



فرائن للعلوم الاقتصادية والإدارية

KHAZAYIN OF ECONOMIC AND
ADMINISTRATIVE SCIENCES

ISSN: 2960-1363 (Print)

ISSN: 3007-9020 (Online)



Analysis of Ozone Concentration in Wasit Governorate – Iraq Using Fuzzy Nonparametric Regression Estimation Methods

Ola Hadi Sadiq

Statistics Department, College of Administration and Economics, Wasit University, Iraq.

ola.sadeq@uowasit.edu.iq

<https://orcid.org/0009-0003-5373-8228>

ABSTRACT: This article aims to analyze the relationship between temperature and ozone concentration in Wasit Governorate using fuzzy nonparametric regression models, given the nonlinear and ambiguous nature of the environmental data. Three main fuzzy nonparametric regression modeling methods were employed: the Nearest Neighbor (KNN) method, the Nadaraya–Watson method, and the Local Polynomial method. The estimation process relied on the Leave-One-Out (LOOCV) cross-validation method to select the optimal parameters for each model, using the Mean Squared Error (AMSE) and the Mean Integral Squared Error (IMSE) criteria. The response variable was treated as a triangular fuzzy number representing the uncertainty in the environmental measurements, and the Diamond squared distance was used to measure the difference between the estimated and actual fuzzy numbers. The results showed that all models performed acceptably, but the local linear regression model with the Epanechnikov kernel outperformed by achieving the lowest AMSE and IMSE values, reflecting its higher accuracy in representing the nonparametric relationship between temperature and ozone concentration. These results confirm that nonparametric fuzzy models represent environmental phenomena characterized by ambiguity.

Keywords: Nonparametric Regression, fuzzy models, Epanechnikov kernel, Nadaraya–Watson, local polynomial regression, Ozone, Wasit Governorate, IMSE

DOI: [10.69938/Keas.2603015](https://doi.org/10.69938/Keas.2603015)

تحليل تركيز الأوزون في محافظة واسط – العراق باستخدام أساليب تقدير الانحدار الضبابي اللامعلمي

علا هادي صادق الوائلي

قسم الإحصاء، كلية الإدارة والاقتصاد، جامعة واسط، العراق

ola.sadeq@uowasit.edu.iq

<https://orcid.org/0009-0003-5373-8228>

المستخلص: يهدف هذا المقال إلى تحليل العلاقة بين درجة الحرارة وتركيز غاز الأوزون في محافظة واسط باستخدام نماذج الانحدار اللامعلمي الضبابي، نظرًا للطبيعة غير الخطية والغامضة للبيانات البيئية. تم توظيف ثلاث طرائق رئيسية لنموذج اندار لا معلمي ضبابي وهي طريقة الجيران الأقرب (KNN)، وطريقة نداريا واتسون (Nadaraya–Watson)، وطريقة متعدد الحدود الانحدار المحلي (Local Polynomial). اعتمدت عملية التقدير على طريقة التحقق المتقاطع من نوع ترك-واحد-خارجًا (LOOCV) لاختيار المعلمات المثلى لكل نموذج، باستخدام معيار متوسط مربع الخطأ (AMSE) ومعيار التكامل الوسطي لمربع الخطأ (IMSE). كما تم التعامل مع متغير الاستجابة على أنه رقم ضبابي مثلثي يمثل عدم اليقين في القياسات البيئية، واعتمدت مسافة دياموند المربعة لقياس الفرق بين الأعداد الضبابية المقدرة والحقيقية. أظهرت النتائج أن جميع النماذج قدمت أداءً مقبولاً، إلا أن نموذج الانحدار المحلي الخطي بنواة Epanechnikov تفوق بتحقيقه أقل قيمتي AMSE وIMSE، مما يعكس دقته الأعلى في تمثيل العلاقة اللامعلمية بين درجة الحرارة وتركيز الأوزون. تؤكد النتائج ان النماذج الضبابية اللامعلمية تمثل الظواهر البيئية التي تتسم بالغموض.

Corresponding Author: E-mail: ahmedrazzaq@uowasit.edu.iq

يواجه الباحثون في المجالات العلمية والتطبيقية العديد من التحديات المتعلقة بعدم دقة البيانات ووجود الضبابية فيها، حيث انه لا تكون المعلومات المتاحة دائماً واضحة أو دقيقة، بل تحمل درجات متفاوتة من عدم الدقة. تمثل هذه المشكلة تحدياً جوهرياً في تحليل البيانات وبناء النماذج الإحصائية، خصوصاً في الدراسات التي تعتمد على المراقبة البيئية، والاقتصاد، والقياسات الطبية حيث يتعين على الباحثين إيجاد طرق للتعامل مع هذه البيانات بفعالية وتوفر نتائج موثوقة. من بين الطرق الحديثة التي أثبتت فعاليتها في التعامل مع عدم الدقة والضبابية هي نماذج الانحدار اللامعلمية، والتي تُعتبر أدوات مرنة لا تفترض شكلاً محدداً مسبقاً للعلاقة بين المتغيرات، مما يجعلها مناسبة لتحليل البيانات المعقدة. في نفس الوقت، أدخلت مفاهيم الانحدار الضبابي (Fuzzy Regression) لتوصيف التقديرات عندما تكون القيم المخرجات غير دقيقة أو تحتوي على ضبابية فيها، مما يعكس الطبيعة الحقيقية للبيانات في العديد من التطبيقات. (انظر: البياتي وعبد الحميد، 2017؛ جمعة وعبد الغني، 2020؛ عيد الله، 2017) مع ذلك، ما زال هناك نقص في الدراسات التي تجمع بين هذين المنهجين، أي دمج الانحدار اللامعلمي مع معالجة البيانات الضبابية، خصوصاً في وجود بيانات غير دقيقة ذات طبيعة شبيهة ضبابية. هذا النقص يخلق فجوة بحثية مهمة، حيث توجد حاجة ماسة إلى تطوير إطار إحصائي تكاملي يجمع مزايا كلا المنهجين لتحسين دقة التقديرات والتعامل مع عدم اليقين بكفاءة. يقتصر هذا البحث على استخدام بيانات ضبابية محددة تمثل في أشكال مثلثية وتربيعية، مع التركيز على النماذج المذكورة آنفاً دون الدخول في نماذج أخرى. ويعتمد البحث على تقييم أداء النماذج باستخدام معايير دقيقة مثل متوسط الخطأ التربيعي (MSE) ومقاييس جودة الضبابية، إضافة إلى دراسة فاعلية الإطار التكاملي المقترح.

1.1. مشكلة البحث :

في العديد من التطبيقات الواقعية، خصوصاً في المجالات البيئية والاقتصادية والطبية، تكون البيانات المتاحة غير دقيقة أو مضطربة، مما يضعف أداء نماذج الانحدار التقليدية في التنبؤ والفهم الإحصائي للعلاقات بين المتغيرات. كما أن النماذج اللامعلمية التقليدية تفترض بيانات دقيقة، في حين أن النماذج الضبابية غالباً ما تكون معلمية وتفترض شكلاً محدداً للعلاقة. وهذا يجاور الحقيقة حيث ان اغلب التطبيقات العلمية ذات علاقات غير معروفة، ومن هنا تظهر الحاجة إلى تطوير نموذج إحصائي تكاملي يجمع بين النماذج اللامعلمية والتعامل مع البيانات الضبابية لتحسين دقة التقدير، والمقارنة بينها لمعرفة أيهما أدق في وصف الظاهرة المدروسة.

1.2. هدف البحث:

يهدف البحث إلى مقارنة كفاءة ثلاثة طرائق تقدير نماذج من الانحدار الضبابي اللامعلمي، وهي: نموذج النواة (Kernel smoothing)، نموذج الجار الأقرب (K-Nearest Neighbors (K-NN))، ونموذج المربعات الصغرى المحلية (Local Linear Smoothing (LLS))، عند التعامل مع بيانات غير ضبابية، واقتراح إطار إحصائي تكاملي يجمع بين هذه النماذج لتحسين أداء التقدير وجودة النتائج.

