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

تحديد اهم أنواع الغازات المسببة لتلوث الهواء باستعمال الخوارزمية الجينية MCD

م.د فاطمة عبد الحميد جواد البيرماني	م.د محمد حسين عبد الحميد جواد البيرماني
fatimah.a@coadec.uobaghdad.edu.iq	mohammed.h@coagri.uobaghdad.edu.iq
كلية الإدارة والاقتصاد - جامعة بغداد، بغداد، العراق	كلية علوم الهندسة الزراعية - جامعة بغداد، بغداد، العراق

المستخلص

يعتبر تلوث الهواء من اهم العناصر الرئيسية التي تؤثر على حياة الانسان، ونظرا للظروف البيئية التي يمر بها البلاد فقد دعت الحاجة الى دراسة لأنواع الغازات المنبعثة والمسببة للتلوث والتي تؤثر على الهواء، وذلك من خلال استعمال الخوارزمية الجينية Minimum Covariance Determinant المختصر MCD. فقد طبقت الدراسة على مجموعة من الغازات والمواد الملوثة المقاسة كمعدلات شهرية في محافظة بغداد لثلاث محطات (الوزيرية، ساحة الاندلس، السيدية) لسنة 2019، اخذت البيانات من الجهاز المركزي للإحصاء لأربعة انواع من الغازات (غاز الميثان CH_4، غاز اول اوكسيد الكاربون CO، غاز النتروز NO_2، غاز ثاني اوكسيد الكبريت SO_2). ولتحديد اهم الغازات المسببة لتلوث الهواء وذلك باستخدام اصغر محدد لمصفوفة التباين المشترك ومن خلال مقارنة التباين لكل نوع وجد ان غاز اول اوكسيد الكاربون CO اكثر الغازات تلويثا للهواء ثم يليه غاز النتروز NO_2 ثم غاز ثاني اوكسيد الكاربون SO_2 ثم غاز الميثان CH_4 الذي يعتبر اقل الغازات تلوث للهواء .

معلومات البحث

تواريخ البحث
تاريخ تقديم البحث: 2022/12/14
تاريخ قبول البحث: 2023/2/28
تاريخ رفع البحث على الموقع: 2023/12/31

الكلمات المفتاحية
الخوارزمية الجينية، الوضوح، المحدودية، امكانية الحل (الفعالية)، اصغر محدد للتباين المشترك

للمراسلة:
م.د فاطمة عبد الحميد جواد البيرماني
fatimah.a@coadec.uobaghdad.edu.iq

https://doi.org/10.55562/jrucs.v54i1.580

الاطار النظري

المقدمة

في كثير من الظواهر تكون عدد المشاهدات قليلة اي ان حجم العينة n صغيرة والمشاهدات ليس لها توزيع، لذا فعند استخدام الطرق التقليدية للتقدير لا تعطي نتائج تفسر الظاهرة بشكل واضح . لذا تم استخدام الطرق اللامعلمية في ايجاد النتائج حيث طبقت الخوارزمية الجينية Genetic Algorithm تكتب بشكل مختصر GA في ايجاد الحل الامثل للظاهرة. من الطرق اللامعلمية المستخدمة الخوارزمية الجينية Minimum Covariance Determinant تكتب بشكل مختصر MCD لتقدير التباين حيث تعتمد هذه الخوارزمية على اصغر محدد لمصفوفة التباين المشترك لايجاد الحل الامثل.

من الباحثين الذين استخدموا الخوارزمية الجينية MCD في عام (2003) قدم الباحثان (Biman chakraborty) و (Probal Chaudhuri) [1] عدة طرق لتقدير موقع الوسيط لأنموذج متعدد المتغيرات اللامعلمي وخاصة فيما يتعلق بالبيانات ذات الابعاد الكبيرة، وتم اللجوء الى الخوارزميات لتقليل الجهد والوقت، ومن هذه الطرق هو التقدير باستعمال اصغر محدد لمصفوفة التباين المشترك (minimum covariance determinant) الصيغة المختصرة (MCD) وقد تمت المقارنة بين خوارزمية (FAST-MCD) والخوارزمية الجينية النخبوية (Elitism) .

