AL- Rafidain University College

PISSN: (1681-6870); EISSN: (2790-2293)

مجلة كلية الرافدين الجامعة للعلوم

Available online at: https://www.jrucs.iq

JRUCS

Journal of AL-Rafidain University College for Sciences

بناء انموذج لدراسة تأثير بعض المتغيرات على تلوث الهواء لمحافظة بغداد للفترة (2017-2019)

أ.م.د. عماد حازم عبودي	أ.د. قتيبة نبيل نايف
emadhazim@coadec.uobaghdad.edu.iq	dr.qutaiba@coadec.uobaghdad.edu.iq
قسم الاحصاء - كلية الادارة والاقتصاد - جامعة بغداد، بغداد، العراق	
م .م ميسم ساجت خضير	
maysam.s@albayan.edu.iq	
جامعة البيان، كلية إدارة الاعمال ، بغداد ، العراق	

المستخلص

يعتبر التلوث احد المشكلات المهمة التي تواجه البشرية في الوقت الحاضر نتيجة للنشاط الإنساني المتزايد في مجالات الحياة المختلفة وتوجد أنواع مختلفة من اشكال التلوث، ويعتبر تلوث الهواء احد مظاهر التلوث خطورة، لذلك اجريت بحثا باستعمال انموذج الانحدار اللوجستي حيث يعتبر من النماذج الكفؤة والملائمة في عملية تحليل البيانات الوصفية ثنائية الاستجابة، وانموذج بروبت الذي يشبه أنموذج اللوجستي الثنائي في طبيعة المتغير التابع اذ يكون متغير نوعي يأخذ صفتين هما الصفر والواحد ويعتمد على دالة الكثافة الاحتمالية ودالة التوزيع التراكمي وانه يتبع التوزيع الطبيعي القياسي، ان هدف البحث هو بحث التلوث البيئي للهواء حيث يتم بحث تأثير المتغيرات المؤثرة التي تمثل الغازات المنبعثة في الهواء وهي ($CO,NO,CH_4,SO_2,NO_2,NO_x$) على متغير الاستجابة والذي يمثل الدقائق العالقة ضمن المحدد المحلي للتلوث خلال 24 ساعة ($PM_{10},PM_{2.5}$) حيث تم اعتماد نسبة التلوث المنخفضة Law والمرتفعة High والتي تأخذ (0,1)، وتم تحويل هذا المتغير الى استجابة ثنائية على أساس النسب المقاسة لهذا المتغير، وتناول البحث ثلاث محطات في محافظة بغداد، هي (الوزيرية، وساحة الاندلس، والسيدية) ولقد توصل البحث الى ان ارتفاع مستوى التلوث بالدقائق العالقة بسبب المتغير المؤثر (CH_4) في محطة الوزيرية، والمتغيرين (SO_2,NO_x) في محطة ساحة الاندلس، والمتغير(NO_2,CH_4) في محطة السيدية.

معلومات البحث

تواريخ البحث

تاريخ تقديم البحث: 2022/12/3
تاريخ قبول البحث: 2023/2/25
تاريخ رفع البحث على الموقع: 2023/12/31

الكلمات المفتاحية

تلوث الهواء، انموذج الانحدار اللوجستي الثنائي، انموذج بروبت.

للمراسلة:

م .م ميسم ساجت خضير

maysam.s@albayan.edu.iq

https://doi.org/10.55562/jrucs.v54i1.572

1. المقدمة

يعد التلوث البيئي من اخطر المشكلات والازمات التي اصبحت تهدد كل اشكال الحياة وعناصر الوجود في عالمنا المعاصر، ولاشك في ان الانسان هو الملوث الاول للبيئة التي يحيا فيها، ان تزايد ظواهر التلوث البيئي وقلة الوعي البيئي وسوء التعامل مع البيئة من قبل اغلب الناس الى جانب انعدام التخطيط البيئي السليم سيؤدي الى تصاعد وتيرة المشكلات البيئية والاجتماعية والاقتصادية، كما ان تزايد اعداد السكان بشكل كبير وانعدام التخطيط العائلي في وضع تتنافر فيها المشكلات البيئية شكل هاجساً عند الكثير. مشكلة البيئة من المشكلات التي تواجه معظم المدن العراقية في وقتنا الحاضر اذ تتفاقم مشكلة التلوث في المدن بفعل نمو المدن الكبرى وتزايد سكانها وارتفاع الحجوم السكانية وسرعة انتشار المناطق الصناعية، ومع تزايد الحجم السكاني وبالأخص في المناطق الفقيرة والاحياء المتدنية حول المدن الكبرى ظهرت علامات التلوث واضحة سواء كان تلوث في الهواء او الماء او التربة او تلوث بصري او تلوث سمعي.

أصبح التلوث البيئي ظاهرة نشعر بها جميعا فلم تعد البيئة قادرة على تجديد مواردها الطبيعية، اذ أصبح تحت ضغط شديد ناتج عن الفعاليات البشرية المتزايدة. والتلوث في نظر بعض المهتمين يعني كافة الطرق التي يتسبب بها النشاط البشري في إلحاق الضرر بالبيئة الطبيعية.

- ❖ التلوث: هو ادخال الملوثات الى البيئة الطبيعية مما يلحق الضرر بها ويسبب الاضطراب في النظام البيئي، وهذه الملوثات اما ان تكون مواد دخيلة على البيئة او مواد طبيعية تجاوزت المستويات المقبولة ولا يقترن التلوث بالمواد الكيميائية وانما يشمل اشكال الطاقة المختلفة.[2]

ويصنف التلوث البيئي إلى نوعين:

1. التلوث المادي: هو تلوث محسوس يحيط بالإنسان فيشعر ويتأثر به ويراه بالعين المجردة، وقد يكون هو المتسبب فيه في معظم الأحيان، وإهمال منه في حق نفسه وانجرافه المستمر وراء التكنولوجيا ما يتسبب في الإخلال بالتوازن البيئي، ويتمثل بعدة أنواع
   - تلوث الهواء
   - تلوث المياه
   - تلوث التربة
   - التلوث الغذائي
   - التلوث بالنفايات
2. التلوث البيئي غير المادي (المعنوي): ويقصد به التلوث غير المحسوس، وغالبا ما تكون اثاره غير مباشرة على الرغم من انها قد تكون قاتلة في بعض الاحيان. فيؤثر على الانسان بطريقة غير مباشرة وذلك لان الانسان يتعامل مع عناصره بطريقة لا ارادية مما يسبب تأثير على اعضاء الانسان الداخلية.

وسنتناول في هذا البحث تلوث الهواء وهو احد أنواع التلوث المادي.

- ❖ تلوث الهواء: يحدث التلوث الهوائي عندما تتواجد جزيئات في الهواء وبكميات - كبيرة عضوية او غير عضوية وتشكل اضرارا على العناصر البيئية.

فيما يتعلق بملوثات الهواء واثارها في الصحة والبيئة فان زيادة تراكيز الغازات او نقصانها عن التراكيز الطبيعية اصلا يعد ظاهرة غير طبيعية ويجب ان يكون لها مسبباتها وتأثيرها في النظم او حياة الانسان. ومثل هذا الامر اصبح شائعا في الوقت الحاضر خصوصا في هواء المدن او المناطق الصناعية .اذ يلاحظ كثرة حالات زيادة الغازات الناتجة عن احتراق الوقود في السيارات والمعامل ومحطات توليد الطاقة الكهربائية فضلا عن حرق الوقود في المنازل لأغراض الطبخ والتدفئة يضاف الى هذا ما يتطاير في الهواء من دقائق ترابية او رملية وغبار مختلف التراكيب ناتج عن العمليات الصناعية مثل صناعة الاسمنت وغيرها ويضاف الى ذلك نوع اخر من الملوثات هو الضوضاء والضجيج حيث يعد تلوث الضوضاء صورة من صور التلوث الهوائي من منطلق ان الضوضاء او الضجيج عبارة عن موجات صوتية تنتقل عبر الهواء ,ومن التأثيرات الصحية الاخرى لتلوث الهواء حدوث حالات الاختناق او التسمم والتأثيرات الصحية نتيجة تركز الملوثات في الهواء والتي في معظمها ناتجة عن تزايد استهلاك الطاقة من مصادرها الملوثة مع حدوث الضباب الذي يتفاعل مع هذه الملوثات منتجة موادا سامة او انها تودي الى حدوث حالات اختناق.

ويمكن تصنيف ملوثات الهواء الى قسمين:

- ✓ اولا: مصادر طبيعية ليس للإنسان دخل فيها مثل الاتربة وغيرها من العوامل الاخرى
- ✓ ثانيا: مصادر صناعية اي انها من صنع الانسان وهو المتسبب الاول فيها لاختراعه وسائل تكنولوجية تؤدي الى انبعاث غازات وجسيمات دقيقة تنتشر في الهواء وتضر بالبيئة.

- ❖ ملوثات الهواء: من اكثر العناصر انتشاراً والتي تسبب تلوث الهواء هي:
  - ثاني اوكسيد الكربون: المصدر الرئيسي لهذا الغاز الضار هو الصناعة.
  - الدقائق العالقة او الجسيمات الدقيقة: وهي الاتربة الناعمة العالقة في الهواء والتي تأتي من المناطق -الصحراوية او تلك الملوثات الناتجة من حرق الوقود.
  - اكاسيد النتروجين: تنتج من حرق الوقود.
  - اول اكسيد الكربون: يوجد بتراكيز عالية وخاصة مع استعمال الغاز في المنازل.

تؤكد الاحصاءات البيئية بان زيادة احتمالية ارتفاع نسبة التلوث للهواء سيؤدي الى الاصابة بعدد من المشاكل الصحية في ضوء ما سبقه تكمن مشكلة البحث في ان الهواء الملوث بالدخان وغازات الاحتراق التي ترسلها مداخن المصانع وعوادم وسائط النقل ومصادر اخرى عديدة يفتقر إلى الأوكسجين وهو مضر بسبب السموم التي يحتويها فقد يتعرض الإنسان وخاصة على المدى الطويل لأمراض القصبات والرئة (التهابات رئوية مزمنة، عجز التنفس، تسممات، سرطان) ويؤثر بصفة سيئة وخطيرة في القلب والدم فالهواء يمكن عده ملوثًا عند اختلال التركيب أو التركيز لواحد أو أكثر من المكونات الطبيعة الغازية للهواء الطبيعي سواء كان هذا التغير أو التركيز لواحد أو أكثر من المكونات الطبيعية الغازية للهواء الطبيعي سواء كان هذا التغير بالزيادة أو النقصان، أو ظهور غازات أو أبخرة أو جسيمات عالقة عضوية وغير عضوية ، أو غيرها تشكل إضرارا على عناصر البيئة وتحدث خلاً في نظامها الأيكولوجي.[4]

ومن هنا يأتي دور الاحصاء وتطبيقاته في هذا المجال لتقديم بحوث تفسيرية لهذا النوع من التلوث وذلك بتوظيف أنموذج مناسب وقد ازدادت تطبيقات بعض الطرائق الاحصائية الحديثة في تحليل البيانات المصنفة ولاسيما في مجالات البحوث الطبية

والاجتماعية وغيرها، ان أنموذج الانحدار اللوجستي من النماذج المهمة التي تدخل في تحليل البيانات التي يكون فيها متغير الاستجابة (Y) بياناته ثنائية الاستجابة (1, 0)، و يوجد هناك نوعان من الانحدار اللوجستي وهما الانحدار اللوجستي الثنائي والانحدار اللوجستي المتعدد، وان موضوع البحث سيهتم ببحث أنموذج الانحدار اللوجستي الثنائي وانموذج بروبت، و يكون الهدف الاساسي هو ايجاد افضل أنموذج يصف الحالة بين متغير الاستجابة والمتغير المؤثر(المتغيرات المؤثرة).