1.3. الدراسات السابقة:

في الفترة الأخيرة، بدأ الباحثون في ربط مفهومي الضبابية والانحدار اللامعلمي لمعالجة المشكلات التي تحتوي على بيانات غير دقيقة. وظهرت دراسات عديدة تقترح نماذج لا معلمية تتعامل مع مدخلات ومخرجات ضبابية. حسب دراسة (Zadeh, L. A. (1965) [13] فإن الانحدار الضبابي هو امتداد للانحدار التقليدي يهدف إلى معالجة البيانات التي تحتوي على غموض أو عدم يقين. في هذا النوع من الانحدار، تكون العلاقة بين المتغيرات المدخلة والمخرجات غير دقيقة، مما يتطلب استخدام المجموعات الضبابية (Fuzzy Sets) للتعبير عن درجات العضوية في البيانات. وحسب دراسة (Tanaka et al., 1982) [10] الانحدار الضبابي هو أسلوب نمذجة إحصائي يُستخدم عندما تكون البيانات غير دقيقة أو ضبابية، ويُفترض أن العلاقة بين المتغيرات غير محددة بدقة. يتيح هذا النوع من الانحدار مرونة أكبر في تمثيل العلاقات، لا سيما في المجالات التي يصعب فيها الحصول على قياسات دقيقة. وفي عام 2007 قدم Wang et al. [11] واقتراح نموذج انحدار ضبابي باستخدام Local Linear Smoothing مع مقياس Diamond بين الأعداد الضبابية. هذا النموذج أثبت قدرته على تقديم نتائج دقيقة ومرنة عند التعامل مع الضوضاء والضبابية في المخرجات.

قدم Khammar وآخرون (2018) ، [7] منهجاً جديداً لتقدير نموذج الانحدار الضبابي الحصين بالاعتماد على مفاهيم جديدة في قياس المسافة بين الأعداد الضبابية باستخدام دالة النواة. اعتمدت طريقتهم على أسلوب المربعات الصغرى لتقدير معالم النموذج، وأظهرت كفاءة عالية في التعامل مع القيم الشاذة والبيانات غير النمطية. من خلال تطبيقات على بيانات حقيقية ومحاكاة، أثبتت الدراسة أن النموذج المقترح يوفر نتائج أكثر استقراراً مقارنة بالطرائق التقليدية عند وجود ضوضاء أو مشاهد متطرفة في البيانات.

وفي عام 2021 Razzaghnia et al. [9] قام باستخدام الباحثون ثلاث نماذج لا معلمية (LLS, KNN, Kernel) لمقارنة أدائها في حالة وجود بيانات ضبابية تربيعية. أظهرت الدراسة أن أداء النماذج يختلف حسب طبيعة الضبابية وحجم البيانات، وأن الانحدار الخطي المحلي أعطى أفضل النتائج في بعض الحالات. أما في عام 2021 قدم الباحثون Yıldız et al. [12] دراسة وناقشوا تحسين أداء الانحدار المحلي عبر اختيار bandwidth الأمثل باستخدام تقنيات التحقق المتقاطع (Cross Validation)، وبينوا أن التحكم في عرض النواة يؤثر بشكل كبير على دقة النماذج الضبابية لامعلمية.

1. المنهجية

The Nonparametric Fuzzy Model

1.1. النموذج اللامعلمي الضبابي:

النموذج اللامعلمي الضبابي (Fuzzy Nonparametric Regression Model) هي نموذج إحصائي يستخدم لتوصيف العلاقة بين متغيرات عندما تكون البيانات ضبابية أو غير دقيقة. في هذه النماذج لا نفترض شكلاً معيناً (كالخطية مثلاً) لمنحنى الانحدار مقدماً، وإنما نستخدم أساليب مرنة تعتمد على البيانات نفسها لاستنتاج شكل العلاقة. ظهرت الحاجة إلى هذه النماذج نتيجة مواجهة بيانات غير واضحة المعالم أو ظواهر تتسم بعدم اليقين، حيث تعجز النماذج التقليدية (المعتمدة على بيانات دقيقة أو افتراضات محددة) عن تفسير النتائج بشكل صحيح. فعلى سبيل المثال، قد تكون القيم المرصودة للمتغيرات على شكل أرقام ضبابية (مثل فترات زمنية أو تقديرات خبير مع درجة تأكيد)، مما يجعل استخدام طرق الانحدار الكلاسيكية (التي تفترض بيانات نقطية قطعية) غير ملائم.

تتميز النماذج الضبابية اللامعلمية بقدرتها على التعامل مع عدم دقة البيانات من خلال مفهوم المجموعات الضبابية الذي قدمه لطفي زاده [13]. في هذا السياق، يكون لكل قيمة درجة عضوية بين 0 و 1 تعكس مدى انتمائها للظاهرة المدروسة. وعادةً ما يتم تمثيل المخرجات الضبابية بوصفها أرقامًا ضبابية ذات دوال انتماء (مثل المثلثية أو شبه المنحرف) لتعكس عدم اليقين في القيمة. الغرض من التقنيات الضبابية هو تضمين جميع حالات عدم اليقين المتأصلة في البيانات في النموذج. لذلك، تحتوي النماذج القائمة على البيانات الضبابية على معلومات أكثر من النماذج التي يتم فيها رفض عدم اليقين الأصلي في البيانات أو حذفه اختياريًا [14]. في هذا الصدد، من الضروري التعبير عن جميع الطرق في بنية ضبابية لاستخدامها في البيانات الضبابية. عند التعبير عنها باستخدام هذه النماذج الضبابية، فإن مجموعات البيانات ذات الضبابية المتأصلة تعطي نتائج أفضل وأكثر فعالية.

وعليه فإن النموذج اللامعلمي الضبابي قيد الدراسة سوف يتكون من متغير استجابة (Y) يمثل الرقم الضبابي المثلثي ومتغير مستقل (X) لامعلمي أي لا يمكن معرفة شكل العلاقة بينه وبين متغير الاستجابة، وفق المعادلة الآتية:

$$\tilde{Y}_i = \mu(x_i) + \varepsilon_i \quad (1)$$

عندما \tilde{Y}_i : يمثل متغير استجابة ضبابي يمكن التعبير عنه وفق الرقم الضبابي المثلثي كل الآتي:

$$\tilde{Y}_i = (L_{Y_i}, C_{Y_i}, R_{Y_i})$$

وان $\mu(x_i)$: تمثل دالة تمهيد إلى المتغير اللامعلمي x_i و ε_i حد الخطأ.

وقد أشار Silverman (1988) و Tastan & Yildiz (2020) إلى أن الأساليب اللامعلمية الضبابية تمثل الأساس الإحصائي الأحدث لتقنيات التقدير البيئي غير الخطية.

1.1.2. طرائق التقدير:

1.1.2.1. طريقة الجار الأقرب الضبابي (Fuzzy KNN Regression) [16] [18] [6]

تعد طريقة K-NN Smoothing إحدى أساليب التنعيم اللامعلمية، حيث تعتمد على مفهوم الجوار القريب من النقطة المراد تقدير قيمتها. تقوم الفكرة الأساسية على تحديد عدد من النقاط القريبة (K) من النقطة x باستخدام مسافة معينة (غالبًا إقليدية)، ثم حساب متوسط القيم التابعة لتلك النقاط المجاورة. [6]

وقد طوّرت هذه الطريقة لأول مرة من قبل Keller et al. (1985). لتقدير المتغيرات الضبابية باستخدام مبدأ الجوار القريب. يُعامل كل جار من الجيران الـ K بنفس الوزن، أي أن جميع النقاط ضمن الجوار تساهم بالتأثير نفسه في عملية التقدير، على عكس بعض الطرق الأخرى مثل التنعيم بالنواة (Kernel Smoothing) التي تعطي أوزانًا متفاوتة تبعًا لبعدها عن النقطة عن المركز.