في عام (2003) قام الباحثون (Peter J.Rousseeuw) و(Van Aelst Stefan) و(Katrien Van Driessen) و(Jose Agullo) [2] بتقديم الطرق الحصينة للتقدير لأنموذج انحدار متعدد المتغيرات مستندا بذلك على الموقع ومصفوفة الانتشار باستعمال المتغيرات التوضيحية والاستجابة في التقدير، وقد استعمل اسلوب المحاكاة لمختلف احجام العينات ولزيادة الكفاءة اقترح عدة اوزان للتقدير، وهي اوزان الى الموقع (reweighting location –L) واوزان للقيم الاولية لأنموذج الانحدار (reweighted regression estimators – R) وكانت اكثر كفاءة بالمقارنة بينها وبين التقديرات الى (MCD)، اما

عند الدمج في حالة الوزن للموقع والانموذج (LR-weighted MCD) اضافت خصائص للتحيز bias متوسط مربع الخطأ(MSE) اقل قيمة بالمقارنة مع التقدير بطريقة (Least square regression) .

في عام (2003) قدم الباحثان (Greet Pison) و (Stefan Van Aelst) [3] اقترحت عدة مقدرات حصينة منها محدد اصغر لمصفوفة التباين المشترك (MCD) الخطوة الاولى في تقدير موقع الوسيط هو اعادة الوزن الى الوسط الحسابي ومصفوفة التباين المشترك وذلك لزيادة الكفاءة وسميت اعادة وزن محدد اصغر لمصفوفة التباين المشترك (Reweighted Minimum Covariance Determinant) وتكتب باختصار (RMCD)، وتم اختيار نقطة التوقف 50% عندما تكون $\gamma = 0.5$ اما نقطة التوقف 25% عندما تكون $\gamma = 0.75$.

في عام (2004) قام الباحثان (Johanna Hardin) و (David Mirocke) [4] بتقسيم العينة الى النصف، ثم تجزئة العينة الى عناقيد (clusters) كل عنقود n_i يتم ايجاد المسافة واختيار اصغر مسافة وقد تم استخدام برنامج (CLUSTER) او (CLUSTER algorithm) اذ يعتبر اول عنقود حصين، اما (MCD) يتم ايجاده لكل عنقود بصورة منفصلة وذلك من خلال ايجاد نصف العينة العنقودية ($h_i=\frac{n_i+p+1}{2}$) وبهذا لا يكون اختيار العينة بصورة عشوائية.

في عام (2009) م قام كل من الباحثين (Ella Roelant) و (Stefan Van Aelst) و (Gert Willems) [5] بتقدير الموقع والتشتت لبيانات متعدد المتغيرات مع مجموعة من الاوزان بالاعتماد على رتبة المسافة (Mahalanobis)، حيث تم استعمال دالة الوزن a_n وقد فسر ان مقدر (MWCD) اكثر عموما من مقدر (MCD) والاعتماد على دالة الوزن يمكن ان يعالج المتوسط لمختلف القيم المتطرفة .

في عام (2009) م استعمل الباحث (Muthukrishnan.R,M.Radha) و (Kanneswari) [6] البرنامج (R) في تقدير (MCD) ولاحظ كفاءة المقدرات الحصينة بالمقارنة مع الطريقة الكلاسيكية .

الخوارزميات الجينية (Genetic Algorithm)

- **مقدمة في الخوارزميات [1]**

الخوارزمية هي مجموعة محددة من التعليمات (خطوات الحل) التي تؤدي الى انجاز وظيفة (مهمة) معينة ويجب ان تتوفر فيها الشروط التالية:

1. المدخلات (INPUT) : واحد او اكثر من قيمة .
2. المخرجات (OUTPUT) : قيمة واحدة على الاقل.
3. الوضوح (DEFINITENESS): كل خطوة في الخوارزمية واضحة المعاني وغير غامضة اي يجب ان تفهم من قبل الجميع.
4. المحدودية (FINITENES): كل خطوات الخوارزميات يمكن حلها في فترة زمنية محددة .
5. امكانية الحل (الفعالية) (EFFECTIVENESS): كل خطوة تكون ممكنة الحل (فعالة) .

