JRUCS
Journal of AL-Rafidain University College for Sciences

PISSN: (1681-6870); EISSN: (2790-2293)
مجلة كلية الرافدين الجامعة للعلوم
Available online at: https://www.jrucs.iq

AL- Rafidain University College

تحديد اهم العوامل المؤثرة في تلوث الهواء باستخدام طريقة معكوس لاسو

أ.م. سيف حسام رحيم	أ.د. جاسم ناصر حسين
saif.hosam@qu.edu.iq	Jasim.nasir@uokerbala.edu.iq
قسم الاحصاء - كلية الادارة والاقتصاد - جامعة القادسية، القادسية، العراق	قسم الاحصاء - كلية الادارة والاقتصاد - جامعة كربلاء، كربلاء، العراق

أ.م.د. ايناس عبد الحافظ محمد
enas.albasri@uokerbala.edu.iq
قسم الاحصاء - كلية الادارة والاقتصاد - جامعة كربلاء، كربلاء، العراق

المستخلص

تلوث الهواء يعتبر من المشاكل المهمة التي يعاني منها بلدنا العراق في الوقت الراهن، وللتلوث مخاطر كثيرة اهمها تؤثر على صحة الانسان بالدرجة الاساس ومخاطر اقتصادية. و لمعرفة اهم العوامل المؤثرة في تلوث الهواء تمت دراسة وتحليل مجموعة من العوامل التي تؤثر في تلوث الهواء باستخدام احد طرق اختيار المتغيرات وهي طريقة معكوس لاسو (Reciprocal Lasso) . حيث تم جمع البيانات الشهرية لمدة خمس سنوات (2016-2020)، وباستخدام برنامج لغة R تم التوصل الى ان اهم العوامل المؤثرة في تلوث الهواء بشكل اساسي هي ثلاثة عوامل (حرق الغاز المصاحب لإنتاج النفط ، انبعاث الغازات من عوادم السيارات، حرق النفايات في مواقع الطمر) .

معلومات البحث

تواريخ البحث

تاريخ تقديم البحث: 2022/12/03
تاريخ قبول البحث: 2023/2/25
تاريخ رفع البحث على الموقع: 2023/12/31

الكلمات المفتاحية

اختيار المتغيرات، تلوث الهواء، لاسو، معكوس لاسو.

للمراسلة:

أ.م. سيف حسام رحيم
saif.hosam@qu.edu.iq

https://doi.org/10.55562/jrucs.v54i1.573

المقدمة

تلوث الهواء يعتبر احد اهم المشاكل التي يعاني منها وطننا العراق. ويمكن اعتبار تلوث الهواء من اخطر انواع التلوث وذلك لسرعة نقل الملوثات عبر الهواء وكذلك سرعة و سهولة دخول الهواء الى المنازل، ويمكن تقسيم تلوث الهواء الى ثلاثة اقسام من حيث شدة تلوثه. التلوث المقبول والخطر والذي نجده في اغلب المدن الصناعية و القسم الثالث هو التلوث المدمر. وكذلك يمكن تقسيم الملوثات الى تلوث طبيعي و تلوث ناتج من النشاطات البشرية، واللتان ستتم دراستهما مع بعض. حيث يمكن اعتبار تلوث الهواء مشكلة عالمية تعاني منها جميع الدول ولا يقتصر فقط على العراق. كما نلاحظ ظهور العديد من المنظمات والمؤسسات المحلية منها والعالمية تختص بدراسة وتحليل اهم العوامل التي تؤدي للتلوث لوضع الخطط المناسبة لتلافي أخطاره وانتشاره. كذلك اثبتت معظم الدراسات التي تختص بدراسة التلوث بان ملوثات الهواء تؤدي بشكل اساسي في ارتفاع درجات الحرارة بسبب ظاهرة الاحتباس الحراري، والذي يؤدي في المستقبل الى تغير مناخ الكرة الارضية مما سيؤثر سلبا على حياة البشرية. وبسبب تلك المخاطر جاءت هذه الدراسة لتحديد اهم تلك العوامل التي تؤثر في تلوث الهواء.

مشكلة البحث

تعتبر مشكلة تلوث الهواء من ابرز واخطر المشاكل التي يعاني منها وطننا العراق في الوقت الراهن و بسبب نتائج هذا التلوث على صحة الانسان و كذلك البيئة، فمن هنا جاءت مشكلة البحث.