يستخدم هذا الأسلوب بشكل واسع في مجالات تقدير الكثافة والانحدار اللامعلمي والتصنيف، لما يتميز به من بساطة ومرونة في التعامل مع البيانات غير الخطية أو غير الدقيقة. ولتقدير القيم الضبابية باستخدام نموذج Fuzzy K-Nearest Neighbors Regression ويُعطى تقدير القيمة الضبابية \tilde{Y}_i عند النقطة x كما يلي:

$$\tilde{Y}_i = \sum_{i=1}^n \omega_i(x) Y_i \quad (2)$$

حيث إن $\omega_j(x)$ تمثل تسلسل الأوزان (Weight Sequence) عند النقطة x ويُعرّف كما يلي:

$$\omega_i(x) = \begin{cases} \frac{1}{k}, & i \in J(x), i = 1, \dots, n \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

هنا:

$J(x)$: مجموعة الملاحظات الأقرب إلى النقطة x أي مجموعة أقرب k نقاط إلى x .

Y_i : القيم الملاحظة (قد تكون ضبابية أو حادة).

وان الاشتقاق الرياضي للتقدير في الحالة الضبابية حيث إن J_x هي مجموعة الملاحظات الأقرب k إلى النقطة x ، فإن \tilde{Y}_i تمثل تقدير القيمة الضبابية للمشاهدة i ، ويُفترض أن القيم \tilde{Y}_i أعداد ضبابية شبه منحرفة (Asymmetric Trapezoidal Fuzzy Numbers)، بالشكل:

$$\tilde{Y}_i = (L_{Y_i}, C_{Y_i}, R_{Y_i}) \quad (4)$$

وبناءً على المعادلتين (2) و(3) نحصل على التقدير التالي:

$$\tilde{Y}_i = \left[\sum_{i=1}^n \omega_i(x) L_{Y_i}, \sum_{i=1}^n \omega_i(x) C_{Y_i}, \sum_{i=1}^n \omega_i(x) R_{Y_i} \right]$$

وبما أن الأوزان $\omega_i(x)$ تُعطى حسب المعادلة (6,7)، فيمكن تبسيطها إلى الشكل:

$$\tilde{Y}_i = \left[\frac{1}{k} \sum_{i \in J(x)} L_{Y_i}, \frac{1}{k} \sum_{i \in J(x)} C_{Y_i}, \frac{1}{k} \sum_{i \in J(x)} R_{Y_i} \right] \quad (5)$$

بناءً على ما سبق، فإن طريقة **Fuzzy KNN Regression** تُقدّر القيمة الضبابية المتوقعة عند النقطة x على أنها المتوسط الحسابي المرجح للقيم الضبابية لأقرب k نقاط، حيث تكون الأوزان متساوية $(1/k)$ لجميع الجيران داخل مجموعة $J(x)$ ، وتكون مساوية للصفر خارجها. سنقدم الآن دالتان مختلفتان لتحديد الأوزان w_i .

i. وزن معكوس المسافة المربع (Inverse Distance Squared Weight)

$$w_i = \frac{1}{d(x_0, x_i)^2 + \epsilon} \quad (6)$$

حيث:

- $d(x_0, x_i)$ هي مسافة مناسبة بين $(x_0$ و $x_i)$ مثل مسافة إقليدية Euclidean أو أي مسافة أخرى مناسبة حسب مقياس الضبابية.
- $\epsilon > 0$ ثابت صغير لمنع القسمة على صفر إذا كانت النقطة نفسها جزءاً من الجيران.

بعد حساب الأوزان، يُعاد قياسها غالباً بحيث مجموعها = 1، أي ان:

$$w_i \leftarrow \frac{w_i}{\sum_{i \in N_k(x_0)} w_j}$$

ii. الدالة الثانية Kernel Weight باستخدام نواة Gaussian (Gaussian Kernel Weight)

وتكتب صيغتها الرياضية بالشكل الآتي:

$$w_i = K\left(\frac{d(x_0, x_i)}{h}\right) = \exp\left(-\frac{d(x_0, x_i)^2}{2h^2}\right)$$

عندما h : تمثل عرض الحزمة، والصيغة الكاملة (مُقيّدة بماسيح نواة Gaussian) تمثل بالشكل الآتي:

$$w_i = \frac{1}{\sqrt{2\pi}h} \exp\left(-\frac{d(x_0, x_i)^2}{2h^2}\right)$$

لكن غالبًا في KNN يُستخدم الوزن غير المقيس الثابت (بدون عامل المقدار الأمامي) لأن المقياس سيكون مقيّدًا بواسطة مجموع الأوزان فيما بعد:

$$w_i = \exp\left(-\frac{d(x_0, x_i)^2}{2h^2}\right) \quad (7)$$

ومرة أخرى يُعاد قياس الأوزان بالشكل الآتي:

$$w_i \leftarrow \frac{w_i}{\sum_{j \in N_k(x_0)} w_j}$$

1.2.2. طريقة تقدير نداريا واتسون الضبابية [16] [17] [19] Fuzzy Nadaraya–Watson Estimation

Method

يُعدّ التمثيل المفاهيمي لتسلسل الأوزان في طريقة المتوسطات المحلية قائمًا على استخدام دالة كثافة احتمالية تحتوي على معامل مقياس يحدد حجم وشكل الأوزان وفقًا لموقع النقطة الملاحظة بالنسبة إلى نقطة التقدير x_0 . وتُسمى هذه الدالة بدالة النواة (Kernel Function)، التي تُستخدم لتوزيع الأوزان حول النقطة المراد تقديرها. وتعتمد تقنيات التنعيم بالنواة (Kernel Smoothing) على هذا المبدأ لتقدير الدوال اللامعلمية. في سياق الانحدار الضبابي (Fuzzy Regression)، تُستخدم هذه التقنية لتقدير العلاقة الضبابية بين المتغيرات، إذ يُمنح لكل ملاحظة وزن يتناقص مع زيادة بعدها عن x_0 ، مما يجعل التقدير يعتمد بدرجة أكبر على النقاط القريبة. وتُعد هذه الطريقة تعميمًا لامعلمي (Nonparametric) لأساليب الانحدار الضبابي، لأنها لا تقترض شكلًا وظيفيًا محددًا مسبقًا، بل تُقدّر الدالة محليًا باستخدام أوزان النواة.

ويختلف بناء مقدر النواة (Kernel Estimate) عن المقدر المستند إلى طريقة أقرب k جيران (k -NN) ففي حين أن مقدر NN- k يُعطي أوزانًا متساوية لجميع المشاهدات الواقعة داخل الجوار المتغير حول النقطة، فإن مقدر النواة يُعرف على أنه متوسط مرجح للمتغير التابع داخل جوار ثابت حول النقطة X ، يتم تحديده شكله من خلال دالة النواة K ومعامل الانعطاف h (أو ما يُعرف بعرض النطاق Bandwidth)، يبقى مقدر النواة $S(x)$ بمثابة متوسط محلي، غير أنه على عكس مقدر NN- k ، فإنه يستخدم دالة النواة لإعطاء أوزان مختلفة للمشاهدات وفقًا لبعدها عن نقطة التقدير X إذ تُمنح المشاهدات الأقرب إلى X أوزانًا أكبر، بينما تُعطى المشاهدات الأبعد أوزانًا أصغر. وفي دراسة (Mohammad & Al-R. (2023) تم توضيح تأثير نوع النواة (Epanechnikov) و (Triangular) على دقة التقدير. ويُعرف مقدر النواة للانحدار الضبابي رياضيًا كما يأتي:

يُعرف وزن المشاهدة i عند نقطة التقدير x_0 وفق العلاقة: [8]

$$\omega_i(x_0) = \frac{K_h(|x_j - x_0|)}{\sum_{i=1}^n K_h(|x_i - x_0|)} = \frac{K\left(\frac{|x_i - x_0|}{h}\right)}{\sum_{i=1}^n K\left(\frac{|x_i - x_0|}{h}\right)} \quad (8)$$

حيث تمثل $K(\cdot)$ دالة النواة (Kernel Function) و h هو معامل التنعيم (Bandwidth) الذي يتحكم بدرجة النعومة.