ومن الخوارزميات التي سيتم في هذا البحث التطرق اليها خوارزمية :

✓ **مقدراصغر محدد للتباين المشترك**

(Minimum Covariance Determinant(MCD) Estimator)

1. يتم تجزئة البيانات الى مجموعات جزئية بحجم h ويمكن ايجاد قيمتها كالاتي: [7]

$$h \simeq \frac{3n}{4} \quad (1)$$

تستعمل هذه الطريقة عندما تكون حجم العينة صغيرة ، يتم تقدير الوسط الحسابي ومصفوفة التباين المشترك من حجم العينة h اما افضل نقطة توقف تكون مساوية الى γ=0.25 .

اما تجزئة البيانات تكون كالاتي: [2]

$x = (x_1, \dots, x_p)^T$

$y = (y_1, \dots, y_q)$

p : تمثل عدد المتغيرات المستقلة

q : تمثل عدد متغيرات الاستجابة

$z_n = \{z_i, i = 1, \cdots, n\} \in \mathbb{R}^{p+q}$

حيث يتم ايجاد (MCD) الى المجموعات الجزئية $\{z_{i1}, \cdots, z_{ih}\}$

لتقدير الوسط الحسابي:

$$t_n = \frac{1}{h}\sum_{j=1}^{h} x_{ij} \quad (2)$$

ومصفوفة التباين المشترك

$$C_n = \frac{1}{h}\sum_{j=1}^{h}(x_{ij} - t_n)(x_{ij} - t_n)' \quad (3)$$

لذا يمكن كتابة الحل الامثل الى ايجاد (MCD) كالاتي :-[8]

Minimize $K(X_{i1}, \cdots, X_{ih}) = \det[cov(X_{i1}, \cdots, X_{ih})]$

X_{ih} : المتغيرات المستقلة للعينة الجزئية h

اذا كانت h=n و p=1 النموذج يمثل (Univariate Model)

$\{X_{i1}, \cdots, X_{ih}\} \subset \{X_i, \cdots, X_n\}$

$X_1, \cdots, X_n \in \mathbb{R}^d$

لايجاد (MCD) يمثل T الوسط الحسابي و S تمثل مصفوفة التباين المشترك للعينة n .

اما عندما $P \geq 2$ يكون (Multivariate Model)، يتم ايجاد المسافة (Mahalanobis Distance) بالاعتماد على المتغيرات التوضيحية كالاتي:

$$MD(X_i) = \sqrt{(X_i - T_0)' S_0^{-1} (X_i - T_0)} \quad (4)$$

T_0 تمثل الوسط الحسابي و S_0 تمثل مصفوفة التباين المشترك

$X_n = \{X_1, \cdots, X_n\}$

$H_1 \subset \{1, \cdots, n\}$

H_1 تمثل العينة الجزئية الاولى

$$T_1 = \frac{1}{h}\sum_{i \epsilon H_1} X_i \quad (5)$$

$$S_1 = \frac{1}{h}\sum_{i \epsilon H_1}(X_i - T_1)(X_i - T_1)' \quad (6)$$

تمثل T_1 الوسط الحسابي وتمثل S_1 مصفوفة التباين المشترك للعينة الجزئية الاولى (H_1) .