وهناك العديد من الباحثين الذين اهتموا بدراسة هذا الأنموذج ومنهم، (علي، اسكندر حسين، (2016)) قدم قياس كفاءة حقول الابقار باستخدام نماذج الاستجابة النوعية وتمت المقارنة بين انموذج بروبت وأنموذج الانحدار اللوجستي الثنائي.(وضاح صبري ابراهيم وهاجر فلاح طاهر، (2019)) قدما أنموذج الانحدار اللوجستي لإيجاد اهم العوامل المؤثرة على مرض سرطان الرئة في العراق . وقدما ((Mohammed, M. A., & Raheem, S. H. (2020) بحثاً لتحديد اهم العوامل المؤثرة على الإصابة بمرض القلب باستخدام الانحدار اللوجستي (دراسة تطبيقية في مستشفى اربيل).

سوف يتناول البحث احد اشكال التلوث البيئي وهو تلوث الهواء، وسيتم تقديم وصف للمتغيرات المؤثرة في تلوث الهواء باستعمال النماذج الإحصائية لتفسير هذه الظاهرة والوصول الى نتائج تحصيلية على أساسها سوف يتم اتخاذ القرار.

2. مفهوم التحليل اللوجستي[1]

لقد ازدادت أهمية استخدام التحليل اللوجستي يوماً بعد آخر، لكونه يهتم بتحليل البيانات ذات الاستجابة الثنائية والتي عادة ما يكون فيها متغير الاستجابة (Response Variable) ثنائياً (Binary)، إذ ان حالة النجاح (Success) يأخذ فيها متغير الاستجابة القيمة (1) وحالة الفشل (Failure) يأخذ القيمة (0). ويستخدم الانموذج اللوجستي (Logistic Model) لوصف العلاقة بين متغير الاستجابة (y) والمتغيرات المؤثرة $x_1, \ldots\ldots, x_n$ ويتم التعبير عن تلك العلاقة بالصيغة الآتية:

$$p(x) = \frac{1}{1 + e^{-\alpha - \beta_i xi}} \qquad (1)$$

إذ ان:

α, β_i: هي معلمتا الانموذج المراد تقديرها و $\beta > 0$

$P(x)$: احتمال الاستجابة

Xi: المتغير المؤثر و $-\infty < \boldsymbol{xi} < \infty$

والصيغة اعلاه تعرف بدالة الاستجابة اللوجستية وتمتاز بان $P(x)$ محددة بين (1,0) وان المعلمتين (α , β) غير مقيدة وهناك نوعان من انموذج الانحدار اللوجستي، الأول يسمى بأنموذج الانحدار اللوجستي ثنائي الاستجابة، والثانية يسمى بأنموذج الانحدار اللوجستي المتعدد الاستجابة.

2.1. أنموذج الانحدار اللوجستي الثنائي Binary Logistic Regression Model [5]

يعرف أنموذج الانحدار اللوجستي بأنه احد نماذج الانحدار اللاخطية والذي تكون فيه العلاقة بين المتغير (y) متغير الاستجابة (Response Variable) والمتغيرات المؤثرة المستقلة ($x_1, x_2, x_3, \ldots, x_k$) غير خطية ، يبنى أنموذج الانحدار اللوجستي على فرض أساسي هو أن متغير الاستجابة (y) ثنائي الاستجابة يأخذ (1,0) أما النجاح (Success) باحتمال (p_i) او الفشل(Failure) باحتمال (1- p_i) لذلك يكون المتغير (y_i) يتوزع حسب توزيع برنولي Ber(p_i).

أي أن:

$$y_i \sim \text{Ber}(p_i) \qquad i = 1,2, \ldots, n$$

ومن ثم فأن دالة الكثافة الاحتمالية تكون وفق الصيغة الآتية:

$$P_r(Y_i = y_i) = p_i^{y_i} (1 - p_i)^{1 - y_i} \qquad , y_i = 0,1 \qquad (2)$$

أذ أن :

y_i متغير تابع ثنائي الاستجابة (0,1)

p_i احتمال حدوث الاستجابة عندما $y_i = 1$

1-p_i أحتمال عدم حدوث الاستجابة عندما $y_i = 0$

لذلك فأن توقع المتغير y_i يمثل أحتمال حدوث الاستجابة (p_i) وكالآتي:

$$E(y_i) = P_r(Y = 1) = p_i \qquad (3)$$

أما تباين المتغير y_i بالنسبة لتوزيع برنولي كالآتي:

$$v(y_i) = p_i(1 - p_i) \qquad (4)$$

اذا كان لدينا $[y_1, y_2, y_3, \ldots\ldots, y_n]$ تمثل عينة عشوائية من المتغير ثنائي الاستجابة $y_i \in [0,1]$، ولتكن $x_{i1}, x_{i2}, x_{i3}, \ldots, x_{ik}$ تمثل مجموعة المتغيرات المؤثرة، وان n تمثل عدد المشاهدات لهذه المتغيرات التي تكون المصفوفة الآتية:

$$X = (x_{ij})_{n+1*k} \quad (5)$$

X تمثل مصفوفة المتغيرات المستقلة بدرجة ((n+1)*k)

$i = 1,2, \dots, n+1$ عدد المشاهدات $n+1$

$j = 1, 2, \dots\dots, k$ تمثل عدد المتغيرات المؤثرة k

اذ ان احتمال الاستجابة لأنموذج الانحدار اللوجستي يأخذ القيمة (1) وحسب الصيغة الآتية:

$$P_r(y = 1|\, x) = \frac{1}{1+(e^{\beta_0+\sum_{j=1}^{k}\beta_j x_{ij}})^{-1}} \quad (6)$$

وأن احتمال الاستجابة عند القيمة (0) يكون كالاتي:

$$P_r(y = 0|\, x) = \frac{1}{1+e^{\beta_0+\sum_{j=1}^{k}\beta_j x_{ij}}} \quad (7)$$

وبما أن مجموع الاحتمالات يساوي (1) وحسب الصيغة الآتية:

$$P_r(y = 1|\, x) + P_r(y = 0|\, x) \quad = 1 \quad (8)$$

فأنه يمكن التعبير عن أنموذج الانحدار اللوجستي بالصيغة الاتية:

$$y_i = p_i + \varepsilon_i \quad , \ i = 1,2, \dots . n \quad (9)$$

اذ ان

p_i : تمثل دالة الانحدار اللوجستي أو تمثل دالة الاستجابة اللوجستية

εi : يمثل الخطأ العشوائي حيث يكون له متوسط يساوي صفر وكالآتي:

$$\varepsilon_i = y_i - p_i$$
$$E(\varepsilon_i) = E(y_i) - E(p_i) = p_i - p_i = 0 \quad (10)$$

اما تباين حد الخطأ العشوائي فأنه يكون مساوياً الى تباين المتغير y_i ثنائي الاستجابة وكالآتي:

$$v(\varepsilon_i) = v(y_i) = p_i(1 - p_i) \quad (11)$$

حيث تكتب دالة الانحدار اللوجستي (احتمال الاستجابة) حسب الصيغة الآتية:

$$p_i = \frac{e^{\beta_0+\sum_{j=1}^{k}\beta_j x_{ij}}}{1+e^{\beta_0+\sum_{j=1}^{k}\beta_j x_{ij}}} \quad (12)$$

وأن دالة الانحدار اللوجستي التقديرية تكون كالآتي:

$$\hat{p}_i = \frac{e^{\hat{\beta}_0+\sum_{j=1}^{k}\hat{\beta}_j x_{ij}}}{1+e^{\hat{\beta}_0+\sum_{j=1}^{k}\hat{\beta}_j x_{ij}}} \quad (13)$$

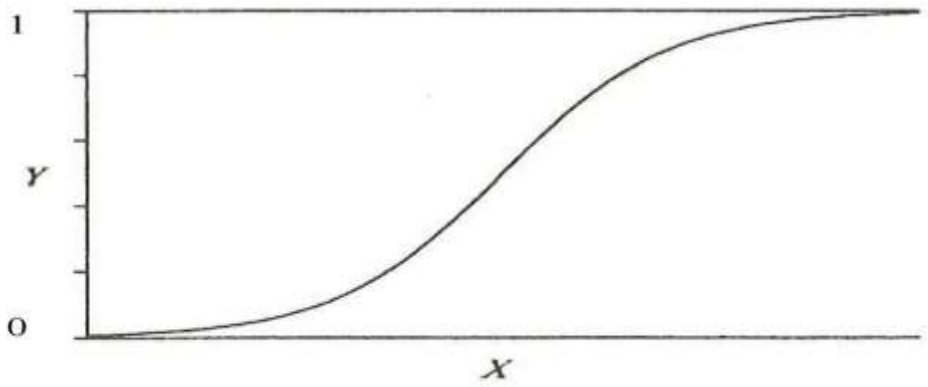
نلاحظ من المعادلة (12) ان شكل العلاقة بين المتغيرات المؤثرة (x_{ij}) واحتمال الاستجابة p_i لا يمكن ان يكون خطياً وإنما تأخذ شكلاً منحنياً أي على شكل حرف (S).

$$1- p_i = \quad 1 - \frac{e^{\beta_0+\sum_{j=1}^{k}\beta_j x_{ij}}}{1+e^{\beta_0+\sum_{j=1}^{k}\beta_j x_{ij}}} \quad (14)$$

حيث أن:

$\beta_0, \beta_1, \beta_2, \dots \dots \dots, \beta_j$: تمثل معلمات مجهولة يراد تقديرها

x_{ij} : متغيرات مؤثرة



شكل (1): العلاقة بين المتغيرين (y,x) بالمنحنى اللوجستي

2.2. طريقة الإمكان الأعظم (Maximum Likelihood Method)[3]

تعتمد هذه الطريقة على ايجاد قيم $\hat{\beta}$ وهي عبارة عن تقديرات للمتجه β التي تجعل الدالة في نهايتها العظمى، وعلى فرض ان لدينا r من المتغيرات المؤثرة $x_1, x_2, \dots, x_r$ وتتوزع ثنائي الحدين بالمعلمتين (n_i, p_i) وان (Y_i) يمثل مجموع حالات النجاح في كل محاولة من (n_i) وان هناك (k) من المتغيرات المؤثرة في كل مجموعة من المجاميع فان دالة الكثافة الاحتمالية لـ (Y_i) هي:

$$P_i(X_i = x_i) = C_{x_i}^{n_i} p_i^{x_i} (1-p_i)^{n_i - x_i} \quad (15)$$

$$i = 1,2, \dots, r \; ; \; x_i = 0,1, \dots, n_i$$

وان

$$\mathrm{E}\,(Y_i) = n_i p_i \quad (16)$$

$$\mathrm{Var}\,(Y_i) = n_i p_i (1-p_i) \quad (17)$$

وان p_i نسبة الاستجابة تقدر كمايأتي:

$$p_i = \frac{x_i}{n_i}$$

$$q_i \; = \; 1 - p_i \; = \; 1 - \frac{x_i}{n_i} \quad = \; (n_i - x_i) \; / \; n_i$$