وبناءً على هذه الأوزان، يُحسب متوسط الاستجابة بعد التنعيم (أو المقدر الضبابي) من خلال مجموعات مرجحة كالتالي:

$$\sum_{i=1}^n K_h(|x_i - x_0|) L_{Y_i}, \sum_{i=1}^n K_h(|x_i - x_0|) C_{Y_i}, \sum_{i=1}^n K_h(|x_i - x_0|) R_{Y_i} \quad (9)$$

ويمكن تمثيلها عمومًا على النحو الآتي:

$$\sum_{i=1}^n K_h(|x_i - x_0|) Y_i^{(r)} = \frac{\sum_{i=1}^n K_h(|x_j - x_0|) Y_i^{(r)}}{\sum_{i=1}^n K_h(|x_i - x_0|)}, r = 1, 2, 3.$$

ويُعرَّف تسلسل الأوزان (Weight Sequence) كما يلي:

$$K_h(x) = \frac{1}{h} K\left(\frac{x}{h}\right) \quad (10)$$

حيث $K(\cdot)$: هي دالة النواة مثل (Gaussian) أو (Epanechnikov)، h : هو معامل التنعيم الذي يحدد درجة انسيابية التقدير، وكل وزن $w_i(x_0)$ يعبر عن مدى تأثير الملاحظة Z على النقطة x_0 . [8].

(Gaussian Kernel Function)

i. دالة النواة الغاوسية [4] [5]

تُعدّ النواة الغاوسية من أكثر الدوال استخدامًا في التقدير اللامعلمي، وذلك لأنها مستمرة، وناعمة (Smooth)، وتمنح وزنًا أعلى للنقاط القريبة من نقطة التقدير x ، بينما يقل الوزن تدريجيًا مع البعد عنها وفق توزيع طبيعي. وان الصيغة الرياضية:

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} \quad (11)$$

حيث:

$$u = \frac{x - x_i}{h}$$

$h > 0$: هو عرض النواة (Bandwidth) الذي يتحكم بدرجة التنعيم، إذا كانت قيمة صغيرة لـ h ⇒ النموذج يتبع البيانات بدقة (Overfitting)، وإذا كانت قيمة كبيرة لـ h ⇒ النموذج يصبح ناعمًا جدًا (Over smoothing) ويمكن تمثيلها في نموذج الانحدار الخطي حيث يتم تقدير الدالة عند نقطة x باستخدام النواة الغاوسية يكون:

$$\hat{m}(x) = \frac{\sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) y_i}{\sum_{i=1}^n K\left(\frac{x-x_i}{h}\right)}$$

وفي النسخة الخطية المحلية:

$$\hat{m}(x) = e_1^T (X^T W X)^{-1} X^T W Y \quad (12)$$

حيث W مصفوفة قطرية تعتمد على النواة الغاوسية كما يلي:

$$W = \text{diag}\left(K\left(\frac{x - x_1}{h}\right), \dots, K\left(\frac{x - x_n}{h}\right)\right)$$

ii. دالة النواة إبانتشكوف Epanechnikov [5] [4]

تُعتبر دالة إبانتشكوف أفضل نواة من حيث الكفاءة الإحصائية (Mean Integrated Squared Error efficiency)، إذ تقلل من الخطأ المربّع التكاملي مقارنة ببعض النوى الأخرى. وهي محدودة النطاق (Compact Support)، أي أنّها تُعطي وزنًا صفرًا للنقاط البعيدة عن نطاقها وان الصيغة الرياضية لها:

$$K(u) = \begin{cases} \frac{3}{4}(1 - u^2), & \text{إذا } |u| \leq 1 \\ 0, & \text{otherwise} \end{cases} \quad (13)$$

حيث:

$$u = \frac{x - x_i}{h}, h: \text{عرض النواة كما في الحالة السابقة، ويتم تمثيلها في التقدير بالشكل الآتي:}$$

$$\hat{m}(x) = \frac{\sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) y_i}{\sum_{i=1}^n K\left(\frac{x-x_i}{h}\right)}$$

وتُستخدم نفس الصيغة في النسخة الخطية المحلية مع W المعرفة بنفس الشكل كما في النواة الغاوسية.

1.2.3 طريقة الانحدار متعدد الحدود المحلي الضبابي [16] [15] [5] Fuzzy Local Polynomial Regression

Method

الهدف من التقنيات الضبابية هو إدخال جميع حالات عدم اليقين الكامنة في البيانات ضمن النموذج. ولذلك، فإن النماذج المبنية على البيانات الضبابية تحتوي على معلومات أكثر من تلك التي يتم فيها إهمال أو تجاهل هذا الغموض. من هذا المنطلق، من الضروري التعبير عن جميع الطرق الإحصائية في بنية ضبابية عند التعامل مع بيانات ضبابية، لأن ذلك يؤدي إلى نتائج أفضل وأكثر دقة. ويُعد Silverman (1986) من أوائل من وضع الأسس الرياضية لتقدير الكثافة اللامعلمي، وهو ما تعتمد عليه الطريقة الحالية. باعتبار أن البيانات تتكون من أزواج مشاهدات (x_i, y_i) ، حيث إن y_i ضبابية من النوع LR، يكون كل عنصر من y_i له مركز c_y وحد يسار $l_y = c_y - \alpha_y$ وحد يمين $r_y = c_y + \beta_y$ ، أي $y_i = (l_y, c_y, r_y)_{LR}$.

تعطى دالة الانتماء (Membership Function) للعدد الضبابي $(l_A, c_A, r_A)_{LR}$ بالعلاقة:

$$\mu_A(x) = \begin{cases} L\left(\frac{c_A - x}{c_A - l_A}\right), & l_A \leq x \leq c_A \\ R\left(\frac{x - c_A}{r_A - c_A}\right), & c_A \leq x \leq r_A \\ 0, & \text{otherwise} \end{cases} \quad (14)$$

حيث إن الدالتين $L(\cdot)$ و $R(\cdot)$ متناقضتان قطعياً ومحققتان للشرط $L(0) = R(0) = 1$ و $L(1) = R(1) = 0$. بناءً على ذلك، يُكتب النموذج بالشكل:

$$\tilde{Y} = \mu(x) + \varepsilon = (l(x), c(x), r(x))_{LR} + \varepsilon \quad (8)$$

إذا كانت الدوال $l(x), c(x), r(x)$ قابلة للاشتقاق من الرتبة $(p + 1)$ ، فيمكن تقريبها باستخدام متسلسلة تايلور من الرتبة p حول النقطة x_0 كما يلي:

$$\begin{aligned} l(x) &\approx l(x_0) + l'(x_0)(x - x_0) + \dots + \frac{l^{(p)}(x_0)}{p!}(x - x_0)^p \\ c(x) &\approx c(x_0) + c'(x_0)(x - x_0) + \dots + \frac{c^{(p)}(x_0)}{p!}(x - x_0)^p \\ r(x) &\approx r(x_0) + r'(x_0)(x - x_0) + \dots + \frac{r^{(p)}(x_0)}{p!}(x - x_0)^p \end{aligned} \quad (15)$$