هدف البحث

يهدف البحث لتحديد اهم العوامل التي تؤثر على تلوث الهواء في محافظة القادسية من خلال جمع البيانات وتحليلها باستخدام احد طرق اختيار المتغيرات و هي طريقة معكوس لاسو.

تلوث الهواء [1,2,3]

يمكن تعريف التلوث بأنه تحميل الهواء الجوي بمواد ضارة كأن تكون غازات منبعثة من المعامل او مصافي النفط او تكون مواد صلبة مثل الاتربة او الغبار. حيث يساهم هذا التلوث بأضرار منها بيئية كتلف المزروعات و موت الحيوانات. والبعض الاخر يساهم في اضرار صحية من خلال انتشار بعض الامراض التي تنتقل بالهواء الجوي و بعض الفايروسات. والتي تكون السبب الرئيسي في اصابة الانسان وحصول حالات الوفاة. و من الاثار الناجمة عن التلوث هي هجرة السكان الذين تكون مناطق سكناهم بالقرب من اسباب الملوثات، كان تكون معامل او مصافي النفط. ومن اهم العوامل التي تؤدي الى تلوث الهواء الجوي هي:

1. انبعاثات السيارات: بسبب كثرة السيارات في المحافظة و كذلك لسوء منتوج البنزين المستخدم لتشغيل السيارات ، فيعتبر الغاز المنبعث من السيارات من اهم الملوثات والاكثر خطورة على البيئة و الهواء. حيث يحوي هذا الغاز على مجموعة من الغازات السامة منها اول اوكسيد الكاربون و الرصاص و اكسيد النيتروجين .
2. الغبار والاتربة: يمكن اعتبار الغبار من الملوثات التي تلوث الهواء خصوصا في المناطق الصحراوية او المناطق غير المعبدة، والتي تكون السبب الرئيسي لبعض الامراض منها التحسس او الربو .
3. المصانع والمعامل: وتعتبر الغازات المنبعثة من المعامل و كذلك النفايات الكيمياوية و المخلفات من اهم ملوثات الهواء، حيث تعتبر مصانع الطابوق التي تستخدم الوقود الاحفوري في عملها سببا رئيسيا للتلوث.
4. الحرائق: من مسببات تلوث الهواء هي الحرائق وخصوصا تلك الحرائق المتعمدة كما في بعض الممارسات الزراعية الخاطئة. واما الحرائق التي تصيب الغابات فتعتبر من الاسباب الخطيرة في تلوث الهواء وتدمير الغابات المهمة للحفاظ على الحياة البشرية في الكوكب .
5. حرق النفايات: يعتبر حرق النفايات في غير مناطق الطمر الصحي هو كذلك من العوامل التي تسبب الهواء. حيث يقوم بعض الاشخاص بحرق تلك النفايات في المناطق السكنية او المناطق القريبة منها بغية التخلص من تلوث النفايات، حيث من الممكن ان تحوي تلك النفايات المحروقة على مواد كيمياوية ضارة و التي بدورها ستلوث الهواء .
6. حرق الغاز المصاحب لإنتاج النفط: يعتبر حرق الغاز لإنتاج النفط من الملوثات الرئيسية للهواء بصورة مباشرة، وكذلك يعتبر من اهم العوامل على تهجير السكان الذين يسكنون بالقرب من مصافي النفط .
7. التصحر: من العوامل الاخرى لتلوث الهواء هو التصحر، حيث نلاحظ في السنوات الاخيرة كثرة العواصف الترابية بسبب تصحر البلاد و قلة المناطق الزراعية و تحول اغلبها الى سكنية الأمر الذي اثر سلبا على الهواء الجوي. حيث اغلب البلدان تقوم بزراعة الاشجار حول المدن لتعمل كحاجز لصد الاتربة و تنقية الهواء حتى ولو بشكل قليل.
8. انبعاثات الطائرات: و تعتبر من اقل الاسباب لتلوث الهواء وذلك لان احتراق الوقود المستخدم للطائرات لا يحوي على كميات كبيرة من الغازات الضارة .
9. غاز الطهي في المنازل: ويعتبر شكلا من اشكال تلوث الهواء لما يحويه من غازات مصاحبة لاستخدام الغاز او النفط الابيض في المنازل .