وان دالة الامكان الاعظم للتوزيع المشترك للبيانات (Y_i) تكون حسب الصيغة:

$$L(P) = \prod_{i=1}^{r} C_{x_i}^{n_i} p_i^{x_i} (1-p_i)^{n_i - x_i} \quad (18)$$

وبالتبسيط

$$L(P) = \prod_{i=1}^{r} C_{x_i}^{n_i} \left[\frac{p_i}{1-p_i}\right]^{x_i} (1-p_i)^{n_i} \quad (19)$$

ولوغاريتم دالة الامكان هو:

$$LnL(P) = \sum_{i=1}^{r} \left[LnC_{x_i}^{n_i} + x_i Ln\left(\frac{p_i}{1-p_i}\right) + n_i Ln(1-p_i) \right] \quad (20)$$

وبعد اجراء بعض العمليات الرياضية للوصول الى انموذج الانحدار اللوجستي نحصل على:

$$P_i = \frac{\exp(\acute{X}_\iota \beta)}{1 + \exp(\acute{X}_\iota \beta)} \quad (21)$$

وان

$$1 - P_i \; = \frac{1}{1 + \exp(\acute{X}_\iota \beta)} \quad (22)$$

وعند تعويض الصيغتين (21) و (22) بالصيغة (20) نحصل على الصيغة الاتية:

$$LnL(P) = \sum_{i=1}^{r} \left\{ LnC_{x_i}^{n_i} + y_i \acute{X}_i \beta + n_i Ln\left(\frac{p_i}{1+\exp(\acute{X}_i\beta)}\right)\right\} \quad (23)$$

ان صيغة الانحدار الثنائي اللوجستي متعدد المتغيرات التوضيحية يوصف بالصيغة الاتية:

$$P(x) = \frac{e^{\beta_0+\sum_{j=1}^{n}\beta_{ij}x_{ij}}}{1+e^{\beta_0+\sum_{j=1}^{n}\beta_{ij}x_{ij}}}$$

ان التحويل $P(x)$ والذي يمثل بحث انموذج انحدار اللوجستك الثنائي ويسمى (logit transformation) ودالة الاحتمالية تمثل بالصيغة:

$$f(x) = ln\left[\frac{P_i}{1-P_i}\right]$$

حيث ان:

$$\frac{P_i}{1-P_i} = \exp\left(\beta_0 + \sum_{j=1}^{n}\beta_j x_{ij}\right)$$

$$Y_{ij} = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}$$

3. انموذج بروبت (Probit) [6]

ان هذا الانموذج يشبه أنموذج *Logit* في طبيعة المتغير التابع اذ يكون متغيرا نوعيا يأخذ صفتين هما الصفر والواحد ويعتمد على دالة الكثافة الاحتمالية $f(x_i\beta)$ ودالة التوزيع التراكمي $F(x_i\beta)$ وانه يتبع التوزيع القياسي بمتوسط $(\mu = 0)$ وتباين $(\sigma^2 = 1)$ فان دالة *PDF,CDF* تكون بالشكل الاتي :

$$f_x = f(x_i\beta) = {\exp(-0.5(x_i\beta)^2)}\Big/{\sqrt{2\pi}} \quad (24)$$

$$F_x = F(x_i\beta) = \int_{-\infty}^{x_i\beta} {\exp(-0.5(x_i\beta)^2)}\Big/{\sqrt{2\pi}}\, dx_i\beta \quad (25)$$

وان الدالة الاحتمالية اللوغاريتمي للأنموذج بروبت تعطى بالصيغة التالية :

$$LLF(\beta) = \sum_{i=1}^{I}(Y_i lnF(x_i\beta) + (1-Y_i)ln[1-F(x_i\beta)])$$

ولو فرضنا ان الانموذج بالشكل الاتي :

$$H = \beta_0 + \beta x_i$$

وبما ان النموذج يتبع التوزيع الطبيعي القياسي فأن يمكن تقدير معلمات وعلى النحو الاتي:

$$P\left(Y = \frac{1}{x}\right) = P(H \le h) = P(Z \le \beta_0 + \beta x_i) = \Phi(\beta_0 + \beta x_i) \quad (26)$$

حيث ان: $P\left(Y = \frac{1}{x}\right)$ يمثل متوسط احتمال حدوث الحدث لمتغير المؤثر x

Z : تمثل Standard Normal

Φ : تمثل دالة التوزيع التراكمي المعيارية وصيغتها يمكن كتابتها بالشكل الاتي:

$$F(H_i) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{H_i} e^{-Z^2/2}\, dZ \;\; = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\beta_0+\beta x_i} e^{-Z^2/2}\, dZ$$

وهنا احتمالية حدوث الحدث تمثل مساحة المنحنى الطبيعي القياسي .

وللحصول على مقدرات المعلمات (β_0, β_i) نأخذ المعكوس لصيغة (26).

$$H = \Phi^{-1}(I) = \Phi^{-1}(P_i) \quad (27)$$

Φ^{-1}: يمثل معكوس دالة *CDF*

4. وصف البيانات وتحليل النتائج

تم جمع البيانات من التقرير السنوي للجهاز المركزي للإحصاءات البيئية للعراق وقد كان حجم عينة البحث 36 شهراً لثلاث سنوات (2017-2019) وتمثلها ثلاث محطات في محافظة بغداد وهي (محطة الوزيرية ومحطة ساحة الاندلس ومحطة السيدية) واخذ البحث متغير الاستجابة المتمثل بالدقائق العالقة بالهواء والتي تضمنت محددات قياس (PM_{10},$PM_{2.5}$)، اما المتغيرات

المؤثرة فهي الغازات ($CO,NO,NO_X,NO2,CH4$) وهذه المؤثرات تم بحث تأثيرها حسب توفر بياناتها لان بعض المؤشرات كانت نسبة الفقدان بالبيانات عالية لذلك تم اخذ المتغيرات المتوفرة بياناتها في كل محطة، اما بالنسبة لبقية المحطات التي تتواجد في بغداد والمحافظات، لم نتمكن من الحصول على البيانات حول المؤشرات التي يتم دراستها بسب العديد من المشاكل والأزمات التي واجهها العراق خل السنوات السابقة واللاحقة وكذلك بسب عطل المحطات وعدم توفر طاقة كهربائية لتشغيلها وغيرها من المشاكل الأخرى.

4.1. وصف المتغيرات

الجسيمات او الدقائق العالقة(SP) Suspended Particulates : يتضمن مصطلح الجسيمات العالقة بالهواء عددا من أنواع الملوثات علي سبيل المثال الدخان الذي يتكون من حبيبات صغيرة من الكربون وتنتج من احتراق غير كامل للمواد الهيدروكربونية وأهمها الفحم والبترول والقطران والأبخرة هي حبيبات صلبة، غالبا تنتج من التكثيف من الحالة الغازية. الضباب يتكون من جزيئات سائلة (ماء حمض نيتريك حمض كبريتيك وغيرها)، او حبيبات الغبار متناهية في الصغر تنشأ عن طريق تكسير وطحن بعض المواد مثل الأحجار والخامات والفحم الخشب والحبوب وغيرها، ومصادرها هي مصدر طبيعي، العواصف التي تهب في المناطق الجافة وشبه الصحراوية وتثير كميات هائلة من الغبار الذي يؤثر بطريقة مباشرة على التنفس، ومصدر صناعي قد يحتوي الغبار الصناعي على مركبات الرصاص والبريليوم والزرنيخ والنحاس والخارصين وذلك يتوقف على نوعية المنشأة الصناعية المسببة للغبار، وأن الوقود الأحفوري هو المصدر الرئيسي للكثير من العوالق في الهواء خاصة الجسيمات الحمضية وتسقط مع الأمطار الحمضية. ان التلوث بالدقائق العالقة واثاره السامة قد تكون ذاتها سامة نتيجة لخواصها الكيميائية أو الفيزيائية وتعمل الدقائق كحوامل لمواد سامة ممتزجة على سطحها، ويعتبر الفحم من المواد القادرة على الامتزاز الكثير من المواد العضوية وغير العضوية بكفاءة عالية وتحملها هي والغازات ذات الآثار السامة مثل ثاني أكسيد الكبريت وأكاسيد النتروجين إلى الرئتين ثم إلى داخل جسم الإنسان وتحدث أضرارا مضاعفة أشد أثرا مما لو كانت موجودة في الهواء غير ممتزجة وان المحدد المحلي للدقائق الأصغر من 10مايكرون PM_{10} هو (100ppm) والدقائق العالقة الأصغر من 2.5 مايكرون هو (25ppm) خلال 24 ساعة.

أول أوكسيد الكربون CO: غاز سام عديم اللون والرائحة مصدره الأساسي في الهواء هو احتراق الوقود الكربوني ويمثل أكبر نسبة من ملوثات الهواء ويختلف تركيز أول أوكسيد الكربون في المناطق العمرانية باختلاف الظروف السائدة وتعتمد أساسا على مدى كثافة حركة المرور الخاصة بالسيارات ويؤثر أول أوكسيد الكربون على الصحة العامة خاصة على هيموجلوبين الدم حيث أن له قابلية شديدة للاتحاد معه ومن ثم فإنه يؤثر تأثيرا خطيرا على عمليات التنفس في الكائنات الحية بما فيها الإنسان ويتسبب في كثير من حالات التسمم اذ ارتفعت نسبته عن المحدد المحلي المقرر له وهو(ppm35) بالساعة.

غاز ثاني أوكسيد الكربون CO_2: زيادته تؤدي إلى صعوبة في التنفس والشعور بالاحتقان مع تهيج للأغشية المخاطية والتهاب القصبات الهوائية وتهيج الحلق، يتكون غاز ثاني أكسيد الكربون من احتراق المواد العضوية كالفحم وزيت البترول، يعتبر غاز ثاني أوكسيد الكربون الناتج من الوقود من أهم الملوثات التي أدخلها الإنسان على الهواء، أن الإسراف في استخدام الوقود أو التقليل من الساحات الخضراء يساهم في ارتفاع نسبة غاز ثاني أوكسيد الكربون في الجو والذي قد يؤدي إلى ارتفاع درجة حرارة الأرض وهو ما يعرف بالاحتباس الحراري.

ثنائي أوكسيد الكبريت SO_2: هو غاز حمضي يعتبر من أخطر ملوثات الهواء فوق المدن والمنشآت الصناعية، يتكون نتيجة عمليات الاحتراق للفحم والبترول والغاز الطبيعي لاحتوائها على كميات كبيرة من الكبريت وهو غاز عديم اللون وذو رائحة كريهة وأثار ضارة إذا ما تواجد بمعدلات تزيد على ثلاثة أجزاء في المليون في الهواء والمحدد الطبيعي له بالساعة هو (0.15ppm).