وباستخدام طريقة المربعات الصغرى الموزونة (Weighted Least Squares)، يُؤخذ بعين الاعتبار بُعد دايموند (Diamond Distance) وهو مقياس مسافة يُستخدم للتعامل مع المجموعات الضبابية لحساب الفرق بين القيم الفعلية والمتوقعة ويُعبر عن تطبيق هذا البعد في تقدير الانحدار متعدد الحدود المحلي بالصيغة الآتية:

$$\begin{aligned} \sum_{i=1}^n d^2(\mu(x_i), y_i) K_h(|x_i - x_0|) &= \sum_{i=1}^n d^2((l(x_i), c(x_i), r(x_i)), (l_{y_i}, c_{y_i}, r_{y_i})) K_h(|x_i - x_0|) = \\ &= \sum_{i=1}^n (l_{y_i} - l(x_0) - l'(x_0)(x_i - x_0) - \dots - \frac{l^{(p)}(x_0)}{p!}(x_i - x_0)^p)^2 K_h(|x_i - x_0|) + \sum_{i=1}^n (c_{y_i} - \\ &= c(x_0) - c'(x_0)(x_i - x_0) - \dots - \frac{c^{(p)}(x_0)}{p!}(x_i - x_0)^p)^2 K_h(|x_i - x_0|) \\ &+ \sum_{i=1}^n (r_{y_i} - r(x_0) - r'(x_0)(x_i - x_0) - \dots - \frac{r^{(p)}(x_0)}{p!}(x_i - x_0)^p)^2 K_h(|x_i - x_0|) \end{aligned} \quad (16)$$

قيمة المعادلة (16). إن تقليل قيمة كل مجموع على حدة يساعد أيضاً في الحصول (Minimize) الهدف الرئيسي هو تصغير، تكون دالة الوزن كما يلي h ومعامل تنعيم $K(\cdot)$ لكل دالة نواة. على الحل الأمثل المطلوب

$$K_h(|x_i - x_0|) = \frac{1}{h} K\left(\frac{|x_i - x_0|}{h}\right), i = 1, 2, \dots, n$$

وهي دالة وزن تُعطي أهمية أكبر للمشاهدات القريبة من النقطة x_0 وأهمية أقل للنقاط البعيدة عنها.

1.3 المقاييس المستخدمة للمقارنة في الانحدار الضبابي [16]

(Integral Mean Squared Error - IMSE)

1.3.1 المعيار التكاملي للخطأ التربيعي

يقيس هذا المعيار الفارق الكلي بين الدالة الحقيقية $m(x)$ ودالة التقدير $\hat{m}(x)$ على كامل المجال الذي تُعرّف عليه البيانات. بمعنى آخر، هو تكامل مربع الخطأ على المجال الكامل وان لصيغة الرياضية:

$$ISE = \int [\hat{m}(x) - m(x)]^2 dx \quad (17)$$

حيث:

$m(x)$: الدالة الحقيقية (أو المرجعية) للانحدار.

$\hat{m}(x)$: الدالة المقدّرة باستخدام الطريقة مثل Local Linear Smoothing.

التكامل يُحسب على مجال المتغير x .

وكلما كانت قيمة الـ ISE صغيرة، دلّ ذلك على أن النموذج يقدر الدالة الحقيقية بدقة عالية ويقبض الانحراف التراكمي للنموذج عبر المجال الكامل ويعتبر معياراً شاملاً لأنه يجمع الخطأ عند جميع النقاط بدلاً من نقطة واحدة فقط.

في التطبيق الضبابي: عندما يكون المتغير التابع آضبابياً، نحسب المعيار على أساس المسافة بين الأعداد الضبابية، فمثلاً إذا استخدمنا مسافة دياموند (Diamond Distance)

$$ISE_f = \int D^2(\tilde{m}(x), \hat{m}(x)) dx \quad (18)$$

حيث $\tilde{m}(x)$ و $\hat{m}(x)$ هما القيم الضبابية الحقيقية والمقدّرة.

(Average Mean Squared Error - AMSE)

1.3.2 متوسط الخطأ التربيعي

هو المتوسط العددي للخطأ التربيعي بين القيم الحقيقية والمقدّرة عبر جميع نقاط العينة. ويُعتبر هذا المعيار نسخة عملية (Discrete) من الـ $IMSE$ عندما تكون البيانات محدودة بعدد من المشاهدات.

وان الصيغة الرياضية لهذه الطريقة كما موضحة ادناه:

$$AMSE = \frac{1}{n} \sum_{i=1}^n [\hat{m}(x_i) - m(x_i)]^2 \quad (19)$$

حيث:

- n : عدد المشاهدات.
- $\hat{m}(x_i)$: القيمة المقدّرة للدالة عند النقطة x_i .
- $m(x_i)$: القيمة الحقيقية (أو المحاكاة المرجعية) عند النقطة نفسها.

في البيانات الضبابية: عند وجود أعداد ضبابية، نستخدم مقياس المسافة Diamond Distance بين العددين الضبابيين بدلاً من الفرق العددي تستخدم المسافة الماسية (Diamond Distance) [7] لقياس التقارب بين عددين ضبابيين $C = (L_c, M_c, R_c)$ و $D = (L_d, M_d, R_d)$ ، نستخدم المسافة الماسية:

$$d(C, D) = \sqrt{(L_c - L_d)^2 + (M_c - M_d)^2 + (R_c - R_d)^2} \quad (20)$$

$$AMSE_f = \frac{1}{n} \sum_{i=1}^n D^2(\tilde{y}_i, \hat{y}_i) \quad (21)$$

حيث:

- \tilde{y}_i : القيمة الضبابية الفعلية (المرصودة).
- \hat{y}_i : القيمة الضبابية المقدّرة من النموذج.

وإذا كانت قيمة صغيرة لـ $AMSE$ تعني أن النموذج يقدر القيم بدقة عالية.

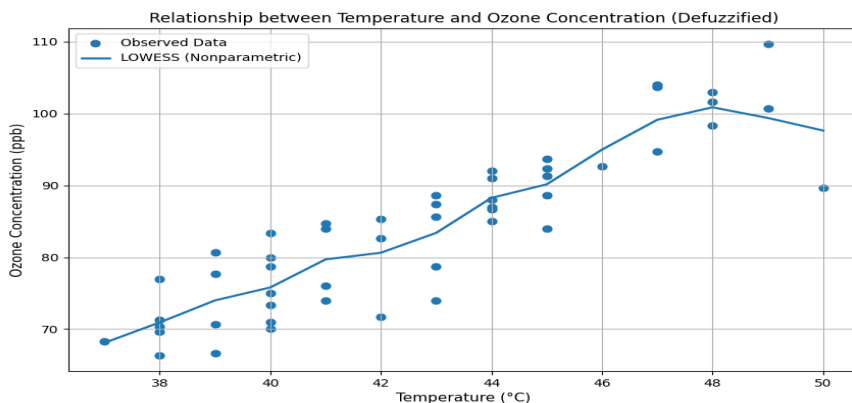
- تُستخدم للمقارنة بين طرق الانحدار الضبابي اللامعلمي.
- أكثر عملية من ISE لأنها لا تتطلب التكامل، بل تعتمد على قيم العينة فقط.

2. النتائج والمناقشة

الأوزون الأرضي (O_3) هو أحد مكونات تلوث الهواء، ويتأثر بشكل كبير بارتفاع درجات الحرارة وكثافة أشعة الشمس. كلما ارتفعت درجة الحرارة، زادت التفاعلات الكيميائية بين أكاسيد النيتروجين (NO_x) والمركبات العضوية المتطايرة (VOCs)، مما يؤدي إلى ارتفاع مستويات الأوزون. البيانات تمثل علاقة غير خطية بين درجة الحرارة ومستويات الأوزون في الجو في محافظة واسط، العراق، خلال فترات ذات درجات حرارة مرتفعة نسبيًا من 36 °C إلى 50 °C كما موضح في الشكل (1)، وقد تم جمع مستويات الأوزون بصيغة أرقام ضبابية مثلثية عند ارتفاع درجات الحرارة في فصل الصيف من عام 2025 (تموز، آب) تمثل الغموض والتقلب في قياسات الأوزون، وهي مناسبة لنمذجتها في نموذج انحدار ضبابي لا معلمي (Fuzzy Non-Parametric Regression) وقد تم الحصول على البيانات من خلال القياسات اليومية لموقع www.accuweather.com.

