالة المراد تقديره

$$m(X) = X\beta \qquad \dots (8)$$

حيث ان X هي مصفوفة التصميم من الرتبة $np + K + 1$ والمتكونة من دوال الاساس وتكون بالشكل الاتي :

$$X = \begin{bmatrix} 1 & \mathrm{X}_1 & \dots & \mathrm{X}_1^{\mathrm{p}} & (\mathrm{X}_1 - \mathrm{k}_1)_+^{\mathrm{p}} & \dots & (\mathrm{X}_1 - \mathrm{k}_{\mathrm{K}})_+^{\mathrm{p}} \\ 1 & \mathrm{X}_2 & \dots & \mathrm{X}_2^{\mathrm{p}} & (\mathrm{X}_2 - \mathrm{k}_1)_+^{\mathrm{p}} & \dots & (\mathrm{X}_2 - \mathrm{k}_{\mathrm{K}})_+^{\mathrm{p}} \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ 1 & \mathrm{X}_{\mathrm{n}} & \dots & \mathrm{X}_{\mathrm{n}}^{\mathrm{p}} & (\mathrm{X}_{\mathrm{n}} - \mathrm{k}_{\mathrm{K}})_+^{\mathrm{p}} & \dots & (\mathrm{X}_{\mathrm{n}} - \mathrm{k}_{\mathrm{K}})_+^{\mathrm{p}} \end{bmatrix}$$

وان β هي معاملات دوال اساس القطع وتكتب

$$\beta = \left[\beta_0, \beta_1, \cdots, \beta_p, \beta_{p+1}, \cdots, \beta_{p+K}\right]^T$$

ومن خلال تقدير β نتمكن من ايجاد المقدرالشرائحي

$$\hat{\mathrm{m}}_{\mathrm{K}}(\mathrm{X}) = \mathrm{X}\hat{\beta} \qquad \dots (9)$$

$$\hat{\beta} = (X^T X)^{-1} X^T X$$

$$\hat{\mathrm{m}}_{\mathrm{K}}(\mathrm{X}) = \mathrm{X}(X^T X)^{-1} X^T Y \qquad \dots (10)$$

6-2 معيار المقارنة

لاختيار افضل الطرق التي تم دراستها في هذا البحث سوف يتم استخدام متوسط مربعات الخطأ (Mean square Error) (MSE) لغرض المقارنة واختيار افضل مقدر ويكتب بالشكل التالي :

$$MSE = \frac{1}{n}\sum_{i=1}^{n}[Y_i - \hat{m}(Y_i)]^2 \qquad \dots (11)$$

3- الجانب التجريبي

لتطبيق طرائق تقدير الانحدار اللامعلمي والتي تم عرضها في الجانب النظري من هذا البحث والمقارنة بينها من خلال معيار المقارنة (متوسط الخطأ MSE) وسيتم استخدام اسلوب المحاكاة (Simulation) لايجاد افضل طريقة للتقدير وتم استخدام برنامج ماتلاب اصدار(2024) لغرض تنفيذ تجارب المحاكاة وقد تم تكرار كل تجربة 5000 مرة وذلك للحصول على نتائج متسقة.والجدول التالي يوضح نتائج تجارب المحاكاة لايجاد افضل طريقة تقدير لانموذج الانحدار اللامعلمي .

جدول (3-1) يبين Mse لكل طريقة

n	σ	NW	SS	KNN	Best
30	0.1	0.9854	0.8701	1.3362	SS
	0.5	0.9677	0.6337	1.5635	SS
	1	3.1111	1.2002	1.2377	SS
50	0.1	0.8812	0.8064	1.0654	SS
	0.5	0.7290	0.5928	0.9971	SS
	1	1.4280	0.6065	0.8510	SS
100	0.1	0.4730	0.4467	0.8460	SS
	0.5	0.4612	0.2108	0.7287	SS
	1	0.9780	0.3422	0.5412	SS
150	0.1	0.2219	0.1253	0.3256	SS
	0.5	0.0854	0.0089	0.5630	SS
	1	0.3265	0.0341	0.4447	SS

من خلال الجدول اعلاه نلاحظ المقارنة بين الطرائق اللامعلمية (NW-KNN-SS) وتلاحظ بان طريقة (SS) هي الافضل لجميع حجوم العينات ولكل مستويات الانحراف المعياري.