أكاسيد النتروجين (Nitrogen dioxide=NOx): هي مجموعة من الغازات عالية التفاعل، والتي تحتوي على النتروجين والأوكسجين بكميات مختلفة. وهذه الاكاسيد هي غازات عديمة اللون والرائحة.

أوكسيد النتروجين NO: وهو ملوث خطر في الهواء وهو ذو لون احمر وذو رائحة سامة وخانقة، ويكون مصاحب لعمليات احتراق الوقود في الهواء عند درجات حرارة عالية عندما يكون التبريد سريعا بحيث يمنع تفكك هذه الغازات، تكون هذه الأكاسيد في الهواء الجوي ذاته وتكون المركبات والأجهزة المولدة للطاقة في محطات القوى الكهربائية هي المصدرين الأساسيين لأكاسيد النتروجين حيث إنها تعمل عند درجات حرارة مرتفعة. وتعد اكاسيد النتروجين وخاصة ثاني اوكسيد النتروجين من الغازات المخدشة، كما انه يساعد على الاشتعال.

غاز ثنائي أوكسيد النتروجين NO_2: هذا الغاز وغيره من أكسيد النتروجين تنتج من احتراق المركبات العضوية ومن عوادم السيارات والشاحنات وبعض العمليات الصناعية، ويكون مع بخار الماء في الجو حمضا قويا هو حمض النتريك ويسبب الأمطار الحامضية. وعند وصوله مع بقية أكاسيد النيتروجين إلى طبقة الأوزون يحدث كثيرا من الضرر لهذه الطبقة، وأهم أضرار غاز ثاني أوكسيد النيتروجين انه يعمل تهيج الأغشية المخاطية للعين وللمجاري التنفسية ويسبب أضرارا في الرئة ويحدث ضررا في طبقة الأوزون ويكون الأمطار الحامضية، ومحدده المحلي المسموح به خلال ساعة هو(0.1ppm).

غاز الميثان CH_4 : هو من الغازات الدفيئة غازات الاحتباس الحراري الطويلة الأمد (LLGHGs). وينتج 40 في المائة تقريبا من الميثان المنبعث في الغلاف الجوي عن مصادر طبيعية (مثل الأراضي الرطبة والنمل الأبيض)، بينما تمثل المصادر البشرية المنشأ (مثل الحيوانات المجترة، وزراعة الأرز، واستغلال الوقود الأحفوري (الغاز الطبيعي وصناعة البترول)، ومدافن القمامة، وحرق الكتلة الأحيائية) حوالي 60 في المائة من انبعاثات الميثان.

1- محطة الوزيرية

جدول (1): بيانات مستويات التلوث في الهواء لمحطة الوزيرية

Y	CH4	CO	SO2
Low	1.828	0.895	0.044
Low	1.65	0.585	0.031
Low	1.503	0.507	0.032
Low	1.632	0.602	0.011
Low	1.706	0.595	0.028
Low	1.794	0.6	0.035
Low	1.667	0.722	0.054
Low	1.728	0.762	0.062
Low	1.843	1.284	0.07
Low	1.897	1.04	0.053
Low	1.764	0.806	0.029
Low	1.867	1.222	0.043
high	1.814	0.961	0.044
high	1.834	1.033	0.035
high	1.707	0.612	0.041
Low	1.718	0.5	0.027
high	1.615	0.339	0.02
high	1.635	0.197	0.028
high	1.697	0.276	0.031
high	1.73	0.296	0.018
high	1.69	0.534	0.045
high	1.66	0.846	0.041
Low	1.674	0.525	0.018
high	1.944	1.292	0.032
high	1.758	1.214	0.038
high	1.731	0.885	0.02
high	1.769	0.71	0.027
Low	1.798	0.632	0.02
high	1.851	0.643	0.025
high	1.833	0.574	0.032
high	1.855	0.521	0.036
high	1.854	0.561	0.034
high	1.866	0.649	0.038
high	1.89	0.908	0.067
high	1.982	1.357	0.061
high	1.896	1.266	0.058

جدول (2): يمثل الإحصاءات الوصفية

	N	Minimum	Maximum	Sum	Mean
CH4	36	1.503	1.982	63.680	1.76889
CO	36	.197	1.357	26.951	.74864
SO2	36	.011	.070	1.328	.03689
Valid N(list wise)	36				

جدول (3): يمثل عدد حالات متغير الاستجابة

Variable	Value	Count
Y	high	21(Event)
	Low	15
	Total	36

هذه النتائج تشير الى الانموذج المستخدم (Logit,Probit) ونوع الاستجابة الثنائية (1,0) والتي تمثل (Low, high) وحجم العينة المستخدمة 36، إذ ان 15 من المؤثرات كانت في حالة القراءة Low و21 في حالة high والسائد هو High .

جدول (4): يمثل نتائج الانحرافات للنموذجين

model	Deviance Table (Logit,Probit)					
	Source	**Chi-Square**	**P-Value**	R-Sq	R-Sq(adj)	AIC
Logit	Regression	4.45	0.217	9.10%	2.96%	52.45
	CH4	**4.44**	**0.035**			
	CO	1.33	0.249			
	SO2	0.09	0.765			
Probit	Regression	4.54	0.209	9.28%	3.14%	52.36
	CH4	**4.53**	**0.033**			
	CO	1.35	0.245			
	SO2	0.09	0.766			

الجدول المذكور انفاً يمثل نتائج الانحرافات للنموذجين (Logit ,Probit) والتي تبين تحليل التباين لمعلمات النموذجين والتي وضحتها قيمة P-value وchi-square لكل من متغير المؤثر (SO2,CH4,CO) فنجد ان P-value لـ(CH4) اقل من 0.05 فأنه يدل على وجود حقائق كافية بان معلمة الانموذج المقدرة لا تتفق مع فرضية العدم التي تنص على ($H_0: \beta_1 = \beta_2 =$ $\beta_3 = 0$ VS $H_1: \beta_i \neq \beta_j$)، لذلك لا نقبل هذه الفرضية ونقبل الفرضية البديلة اي ان المعلمة المقدرة معنوية اي هناك تأثير معنوي لغاز CH4.

وان معامل التحديد (R-Sq) نسبته تمثل المتغيرات المؤثرة بالأنموذج، كذلك هناك معامل التحديد المعدل ((R-Sq(adj)) والذي يمثل القيمة التفسيرية لمدى تأثير المتغيرات المؤثرة بالأنموذج وكانت قيمته 2.96% لأنموذج logit وقيمته 3.14% لأنموذج Probit وهذه النسبة طبيعية للعينة المأخوذة ،وكذلك لدينا مقياس اكايكي والذي يقيس كفاءة النموذجين وكلما كانت قيمته اقل فأن الانموذج افضل وان قيمة أنموذج بروبت هي الاقل وهذا يدل على ان الأنموذج كفؤ.

جدول (5): يمثل نتائج معاملات الانموذج المقدرة ومعامل الاختلاف

model	Term	Coef	SE Coef	VIF
Logit	**Constant**	-15.24	8.13	
	CH4	9.87	5.09	1.99
	CO	-2.02	1.79	2.34
	SO2	-9.2	30.7	1.57
Probit	**Constant**	-9.56	4.90	
	CH4	6.18	3.06	1.88
	CO	-1.25	1.09	2.24
	SO2	-5.7	19.2	1.57

الجدول المذكور انفاً يمثل نتائج قيم المعلمات المقدرة والقيمة الثابتة للأنموذج وعامل تضخم التباين للنموذجين وأن المعلمات المقدرة للمتغيرات (CH4,CO,SO2) تمثل التغير في لوغاريتم الاحتمال المقارن ((P(low)/P(high) عندما يزداد المتغير المؤثر وحدة واحدة مع وجود الحد الثابت للدقائق العالقة (اي بمعنى زيادة احد المتغيرات يؤدي الى زيادة احتمال ارتفاع التلوث بالدقائق العالقة)، وكذلك نلاحظ عامل تضخم التباين لمعلمات المقدرة كان اقل من (5) فأن ذلك يشير الى عدم وجود مشكلة التعدد الخطي .

جدول (6): يمثل معدل المخاطرة لأنموذج Logit

Odds	Ratios for Odds Ratio	Continuous Predictors95% CI
CH4	19320.8369	(0.8905,4.19217E+08)
CO	0.1331	(0.0040,4.4413)
SO2	0.0001	(0.0000,1.38791E+22)

هذه النتائج توضح معدل المخاطرة (odd Ratio) وحدود الثقة لمتغيرات المؤثرة وكان اقل معدل المخاطرة لمتغيرين (CO,SO_2) وهو اقل من (1) فان هذا يعني المخاطرة تكون منخفضة عند المقارنة للمستوى المرجعي للعامل وان اعلى معدل مخاطرة لمتغير (CH4) وهذا يدل على هذا المتغير له تأثير عال في تلوث الهواء.

جدول (7): المعادلة التقديرية للاحتمال ارتفاع مستوى التلوث بالدقائق العالقة في الهواء للنموذجين