نجد المسافة (Mahalanobis Distance) (MD) الى H_1

$$MD_1(i) = \sqrt{(X_i - T_1)' S_1^{-1} (X_i - T_1)} \quad (7)$$

$\{MD_1(i) : i \in H_2\} = \{(MD_1)_{1:n}, \cdots, (MD_1)_{h:n}\} \qquad i = 1, \cdots, n$

لايجاد المسافة الى H_2 تمثل العينة الجزئية الثانية

$(MD_1)_{1:n} \leq (MD_1)_{2:n} \leq \cdots \leq (MD_1)_{n:n}$

نجد T_2 و S_2 للعينة H_2

اذا كان $|S_1| \geq |S_2|$

$T_2 = T_1$ و $S_2 = S_1$ و $|S_1| \neq 0$

اذن لتقدير موقع الوسيط يتم الاعتماد على اصغر مسافة ومحدد لمصفوفة التباين المشترك وبتطبيق الخوارزمية (C – Step) كالاتي: [9]

تقسم العينة الى مجموعات جزئية (H_i) لكل عينة نجد h الموضحة في المعادلة (1) ثم نجد الوسط الحسابي ومصفوفة التباين المشترك الموضحة في المعادلة (2) و(3) على التوالي ثم يتم ايجاد المسافة، ترتب المسافة تصاعديا ويتم اخذ العينة التي لها اصغرمسافة ومحدد وفيما يلي خطوات الخوارزمية (C-Step) : .[10]

1. نجد المسافة

$MD_{old}(i) \quad , \quad i = 1, \cdots, n$

2. نأخذ التباديل الى المسافة (ترتيب المسافة تصاعديا)

$MD_{old}(\pi_1) \leq MD_{old}(\pi_2) \leq \cdots \leq MD_{old}(\pi_n)$

3. $H_{new} = \{\pi_1, \cdots, \pi_h\}$

4. $T_{new} = ave(H_{new}), S_{new} = cov(H_{new})$

لكل $MD_{old}(i)$ نكرر (C – STEP) لحين الحصول على $|S_2| = 0$ او $|S_1| = |S_2|$ عدا ذلك يتم ايجاد $|S_3|$ فتكون السلسلة $|S_1| \geq |S_2| \geq |S_3|$ قد تتكرر(C- STEP) الى m من المرات لحين الحصول على $|S_m| = 0$ او $|S_m| = |S_{m-1}|$ اي لحين الحصول على التقارب (CONVERGE) بعد ذلك نكرر C- Step الى (T_m, S_m) فتكون السلسلة $|S_1| \geq |S_2| \geq \cdots \geq |S_{m-1}| \geq |S_m|$ ويكون الحل لأصغر محدد للتباين المشترك .

الاطار العملي

في هذا الفصل سوف يتم تطبيق ما تم التطرق اليه في الجانب النظري باستخدام بيانات تشمل اربعة انواع من الغازات التي تسبب تلوث الهواء تمثلت بـ (غاز الميثان CH_4، غاز اول اوكسيد الكاربون CO، غاز النتروز NO_2، غاز ثاني اوكسيد الكبريت (SO_2)، اخذت البيانات من 3 محطات وهي (محطة الوزيرية، محطة ساحة الاندلس، محطة السيدية) لسنة 2019 تمثل المعدلات الشهرية لتلوث الهواء في محافظة بغداد وقد اخذت من الجهاز المركزي للإحصاء، باستخدام برنامج Matlab لإيجاد النتائج (البيانات موضحة في الملحق). وقد تم استخدام خوارزمية MCD لتحديد اكثر الغازات المسببة لتلوث الهواء من خلال تحديد اصغر محدد لمصفوفة التباين المشترك .

خوارزمية اصغر محدد لمصفوفة التباين المشترك MCD

$$y_{ij}=x'_{ij}\beta(t_{ij}) + \varepsilon_i(t_{ij})$$

اذ ان:

Y_i : يمثل المتغير المعتمد (يمثل تلوث او عدم تلوث الهواء)

X : يمثل المتغير المستقل (الغازات المسببة للتلوث) .