اختيار المتغيرات [4,5]

يعتبر اختيار المتغيرات (Variables selection) من المواضيع التي لاقت اهتماماً واسعاً في الوقت الحالي، لكونها تقلص عدد المتغيرات الداخلة في الانحدار و استبعاد غير المهمة، وتقليل الجهد والتكلفة. حيث يمكن استخدام اختيار المتغيرات في العديد من المجالات ومنها المجالات الطبية والهندسة الوراثية والزراعة والبيئة وغيرها من المجالات .

وهنالك نوعان من طرق اختيار المتغير، النوع الاول يسمى الطرق التقليدية المتمثلة بطريقة الحذف العكسي التي تبدا باعتماد جميع المتغيرات في معادلة الانحدار ثم يتم حذف تلك المتغيرات تباعاً من خلال الاعتماد على قيمة F الجدولية . اما الطريقة الثانية فهي طريقة الاختيار الامامي المباشر التي يتم من خلالها بأدراج المتغيرات واحداً تلو الاخر الى معادلة الانحدار . اما الطريقة الثالثة فهي طريقة الاختيار التدريجي و كذلك معيار اكاكي للمعلومات و معيار بيز للمعلومات .

اما النوع الثاني من الطرق يسمى بطرق التنظيم حيث تفرض هذه الطرق حداً للجزاء على المعلمات لتسهيل عملية تقدير المعلمات وتقليص المتغيرات. وبعض تلك الطرق هي طريقة لاسو التي اقترحها العالم (Tibshirani 1996)، و طريقة مجموعة لاسو (Yuan and Lin 2006)، وكذلك طريقة لاسو التكيفي (Zou 2006)، و طريقة معكوس لاسو (Song and Liang 2015).

طريقة لاسو [6,7] Lasso Method [6،7]

تعتبر طريقة لاسو (Lasso) كإحدى طرق التنظيم والتي تعني (اضافة دالة جزاء لغرض التحكم بتعقيد النموذج و الذي يعبر عنة Penalty بحيث يكون عمل هذا الجزء هو تقليل تباين النموذج مع الزيادة الكبيرة في التحيز). ومختصر كلمة Lasso هو (Least Absolute Shrinkage and Selection Operator) الذي تم اقتراحه من قبل العالم Tibshirani عام (1996)، وتعتبر من الطرق الفعالة بحيث انها تعمل على اختيار المتغيرات و كذلك تقدير المعالم . وتعتبر طريقة لاسو من عائلة المربعات الصغرى الجزائية . لكن في العام 2001 وضح الباحثان Fan & Li بان طريقة لاسو لا تتمتع بخصائص اوراكل.

خصائص اوراكل وهي مجموعة من الخصائص التي يمتلكها المقدر منها (الاتساق، التناثر، تقارب الامثلية) حيث يمكن جمع تلك الخصائص بقول بان الطريقة تمتلك تلك الخصائص اذا كانت لديها القدرة على اختيار النموذج الحقيقي باحتمال يساوي 1.

من العيوب الاخرى لطريقة لاسو في حالة كون عدد المتغيرات المستقلة P اكبر من حجم العينة n وهذا يعني P > n حيث سيتم اختيار n من المتغيرات المستقلة و اهمال الباقي منها . حيث يتم الحصول على مقدر لاسو من خلال اضافة دالة جزاء لاسو الى دالة الخسارة لطريقة المربعات الصغرى وكما في المعادلة التالية :

$$\hat{\beta}_{lasso} = (\mathbf{y} - \mathrm{X}\boldsymbol{\beta})'(\mathbf{y} - \mathrm{X}\boldsymbol{\beta}) + \lambda \sum_{k=1}^{p} |\beta_k|$$