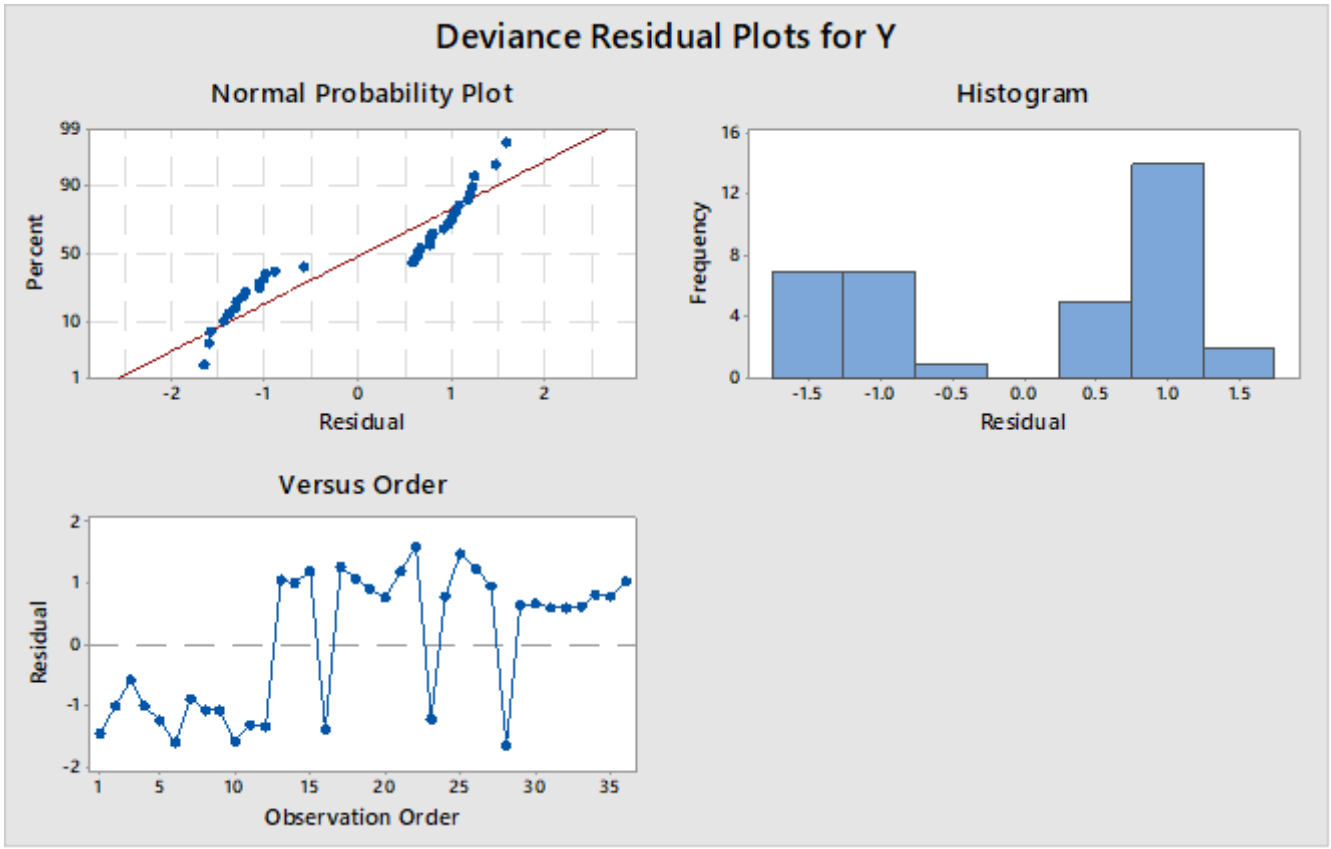
Regression Equation (Logit, Probit)			
P(high)= exp(Y')/(1 + exp(Y'))		P(high)= Φ(Y')	
$Y_L' = -15.24 + 9.87\ CH_4 - 2.02\ CO - 9.2\ SO_2$		$Y_P' = -9.56 + 6.18\ CH_4 - 1.25\ CO - 5.7\ SO_2$	
$PM_{2.5}$	**PM_{10}**	**$PM_{2.5}$**	**PM_{10}**
CH_4=3.5⟶ Y_L' =15.9199	CH_4=11 ⟶ Y_L' =89.9449	CH_4=5⟶ Y_P'= 19.24475	CH_4=18 ⟶ Y_P'=99.58475
CH_4=4⟶ Y_L' =20.8549	CH_4=11.5 ⟶ Y_L' =94.8799	CH_4=6⟶ Y_P'=25.42475	CH_4=18.2 ⟶ Y_P'=100.8207
CH_4=4. 3 ⟶ Y_L' = 26.77686	**CH_4=12. 2 ⟶ Y_L' = 101.7889**	**CH_4=6.2⟶ Y_P'= 26.66075**	**CH_4=18. 3 ⟶ Y_P'=101.4387**

الجدول المذكور انفاً يمثل المعادلة التقديرية لنموذجي الانحدار (Logit ,Probit)، واستعمال النموذجين للتنبؤ بتأثير متغير الاستجابة والذي يمثل الدقائق العالقة لمستوى ($PM_{2.5}$,PM_{10}) وذلك بأخذ اعظم قيمة بالبيانات للمتغيرات المؤثرة وايضاً تم تثبيت المتغيرين المؤثرين (CO,SO_2) للقيم المأخوذة وتغير قيمة المتغير CH_4 نجد ان نسبة التلوث بالدقائق العالقة تزداد عن المحدد المحلي لكل من($PM_{2.5}$,PM_{10}) وحسب النتائج المحددة في الجدول المذكور انفاً.

جدول (8): يمثل نتائج اختبارات حسن المطابقة للنموذجين

model	Test	DF	Chi-Square	P-Value
Logit	**Deviance**	32	44.45	0.071
	Pearson	32	34.97	0.329
	Hosmer-Lemeshow	8	2.46	0.964
Probit	**Deviance**	32	44.36	0.072
	Pearson	32	35.03	0.326
	Hosmer-Lemeshow	8	2.44	0.965

الجدول المذكور انفاً يمثل نتائج اختبارات حسن المطابقة النموذجين ومن هذه الاختبارات اختبار(Pearson) واختبار (Deviance) واختبار (Hosmer-Lemeshow) وهذه الاختبارات تبين مدى توفيق النموذج لبيانات وذلك بالاعتماد على قيمة P-value المرتفع عن مستوى المعنوية المحدد (0.05) وهذا يدل على ان النموذجين كفؤ للبيانات.



شكل (2): يمثل المدرج التكراري والاحتمال الطبيعي لأخطاء الانموذج.

2- محطة ساحة الاندلس

جدول (9): يمثل البيانات للمستويات التلوث في الهواء لمحطة ساحة الاندلس

Y	Nox	NO2	NO	SO2
Law	0.109	0.04	0.077	0.045
Law	0.115	0.037	0.081	0.025
Law	0.045	0.024	0.022	0.03
Law	0.022	0.022	0.017	0.009
Law	0.031	0.031	0.028	0.009
Law	0.037	0.037	0.033	0.012
Law	0.047	0.047	0.038	0.009
Law	0.46	0.046	0.068	0.01
Law	0.058	0.058	0.063	0.012
Law	0.029	0.029	0.03	0.014
Law	0.019	0.019	0.033	0.036
Law	0.029	0.029	0.041	0.043
high	0.067	0.028	0.039	0.05
high	0.057	0.027	0.03	0.029
high	0.049	0.014	0.035	0.012
high	0.03	0.012	0.019	0.013
high	0.034	0.009	0.025	0.01
high	0.03	0.011	0.019	0.01
high	0.039	0.016	0.022	0.011
high	0.036	0.023	0.013	0.012
high	0.052	0.034	0.019	0.016
high	0.049	0.017	0.032	0.015
high	0.068	0.01	0.058	0.012
high	0.045	0.01	0.035	0.012
high	0.058	0.012	0.046	0.034
high	0.072	0.016	0.056	0.027
Law	0.02	0.012	0.008	0.023
Law	0.015	0.009	0.006	0.021
high	0.025	0.015	0.01	0.03
high	0.032	0.019	0.014	0.035
high	0.034	0.016	0.018	0.036
high	0.044	0.018	0.026	0.035
high	0.047	0.026	0.021	0.042
high	0.049	0.032	0.02	0.052
high	0.09	0.022	0.069	0.062
high	0.082	0.011	0.071	0.053

جدول (10): يمثل الإحصاءات الوصفية للمتغيرات المؤثرة

	N	Minimum	Maximum	Sum	Mean
NOx	36	.015	.460	2.125	.05903
NO2	36	.009	.058	.838	.02328
NO	36	.006	.081	1.242	.03450
SO2	36	.009	.062	.906	.02517
Valid N (list wise)	36				

جدول (11): يمثل عدد حالات متغير الاستجابة

Variable	Value	Count
Y	high	22 (Event)
	Low	14
	Total	36

هذه النتائج تشير الى الانموذج المستخدم (Logit,Probit) ونوع الاستجابة الثنائية (1,0) والتي تمثل(Low , high) وحجم العينة المستخدمة 36 ، إذ ان14 من المؤثرات كانت في حالة القراءة Low و22 في حالة high والسائد هو High .

جدول (12): يمثل نتائج الانحرافات للنموذجين

model	Deviance Table(Logit, Probit)					
	Source	**Chi-Square**	**P-Value**	R-Sq	R-Sq(adj)	AIC
Logit	Regression	**14.14**	**0.007**	29.40%	21.09%	43.97
	NOx	0.50	0.481			
	NO_2	10.53	0.001			
	NO	0.11	0.743			
	SO_2	2.03	0.154			
Probit	Regression	13.99	0.007	29.07%	20.76%	44.13
	NOx	0.41	0.522			
	NO_2	10.34	0.001			
	NO	0.04	0.838			
	SO_2	1.89	0.169			

الجدول المذكور انفاً يمثل نتائج الانحرافات للنموذجين (Logit , Probit) والتي تبين تحليل التباين لمعلمات النموذجين والذي وضحتها قيمة P-value وchi-square لكل من متغير المؤثر ($SO2,NO,NO_X,NO_2$) فنجد ان P-value لـ(NO_2) اقل من 0.05 فأنه يدل على وجود حقائق كافية بان معلمة الانموذج المقدرة لا تتفق مع فرضية العدم التي تنص على ($H_0: \beta_1 = \beta_2 =$ $\beta_3 = \beta_4 = 0$ VS $H_1: \beta_i \neq \beta_j$)، لذلك لا نقبل هذه الفرضية ونقبل الفرضية البديلة اي ان المعلمة المقدرة معنوية اي هناك تأثيرا معنويا لغاز NO_2.

وان معامل التحديد (R-Sq) نسبته تمثل المتغيرات المؤثرة بالأنموذج، كذلك هناك معامل التحديد المعدل ((R-Sq(adj)) والذي يمثل القيمة التفسيرية لمدى تأثير المتغيرات المؤثرة بالأنموذج وكانت قيمته 21.09% لأنموذج logit وقيمته 20.76% لأنموذج Probit وهذه النسبة طبيعية للعينة المأخوذة ،وكذلك لدينا مقياس اكايكي والذي يقيس كفاءة النموذجين وكلما كانت قيمته اقل فأن الانموذج افضل وان قيمته لأنموذج بروبت اقل مما يدل على انه كفؤ.

جدول (13): يمثل نتائج معاملات النموذجين المقدرة ومعامل الاختلاف

model	Term	Coef	SE Coef	VIF
Logit	**Constant**	2.70	1.30	
	NOx	6.22	7.90	1.92
	NO_2	-144.1	55.5	1.52
	NO	-9.5	29.2	1.63
	SO_2	42.7	31.4	1.21
Probit	**Constant**	1.518	0.715	
	NOx	3.32	4.79	1.85
	NO_2	-84.2	30.7	1.48
	NO	-3.4	16.8	1.70
	SO_2	25.0	18.8	1.18

الجدول المذكور انفاً يمثل نتائج قيم المعلمات المقدرة والقيمة الثابتة للأنموذج وعامل تضخم التباين للنموذجين وأن المعلمات المقدرة لمتغيرات ($SO2,NO,NO_X,NO_2$) تمثل التغير في لوغاريتم الاحتمال المقارن ((P(low)/P(high)) عندما يزداد المتغير المؤثر وحدة واحدة مع وجود الحد الثابت للدقائق العالقة (اي بمعنى زيادة أحد المتغيرات يؤدي الى زيادة احتمال ارتفاع التلوث بالدقائق العالقة) ، وكذلك نلاحظ عامل تضخم التباين للمعلمات المقدرة كان اقل من (5) فأن ذلك يشير الى عدم وجود مشكلة التعدد الخطي.

جدول (14): يمثل معدل المخاطرة لأنموذج Logit

Odds	Ratios for Odds Ratio	Continuous Predictors95% CI
NOx	501.6355	(0.0001, 2.65934E+09)
NO_2	0.0000	(0.0000,0.0000)
NO	0.0001	(0.0000, 5.19703E+20)
SO_2	3.68069E+18	(0.0000, 1.94782E+45)

هذه النتائج توضح معدل المخاطرة (odd Ratio) وحدود الثقة لمتغيرات المؤثرة وكان اقل معدل المخاطرة لمتغيرين(NO_2, NO) وهو اقل من (1) فان هذا يعني المخاطرة تكون منخفضة عند المقارنة للمستوى المرجعي للعامل وان اعلى معدل مخاطرة لمتغيرين (NOx,SO2) وهذا يدل على ان المتغيرين لهما تأثير عال في تلوث الهواء.

جدول (15): المعادلة التقديرية للاحتمال ارتفاع مستوى التلوث بالدقائق العالقة في الهواء للنموذجين

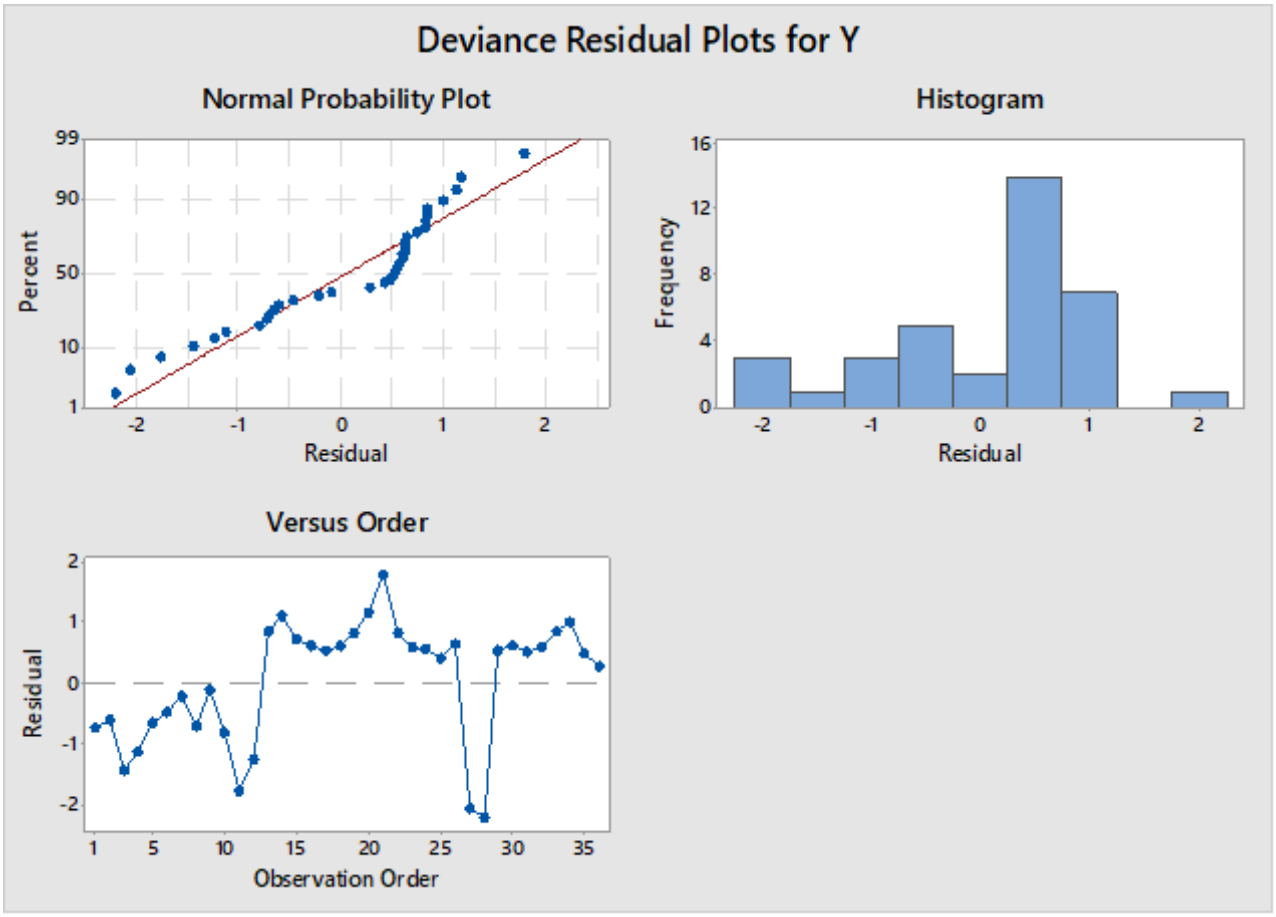
Regression Equation (Logit, Probit)							
P(high)= exp(Y')/(1 + exp(Y'))				P(high)= Φ(Y')			
Y_L' = 2.70 + 6.22 NOx - 144.1 NO_2 - 9.5 NO+ 42.7 SO_2				**Y_P' = 1.518 + 3.32 NOx - 84.2 NO_2 - 3.4 NO + 25 SO_2**			
$PM_{2.5}$		**PM_{10}**		**$PM_{2.5}$**		**PM_{10}**	
SO_2=0.65 Y_L' =24.1889	NOx=4.5 Y_L' =24.2101	SO_2=2.2 Y_L' = 90.3739	NOx= 16.5 Y_L' =98.8501	SO_2=0.9 Y_P'=20.3862	NOx=8 Y_P'=24.469	SO_2=3.8 Y_P'= 92.8862	NOx=29 Y_P'= 94.189
SO_2=0.68 Y_L' =25.4699	NOx=4.7 Y_L'=25.4541	SO_2=2.4 Y_L' = 98.9139	NOx=16.8 Y_L'=100.7161	SO_2=1.1 Y_P'=25.3862	NOx=8.2 Y_P'=25.133	SO_2=4 Y_P'= 97.8862	NOx=31 Y_P'=100.829
SO_2=0.70 Y_L'=26.3239	**NOx=4.8 Y_L'=26.0761**	**SO_2=2.5 Y_L'=103.1839**	**NOx=17 Y_L'=101.9601**	**SO_2=1.2 Y_P'=27.8862**	**NOx=8.5 Y_P'=26.129**	**SO_2=4.2 Y_P'=102.8862**	**NOx=31.1 Y_P'= 101.161**
NOx=0.3 , SO_2=0.7 Y_L'=25.3287		NOx=2.1 , SO_2=2.2 Y_L'= 100.5747		NOx=1.6 , SO_2=0.8 Y_P'= 21.671		NOx=4.3 ,SO_2=3.5 Y_P'= 98.135	
NOx=0.5 , SO_2=0.7 Y_L'=26.5727		**NOx=2.3, SO_2=2.2 Y_L'= 101.8187**		**NOx=2.2 , SO_2=0.9 Y_P'= 26.163**		**NOx=4.7, SO_2=3.6 Y_P'=101.963**	

الجدول المذكور انفاً يمثل المعادلة التقديرية لنموذجي الانحدار (Logit,Probit)، واستعمال النموذجين للتنبؤ بتأثير متغير الاستجابة والذي يمثل الدقائق العالقة لمستوى ($PM_{2.5}$,PM_{10}) وذلك بأخذ اعظم قيمة بالبيانات للمتغيرات المؤثرة وايضاً تم تثبيت المتغيرين المؤثرين (NO_2, NO) على اعظم قيمة بالبيانات وتغير قيمة المتغيرين (NOx,SO2) لكل متغير على حدة وايضاً في حالة تغيرهما معاً، نجد ان نسبة التلوث بالدقائق العالقة تزداد عن المحدد المحلي لكل من($PM_{2.5}$,PM_{10}) حسب النتائج المذكور في الجدول آنفاً.

جدول (16): يمثل نتائج اختبارات حسن المطابقة للنموذجين

model	Test	DF	Chi-Square	P-Value
Logit	**Deviance**	31	33.97	0.326
	Pearson	31	37.40	0.199
	Hosmer-Lemeshow	8	6.04	0.643
Probit	**Deviance**	31	34.13	0.320
	Pearson	31	36.14	0.241
	Hosmer-Lemeshow	8	6.33	0.610

الجدول المذكور انفاً يمثل نتائج اختبارات حسن المطابقة للنموذجين ومن هذه الاختبارات اختبار(Pearson) واختبار (Deviance) واختبار (Hosmer-Lemeshow) وهذه الاختبارات تبين مدى توفيق النموذج للبيانات وذلك بالاعتماد على قيمة P-value المرتفع عن مستوى المعنوية المحدد (0.05) وهذا يدل على ان النموذجين كفؤ للبيانات.



شكل (3): يمثل المدرج التكراري والاحتمال الطبيعي لأخطاء الانموذج

3- محطة السيدية

جدول (17): يمثل البيانات لمستويات التلوث في الهواء لمحطة السيدية

Y	CH4	CO	Nox	NO2	NO	SO2
Low	1.93	0.861	0.113	0.038	0.075	0.036
Low	1.913	0.547	0.067	0.03	0.037	0.017
Low	1.719	0.405	0.033	0.019	0.014	0.015
Low	1.802	0.382	0.036	0.022	0.015	0.031
Low	1.803	0.321	0.042	0.025	0.017	0.025
Low	0.799	0.368	0.066	0.034	0.032	0.038
Low	1.753	0.389	0.073	0.038	0.035	0.044
Low	1.766	0.41	0.075	0.042	0.033	0.047
Low	1.799	0.599	0.091	0.053	0.039	0.062
Low	1.709	1.456	0.048	0.032	0.015	0.031
Low	1.764	0.392	0.04	0.021	0.018	0.011
Low	1.8	0.512	0.064	0.027	0.037	0.021
high	1.845	0.431	0.061	0.025	0.035	0.019
high	1.764	0.477	0.057	0.03	0.027	0.034
high	1.671	0.31	0.028	0.018	0.01	0.012
high	1.62	0.304	0.028	0.019	0.009	0.016
high	1.586	0.311	0.044	0.026	0.018	0.024
high	1.561	0.321	0.065	0.037	0.028	0.049
high	1.569	0.286	0.075	0.037	0.038	0.032
high	1.543	0.289	0.089	0.036	0.053	0.031

high	1.634	0.516	0.084	0.052	0.032	0.054
high	1.716	0.601	0.083	0.045	0.038	0.052
high	1.747	0.434	0.053	0.028	0.025	0.019
Low	1.643	0.352	0.045	0.021	0.026	0.006
Low	1.943	0.493	0.062	0.025	0.037	0.024
Low	1.965	0.427	0.055	0.026	0.03	0.018
Low	2.013	0.275	0.026	0.016	0.009	0.011
Low	1.88	0.254	0.023	0.015	0.008	0.016
high	1.904	0.362	0.06	0.033	0.027	0.035
high	1.944	0.349	0.061	0.039	0.022	0.035
high	1.817	0.299	0.061	0.035	0.026	0.024
high	1.938	0.464	0.088	0.048	0.04	0.043
high	1.896	0.429	0.078	0.043	0.035	0.046
high	1.957	0.451	0.048	0.03	0.019	0.042
high	2.055	0.49	0.048	0.032	0.017	0.034
high	2.069	0.551	0.076	0.027	0.049	0.026

جدول (18): يمثل الإحصاءات الوصفية للمتغيرات المؤثرة

	N	Minimum	Maximum	Sum	Mean
CH4	36	.799	2.069	63.837	1.77325
CO	36	.254	1.456	16.118	.44772
NOx	36	.023	.113	2.146	.05961
NO	36	.008	.075	1.025	.02847
NO2	36	.015	.053	1.124	.03122
SO2	36	0.006	0.062	1.08	0.03
Valid N (listwise)	36				

جدول (19): يمثل عدد حالات متغير الاستجابة

Variable	Value	Count
Y	high	19 (Event)
	Law	17
	Total	36

هذه النتائج تشير الى الانموذج المستخدم (Logit,Probit) ونوع الاستجابة الثنائية (1,0) والتي تمثل(Low , high) وحجم العينة المستخدمة 36 , إذ ان17 من المؤثرات كانت في حالة القراءة Low و19 في حالة high والسائد هو High ،تم استعمال أنموذج الانحدار المتسلسل للنموذجين (Logit,Probit) للتحديد المتغيرات الأكثر تأثيراً.

جدول (20): يمثل نتائج الانحرافات للنموذجين

model	Deviance Table (Logit,Probit)					
	Source	**Chi-Square**	**P-Value**	R-Sq	R-Sq(adj)	AIC
Logit	Regression	7.20	0.027	14.47%	10.45%	48.59
	CO	4.51	0.034			
	NO2	5.17	0.023			
Probit	Regression	8.37	0.039	16.80%	10.78%	49.43
	CH4	1.11	0.293			
	CO	5.60	0.018			
	NO2	6.07	0.014			

الجدول المذكور انفاً يمثل نتائج الانحرافات للنموذجين (Logit,Probit) والتي تبين تحليل التباين لمعلمات النموذجين والذي وضحتها قيمة P-value وchi-square لكل متغير مؤثر بالنسبة لأنموذج الانحدار اللوجستي (NO_2,CO) اما أنموذج بروبت فكانت المتغيرات المؤثرة هي (CH_4,CO,NO_2) فنجد ان P-value لمتغيرات اقل من 0.05 فأنه يدل على وجود حقائق كافية

بان معلمات النموذجين المقدرة لا تتفق مع فرضية العدم التي تنص على ($H_0: \beta_i = \beta_j = 0$ VS $H_1: \beta_i \neq \beta_j$)، لذلك لا نقبل هذه الفرضية ونقبل الفرضية البديلة اي ان المعلمات المقدرة معنوية اي أن هناك تأثيرا معنويا لغاز (CH_4,CO,NO_2).
وان معامل التحديد (R-Sq) نسبته تمثل المتغيرات المؤثرة بالأنموذج، كذلك هناك معامل التحديد المعدل (R-Sq(adj)) والذي يمثل القيمة التفسيرية لمدى تأثير المتغيرات المؤثرة بالأنموذج وكانت قيمته 10.45% لأنموذج logit وقيمته 10.78% لأنموذج Probitوهذه النسبة طبيعية للعينة المأخوذة ،وكذلك لدينا معيار اكايكي والذي يقيس كفاءة النموذجين وكلما كانت قيمته اقل فأن الانموذج افضل وان قيمته لأنموذج بروبت اقل وهذا يدل على انه كفؤ.

جدول (21): يمثل نتائج معاملات الانموذج المقدرة ومعامل الاختلاف

model	Term	Coef	SE Coef	VIF
Logit	Constant	-0.18	1.54	
	CO	-5.95	3.74	1.28
	NO_2	92.2	43.8	1.28
Probit	Constant	-1.98	2.14	
	CH4	1.17	1.18	1.19
	CO	-4.77	2.60	1.56
	NO_2	64.7	27.6	1.46

الجدول المذكور انفاً يمثل نتائج قيم المعلمات المقدرة والقيمة الثابتة للأنموذج وعامل تضخم التباين للنموذجين وأن المعلمات المقدرة لمتغيرات (CH4,CO,NO2) تمثل التغير في لوغاريتم الاحتمال المقارن (P(low)/P(high)) عندما يزداد المتغير المؤثر وحدة واحدة مع وجود الحد الثابت للدقائق العالقة (اي بمعنى زيادة احد المتغيرات يؤدي الى زيادة احتمال ارتفاع التلوث بالدقائق العالقة)، وكذلك نلاحظ عامل تضخم التباين لمعلمات المقدرة كان اقل من (5) فأن ذلك يشير الى عدم وجود مشكلة التعدد الخطي.

جدول (22): يمثل معدل المخاطرة لأنموذج Logit

Odds	Ratios for Odds Ratio	Continuous Predictors95% CI
CO	0.0026	(0.0000,3.9967)
NO2	1.11562E+40	(602.9143, 2.06431E+77)

هذه النتائج توضح معدل المخاطرة (odd Ratio) وحدود الثقة لمتغيرات المؤثرة وكان اقل معدل المخاطرة لمتغيرين (CO) وهو اقل من (1) فان هذا يعني المخاطرة تكون منخفضة عند المقارنة للمستوى المرجعي للعامل وان اعلى معدل مخاطرة لمتغير (NO_2) وهذا المتغير هو احد مركبات المتغير (NO_x)، وهذا يدل على ان هذا المتغير له تأثير عال في تلوث الهواء.

جدول (23): المعادلة التقديرية للاحتمال ارتفاع مستوى التلوث بالدقائق العالقة في الهواء للنموذجين

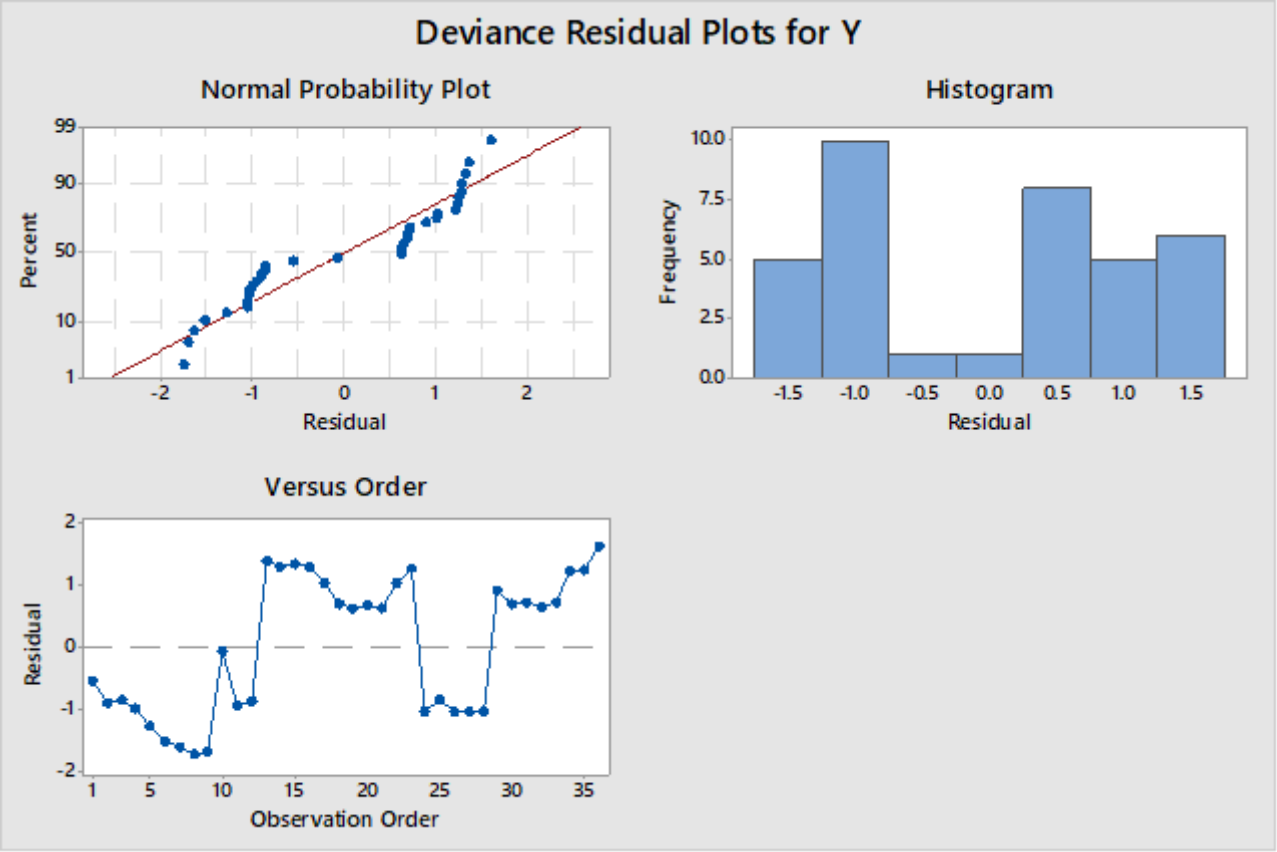
Regression Equation (Logit, Probit)					
P(high)= Φ(Y')				P(high)= exp(Y')/(1 + exp(Y'))	
Y_P' = -1.98 + 1.17 CH_4 - 4.77 CO + 64.7 NO_2				**Y_L' = -0.18 - 5.95 CO + 92.2 NO_2**	
$PM_{2.5}$		**PM_{10}**		**$PM_{2.5}$**	**PM_{10}**
NO_2=0.45 Y_P'=22.61061	CH_4=26 Y_P'= 24.92398	NO_2=1.5 Y_P'= 90.54561	CH_4=90 Y_P'=99.80398	NO_2=0.35 Y_L'=23.4268	NO_2=1 Y_L'=83.3568
NO_2=0.49 Y_P'=25.19861	CH_4=26.5 Y_P'= 25.50898	NO_2=1.6 Y_P'= 97.01561	CH_4=91 Y_P'= 100.974	NO_2=0.37 Y_L'=25.2708	NO_2=1.1 Y_L'=92.5768
NO_2=0.51 Y_P'=26.49261	**CH_4=27 Y_P'=26.09398**	**NO_2=1.7 Y_P'= 103.4856**	**CH_4=91.5 Y_P'= 101.559**	**NO_2=0.38 Y_L'=26.1928**	**NO_2=1.2 Y_L'=101.7968**
NO_2=0.50 CH_4=0.52 Y_P'= 24.03328		NO_2=1.6 CH_4=0.4 Y_P'= 95.06288			
NO_2=0.53 CH_4=0.55 Y_P'= 26.00938		**NO_2=1.7 CH_4 =0.5 Y_P'= 101.6499**			

الجدول المذكور انفاً يمثل المعادلة التقديرية للنموذجين (Logit,Probit)، واستعمال النموذجين لتنبؤ بتأثير متغير الاستجابة والذي يمثل الدقائق العالقة لمستوى ($PM_{2.5}$,PM_{10})، وذلك بأخذ اعظم قيمة بالبيانات للمتغيرات المؤثرة وايضاً تم تثبيت المتغير المؤثر(CO) لأعظم قيمة في البيانات وتغير قيمة المتغيرين (CH_4,NO_2) لكل متغير على حدة وايضاً في حالة تغيرهما معاً، نجد ان نسبة التلوث بالدقائق العالقة تزداد عن المحدد المحلي لكل من ($PM_{2.5}$,PM_{10}) حسب النتائج المذكور في الجدول آنفاً.

جدول (24): يمثل نتائج اختبارات حسن المطابقة للنموذجين

Model	Test	DF	Chi-Square	P-Value
Logit	Deviance	33	42.59	0.122
	Pearson	33	34.42	0.400
	Hosmer-Lemeshow	8	9.24	0.323
Probit	Deviance	32	41.43	0.123
	Pearson	32	33.88	0.377
	Hosmer-Lemeshow	8	8.49	0.387

الجدول المذكور آنفاً يمثل نتائج اختبارات حسن المطابقة للنموذجين ومن هذه الاختبارات اختبار(Pearson) واختبار (Deviance) واختبار (Hosmer-Lemeshow) وهذه الاختبارات تبين مدى توفيق النموذج للبيانات وذلك بالاعتماد على قيمة P-value المرتفع عن مستوى المعنوية المحدد (0.05) وهذا يدل على ان النموذجين كفوءان للبيانات.



شكل (4): يمثل المدرج التكراري والاحتمال الطبيعي لأخطاء الانموذج.

5. الاستنتاجات

تم اجراء بحث نظرية لنماذج الاستجابة الثنائية من خلال أنموذج اللوجت وبروبت (Logit ,Probit) بالاعتماد على التوزيع اللوجستي والطبيعي، وتم تطبيقها على بيانات لعينة حجمها 36 شهراً للفترة (2017-2019) متكونة من ثلاث محطات (الوزيرية، ساحة الاندلس، السيدية)، ومن خلال بناء النماذج المتعددة، تم التوصل الى عدد من الاستنتاجات:

1. اظهرت نتائج محطة الوزيرية من خلال تحليل التباين لمتغيرات النموذجين وبشكل عام نجد ان تأثير المتغير CH_4 كان معنوياً وذو دلالة احصائية فنجد ان P-value له اقل من 0.05 بالإضافة الى اختبار مربع كاي فأنه يدل على وجود حقائق كافية بان معلمة الانموذج المقدرة معنوية اي هناك تأثير لمتغير CH_4 على متغير الاستجابة الذي يمثل الدقائق العالقة، وكذلك اكدت نتائج المقارنة وفق معيار اكايكي ان انموذج (Probit) كفؤ، وكذلك أوضح انموذج Logit ان معدل المخاطرة لمتغير CH_4 كانت عالية وهذا يشير الى ان هذا المتغير له تأثير عال في تلوث الهواء.
2. اظهرت نتائج محطة ساحة الاندلس من خلال تحليل التباين لمتغيرات النموذجين وبشكل عام نجد ان تأثير المتغير NO_2 كان معنوياً وذو دلالة إحصائية، فنجد ان P-value له اقل من 0.05 بالإضافة الى اختبار مربع كاي فأنه يدل على وجود حقائق كافية بان معلمة الانموذج المقدرة معنوية اي هناك تأثير لمتغير NO_2 على متغير الاستجابة الذي يمثل

الدقائق العالقة، وكذلك اكدت نتائج المقارنة وفق معيار اكايكي ان انموذج (Logit) كفؤ ، وكذلك أوضح انموذج Logit ان معدل المخاطرة لمتغيرين (NOx,SO_2) كانت عالية وهذا يوضح ان المتغيرين لهما تأثير في تلوث الهواء.

3. اظهرت نتائج محطة السيدية من خلال تحليل التباين للمتغيرات النموذجين وبشكل عام نجد ان تأثير المتغيرين (NO_2,CO) لأنموذج انحدار Logit والمتغيرات (NO_2,CO,CH_4) لأنموذج Probit ، كانت معنوية وذو دلالة إحصائية فنجد ان P-value لهما اقل من 0.05 بالإضافة الى اختبار مربع كاي فأنه يدل على وجود حقائق كافية بان معلمات الانموذج المقدرة معنوية اي هناك تأثير لكل من (NO_2,CO,CH_4) على متغير الاستجابة الذي يمثل الدقائق العالقة، وكذلك اكدت نتائج المقارنة وفق معيار اكايكي ان انموذج (Logit) كفؤ ، وكذلك أوضح انموذج Logit ان معدل المخاطرة لمتغير (NO_2) كانت عالية ويوضح ان NO_2 له تأثير في تلوث الهواء.
4. أظهرت النتائج التقديرية للمحطات الثلاث باستعمال النموذجين (Logit,Probit) التي تم تقديرها بفرض قيم للمتغيرات المؤثرة التي كانت نسبة المخاطرة لها عالية يلاحظ ان مستوى التلوث بالدقائق العالقة يأخذ بالزيادة عند تلك القيم، حيث ان

- محطة الوزيرية يرتفع مستوى التلوث بالدقائق العالقة عندما تكون حدود مستوى CH_4 لأنموذج Logit ، ولأنموذج Probit .
- محطة ساحة الاندلس يرتفع مستوى التلوث بالدقائق العالقة عندما تكون حدود مستوى (NOx,SO_2) على التوالي لأنموذج Logit ،ولأنموذج Probit .
- محطة السيدية يرتفع مستوى التلوث بالدقائق العالقة عندما تكون حدود مستوى (NO_2) لأنموذج Logit ، اما أنموذج Probit فهناك متغيران مؤثران (NO_2,CH_4) على التوالي هما سبب ارتفاع مستوى التلوث بالدقائق العالقة.

6. التوصيات

1. ان الصعوبة في الحصول على البيانات التي تعد الاساس في التحليل الاحصائي تمثل العائق الاساس والتي تحتوي على عدم التصنيف الملائم او اخطاء التسجيل او عدم انتظامها او نسبة فقدان بالبيانات عالية وغيرها من المعوقات، والتي غالبا ما تقوض البحوث الميدانية، لذا فان التوصية الاساسية تكون من خلال التأكيد على دوائر البيئة في المحافظات والوزارة (وزارة البيئة) العمل على تدوين وتبويب البيانات بصورة علمية ، لكل العراق لما لها من اهمية تمكن الدولة من التخطيط السليم الى مجمل الواقع البيئي في العراق والعمل على تطويره.
2. ضرورة الاعتماد على النماذج التي تم بنائها في البحث لأجل وضع الخطط والسياسات العلمية البيئية لتقليل التلوث والتركيز على المؤثرات البيئية التي تسبب اعلى نسبة تلوث.
3. لابد من العمل على تحديث البحوث في المجال البيئي واستعمال وسائل واساليب رياضية واحصائية تتابع الملوثات البيئية بأنواعها وتلاحظ مدى التطور البيئي في العراق.
4. لدراسة البيانات التي تخص واقع البيئة تتطلب نماذج إحصائية كفؤة فلابد من استعمال نماذج أخرى كنماذج البقاء والمخاطرة لتحديد مستوى المخاطرة بشكل ادق.

المصادر

[1] الراوي، اريج خيري، (2013)، "بحث الاثار البيئية للمولدات الكهربائية في مدينة بغداد-منطقة الكرادة-محلة 903"، مجلة كلية التربية، جامعة واسط، مجلد1، العدد (14)، 314-294.

[2] السراي، احمد عبد الرضا كاظم (2015)، "دراسة تطبيقية لتقدير انموذج الانحدار اللوجستي على استمرار الحياة الزوجية في مدينة الكوت"، بحث دبلوم عالي في الاحصاء التطبيقي، كلية الادارة والاقتصاد، جامعة بغداد.

[3] السعدي، حسين علي، علم البيئة والتلوث، دار اليازوري للنشر، 2002.

[4] أ. م. د. وضاح صبري ابراهيم & ,هاجر فلاح طاهر. (2019). "استعمال أنموذج الانحدار اللوجستي لإيجاد اهم العوامل المؤثرة على مرض سرطان الرئة في العراق للعام 2017"، مجلة الادارة والاقتصاد، المجلد 42، العدد 121، 376-365.

[5] حنوش، علي حسين، البيئة العراقية وسبل حمايتها، ط 1، دار الشؤون الثقافية العامة، 2013.

[6] علي، اسكندر حسين، (2016)، " قياس كفاءة حقول الابقار باستخدام نماذج الاستجابة النوعية "، مجلة العلوم الزراعية العراقية، المجلد47، العدد (2) :PP611-600.

[7] مختار، غزالة و مختار، هاجر الهادي،(2023) التلوث البيئي المشاكل والحلول، بحث الليسانس، جامعة سبها.

[8] Saunois, M., A.R. Stavert, B. Poulter et al., 2019, "The Global Methane Budget 2000–2017", Earth System Science Data, Vol. (12), No. (3)..

[9] Mohammed, M. A., & Raheem, S. H. (2020). Determine of the Most Important Factors that Affect the Incidence of Heart Disease Using Logistic Regression Model (Applied Study in Erbil Hospital). Economic Sciences, Vol. (15), No.(56).

Journal of AL-Rafidain University College for Sciences (2023); Issue 54; 11-29

AL- Rafidain University College

PISSN: (1681-6870); EISSN: (2790-2293)

Journal of AL-Rafidain University College for Sciences

Available online at: https://www.jrucs.iq

JRUCS

Journal of AL-Rafidain University College for Sciences

Constructing A Model to Study the Effect of Some Variables on Air Pollution in Baghdad Governorate for the Period (2017-2019)

Emad H. Abody	Qutaiba N. Nayef
emadhazim@coadec.uobaghdad.edu.iq	dr.qutaiba@coadec.uobaghdad.edu.iq
Statistics Department - College of Administration and Economics - University of Baghdad, Baghdad, Iraq	
Assist. Lect. Maysam S. Khudayir	
maysam.s@albayan.edu.iq	
Department of Accounting - College of Business Administration - Albayan University, Baghdad, Iraq	

Article Information

Article History:
Received: December, 3, 2022
Accepted: February, 25, 2023
Available Online: December, 31, 2023

Keywords:
Air pollution, Binary logistic Regression model, Probit model..

Correspondence:
Maysam S. Khudayir
maysam.s@albayan.edu.iq

Abstract

Pollution is one of the important problems facing humanity at the present time as a result of the increase in human activity in various fields of life, and there are different types of forms of pollution, and air pollution is one of the serious aspects of pollution, so a study was conducted using the logistic regression model, as it is considered one of the effective and appropriate models in The process of descriptive data analysis of the binary response, and the Probit model, which is similar to the binary logistic model in the nature of the dependent variable, because it is a qualitative variable that takes two properties, zero and one, and depends on the probability density function and the cumulative distribution function, and follows the standard normal distribution. The effect of influential variables that represent gases emitted into the air (CO, NO, CH4, SO2, NO2, NOx) on the response variable, which represents suspended particles within the local determinant of pollution within 24 hours (PM10, PM2.5), where the law of low pollution rate was adopted and elevated which takes (0,1) and this variable was converted into a binary response based on the measured proportions of this variable. The study took three stations in the Governorate of Baghdad, which are (Al-Waziriyah, Al-Andalus Square, and Al-Saidiah). The research found that the high level of pollution in suspended particles is due to the influencing variable (CH4) in Waziriya station and the two variables (SO2, NOx) in Meydan Al-Andalus station, and the variable (NO2, CH4) at Saidia station.

https://doi.org/10.55562/jrucs.v54i1.572