$X=(X_1, X_2, X_3)$

$X_i=x_{ijk}$,

حيث ان:

تمثل عدد المحطات i=1,2,3

تمثل الاشهر J=1,2,…,n_i $(n_i=12)$

تمثل الغازات الملوثة K=1,2,3,4

ويمكن اعادة كتابتها كالاتي:

$Y_{1j}=(y_{11}, \dots, y_{1n_i})$,$Y_{2j}=(y_{21}, \dots, y_{2n_i})$, $Y_{3j}=(y_{31}, \dots, y_{3n_i})$

$X_1=(x_{11k}, \dots, x_{1n_ik})$,$X_2=(x_{21k}, \dots, x_{2n_ik})$, $X_3=(x_{31k}, \dots, x_{3n_ik})$

- الخطوة الاولى: تتم تجزئة حجم العينة N=36 الى 4 مجموعات جزئية وحسب صيغة المعادلة (1) كالاتي:

h=9 حيث ان N=n_1, n_2, n_3, n_4 والتي تمثل المحطات التي تقيس تلوث الهواء .

- الخطوة الثانية : نقدر الوسط الحسابي للمشاهدات h الموضحة في المعادلة (2) ثم نجد مصفوفة التباين المشترك بموجب المعادلة (3) والموضحة نتائجها كالاتي:

M_1 = [1.8128 , 0.7099 , 0.0260 , 0.0300] , M_2 = [1.9250 , 0.6377 , 0.0251 , 0.0396] ,
M_3 = [1.8426 , 0.5270 , 0.0213 , 0.0370] , M_4= [1.9400 , 0.4054 , 0.0336 , 0.0343]

حيث ان M_i و i= 1,2,3,4 الوسط الحسابي لمجموعات العينات المجزأة .

اما التباين موضحة نتائجه كالاتي:

S_1=[0.0025 , 0.0469 , 305500e-05 , 5.2250e-05] , S_2=[0.0022 , 0.1840 , 2.9561e-04 , 3.0903e-04] ,
S_3=[0.0487 , 0.0609 , 4.3250e-05 , 2.9275e-04] , S_4=[0.0065 , 0.0092 , 9.1528e-05 , 1.0525e-04]

ولإيجاد الحل الامثل من خلال ايجاد اصغر محدد لمصفوفة التباين المشترك والموضحة نتائجها كالاتي:

ويمكن ترتيب محدد مصفوفة التباين المشترك $S_2 \leq S_3 \leq S_4 \leq S_1$

Min(S_i) = [3.3378e-14 , 1.4005e-13 , 2.3796e-12 , 3.1570e-14]

يتضح من النتائج اعلاه ان اصغر محدد لمصفوفة التباين المشترك تمثل في S_2 = 1.4005e-13 ولتحديد اهم الغازات المسببة للتلوث تمثل بغاز النتروز NO_2 وهو اقل الغازات تلوث للهواء حيث كانت قيمة تباينه (2.9561e-04) يليه غاز ثاني اوكسيد الكبريت SO_2 وقيمة تباينه (3.0903e-04) ثم غاز الميثان CH_4 وقيمة تباينه (0.0022) اما اكثر الغازات تلويثا للهواء هو غاز اول اوكسيد الكاربون CO و قيمة تباينه (0.184). اما المرتبة الثانية لمحدد مصفوفة التباين المشترك -2.3796e S_3 = 12 وقد كانت نتائجه غاز النتروز NO_2 وهو اقل الغازات تلوث للهواء حيث كانت قيمة تباينه (4.3250e-05) يليه غاز ثاني اوكسيد الكاربون SO_2 وقيمة تباينه (2.9275e-04) ثم غاز الميثان CH_4 وقيمة تباينه (0.0487) اما اكثر الغازات تلويثا للهواء هو غاز اول اوكسيد الكاربون CO و قيمة تباينه (0.0609). اما المرتبة الثالثة لمحدد مصفوفة التباين المشترك -3.1570e S_4= 14 وقد كانت نتائجه غاز النتروز NO_2 وهو اقل الغازات تلويثا للهواء حيث كانت قيمة تباينه (9.1528e-05) يليه غاز ثاني اوكسيد الكاربون SO_2 وقيمة تباينه (1.0525e-04) ثم غاز الميثان CH_4 وقيمة تباينه (0.0065) اما اكثر الغازات تلويثا للهواء هو غاز اول اوكسيد الكاربون CO و قيمة تباينه (0.0092).