حيث ان λ تمثل معلمة التنظيم و التي تكون قيمتها اكبر او يساوي صفر $\lambda \geq 0$ حيث انها تتحكم او تنظم بكمية الجزاء او العقوبة، فكلما زادت قيمة λ يؤدي هذا الى اعلى مستوى للانكماش. في حالة كون قيمة الجزاء صغيرة فان هذا يؤدي الى ان النماذج المختارة لها تباين كبير و بالمقابل تحيز صغير. اما في حالة كبر مقدار قيمة الجزاء سيتم اختيار نماذج بها متغيرات اقل و بتباين اقل . اكد الباحثون على ضرورة الاختيار الدقيق لقيمة λ وذلك لغرض الحصول على نتائج دقيقة و يعول عليها، حيث يمكن تحديد قيمتها من خلال GCV (Generalized Cross Validation) و الموضحة في المعادلة التالية:

$$\mathrm{GCV}=\frac{RSS}{n\{1-p(\lambda)/nx\}^2\}} \quad (1)$$

حيث ان :

$$\mathrm{RSS}=\sum_{i=1}^{n}(y_{i-\theta^T x_i})^2$$

P(λ) تمثل عدد المعلمات ، حيث ان كبر قيمتها يؤدي الى المزيد من الجزاء .
P : تمثل عدد المتغيرات.
n يمثل حجم العينة.

طريقة معكوس لاسو [9,10] Reciprocal Lasso Method [9,10]

تم اقتراح طريقة معكوس (rLasso) لاسو في العام 2015 من قبل Song & Ling وذلك لتجنب النماذج التي تعاني من overfitted، حيث تعتبر هذه الطريقة من طرق اختيار المتغيرات و تقدير المعالم في نفس الوقت . ومن مزايا طريقة معكوس لاسو بانها مستمرة عند الصفر، تتناقص في الفترة $(0،\infty)$. ولتلك المزايا للطريقة فأنها اصبحت مرغوبة لاختيار النماذج، ويمكن تمثيل مقدر الطريقة من خلال الصيغة التالية:

$$\hat{\beta}_{r\mathrm{lasso}} = (\mathbf{y} - \mathrm{X}\boldsymbol{\beta})'(\mathbf{y} - \mathrm{X}\boldsymbol{\beta}) + \lambda \sum_{k=1}^{p} \frac{1}{|\beta_k|} I\{\beta_k \neq 0\} \quad (2)$$

حيث ان :
λ : تمثل معلمة الضبط و التي تتحكم بقدر درجة العقوبة او الجزاء .
$I(.)$: تمثل indicator function .
قيمة λ اذا كانت في قيمتها الدنيا سوف تعطي اعلى مستوى من الانكماش و بذلك تعطي ما يقارب صفر من المعاملات. وبالمقارنة مع دالة جزاء لاسو فأنها لا تتناقص في الفترة $(0،\infty)$. وعندما تكون λ مساوية للصفر في هذه الحالة تعطي ما يقارب صفر من المعاملات ، بعكس طريقة المعكوس التي تعطي صفرا من المعاملات عند العقوبات التي تقترب من اللانهاية. فمن الناحية النظرية تمتلك طريقة معكوس لاسو خاصية اوراكل، وعلية فأنها افضل من حيث الاداء من طريقة لاسو.

طريقة الشبكة المرنة [12,13] Elastic Net Method [12,13]

طريقة الشبكة المرنة (EN) تعتبر من طرق التنظيم، وهي عبارة عن مزيج بين حد الجزاء لطريقة لاسو و حد الجزالة لطريقة طريقة الحرف (ridge model). حيث تغلبت هذه الطريقة على طريقة لاسو التي تستخدم دالة الجزاء التالية $|\beta|^1 = \sum_{j=1}^{p}|\beta_j|$. ففي حالة كون المتغيرات مرتبطة فيما بينها ارتباطاً قويا فان طريقة لاسو تختار احد تلك المتغيرات وتهمل الباقية . وللتغلب على تلك العيوب تم اضافة حد الجزاء $((\|\beta\|^2))$ والذي يمثل حد جزاء انحدار الحرف . فانحدار الحرف يعتبر من اشهر نماذج الانحدار الخطي المنتظم، فتمت من خلال اضافة حد الجزاء l_2 . والتي يمكن توضيحها بالمعادلة التالية :

$$\hat{\beta}_{Ridge} = arg\min_{\beta} \sum_{i=1}^{N}(Y_i - X\beta) - \lambda \sum_{j=1}^{P} \beta_j^2 \quad (3)$$

لا يمكن استخدام انحدار الحرف كطريقة لاختيار المتغيرات وذلك لان تلك الطريقة لا يتم من خلالها تقليص المعاملات الى الصفر .فمن خلال طريقة الشبكة المرنة يتم اختيار المتغيرات. وصيغة هذه الطريقة يمكن توضيحها بالمعادلة الاتية :

$$\hat{\beta} = arg\min_{\beta}(\|Y - XB\|^2 + \lambda_2\|\beta\|^2 + \lambda_1\|\beta\|_1) \quad (4)$$

حيث ان

λ_1 : تمثل معلمة الضبط لدالة جزاء طريقة لاسو .
λ_2 : تمثل معلمة الضبط لحد الجزاء لطريقة انحدار الحرف.

Simulation المحاكاة

تناول هذا الجزء من البحث اجراء دراسة المحاكاة لاختبار اداء طريقة معكوس لاسو و المقارنة مع الطرق الاخرى لغرض تقييم اداء الطريقة، حيث تمت المقارنة من طريقة لاسو و طريقة الشبكة المرنة. حيث تم توليد البيانات بواسطة :

$$y_i = x_i'\beta + \varepsilon_i ، i = 1, 2, \ldots, n$$

كما تم اعتماد توزيع الخطأ على انه يتوزع وفق التوزيع الطبيعي بمتوسط يساوي صفر وتباين σ^2 اي ان $\varepsilon_i \sim N(0، \sigma^2)$. اما بالنسبة لصفوف مصفوفة التصميم X تم توليدها بالاعتماد على $N(0، \Sigma)$ ، حيث ان Σ هي autoregressive correlation matrix حيث ان $\Sigma ij = 0.50|i-j|$ لكل $1 \le i \le j \le p$. كما تم استخدام معيارين للمفاضلة بين الثلاثة طرق وهو متوسط مربع الاخطاء Mean squared errors (MSE) و الانحراف المعياري standard deviations (SD) تم افتراض ثلاثة حالات ل β وهي كما مبين أدناه :

- الحالة الاولى $\beta = (1,1,1,1,1,1,1,1,1,1,1,1)$

جدول (1): يبين قيمة المعيارين (MSE) و (SD) للحالة الاولى

المعيار	قيمة σ^2	EN	Lasso	rLasso
MSE	$\sigma^2 = 1$	1.79	1.88	1.7
SD		0.52	0.54	0.5
MSE	$\sigma^2 = 4$	5.86	6.17	5.55
SD		1.03	1.14	0.92
MSE	$\sigma^2 = 8$	12.72	13.11	12.33
SD		2.4	2.28	2.52

- الحالة الثانية $\beta = (1,1,1,1,1,0,0,0,0,1,1,1)$

جدول (2): يبين قيمة المعيارين (MSE) و (SD) للحالة الثانية

المعيار	قيمة σ^2	EN	Lasso	rLasso
MSE	$\sigma^2 = 1$	1.26	1.45	1.64
SD		0.22	0.23	0.24
MSE	$\sigma^2 = 4$	5.5	6.27	7.04
SD		1.22	1.21	1.2
MSE	$\sigma^2 = 8$	11.78	13.43	15.08
SD		1.95	2.25	2.55

- الحالة الثالثة :

$\beta = (1، 1، 1، 1، 0، 0، 0، 0، 0، 0، 0، 0)$

جدول (3): يبين قيمة المعيارين (MSE) و (SD) للحالة الثالثة

المعيار	قيمة σ^2	EN	Lasso	rLasso
MSE	$\sigma^2 = 1$	1.51	1.75	1.27
SD		0.3	0.34	0.26
MSE	$\sigma^2 = 4$	6.16	7.22	5.1
SD		1.81	2.56	1.06
MSE	$\sigma^2 = 8$	13.68	16.21	11.15
SD		2.3	2.77	1.83

نلاحظ من خلال الجداول (1، 2 و3) بان النتائج اغلبها متقاربة لطريقة البحث معكوس لاسو والطرق التي تمت المقارنة بها (لاسو والشبكة المرنة)، حيث ان جميع الطرق تؤدي اداءً نسبياً في اختيار الانموذج في دراسة المحاكاة. و يمكن ترتيب الطرق بالاعتماد على معايير المفاضلة بان طريقة معكوس لاسو جاءت بالترتيب الاول وذلك لان في جميع الامثلة كانت قيمة المعيارين هي اقل من باقي طرق المقارنة. اما في المرتبة الثانية جاءت طريقة الشبكة المرنة و في المرتبة الثالثة كانت من نصيب طريقة لاسو.

Real Data البيانات الحقيقة

تم تطبيق طرق اختيار المتغيرات (معكوس لاسو و الشبكة المرنة و لاسو) على بيانات حقيقية معتمدة ومسجلة في مديرية بيئة محافظة القادسية للفترة من (2016-2020) ، حيث كانت تلك البيانات الشهرية تمثل العوامل المؤثرة على تلوث الهواء . ويمكن توضيح تلك العوامل بما يلي :

جدول (4): يبين ترميز العوامل المؤثرة على تلوث الهواء

ت	الترميز	العوامل
1	X1	انبعاثات السيارات
2	X2	الغبار و الاتربة
3	X3	المصانع و المعامل
4	X4	الحرائق
5	X5	حرق النفايات
6	X6	حرق الغاز المصاحب لإنتاج النفط
7	X7	التصحر
8	X7	انبعاثات الطائرات
9	X9	غاز الطهي في المنازل

فمن خلال الدراسة سيتم تحديد اهم العوامل التي تؤثر في زيادة التلوث في الهواء مما يساعد اصحاب القرار او المنظمات على معرفة تلك العوامل و اتخاذ القرارات او الحلول للحد من تلك المشكلة، وتلك المتغيرات يمكن اعتبارها المتغيرات المستقلة للدراسة، واما المتغير المعتمد فهو كمية التلوث في الهواء. لتحليل البيانات الحقيقية المتمثلة بتسعة متغيرات مستقلة ومتغير معتمد تم استخدام طريقة معكوس لاسو لتحديد اهم تلك العوامل، وذلك لان الطريقة اثبتت افضليتها بالمقارنة مع الطرق البقية. تم استخدام برنامج لغة R حيث تم التوصل للنتائج في الجدول ادناه .

جدول (5): يبين نتائج النموذج المقدر

Variable		Estimate	S.E.	t	P-Value
انبعاثات السيارات	x1	1.317	0.277	4.743	0.00000209
الغبار و الاتربة	x2	0.000	0.159	0.000	1.000
المصانع و المعامل	x3	0.000	0.175	0.000	1.000
الحرائق	x4	0.000	0.215	0.000	1.000
حرق النفايات	x5	-1.427	0.164	-8.690	0.000384
حرق الغاز المصاحب لإنتاج النفط	x6	6.442	0.179	3.590	0.00000719
التصحر	x7	0.000	0.139	0.000	1.000
انبعاثات الطائرات	x8	0.000	0.129	0.000	1.000
غاز الطهي في المنازل	x9	0.000	0.131	0.000	1.000

يوضح الجدول رقم (5) القيم التقديرية للنموذج والاخطاء المعيارية واختبار t وكذلك قيمة (P-Value)، حيث نلاحظ ان المتغيرات (x1: انبعاثات السيارات) و (x5: حرق النفايات) وكذلك (x6: حرق الغاز المصاحب لإنتاج النفط) لها مستوى تأثير معنوي عند مستوى معنوية ($\alpha = 0.05$) وهذا واضح من خلال ملاحظة قيم (P-Value) من الجدول اعلاه. وهذا يوضح ان تلك المتغيرات لها اثر كبير في نموذج الدراسة.

الاستنتاجات

1. تم التوصل الى ان طريقة معكوس لاسو لها افضلية بعد مقارنتها مع طريقة الشبكة المرنة و طريقة لاسو .
2. باستخدام طريقة معكوس لاسو حصل الباحثون على اربعة عوامل لها تأثير مباشر في تلوث الهواء .
3. حرق الغاز المصاحب لإنتاج النفط الاثر الاكبر في تلوث هواء المحافظة .
4. كثرة اعداد السيارات في العراق بشكل عام و المحافظة بشكل خاص اسهم بشكل مباشر في تلوث الهواء من خلال الغازات المنبعثة من عوادم السيارات .
5. قلة اماكن الطمر الصحي و كثرة اماكن حرق النفايات و خصوصا التي تحوي على مواد بلاستيكية كان لها اثر رئيسي في تلوث الهواء.
6. في الآونة الاخير لوحظ بان اكثر المناطق الزراعية تم هجرها وعدم زراعتها و كذلك تحويلها الى مناطق سكنية مما ادى الى تحول اغلب الاراضي الزراعية الى صحراء السبب الذي ادى الى سرعة انتشار التلوث.

التوصيات

1. التقليل من حرق الغاز المصاحب لإنتاج النفط او استخدام تلك الغازات في انتاج الطاقة الكهربائية .
2. استخدام سيارات النقل الجماعي لتقليل اعداد السيارات في الشوارع ، التوجه الى استيراد السيارات الكهربائية .
3. فرض السيطرة التامة على اماكن حرق النفايات ، وفرض غرامات مالية عالية و العقوبات القانونية على المخالفين .

المصادر

[1] Brunekreef, Bert, and Stephen T. Holgate. "Air pollution and health." The lancet Vol. (360), No. (9341), (2002): 1233-1242.

[2] Kampa, Marilena, and Elias Castanas. "Human health effects of air pollution." Environmental pollution 151 (2008): 362-367.

[3] Stern, Arthur C. Air Pollution: The effects of air pollution. Vol. 2. Elsevier, 1977.

[4] Andersen, Charlotte Møller, and Rasmus Bro., "Variable selection in regression- A tutorial.", Journal of chemometrics Vol. (24), No. (11-12), (2010): 728-737.

[5] Allen, David M., "The relationship between variable selection and data augmentation and a method for prediction." technometrics Vol. (16), No. (1), (1974): 125-127.

[6] Zou, Hui. "The adaptive lasso and its oracle properties." Journal of the American statistical association Vol. (101), No. (476), (2006): 1418-1429.

[7] Roth, Volker. "The generalized LASSO.", IEEE transactions on neural networks, Vol. (15), No. (1), (2004): 16-28.

[8] Tibshirani, Robert. "Regression shrinkage and selection via the lasso.", Journal of the Royal Statistical Society: Series B (Methodological), Vol.(58), No. (1), (1996): 267-288.

[9] Song, Qifan, and Faming Liang. "High-dimensional variable selection with reciprocal L1-regularization." Journal of the American Statistical Association, Vol. (110), No. (512), (2015): 1607-1620.

[10] Mallick, Himel, et al., "The reciprocal Bayesian lasso.", Statistics in Medicine, Vol. (40), No. (22), (2021): 4830-4849.

[11] Paul, Erina, and Himel Mallick. "Unified reciprocal LASSO estimation via least squares approximation.", Communications in Statistics-Simulation and Computation (2022): 1-11.

[12] Zou, Hui, and Trevor Hastie. "Regularization and variable selection via the elastic net.", Journal of the royal statistical society: Series B (Statistical Methodology), Vol. (67), No. (2), (2005): 301-320.

[13] Li, Qing, and Nan Lin., "The Bayesian elastic net.", Bayesian analysis Vol. (5), No. (1), (2010): 151-170.

Journal of AL-Rafidain University College for Sciences (2023); Issue 54; 30-36

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 Factors Affecting Air Pollution Using the Reciprocal Lasso Method

Saif H. Raheem	Jassim N. Hussain
saif.hosam@qu.edu.iq	Jasim.nasir@uokerbala.edu.iq
Statistics Department - College of Administration and Economics - University of Al-Qadisiyah, Al-Qadisiyah, Iraq	Statistics Department - College of Administration and Economics - University of Karbala, Karbala, Iraq

Enas A. Mohamed

enas.albasri@uokerbala.edu.iq

Statistics Department - College of Administration and Economics - University of Karbala, Karbala, Iraq

Article Information

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

Keywords:
Variables Selection, air pollution, lasso, Reciprocal lasso.

Correspondence:
Saif H. Raheem
saif.hosam@qu.edu.iq

Abstract

Air pollution is one of the important problems that our country, Iraq, suffers from at the present time . Pollution has many risks the most important being that it affects human health primarily and some economic risks. In order to find out the most critical factors affecting air pollution, a group of factors that affect air pollution was studied and analyzed using one of the methods of selecting variables, the Reciprocal Lasso method. from collected monthly data for a period of five years (2016-2020), And using the R language program, it was concluded that the most important factors affecting air pollution mainly are three factors (burning gas associated with oil production, gas emissions from car exhausts, burning waste in landfill sites).

https://doi.org/10.55562/jrucs.v54i1.573