4- الجانب التطبيقي

تم في هذا الجانب اختيار الطريقة الأكثر كفاءة من بين طرائق التقدير التي تم تطبيقها في الجانب التجريبي، وهي طريقة التمهيد الشرائحي (SS)، وذلك بتطبيقها على مجموعة من البيانات الحقيقية المرتبطة بجودة مياه نهر الفرات والتي تم الحصول عليها من وزارة البيئة العراقية. وقد تم اعتماد المتغير(TDS) كمؤشر رئيسي لجودة المياه، والذي يمثل المتغير التابع في النموذج

الإحصائي، بهدف توضيح العلاقة بينه وبين المتغيرات التوضيحية المستقلة الثلاثة (Cl، TH،SO_4) وذلك باستخدام بيانات مأخوذة من عدة محافظات عراقية يمر بها نهر الفرات. وقد تم تحليل هذه البيانات الفعلية باستخدام برنامج MATLAB، لما يوفره من أدوات متقدمة في تنفيذ النماذج اللامعلمية.

وجاء إختيار المتغيرات المذكورة أعلاه بناءاً على دورها الفيزيائي والكيميائي المباشرفي تحديد جودة المياه وتفصيلها كالاتي:

1- العسرة الكلية (Total hardness)(TH)

وتعرف بانها تراكيز ايونات الكالسيوم (Ca) والمغنيسيوم (Mg) الذائبة في الماء والتي لها دوراً مهماً ومباشراً في صلابة المياه وتعد من المؤشرات الكيميائية التي لها تأثير في جودة المياه.

2-**الكبريتات** Sulfates (So_4)

وهي احدى الايونات الرئيسية السالبة الذائبة في المياه الطبيعية وتتكون من خلال تحلل الصخور الرسوبية الغنية بالكبريت او من خلال مصادر الزراعة والصناعة ، وتعتبر الكبريتات من العناصر المهمة والمؤثرة في جودة المياه.

3-**الكلوريد** Chloride (cl)

وهو أحد الايونات السالبة المتواجدة في المياه الطبيعية وهو من العناصر عالية الذوبان ويكون على الاغلب على شكل املاح الكلوريد مثلاً كلوريد الكالسيوم وكلوريد الصوديوم ويتكون (cl) بفعل التحليل الطبيعي للصخور اووبفعل تلوث المياه بمياه الصرف الصحي او نتيجة تسريب الماء المالح.

4-**المواد الصلبة الذائبة الكلية** (TDS) Total Dissolved Solids

وهو المتغير الذي يعد مؤشراً يضم جميع المواد غير العضوية والمواد العضوية الذائبة بالماء كلايونات السالبة كـ (cl) ,So_4 ,HCo_3 والايونات الموجبة ك (Na , Mg , Ca ,K). ويعد TSD مؤشراً مهماً في جودة المياه ويستعمل في تحديد مدى صلاحية المياه ان كان للشرب ام للري.

ويعرض الجدول التالي القيم الجقيقية والتقديرية للأنموذج وفقاً لطريقة (SS) مما يساعد في تقييم كفاءة الطريقة ودقة الانموذج في تمثيل العلاقات بين المتغيرات.

الجدول (3-2)يوضح القيم الحقيقية والتقديرية للانموذج حسب طريقة SS

الطريقة		طريقة SS	
المناطق	ت	القيم الحقيقية Y	القيم التقديرية $\hat{Y}$
الانبار	1	633.93	738.70293
	2	715.64	757.88912
	3	749.56	747.89052
	4	755.78	768.66473
	5	739.03	766.38398

777.72015	742.38	6
830.58991	786.66	7
700.63133	626.54	8
678.82712	637.67	9
681.14496	616.95	10
689.38711	630.32	11
712.58626	658.59	12
735.82752	679.14	13
729.67562	702.92	14
618.29318	572.90	15
605.71308	579.20	16
625.48436	599.20	17
647.42284	602.00	18
630.52184	675.60	19
668.60835	677.50	20
641.76624	598.20	21
617.52450	559.29	22
598.52298	544.99	23
604.49974	544.40	24
613.90598	553.45	25
631.74747	566.47	26
643.58515	580.41	27
649.85134	590.09	28
598.52298	544.99	29
604.49974	544.40	30
649.85134	590.09	31

625.48436	599.20	32
647.42284	602.00	33
630.52184	675.60	34
729.67562	702.92	35
604.49974	544.40	36
649.85134	590.09	37
625.48436	599.20	38
647.42284	602.00	39
630.52184	675.60	40
747.89052	749.56	41
768.66473	755.78	42
642.93520	452.79	43
629.92974	441.08	44
631.31057	445.29	45
633.92618	448.08	46
642.93520	452.79	47
650.81191	461.79	48
650.81191	461.79	49
642.93520	452.79	50
650.81191	461.79	51
661.71936	479.67	52
698.94808	708.28	53
755.63891	570.75	54
778.09023	579.83	55
813.36558	611.75	56
642.93520	452.79	57

629.92974	441.08	58
631.31057	445.29	59
633.92618	448.08	60
642.93520	452.79	61
650.81191	461.79	62
661.71936	479.67	63
548.00923	401.27	64
618.52723	420.28	65
619.91761	424.82	66
650.87524	450.45	67
731.74521	454.18	68
714.79385	491.82	69
745.46737	494.45	70
582.13843	400.88	71
654.25886	438.89	72
655.51029	454.25	73
680.99568	492.38	74
784.77380	516.25	75
750.82485	538.63	76
801.61910	537.38	77
545.23769	409.50	78
587.46591	423.92	79
588.17868	437.92	80
611.03567	446.92	81
624.38840	448.75	82
619.15418	461.67	83

698.19888	505.91	84	
706.85730	522.00	85	
755.63891	570.75	86	
778.09023	579.83	87	
813.36558	611.75	88	
853.89286	645.17	89	
851.27142	660.00	90	
849.81039	662.50	91	
782.31878	936.00	92	بابل
982.56970	1035.09	93	
988.38397	972.70	94	
774.99609	970.24	95	
752.59765	928.26	96	
770.22268	981.58	97	
770.22268	981.58	98	
631.54165	678.40	99	
656.57346	695.70	100	
726.24905	784.68	101	
722.86524	776.53	102	
900.61716	994.79	103	
720.50753	707.26	104	
723.02562	721.15	105	
747.71714	796.16	106	
837.62993	902.25	107	
846.75097	885.46	108	
855.40352	941.37	109	

835.97798	866.77	110	
836.52294	881.18	111	
872.13526	911.34	112	
665.71031	662.23	113	
752.38650	662.32	114	
659.61643	664.58	115	
698.94808	708.28	116	
702.17332	713.56	117	
725.46337	745.69	118	
657.10036	659.91	119	
655.14693	656.51	120	
660.76002	671.88	121	
705.10419	707.14	122	
710.75819	706.09	123	
711.09374	719.65	124	
688.17473	706.42	125	
690.17243	717.87	126	
692.79249	722.23	127	
826.42885	781.95	128	
828.32459	787.16	129	
840.43531	1073.11	130	
788.29448	943.71	131	كربلاء
751.10602	816.54	132	
690.31620	781.50	133	
705.94145	762.50	134	
654.29913	748.00	135	

828.66943	951.00	136	
933.96771	870.50	137	
649.78503	687.26	138	
719.47498	780.40	139	
663.27857	706.25	140	
730.88362	762.90	141	
698.47104	694.46	142	
892.36731	906.92	143	
864.11102	976.65	144	النجف
919.49605	942.29	145	
916.00742	902.75	146	
944.76679	823.49	147	
697.39003	745.00	148	
704.80463	744.00	149	
897.24032	970.21	150	
907.90008	984.29	151	
834.54596	915.75	152	
843.10496	918.21	153	
923.87895	956.01	154	
955.95760	996.89	155	
960.75044	1021.40	156	
1032.58814	1058.45	157	
740.46172	744.89	158	
724.22710	819.94	159	
690.80254	774.83	160	
792.75618	868.75	161	

703.30844	725.85	162	
781.89218	850.50	163	
745.83019	787.09	164	
856.71855	882.25	165	
716.81105	857.32	166	
739.34501	883.82	167	
852.43005	1037.42	168	
878.87147	1061.33	169	
980.69014	952.15	170	القادسية
3283.75042	3004.25	171	
898.30448	773.60	172	
2541.76288	2592.55	173	
812.52384	834.40	174	
811.50659	835.40	175	
785.41600	826.95	176	
2524.92682	2573.80	177	
832.90208	874.50	178	
2516.61464	2424.36	179	
986.66922	1090.15	180	
3062.75616	3143.83	181	
1009.90702	1014.50	182	
2391.62400	3356.50	183	
858.10807	869.61	184	
2326.39587	1963.83	185	
871.96753	919.18	186	
2482.92324	1707.58	187	

899.02755	979.53	188	
2472.52462	2083.00	189	
937.46852	968.71	190	
2502.02027	2057.50	191	
1257.41511	1144.17	192	
2798.56651	2318.67	193	
979.45568	969.42	194	
2110.98196	1950.92	195	
2309.28275	2393.88	196	المثنى
2331.15270	2560.00	197	
2556.42971	2696.32	198	
3020.69303	2739.25	199	
2452.73495	2457.00	200	
2164.25530	2638.00	201	
2414.92487	2706.08	202	
2527.17137	2683.96	203	
2074.19147	2303.23	204	
2154.88686	2331.55	205	
2530.24369	2763.41	206	
2833.41986	3038.02	207	
2092.71616	2714.20	208	
2249.50616	2885.50	209	
1995.94096	2078.61	210	
2182.80580	2206.72	211	
2003.04656	1802.66	212	
2166.94393	2005.28	213	

1824.74499	1515.85	214	
1780.81568	1532.34	215	
1824.55768	1632.92	216	
1817.88918	1647.31	217	
2015.67096	1996.08	218	
2097.44569	2056.36	219	
2013.11595	2041.75	220	
2139.24936	2147.75	221	
2415.05052	2558.08	222	ذي قار
2482.08309	2701.46	223	
2367.18974	2583.67	224	
2901.51442	3039.29	225	
2805.13385	2969.06	226	
4472.15070	4079.83	227	
2433.32487	2820.00	228	
2468.04166	2567.00	229	
2471.46186	2545.00	230	
2438.52224	2701.96	231	
2428.72017	2614.58	232	
2479.90912	2685.13	233	
2520.11439	2683.95	234	
2495.15389	2673.91	235	
2590.62449	2794.27	236	
3983.39728	4107.35	237	
3849.27707	4023.65	238	
3869.56713	4030.05	239	

2712.92308	4167.60	240	
2654.91786	4035.80	241	
2716.63329	3845.40	242	
2170.23172	2248.89	243	
2196.45116	2270.00	244	
2282.54696	2245.11	245	
2790.08337	2742.88	246	
2766.56304	2742.50	247	
2872.91430	2857.83	248	
2605.46107	2339.95	249	
2630.87740	2390.50	250	
2707.50438	2440.80	251	
2741.57185	2432.31	252	
2722.17414	2475.25	253	
2875.61130	2601.50	254	
2887.05225	2977.32	255	
3014.29801	3092.32	256	
2938.39970	3001.91	257	
2844.46823	2962.29	258	
2990.06688	3087.71	259	
2932.07229	2908.71	260	
1567.99023	1713.83	261	البصرة
1803.24602	1942.42	262	
1172.68701	1130.96	263	
1624.46657	1663.33	264	
2451.14267	2439.00	265	

1180.05835	1257.00	266
1223.77971	1301.70	267
3069.94459	3387.20	268
1393.03194	1497.17	269
2446.01252	2677.46	270
1581.55313	1633.82	271
3820.68542	4219.15	272
1445.09689	1681.00	273
3226.99557	3048.10	274
1489.56635	1535.93	275
631.31057	2633.56	276
1611.80915	1852.83	277
9438.73346	8831.75	278
1549.30792	1523.17	279
2771.43454	2817.45	280
1693.07363	1712.33	281
2574.47486	2643.56	282
1452.17629	1460.40	283
2834.03691	2815.60	284
1467.16101	1409.92	285
5616.78783	5222.42	286

تفسير النتائج

أظهرت نتائج التقدير باستخدام طريقة التمهيد الشرائحي (SS) تطابقًا جيدًا في العديد من النقاط بين القيم التقديرية والقيم الحقيقية، مما يشير إلى فعالية هذه الطريقة في تمثيل العلاقة بين المتغيرات دون افتراض شكل دالي معيَّن. وقد أتاحت SS الحصول على منحنيات تقديرية ناعمة تمثل الاتجاه العام للبيانات وتعكس المرونة المطلوبة في التعامل مع العلاقات المعقدة.

وعلى الرغم من ظهور بعض الانحرافات البسيطة في بعض المواضع، لكنه يعتبر ضمن الحدود المقبولة، مما يعكس كفاءة هذه الطريقة في تقدير النماذج اللامعلمية.

5- الاستنتاجات والتوصيات

5-1 الاستنتاجات

1-تمثل الطريقة (SS) أداة فعالة لتقدير النماذج اللامعلمية .

2-تفوقت طريقة (SS) على باقي الطرق من حيث تحقيقها اقل (MSE) مما يعكس كفاءتها العالية في التقدير.

3-في الجانب التطبيقي أظهرت (SS) قدرتها على تمثيل العلاقة بين المتغير التابع (TDS) والمتغيرات التوضيحية (Cl, SO_4, TH) وبدقة عالية بأستخدام البيانات الحقيقية مما يعزز أعتمادها في تحليل بيانات بيئية مماثلة.

5-2 التوصيات

1-يوصى بأتخاذ طريقة(SS) كوسيلة فعالة في تقدير البيانات البيئية المتصفة بألاعتمادية المكانية والزمانية نظراً لدقة وكفائة هذه الطريقة وتفوقها على الطرائق اللامعلميةالاخرى .

2-تشجيع المؤسسات البيئية والبحثية لأعتماد الطريقة (SS) للتعامل مع طبيعة البيانات الحقيقية المعقدة وتحليلها مما له دوراً فعالاً للاسهام في اعتماد الدقة في القرارات البيئية وتوصياتها.

3- التوصية بأجراء دراسات مستقبلية لتطبيق الطريقة الشرائحية (SS) في الميادين الصحية والبيئية المختلفة المتضمنة متغيرات اخرى كالملوحة والعكارة او ملوثات الهواء.

4- يفضل اعتماد النماذج اللامعلمية في الدراسات التي تعتمد على بيانات جودة المياه لقدرة هذه النماذج على احتواء بنية مكانية وزمانية معقدة .

المصادر

1.Araveeporn, A., & Chaikajonwat, T. (2025). The use of spline techniques in nonparametric regression analysis for sequence data with a random walk process. WSEAS Transactions on Systems and Control, 20, 81–91.

2.Armando, L. (1999). A Simple Consistent Non-parametric Estimator for the Regression Function in a Truncated Sample. Dept. of Economics, North Carolina State University.

3. Aydin, D. (2007). A Comparison of the Nonparametric Regression Models using Smoothing Spline and Kernel Regression. World Academy of Science, Engineering and Technology.

4.Bosch, R. J., Ye, Y., & Woodworth, G. G. (1995). A Convergent Algorithm for Quantile Regression with Smoothing Splines.

5. Burba, F., Ferraty, F., & Vieu, P. (2009). k-Nearest Neighbour Method in Functional Nonparametric Regression. Journal of Nonparametric Statistics, 21(4), 453–469.

6.Denis, H., & Yan, L. (2005). Cross-Validation in Nonparametric Regression with Outliers. The Annals of Statistics, 33(5), 2291–2310.

7.Hamoud, Manaf Yousef (2000). "Comparison of Nonparametric Kernel Estimators for Estimating Regression Functions". Master's Thesis, University of Baghdad.

8. Härdle, W. (1994). Applied Nonparametric Regression. Cambridge University Press.

9.Khamo, Kholoud Yousef (2004). "Comparing Bayesian methods with other methods for estimating nonparametric regression curves." PhD thesis, University of Baghdad.

10. Lee, Thomas C.M. (2004). Improved Smoothing Spline Regression by Combining Estimates of Different Smoothness. Statistics & Probability Letters, 67.

11. Nadaraya, E. A. (1964). On Estimating Regression. Theory of Probability and its Applications, 10, 186–190.

12.Srisuradetchai, P., & Suksrikran, S. (2024). Random kernel k-nearest neighbors regression. Computational Statistics & Data Analysis, (In Press).

13.Trottner, L. (2023). Covariate shift in nonparametric regression with Markovian design. Proceedings of a peer-review conference.

14. Watson, G. S. (1964). Smooth Regression Analysis. Sankhya, 26(A):pp.72–359.

15. Wu, Hulin & Zhang, Jin-Ting (2006). Nonparametric Regression Methods for Longitudinal Data Analysis. Wiley Series in Probability and Statistics,John wiley&Sons.