اما المرتبة الرابعة لمحدد مصفوفة التباين المشترك S_1= 3.3378e-14 وقد كانت نتائجه غاز النتروز NO_2 وهو اقل الغازات تلويثا للهواء حيث كانت قيمة تباينه (3.5500e-05) يليه غاز ثاني اوكسيد الكبريت SO_2 وقيمة تباينه (5.2250e-05) ثم غاز الميثان CH_4 وقيمة تباينه (0.0025) اما اكثر الغازات تلويثا للهواء هو غاز اول اوكسيد الكاربون CO و قيمة تباينه (0.0469)

الاستنتاجات

من خلال ايجاد اصغر محدد لمصفوفة التباين المشترك تصاعديا نستنتج ان اكثر الغازات المسببة لتلوث الهواء كان غاز اول اوكسيد الكاربون CO ولجميع المصفوفات المجزأة ، حيث ان مصادر هذا الغاز ناتج من عوادم السيارات والشاحنات وكذلك من المولدات وغيرها من المعدات التي تعمل بالبنزين والتي تعتبر المسبب الرئيسي لانبعاث غاز اول اوكسيد الكاربون وتلوث الهواء. اما اقل الغازات تلويثا للهواء كان غاز النتروز NO_2 ، ويليه غاز ثاني اوكسيد الكبريت SO_2 ، ثم غاز الميثان CH_4 .

التوصيات

1. تخفيف المحروقات وتنظيم عمليات الاستيراد للسيارات.
2. اجراء دراسة لإيجاد اكثر المحافظات تلوثا للهواء.

المصادر

[1] Chakraborty, B. and Chaudhuri, P.,(2003),"On The Use Genetic Algorithm With Elitism in Robust and Nonparametric Multivariate Analysis", Austrian Journal of Statistics,Vol.32, pp,13-27.

[2] Peter J. Rousseeuw, Stefan Van Aelst, Katrien Van Driessen, and Jose Agullo, "Robust Multivariate Regression", Vol. 17 , September, pp.1-31,2003.

[3] Greet Pison and Stefan Van Aelst, " Diagnostic Plots for Robust Multivariate Methods ", Journal of Computational and Graphical Statistics, April. 24, pp.1-21,2003.

[4] Johanna Hardin, David Mirocke , "Outlier detection in the multiple cluster setting using the minimum covariance determinant estimator", computational Statistics and Data Analysis, Vol. 44, August, pp.625-638,2004 .

[5] Ella Roelant, Stefan Van Aelst., Gert Willems, "The Minimum Weighted Covariance Determinant Estimator", DOI10.1007/s00184-0186-3,pp.70:177-204, 2009.

[6] Muthukrishnan R. , M. Radha and T. kanneswari, "Robust Estimators in Multivariate Models using R", Interstat. journal, 2009 .

[7] طعمة، د. حسن ياسين، شعبان، هند رستم محمد و حسن ثابت كرماشة، تحليل وتصميم الخوارزميات،

[8] Peter J. Rousseeuw, "Least Median of Squares Regression", Journal of the American statistical Association, December ,Vol. 79, No. 388, pp.871-880 .1984.

[9] Eric A. Cator, Hendrik P. Lopuhaä, "Asymptotic expansion of the minimum covariance determinant estimators", Journal of Multivariate Analysis, Volume 101, Issue 10, 2010.

[10] Rousseeuw, Peter J., and Katrien van Driessen. "A Fast Algorithm for the Minimum Covariance Determinant Estimator." *Technometrics*, Vol. 41, No. 3, 1999, pp. 212–23. *JSTOR*, https://doi.org/10.2307/1270566.