2.1 اختبار شكل العلاقة بين درجة الحرارة وتركيز الأوزون:

لتحديد ما إذا كانت العلاقة بين المتغير المستقل (درجة الحرارة) والمتغير التابع (تركيز الأوزون) خطية أم لا، تم إجراء تحليل بياني باستخدام تقنية LOWESS (Locally Weighted Scatterplot Smoothing) بعد إجراء عملية Defuzzification للقيم الضبابية. باستعمال لغة البرمجة Python 3.11. كما هو موضح في الشكل (1)، يتبين أن منحنى LOWESS غير خطي بوضوح، حيث يظهر نمط تصاعدي متزايد لتركيز الأوزون مع ارتفاع درجات الحرارة وصولاً إلى حوالي 46°C، ثم يميل إلى الاستقرار. وبناءً على ذلك، يمكن الاستنتاج أن العلاقة بين المتغيرين لا معلمية (Nonlinear)، مما يبرر استخدام طرائق الانحدار اللامعلمي الضبابي مثل نموذج Kernel، و KNN، و Local Linear Smoothing في الدراسة التحليلية اللاحقة.



الشكل 1: العلاقة بين درجة الحرارة وتركيز الأوزون في محافظة واسط (العراق) استناداً إلى البيانات المحولة من الصيغة الضبابية باستخدام أسلوب LOWESS اللامعلمي.

يوضح الشكل وجود علاقة تصاعدية غير خطية بين درجة الحرارة وتركيز الأوزون، إذ يزداد تركيز الأوزون تدريجياً مع ارتفاع درجة الحرارة حتى يبلغ ذروته عند نحو 46°C، ثم يميل المنحنى إلى الاستقرار. هذا السلوك يؤكد أن العلاقة بين المتغيرين لا معلمية، مما يبرر استخدام نماذج الانحدار اللامعلمي الضبابي في التحليل الإحصائي للظاهرة.

الجدول 1: وصف المتغيرات المستخدمة في نموذج الانحدار اللامعلمي الضبابي لتحليل العلاقة بين درجة الحرارة وتركيز الأوزون في محافظة واسط.

المتغيرات	الوصف	المدى التقريبي
Temperature: X_1	درجة الحرارة اليومية المسجلة (بدرجة مئوية) في مدينة الكوت بمحافظة واسط، خلال أيام الصيف الحارة.	من 36 إلى 50
Ozone Left: Y	الحد الأدنى للقيمة الضبابية للأوزون (يمثل التقدير الأدنى لتركيز الأوزون في الهواء).	من حوالي 40 إلى 120
Ozone Center: Y	القيمة المركزية (النقطة الأكثر احتمالاً) لتركيز الأوزون المقاس، تمثل القيمة النموذجية أو "القلب" للرقم الضبابي المثلثي.	من حوالي 50 إلى 130
Ozone Right: Y	الحد الأعلى للقيمة الضبابية للأوزون (يمثل التقدير الأعلى لتركيز الأوزون).	من حوالي 60 إلى 140

2.2. التقدير والمقارنة:

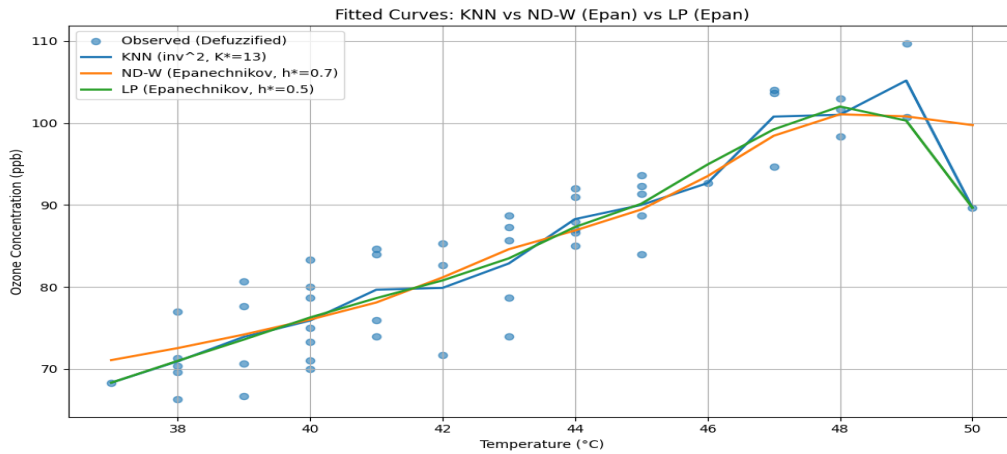
هنا يتم تقدير انحدار لا معلمي ضبابي لبيانات الأوزون وفق طرائق التقدير التي تم عرضها وبالاتماد على معايير المقارنة باستعمال Python 3.11، سوف نبين أي الطرق اللامعلمية الضبابية هي الأنسب لتحليل ظاهرة غاز الأوزون في محافظة واسط العراق، وباستعمال لغة البرمجة بايثون وكانت النتائج كما في الجدول 2:

وتجدر الإشارة هنا اعتماد أسلوب التحقق المتقاطع من نوع ترك-واحد-خارجًا (Leave-One-Out Cross-Validation, LOOCV) لاختيار المعلمات المثلى لكل نموذج لامعلمي ضبابي. يقوم هذا الأسلوب على حذف مشاهدة واحدة من العينة في كل مرة، ثم تقدير النموذج باستخدام باقي المشاهدات، وبعدها حساب الخطأ التنبؤي للمشاهدة المستبعدة. تتكرر هذه العملية لجميع قيم X ثم يحسب المتوسط العام لمربعات الأخطاء الناتجة (AMSE). تُعد هذه الطريقة من أكثر طرق التحقق المتقاطع دقة واستقرارًا، خصوصًا عند أحجام العينات الصغيرة، إذ تقلل من الانحياز في تقدير الخطأ وتضمن استغلال جميع البيانات المتاحة في عملية التدريب والاختبار. تم اعتماد القيم المثلى لكل من عدد الجيران K في نموذج KNN، و عرض النواة h في نماذج Nadaraya-Watson و Local Polynomial بناءً على أقل قيمة لمعيار AMSE المحسوب بطريقة LOOCV.

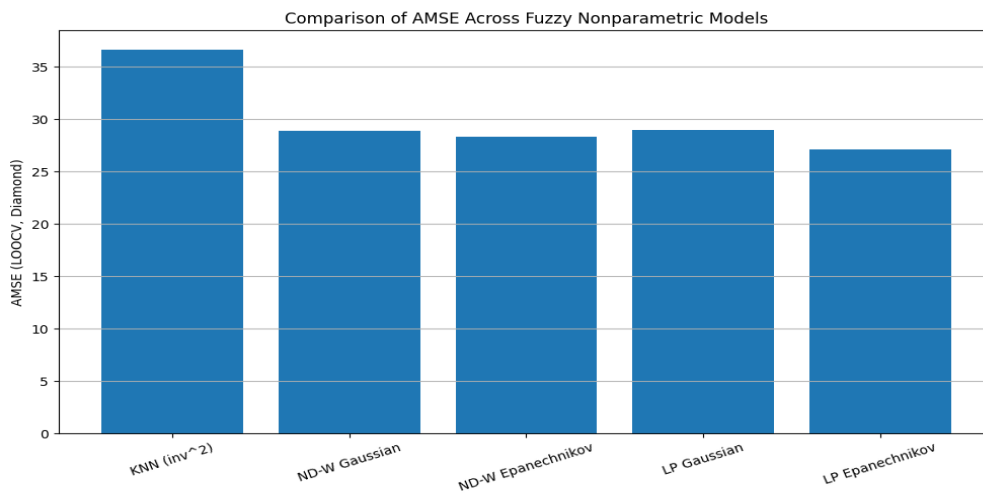
الجدول 2: نتائج المقارنة بين نماذج الانحدار اللامعلمي الضبابي باستخدام معايير AMSE و IMSE وفق أنواع النواة المختلفة.