ملحق البيانات

month	methane gas CH_4	carbon monoxide CO	nitrous gas NO_2	carbon dioxide gas SO_2	STATION
1.000	1.758	1.214	023.	038.	1.000
2.000	1.731	885.	032.	020.	1.000
3.000	1.769	710.	020.	027.	1.000
4.000	1.798	632.	019.	020.	1.000
5.000	1.851	643.	020.	025.	1.000
6.000	1.833	574.	025.	032.	1.000
7.000	1.855	521.	028.	036.	1.000
8.000	1.854	561.	033.	034.	1.000
9.000	1.866	649.	034.	038.	1.000
10.000	1.890	908.	045.	067.	1.000
11.000	1.982	1.357	050.	061.	1.000
12.000	1.896	1.266	048.	058.	1.000
1.000	1.882	459.	012.	034.	2.000
2.000	1.891	498.	016.	027.	2.000
3.000	1.990	269.	012.	023.	2.000
4.000	1.912	245.	009.	021.	2.000
5.000	1.987	383.	015.	030.	2.000
6.000	1.895	354.	019.	035.	2.000
7.000	1.752	360.	016.	036.	2.000
8.000	1.819	398.	018.	035.	2.000
9.000	1.959	443.	026.	042.	2.000
10.000	2.058	553.	032.	052.	2.000
11.000	1.737	1.105	022.	062.	2.000
12.000	1.337	689.	011.	053.	2.000
1.000	1.943	493.	025.	024.	3.000
2.000	1.965	427.	026.	018.	3.000
3.000	2.013	275.	016.	011.	3.000
4.000	1.880	254.	015.	016.	3.000
5.000	1.904	362.	033.	035.	3.000
6.000	1.944	349.	039.	035.	3.000
7.000	1.817	299.	035.	024.	3.000
8.000	1.938	464.	048.	043.	3.000
9.000	1.896	429.	043.	046.	3.000
10.000	1.957	451.	030.	042.	3.000
11.000	2.055	490.	032.	042.	3.000
12.000	2.069	551.	027.	026.	3.000

المعدلات الشهرية لمجموعة من الغازات الملوثة في محافظة بغداد لسنة 2019

الجهاز المركزي للاحصاء

المصدر : وزارة الصحة والبيئة /القطاع البيئي /دائرة التخطيط والمتابعة

1 : يمثل محطة الوزيرية ،2 : يمثل محطة ساحة الاندلس ، 3 : يمثل محطة السيدية

Journal of AL-Rafidain University College for Sciences (2023); Issue 54; 100 – 106

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

Determine the Most Important Types of Gases That Cause Air Pollution Using the MCD Genetic Algorithm

Dr. Fatimah A. Jawad Al-Bermani	Dr. Mohammad H. Jawad Al-Bermani
fatimah.a@coadec.uobaghdad.edu.iq	mohammed.h@coagri.uobaghdad.edu.iq
College of Administration and Economics - Baghdad University, Baghdad, Iraq	College of Agricultures Engineering sciences - Baghdad University, Baghdad, Iraq

Article Information

Article History:
Received: December, 14, 2022
Accepted: February, 28, 2023
Available Online: December, 31, 2023

Keywords:
Genetic Algorithm, Definiteness, Finitenes, Effectiveness, Minimum Covariance Determinant

Correspondence:
Dr. Fatimah A. Jawad Al-Bermani
fatimah.a@coadec.uobaghdad.edu.iq

Abstract

One of the primary factors affecting human life is air pollution, and due to the environmental conditions that the country is going through, it was necessary to investigate the different gases that are released into the atmosphere and contribute to air pollution by using the genetic algorithm Minimum Covariance Determinant (MCD).

The study was conducted on a set of gases and pollutants that were measured as monthly averages at three sites (Al-Waziriyah, Al-Andalus Square, and Al-Saydiya) in Baghdad Governorate in 2019. Data were collected from the Central Statistical Organization for four different types of gases: (Methane gas CH_4 , Carbon Monoxide gas CO , Nitrous gas NO_2 , and Sulfur Dioxide gas SO_2) .

To identify the most significant gases that contribute to air pollution, use the smallest determinant of the covariance matrix and by comparing the variance for each type. The most polluting gas was discovered to be carbon monoxide (CO), which was followed by nitrous gas NO_2, carbon dioxide gas SO_2, and methane gas CH_4, which is regarded as the least polluting gas.

https://doi.org/10.55562/jrucs.v54i1.580