Methods	Kernel Function	IMSE	AMSE
KNN (K=13)	Inverse Distance Squared Weight	476.280518	36.636963
ND-W	Gaussian (h=0.3)	375.436755	28.879750
	Epanechnikov (h=0.7)	368.642431	28.357110
LP	Gaussian (h=2)	376.685042	28.975772
	Epanechnikov (h=0.5)	352.490248	27.114634

يُلاحظ أن جميع النماذج اللامعلمية غير الخطية قدّمت أداءً مقبولاً بقيم خطأ تقع ضمن النطاق (27-37)، في حين تفوق نموذج Local Polynomial (Epanechnikov) بتحقيقه أدنى قيمتي $AMSE = 27.11$ و $IMSE = 352.49$ ، مما يعكس دقة تنبؤية أعلى واستجابة أكثر سلاسة لتغيرات درجة الحرارة. وقد تبين أن نموذج KNN سجّل أعلى خطأ نسبيًا بسبب حساسيته العالية لتوزيع الجيران وقلة مرونته في التعامل مع التغيرات المحلية، إذ إن استخدام وزن المسافة المربعة المعكوسة لم يكن كافيًا لتعويض التأثيرات الناتجة عن التوزيع الضبابي للبيانات، مما جعل بعض الجيران يبتعدون عددًا في المسافة المثلثية. أما نموذج Nadaraya-Watson (ND-W) فقد أظهر أداءً جيدًا بفضل خاصية النواة الملساء التي تقلل التذبذب في الحواف، وتوفقت نواة Epanechnikov قليلاً على نواة Gaussian لكونها ذات دعم محدود (Compact Support)، وهو ما جعلها أكثر كفاءة في المساحات المحلية. في المقابل، قدّم نموذج Local Polynomial (LP) أفضل أداء عام، وخاصة عند استخدام نواة Epanechnikov، إذ يتميز هذا النموذج بقدرته على الموازنة بين النعومة (smoothing) والدقة المحلية من خلال تقدير انحدار خطي بسيط حول كل نقطة، ما يقلل من انحراف الحافة (Boundary Bias) الموجود عادة في نماذج ND-W، وكانت القيمة المثلى لمعامل النعومة $h=0.5$ قد وُفرت مستوى مناسبًا من التنعيم دون إفراط أو نقص. وانخفاض قيمتي AMSE و IMSE يؤكد أن الانحدار اللامعلمي المحلي الخطي (LLS) قادر على التقاط التغير غير الخطي بين درجة الحرارة وتركيز الأوزون بدقة عالية، وهو ما يتسق مع الاختبار البياني المسبق الذي أظهر منحنى LOWESS غير خطي بوضوح. كما أن فروق القيم المطلقة بين النماذج لم تتجاوز 10 وحدات، مما يدل على أن جميع النوى تقدم أداءً متقاربًا، مع تفوق Epanechnikov LP بتحقيق أقل متوسط خطأ. من الناحية الفيزيائية، يمكن تفسير هذا النمط بأن ارتفاع درجة الحرارة يؤدي إلى تسارع التفاعلات الكيميائية المسؤولة عن تكوين الأوزون الأرضي حتى يصل النظام إلى درجة حرارية معينة، بعدها تبدأ معدلات التفاعل بالاستقرار أو التناقص نتيجة تفكك بعض المركبات أو انخفاض نسبة الأكسجين الفعال، وهو ما يعكس الشكل اللامعلمي للعلاقة الذي التقطته النماذج. وبناءً على هذه النتائج التحليلية والبيانية، يتضح أن نموذج Local Polynomial (Epanechnikov) هو الأكثر كفاءة في تمثيل العلاقة بين درجة الحرارة وتركيز الأوزون في محافظة واسط، إذ يقدم أفضل توازن بين الدقة والمرونة. وتشير هذه النتائج إلى أن الانحدار اللامعلمي الضبابي يُعد أداة تحليلية دقيقة وفعالة لدراسة العلاقات البيئية غير الخطية، مقارنةً بالنماذج الخطية أو الأساليب الكلاسيكية التقليدية.

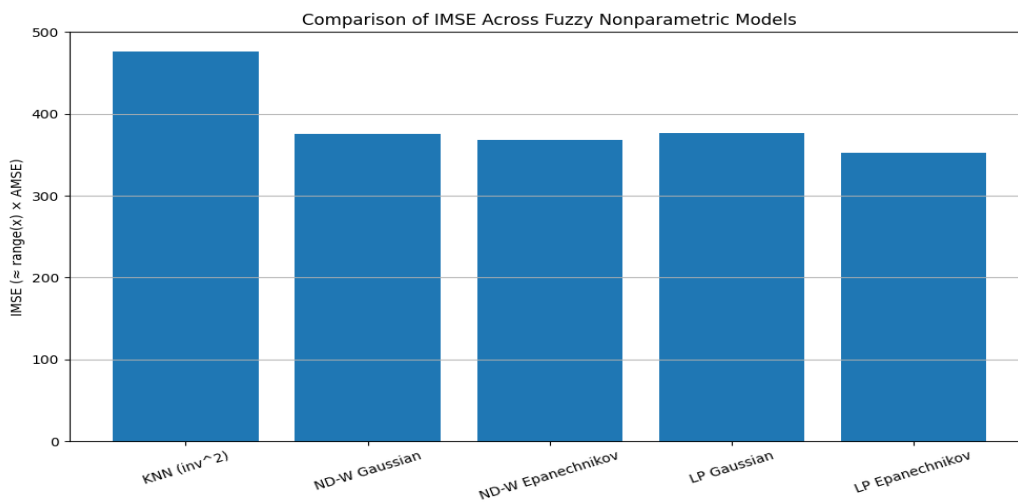


الشكل 2: منحنيات التقدير فوق النقاط ((KNN vs ND-W (Epan) vs LP (Epan))



الشكل 3: مقارنة قيم معيار AMSE بين النماذج اللامعلمية الضبابية المختلفة.

يُظهر الشكل تفوق نموذج **Local Polynomial (Epanechnikov)** بتحقيقه أقل قيمة لمتوسط الخطأ التربيعي، مما يشير إلى دقته الأعلى في تمثيل العلاقة بين درجة الحرارة وتركيز الأوزون مقارنة ببقية النماذج.



الشكل 4: مقارنة قيم معيار IMSE بين النماذج اللامعلمية الضبابية المختلفة.

يبين الشكل أن نموذج **Local Polynomial (Epanechnikov)** حافظ أيضًا على أقل قيمة لمعيار التكامل الوسطي لمربع الخطأ، ما يعزز كفاءته في التقدير اللامعلمي ويؤكد اتساق نتائجه مع تحليل معيار AMSE .

3. الاستنتاجات

1. أظهرت نتائج التحليل أن العلاقة بين درجة الحرارة وتركيز الأوزون في محافظة واسط علاقة غير خطية (اللامعلمية)، وقد تم التأكد من ذلك من خلال الاختبار البياني باستخدام منحني LOWESS الذي أوضح سلوكًا منحنيًا واضحًا.
2. أثبتت المقارنات الكمية أن النماذج اللامعلمية الضبابية توفر دقة أعلى في تمثيل الظواهر البيئية مقارنة بالنماذج الخطية، إذ أظهرت قدرتها على التعامل مع الغموض وعدم اليقين في البيانات البيئية.
3. من بين النماذج التي تم اختبارها، تفوق نموذج **Local Polynomial (Epanechnikov)** بتحقيقه أدنى قيم لمؤشري الخطأ $AMSE = 27.11$ و $IMSE = 352.49$ ، ما يدل على أنه النموذج الأكثر كفاءة في تمثيل العلاقة المدروسة.
4. أظهر نموذج **Nadaraya–Watson** أداءً قريبًا من نموذج **LP**، خصوصًا عند استخدام نواة **Epanechnikov**، مما يؤكد أن نوع النواة (**Kernel Function**) له تأثير مباشر على دقة النموذج.
5. نموذج **KNN** سجّل أعلى قيم خطأ نسبيًا، بسبب حساسيته لتوزيع الجيران ضمن المسافات الضبابية، الأمر الذي يجعله أقل ملاءمة في الحالات التي تحتوي على غموض أو تفاوت في القياسات.
6. تؤكد النتائج أن استخدام المسافة الضبابية (**Diamond Distance**) في التقدير يوفر قياسًا أكثر واقعية للخطأ في النماذج التي تتعامل مع متغيرات غامضة أو ذات طبيعة ضبابية.
7. النتيجة العامة تشير إلى أن الانحدار اللامعلمي الضبابي هو أداة فعالة ومرنة في دراسة العلاقات البيئية غير الخطية، مثل العلاقة بين درجة الحرارة وتكوّن الأوزون، لقدرته على التقاط الأنماط المعقدة والتقلبات غير المنتظمة في البيانات.
8. لوحظ أن المنهج المتبع الذي يجمع بين النماذج اللامعلمية والضبابية ضمن منهج واحد، عالج القصور القائم في الدراسات السابقة المقدمة التي ركزت على نموذج واحد فقط وطريقة تقدير واحدة. كما يُعد هذا خطوة عملية نحو تطوير أساليب تحليل أكثر واقعية للبيانات الضبابية في المجالات البيئية والاقتصادية والطبية، بما ينتج تحسين دقة التنبؤات واتخاذ القرارات المبنية على بيانات ضبابية.

3.1. التوصيات:

1. اعتماد على النماذج اللامعلمية الضبابية في الدراسات البيئية المستقبلية، خصوصًا تلك التي تتضمن بيانات تحتوي على عدم دقة في القياس مثل ملوثات.
2. توسيع نطاق الدراسة لتشمل متغيرات إضافية مؤثرة في تركيز الأوزون، مثل سرعة الرياح، الإشعاع الشمسي، والرطوبة النسبية، من أجل بناء نموذج لامعلمي تجميعي أكثر دقة.
3. اختبار تغير قيم عرض الحزمة (h) لمعرفة تأثيرها على جودة التقدير.
4. استعمال طرائق ضبابية جديدة تدمج بين النماذج اللامعلمية وحصانة التقدير لتحسين التقدير عند وجود قيم شاذة.
5. دراسة النموذج لملاحظات أخرى من أجل المقارنة بين الأنماط البيئية ومراقبة العوامل المؤثرة على مستويات الأوزون في مختلف العراق.

الشكر والتقدير

تتقدم الباحثة بخالص الشكر والتقدير إلى **جامعة واسط / كلية الإدارة والاقتصاد** على دعمها العلمي والإداري لإنجاز هذا البحث، كما تعبر الباحثة عن امتنانها لأعضاء قسم الإحصاء في كلية الإدارة والاقتصاد – جامعة واسط، لما قدموه من دعم علمي وملاحظات قيمة ساعدت في تطوير هذا البحث.

المصادر

1. البياتي، عبد الكريم عبد الله، وعبد الحميد، أحمد خليل. (2017). النماذج الإحصائية الضبابية واستخدامها في تحليل البيانات الاقتصادية. مجلة تنمية الرافدين، 36(118)، 1-22.
2. جمعة، علاء كامل، وعبد الغني، ندى عبد الوهاب. (2020). تحسين التنبؤ باستخدام نماذج الانحدار الضبابي: دراسة تطبيقية. مجلة البحوث الإحصائية، جامعة بغداد، 15(2)، 45-66.
3. عبد الله، فاطمة حسين. (2019). مقارنة نماذج انحدار لا معلمية ضبابية باستخدام بيانات بيئية غير دقيقة. المجلة العراقية للعلوم، 60(3)، 748-732.
4. Epanechnikov, V. A. (1969). *Non-parametric estimation of a multivariate probability density. Theory of Probability & Its Applications*, 14(1), 153-158.
<https://doi.org/10.1137/1114019>
5. Silverman, B. W. (2018). *Density estimation for statistics and data analysis*. Routledge.
6. Keller, J. M., Gray, M. R., & Givens, J. A. (1985). A fuzzy k-nearest neighbor algorithm. *IEEE Transactions on Systems, Man, and Cybernetics*, (4), 580-585.
7. Khammar, A., et al. (2018). *A robust least squares fuzzy regression model based on kernel functions. Iranian Journal of Fuzzy Systems*, 15(6), 27-45.
8. Mohammad, J. J., & Al-R. S. S. (2023). *Comparison between the kernel functions used in estimating the fuzzy regression discontinuous model*. [Uses Epanechnikov and Triangular kernels].
9. Razzaghnia, M., Lotfi, F. H., & Dehghani, M. (2021). *A comparison of several nonparametric fuzzy regressions with trapezoidal data. Journal of Intelligent & Fuzzy Systems*, 41(4), 5581-5593.
<https://doi.org/10.3233/JIFS-210394>
10. Tanaka, H., Uejima, S., & Asai, K. (1982). *Linear regression analysis with fuzzy model. IEEE Transactions on Systems, Man, and Cybernetics*, 12(6), 903-907.
<https://doi.org/10.1109/TSMC.1982.4308925>
11. Wang, Y. J., & Tsaor, R. C. (2007). *A fuzzy regression approach with nonparametric estimation. European Journal of Operational Research*, 176(1), 54-68.
<https://doi.org/10.1016/j.ejor.2005.07.040>
12. Yıldız, B., & Alparslan Gök, S. Z. (2021). *Bandwidth selection in local linear fuzzy regression. Fuzzy Sets and Systems*, 417, 1-18.
<https://doi.org/10.1016/j.fss.2020.06.014>
13. Zadeh, L. A. (1965). *Fuzzy sets. Information and Control*, 8(3), 338-353.
[https://doi.org/10.1016/S0019-9958\(65\)90241-X](https://doi.org/10.1016/S0019-9958(65)90241-X)
14. Klir, G. J., & Yuan, B. (1996). *Fuzzy sets and fuzzy logic: theory and applications. Possibility Theory versus Probab. Theory*, 32(2), 207-208.
15. Tastan, H., & Yıldız, B. (2020). *Nonparametric fuzzy regression based on local polynomial fitting. Fuzzy Sets and Systems*, 402, 78-95.
16. Chen, C. T. (2009). *Fuzzy multiple regression analysis for data with multicollinearity. Fuzzy Sets and Systems*, 160(13), 1909-1922.

Y: Ozone Left	Y: Ozone Center	Y: Ozone Right	X: Temperature
78	85	89	41
97	103	109	48
87	94	100	45
80	88	92	44
61	66	72	38
66	70	75	38
63	68	74	37
99	104	109	47
84	91	98	44
77	85	90	45
66	70	78	38
95	101	106	49
89	95	100	47
74	79	83	40
75	80	87	39
74	79	87	40
67	75	80	41
80	85	92	43
79	85	92	42
77	83	90	40
82	88	94	44
64	72	77	40
65	73	77	42
70	75	83	41
82	88	96	43
98	103	110	47
70	75	80	40
87	92	97	44
85	91	98	45
60	66	74	39
87	92	99	46
77	82	89	42
63	70	77	40
105	110	114	49
85	90	94	50
95	101	109	48
71	79	86	43
65	70	74	38
86	93	98	45
67	74	81	43
66	71	75	39
81	88	92	44
71	77	83	38
92	98	105	48
67	73	80	40
83	89	94	45
79	84	91	41
81	88	93	43
80	85	90	44
72	77	84	39

الملحق 1: