

Determining the most important factors of air pollutants on respiratory patients: a statistical study using the multiple linear regression model based on Genetic Algorithm

Raed Sabeeh Karyakos

Department of Mathematics, Education College for Pure Sciences, University of
AL-Hamdaniya, Mosul, Iraq
Orcid: 0000-0002-2278-7938
raed_sabeeh@uohamdaniya.edu.iq

Sarah Ghanim Mahmood Al- Kababchee

Department of Mathematics, Education College for Pure Sciences, University of
AL-Hamdaniya, Mosul, Iraq
Orcid: 0000-0002-0574-985X
sarahghanim@uohamdaniya.edu.iq

Ilham M. Yacoob

Department of Mathematics, Education College for Pure Sciences, University of
AL-Hamdaniya, Mosul, Iraq
Orcid: 0000-0003-3994-8619
ilhammatta@uohamdaniya.edu.iq

Abstract

Atmospheric air pollutants contain particles and gases that are harmful to human health. These pollutants cause chronic diseases and allergies, including bronchitis and bronchial asthma. Accordingly, the research used one of the traditional estimation methods for the normal multiple linear regression model parameters, represented by the ordinary least squares and genetic algorithm methods, to determine the most important influencing variables (dust falling, lead, carbon monoxide, sulfur dioxide, total suspended particles) on patients with bronchial asthma as a response variable in Baghdad, 2004. It was shown through data analysis that the variable carbon monoxide and total suspended particles do not have any significant effects on the number of people from bronchial asthma.

Keywords: Multiple Linear Regression, Air Pollutants, Ordinary Least Square Method, Bronchial Asthma, Genetic Algorithm.

1. Introduction

Air pollution refers to the contamination of the indoor or outdoor environment by any factor, whether chemical, physical, or biological. Air pollutants are a major risk factor for many pollution-related

diseases, such as respiratory infections, heart disease, lung disease, and stroke. Furthermore, pollutants can affect an individual's intelligence and contribute to mental health problems. These pollutants are typically found in the atmosphere and have harmful effects on human health. They can take the shape of gasses, droplets, or particles.

Al-Sarifi (2022) studied the effect of the dust falling over the Safwan district and the city center of Basra Governorate and concluded that the falling dust consists of many minerals, including quartz, gypsum and kaolin, and that these particles have mechanical effects of causing respiratory diseases, as well as eye, skin and ear damage [3].

Al-Hashemi and Al-Shammari (2021) valued total suspended particulates and heavy elements in the district centers of Wasit Governorate. The concentrations of total particulates and some heavy elements such as (Pb, Cr, Fe, Zn, Cd, Cu, Ni) were studied, in addition to secondary pollutants (NO₃, SO₄), using an air extraction device during January and July of 2019 [5]. The results showed that the concentrations of the studied pollutants exceeded the environmental limits during the summer season.

(Al-Bu'aishi, 2021) Presented a study on the effect of sulfur dioxide smoke at specific concentrations on wheat and barley plants, aiming to investigate the gas pollution's impact on some of the plants' morphological characteristics [1]. The smoking process was carried out two weeks after planting, and the plants' height and number of leaves were measured. A statistical analysis was conducted using a two-way ANOVA table, comparing the means at a significance level of (0.05). The results showed significant differences in the average leaf area of wheat plants, and the effect of sulphur dioxide gas slowed wheat plant growth.

The paper included a general introduction to air pollutants in the first section, and the second section addressed the study's objective. The third and fourth sections presented the concept of bronchial asthma and some airborne factors affecting the disease. The fifth section described the normalized multiple linear regression models and the sixth section estimated the model parameters using the ordinary least square method. The seventh section presented the statistical analysis of real data related to the quantity of individuals suffering from bronchial asthma in Baghdad in 2004 and some factors affecting the disease. Finally, the eighth section presented the most important conclusions and suggested future studies.

2. Paper objective

The main objective of this study is to identify the most important air pollutant factors affecting the quantity of individuals suffering from bronchial asthma in Baghdad 2004, based on a normalized multiple linear regression model.

3. Concept of bronchial asthma

Bronchial asthma is defined as a chronic inflammation that causes narrowing of the airways, leading to hypersensitivity and affecting the air cilia, as well as the oxygen and carbon dioxide ratios in the

blood. Bronchial asthma occurs at all ages, but in young children, approximately half of the cases occur before the age of ten. To study the diagnosis of patients suffering from the same symptoms as asthma, such as chronic bronchitis and emphysema, there is an electronic device called a spirometer that measures the volume of air in the lungs and the force of breathing. Therefore, the initial diagnosis depends on measuring the volume of expiratory air (FEV1) in the first second, which is less than (0.80) of the volume of expiratory air in the first second of a normal person. To increase the (FEV1) measurement by (0.15), Ventolin is administered via nebulizer. The following diagram illustrates the movement of air within the bronchi. [2]

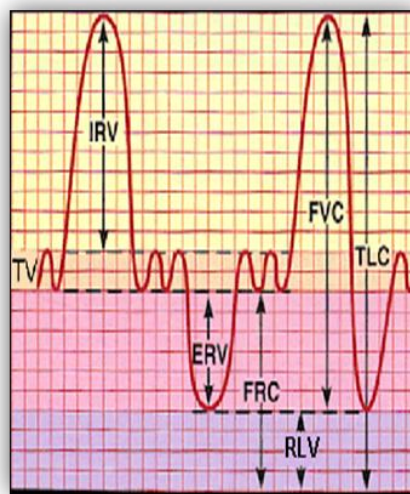


Fig. (1): movement of air within the bronchi

4. Air pollutant factors affecting the number of people with bronchial asthma:

This section will review some air pollutants that affect bronchial asthma, including:

4.1 Effects of falling dust:

Dust storms are a contributing factor to the development of bronchial asthma. They result from road construction and excavation work, factory dust, and volcanic ash dispersed by wind, as well as dust storms caused by desertification and drought. In 1996, the World Health Organization identified dust storms in sub-Saharan Africa as a cause of meningitis outbreaks. Furthermore, recent scientific studies have shown that dust particles can carry dangerous types of bacteria and fungi that cause severe pneumonia in humans. [8]

4.2 Effects of Lead:

Lead exposure occurs through various means, including inhalation of air. Lead can also be ingested through food or water. Its chemical formula is Pb. Lead is a toxic metal that occurs naturally in the Earth's crust and accumulates in the body through the blood and bones. It affects various organs,

including the liver, kidneys, and nervous system. Excessive lead exposure can cause neurological damage, such as intellectual disability and behavioral disorders. Additionally, a recent study found that lead may be linked to middle-aged men's high blood pressure and heart disease. [6]

4.3 Effects of Carbon Monoxide:

Carbon monoxide is a colorless and odorless toxic gas with the chemical formula (CO). It is produced from the incomplete combustion of organic materials, including leaks from chimneys, furnaces, and gas-fired heating appliances, as well as cigarette smoke. When carbon monoxide mixes with hemoglobin in the blood, it lowers oxygen absorption. Chest ache, exhaustion, nausea, blurred vision, and diminished cognitive function are among its health impacts. [10]

4.4 Effects of sulfur dioxide:

Sulfur dioxide is a colorless, odorous gas. It is considered a heavy gas and an air pollutant. Its chemical formula is (SO₂). It is produced by the industrial combustion of coal, gas, and oil, and naturally by volcanoes. Increased inhalation of it poses a risk of heart disease, stroke, premature death, and lung cancer. It is linked to bronchitis and, due to smog, is one of the worst factors, in addition to being one of the causes of acid rain. [7]

4.5 Effects of Total Particulate Matter:

Total particulate matter consists of particles that affect human health and are air pollutants. These particles are similar to sulfur dioxide, and prolonged exposure to these particles increases the incidence of respiratory diseases, especially asthma and emphysema. The organic components of these particles are also a cause for concern as they are carcinogenic. The people most affected by these particles are children and the elderly who suffer from chronic diseases such as heart disease. [4]

5. Description of the Multiple Linear Regression Model

The multiple linear model is predicated on the idea that the response (dependent) variable Y_i and the number of influencing (explanatory) variables $X_1, X_2, \dots, X_P$ and the random error term e_i have a linear relationship. For n observations and P of the explanatory variables, this connection can be written as follows: [9]

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_P X_{Pi} + e_i \quad \dots (1)$$

In matrix form, the above model is written in the following format:

$$\underline{Y} = \underline{X}\underline{\beta} + \underline{e} \quad \dots (2)$$

As:

Y: The observations vector represents the response variable, with an amplitude of $(n \times 1)$.

X : A non-random array representing observations of an explanatory variables with capacity of $(n \times p)$, where P is the number of independent variables.

$\underline{\beta}$: Vector parameter of a regression model with an amplitude of $(p \times 1)$.

$\underline{e}$: Random errors vector has an amplitude $(n \times 1)$, and a probability distribution of a random errors vector is expressed descriptively as.

$$\underline{e} \sim N_n(\underline{0}, \sigma^2 I_n) \quad \dots (3)$$

6. Ordinary least squares method

The ordinary least squares method is an important technique for estimating regression model parameters. This method is characterized by its ease of use, simplicity, and accuracy in results when its underlying assumptions are met. These assumptions are crucial for obtaining estimates of a particular phenomenon. If any of these assumptions are not met, the method loses its status as a best linear unbiased estimator (BLUE). One of the most important assumptions is the absence of a strong linear relationship between the explanatory variables. [9]

Ordinary least squares estimators using matrices are obtained using the following formula:

$$\hat{\beta}_{OLS} = (X'X)^{-1} X'Y \quad \dots (4)$$

7. Genetic Algorithms

GA is developed by observing Darwin's theory of evolution. The algorithm's flowchart is displayed in Fig. 2[11]. GA encodes the latter as an unmarried chromosome for a particular disorder. The initial mass of those that are part of the second site of the disturbance is then described. Therefore, an answer space where each possible answer is represented by a type of chromosome is a search space. Before searching off advanced organisms begins, the initial population is created by tough and quick random selection of chromosomes from the hunt website. The utility of the computations through a particular goal specification is then determined by means of competitively deciding on coefficients based totally on their benefit. Genetic search operators, which include selection, mutation, and crossover, are again performed to produce a new availability of chromosomes, where generally pleasant earlier periods are expected to be surpassed This process is repeated until the end requirements are satisfied, and the best chromosome from the last lineage is resolved. [12].

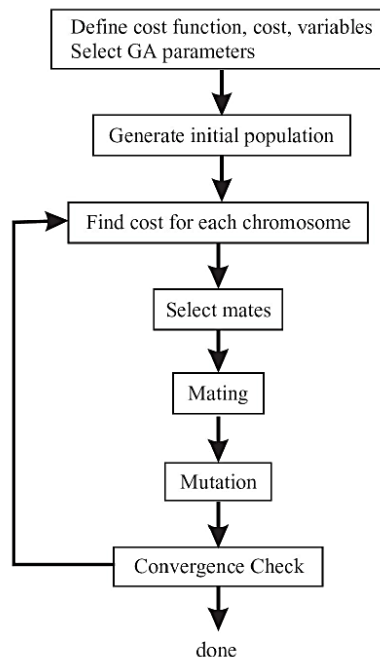


Fig. (2). chart Gen. Algo.

Employing genetic algorithms to estimate parameters of a multilinear regression model.

In this element, we used (GA) to determine the model parameters in the regression version.

Step 1: Initialize the parameters of the GA algorithm by setting $N_d=0$.

Step 2:

Specify the parameters of the regression model → Create the initial population → Determine the fitness value of the genuine population

Step 3: $N_d=N_d+1$ decade loop.

Step 4:

selection → crossover → mutation → Determine each chromosome's fitness value

Step (5): Terminating Criterion Control; Repeat Steps 3-5 until a terminating criterion is satisfied;

Input -Information:

The training process for the multilinear regression model begins with data input, divided into two segments: the first segment represents the independent set (input variables), while the second segment

represents the dependent set (target variables). These two segments are combined into a matrix that encapsulates the model's input data. The matrix form can be expressed as follows:

$$X = \begin{bmatrix} x_1 & x_2 & \cdots & x_p \\ x_2 & x_3 & \cdots & x_{p+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n-p} & x_{n-p+1} & \cdots & x_{n-1} \end{bmatrix} \quad \dots (5)$$

It ought to be mentioned that each one input education information should be normalized and a suitable embedding size ought to be selected prior to creating and normalizing education and testing samples. To ensure the effectiveness of the model, it is crucial to standardize the input training data prior to the establishment and normalization of training and testing samples. This normalization process helps in mitigating biases and enhancing the model's performance. Additionally, selecting an appropriate embedding dimension is vital, as it directly influences the representation of the data in the model. Equation (6) illustrates the normalized approach.

$$\hat{x} = 2 * \left(\frac{x - x_{\min}}{x_{\max} - x_{\min}} \right) - 1 \quad \dots (6)$$

Max. and min. values of a measured data are denoted by $x_{\max}$, $x_{\min}$, where x is measured data and Y is normalized data. High values are frequently found in the training data. Therefore; it is usually advisable to normalize the model's inputs and outputs to fall inside the interval $(-1, 1)$. This normalization has numerous benefits, including:

1. As all inputs converge to certain values, they do not regulate the entry to one another.
2. The normalization values expedite training [13].

Encoding

The regression model has many parameters & every parameter is involved, which parameters. each chromosome is encoded as:

$$\begin{bmatrix} \beta_1 & \beta_2 & \cdots & \beta_p & \sigma^2 \end{bmatrix}$$

Fig. (3). Chromosome representation

First Population:

The GA started a search by generate the population of candidate solutions. In our application, this population is randomly generated according to the uniform distributions. Initially, we produce a random value (β_i) for each independent variable i from UD $[-1,1]$.

Fitness Evaluation:

The reciprocal of the mean error sum of squares (MSE) is a fitness function value. Gen. Algo. fitness function can be expressed as follows:

Fun. $M = F$ (parameters, x , y)

$$\hat{y} = x\hat{\beta}$$

M =mean square error ($\hat{y}-y$)

end

Selection:

Based on their rank value or fitness function value, chromosomes are chosen. In this study, a highest rank selection is used. This method selects the first chromosome with the highest fitness (minimum MSE), whereas the second chromosome is selected at random. This method can be used to enhance the most fit chromosomes to carry some advantageous features.

Crossover:

During the crossover, parent individuals produce new offspring chromosomes. Two new chromosomes are created by exchanging specific genes from two parent chromosomes. In our implementation, we use the arithmetic crossover. These formulas explain how this kind of crossover produces two new offspring by linearly combining two parent chromosome vectors:

$$\text{offspring I} = B * \text{Parent I} + (1-B) * \text{Parent II} \quad \dots (7)$$

$$\text{offspring II} = (1-B) * \text{Parent I} + B * \text{Parent II} \quad \dots (8)$$

Where B is a weighting factor that is selected at random prior to each crossover operation.

Mutation:

In this paper, a basic static standard mutation is called a mutation. An individual's i -th parameter, z_i , is modified by $z_i = z_i \pm U(l_i, u_i)/10$ with a mutation probability p_m . The tenth part of a random number in a particular domain is represented by $U(l_i, u_i)/10$ in our study. There is a 50%–50% chance in the “+/-” portion. To fit inside the boundaries, the new offspring might be modified as needed. The lower bound of z_i is denoted by “ l_i ,” whereas the upper bound is denoted by “ u_i .”

Conditions for Stopping:

For genetic algorithms, there are no universally recognized stopping criteria. In this study, our method is simply terminated after a predetermined number of iterations (N_{gen}).

8. Practical aspect

This section analyzes the data on the number of individuals in Baghdad who had bronchial asthma in 2004 and the factors affecting it statistically.

8.1 Defining variables and formatting data:

Paper data represents the number of individuals suffering from bronchial asthma in Baghdad Governorate in 2004 as the response variable (Y), and the variables affecting it as explanatory variables: particulate matter (X1), lead (X2), carbon-monoxide (X3), sulfur-dioxide (X4), as well as total suspended particle matter (X5). Table (1) shows the monthly concentrations of air pollutants and the monthly number of patients with bronchial asthma, collected by the Ministry of Environment/Department of Planning and Technical Follow-up/Air Quality Section.

Table (1): Monthly concentration measurements of particulate matter, lead, carbon- monoxide, sulfur -dioxide, and total suspended particle matter affecting the number of individuals suffering from bronchial asthma in Baghdad in 2004

Month	Number of individuals suffering from bronchial asthma	falling dust gm/m2	lead mg/m3	Carbon monoxide ppm	sulfur dioxide ppm	Total stuck minutes mg/m3
January	9266	19	0.79	1.92	0.03	391.2
February	13310	23	0.69	1.53	0.02	516.2
March	6268	9.32	0.48	2.36	0.03	235.2
April	6691	36.38	4.43	1.59	0.07	307.9
May	7341	21.03	2.39	1.35	0.03	490.3
June	7661	19.98	2.76	1.78	0.03	399.4
July	7949	26.05	1.59	2.73	0.03	401.3
August	9145	13.8	1.74	1.79	0.01	459.8
September	6875	12.52	1.29	1.66	0.01	363.6
October	8857	43.67	9.63	1.68	0.02	309.5
November	9112	9.32	3.65	0.97	0.002	175.3
December	6325	0	1.11	1.18	0.01	147

The following figure illustrates the behavior of the variable number of individuals suffering from bronchial asthma in Baghdad for the year 2004.

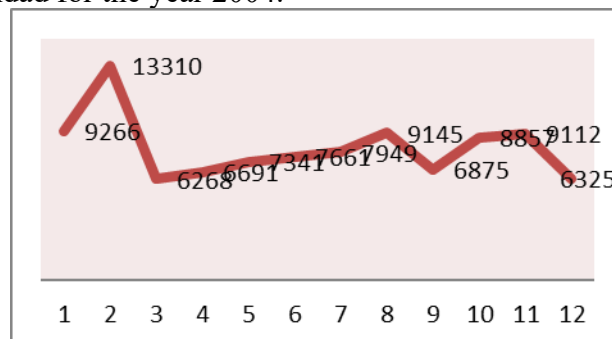


Figure (4): Data behavior regarding the number of individuals suffering from bronchial asthma

To determine if a study data conformed to a normal distribution, Microsoft Excel was used as follows:

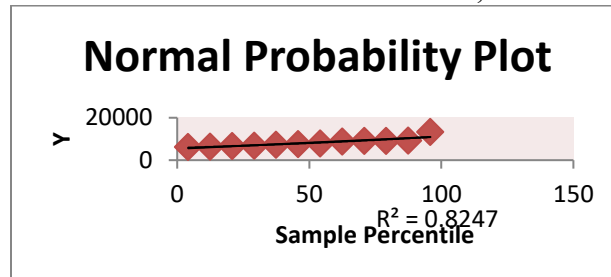


Figure (5): Diagram of the normal probability of the number of people with bronchial asthma

We observe from Figure (5) that the behavior of the data lies on a straight line, which indicates that the study data are normally distributed, in addition to the value of the coefficient of determination, which reached (0.91), meaning that 91% of this data are normally distributed and 9% of this data follow a non-normal distribution.

8.2 Estimating the parameters of a normalized multiple linear regression model:

Equation (4) was applied to estimate the model parameters using the ordinary least squares method, as displayed in the table (2):

Table (2): Show results of estimating the parameters of the normal multiple linear regression model - Source: Using Excel

Var.	Par.	Coeff.	Stand.err	t -Stat	p-value
Int.	β_0	13084.82008	3094.302515	4.228681591	0.005508154
X 1	β_1	383.1189633	141.1184589	2.714874909	0.034879399
X 2	β_2	-1336.187689	546.8577817	-2.443391562	0.049128458
X 3	β_3	-2319.580582	1111.353274	-2.087167633	0.081918554
X 4	β_4	-128951.0822	42830.5647	-3.010725708	0.023678355
X 5	β_5	-5.188853539	6.715012815	-0.772724294	0.469021936

From the results in Table (2), regarding the effect of the explanatory variables on the number of individuals suffering from bronchial asthma, we find the following:

- p-value of t-test corresponding to a carbon-monoxide variable (X3) was (0.0819), This above the significance level of 0.05, meaning that the carbon- monoxide variable has no significant effect on the number of individuals suffering from bronchial asthma.
- p-value for t-test corresponding to the total suspended particule matter variable (X5) was (0.469), This above the significance level of 0.05, meaning that a total suspended particule matter variable has no significant effect on the number of individuals suffering from bronchial asthma.
- p-value for t-test corresponding to the dust storm variable (X1) was 0.035, which is less than the significance level of 0.05, meaning that the dust storm variable has a significant effect on the incidence of bronchial asthma.

- p-value for t-test statistic corresponding to the lead variable (X2) was 0.049, which is less than the significance level 0.05, meaning that the lead variable has a significant effect on the incidence of bronchial asthma.
- p-value for t-test statistic corresponding to the lead variable (X4) was 0.024, which is less than the significance level (0.05), meaning that the sulfur dioxide variable has a significant effect on the incidence of bronchial asthma.

Therefore, the estimated regression equation is as follows:

$$\hat{Y} = 13084.82 + 383.12 X_1 - 1336.19 X_2 - 2319.58 X_3 - 128951.08 X_4 - 5.19 X_5 \quad \dots (9)$$

The normalized multiple linear regression model using the least squares method was found to be 74% suitable for the data, as it showed statistical significance based on the coefficient of determination. This means that the explanatory variables were able to explain 74% of the changes in the number of individuals suffering from bronchial asthma, while the remaining 26% is attributed to other factors. Therefore, we conclude that the explanatory variables that had a significant effect on the incidence of bronchial asthma were dust mites, lead, and sulfur dioxide.

The Multi-Linear Normal Regression Model is tested using data sets. For comparison, and Artificial Neural Network [14] are also executed on these data sets.

Table (3). shows forecasting comparison among these models.

Data	Measure	(RM)Regression Model		(ANN)Artificial Neural Network	
		Train.	Testing	Train.	Testing
Number of people with bronchial asthma	MSE	0.360	1.800	0.378	1.800
	MAD	0.265	1.800	0.270	1.800
	R ²	0.740	1.001	0.753	1.001

From Table (3). The results unequivocally demonstrate that the genetic algorithm-based regression model outperforms the artificial neural network in terms of best (mse, mad, and r2) across all data sets, in terms of the number of individuals with bronchial asthma. Additionally, the results showed how well the regression model balanced remembering (training) and forecasting (testing).

Table (4) shows actual and estimated values for the number of individuals suffering from bronchial asthma in Baghdad in 2004, using the estimated regression equation defined in Equation (9).

Table (4): Displays both estimated and actual values of the number individuals suffering from bronchial asthma

Estimated $\hat{Y}$ values	Y Real values	NO.
8956.485422	9266	1
12168.1206	13310	2
5450.957737	6268	3
6791.019617	6691	4
8404.26216	7341	5
6981.84494	7661	6
8657.256269	7949	7
8219.500275	9145	8
9131.105648	6875	9
8866.270566	8857	10
8360.902399	9112	11
6812.274368	6325	12

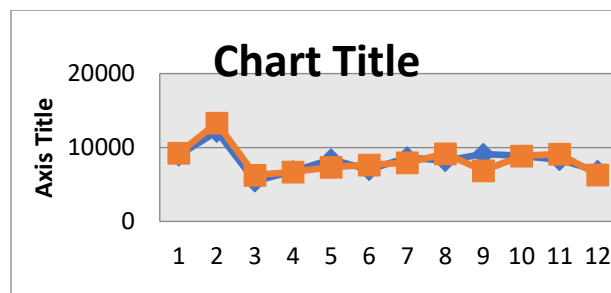
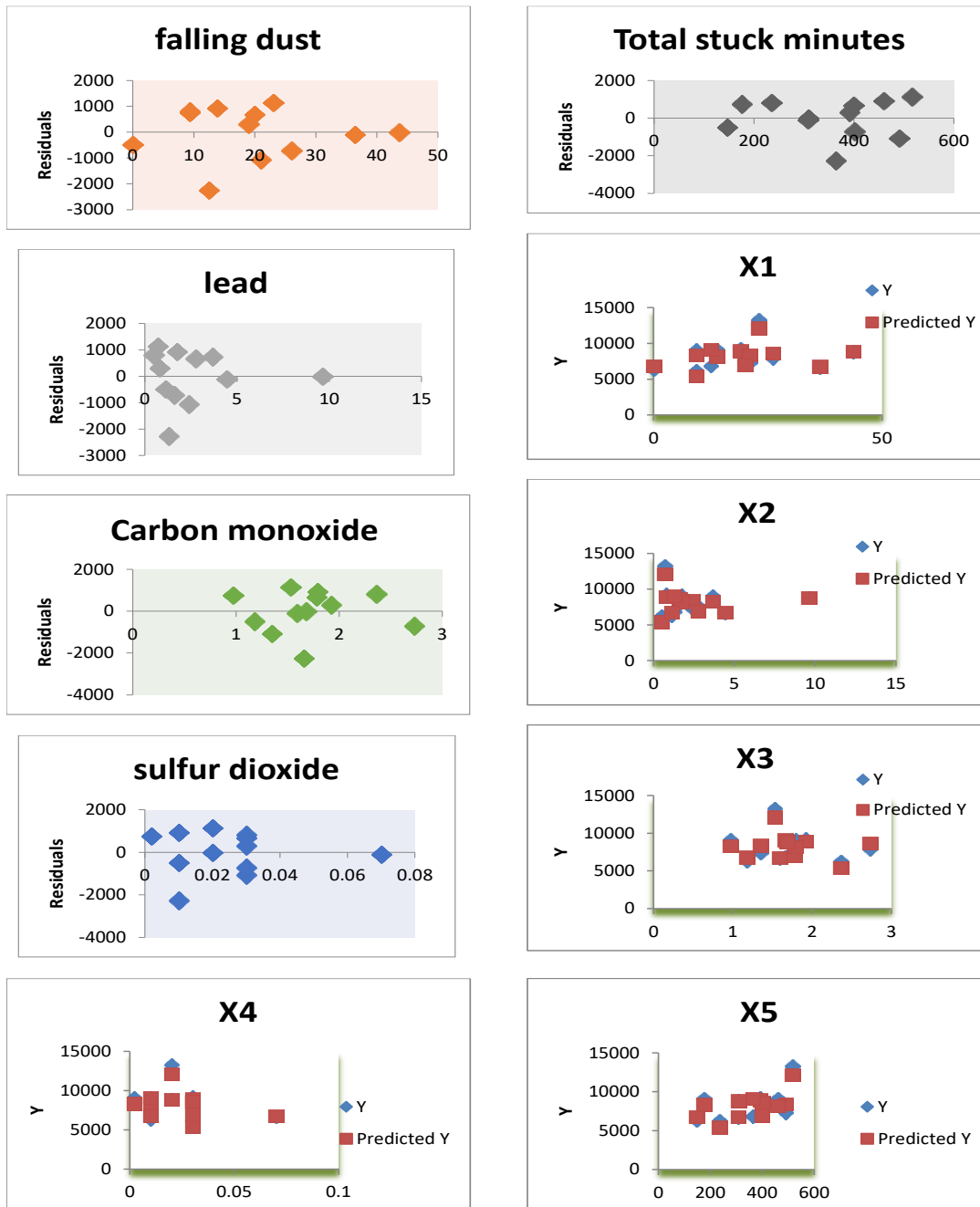


Figure (6) below shows the real and estimated values of Table (4).

As seen in the graph above, the response variable vector's estimated values follow the same pattern as the actual values, suggesting that the estimated model was a good fit for the research data.

The following figures illustrate the air pollutant factors and the number of people with bronchial asthma associated with each factor (decomposing the multiple linear regression model into a simple linear regression model).



Therefore, the values of the correlation coefficient between the influencing variables and the variable of those suffering from bronchial asthma are as shown in the following table:

Table (5): Shows the correlation coefficient values between the influencing variables and the variable of those with bronchial asthma

	Y	X1	X2	X3	X4	X5
Y	1	0.719318	0.570033	-0.12800	0.511172	-0.25534
X1	0.719318	1	0.214394	0.21768	-0.01272	0.358444
X2	0.570033	0.214394	1	-0.20736	0.126433	-0.17416
X3	-0.12800	0.21768	-0.20736	1	0.30360	0.25699
X4	0.511172	-0.01272	0.126433	0.30360	1	0.153714
X5	-0.25534	0.358444	-0.17416	0.25699	0.153714	1

We note from Table (5) that the correlation between the falling dust variable and the variable of the number of people with bronchial asthma is a positive correlation, as it reached (0.7193), meaning there is a strong direct relationship. There is also a positive correlation between the lead variable and the variable of the number of people with bronchial asthma, as it reached (0.570033), meaning there is a good direct relationship. There is a negative correlation between the carbon monoxide variable and the variable of the number of people with bronchial asthma, as it reached (-0.12800), meaning there is an inverse relationship. There is a positive correlation between the sulfur dioxide variable and the variable of the number of people with bronchial asthma, as it reached (0.511172), meaning there is a good direct relationship. There is an inverse relationship between the total suspended particulate matter variable and the variable of the number of individuals suffering from bronchial asthma, as a correlation reached (-0.25534).

Table (6): Comparative Performance of OLS, Genetic Algorithm-Based Regression, and Artificial Neural Network Models for Predicting Bronchial Asthma Cases

Method	MSE	RMSE	MAE	R ²
OLS	0.360	0.600	0.265	0.740
GA-MLR	0.285	0.534	0.211	0.812
ANN	0.378	0.615	0.270	0.753

The GA-based regression model achieved the lowest prediction error and the highest coefficient of determination. Compared with OLS, the MSE was reduced by 20.8% and the coefficient of determination increased from 0.740 to 0.812, indicating that the GA was able to identify a better set of regression coefficients.

9. Conclusions

This section will present the most important conclusions reached in the research, as well as future work:

1. The estimated multiple linear regression model was 74% suitable for the study data.
2. The carbon monoxide and total suspended particulate matter variables do not have significant effects on the number of people with bronchial asthma.

3. Dust storms, lead, and sulfur dioxide are variables that have significant effects on the number of people with bronchial asthma.
4. The correlation table between the number of people with bronchial asthma and the variables affecting it shows that dust storms have a strong positive correlation with the number of people with asthma, followed by lead and sulfur dioxide.

Based on these findings, researchers recommend analyzing the data on the number of individuals suffering from bronchial asthma using a non-normalized multiple linear regression model and robust estimation methods, as well as analyzing the data using modern artificial intelligence algorithms.

5. The genetic algorithm-based regression model outperformed artificial neural networks in terms of comparison criteria and in balancing training and testing data.

References

- [1] Al-Bu'ayshi, Fathi Ali Al-Tahir (2021) .The Effect of Sulfur Dioxide Fumigation on Some Morphological Characteristics of Wheat and Barley Seedlings. International Multilingual Journal of Science and Technology, Vol. 6, No. 6, pp. 3324-3327.
- [2] Al-Harkani, Zaidan Khalaf (2007). An Electrocardiogram Study of Patients with Acute Bronchial Asthma Accompanied by Chest Pain. Iraqi Journal of Medical Sciences, Vol. 5, No. 2, pp. 18-22, Iraq.
- [3] Al-Suraifi, Ali Nasser Abdullah (2022) .Environmental Effects of Dust Storms over Safwan District and the Center of Basra Governorate. Diyala Journal of Human Research, Vol. 4, No. 94, Iraq. 4. Al-Maamouri, Bader Jadou' Ahmed, Al-Alali, [4]
- [4] Israa Adel Rasoul, and Afaj, Adnan Hassan (2016) .The Effect of Some Air Pollutants (Particulate Matter and Lead) on Human Health. Journal of the College of Education for Women, Vol. 27, No. 2, pp. 455-460, Iraq.
- [5] Al-Hashemi, Abdul-Redha Matar Abdul-Redha and Al-Shammari, Ali Karim Hamid (2021) .Spatial Analysis of Air Pollution by Particulate Matter and Heavy Metals in Urban Centers in Wasit Governorate. Al-Qadisiyah Journal of Human Sciences, <http://dx.doi.org/10.13140/RG.2.2.12501.65763>
- [6] Hassan, Bushra Khalid (2012) "Measuring Lead Pollution in Air, Humans, Soil, and Plants in Al-Dora District, Baghdad," Al-Taqni Journal, Vol. 5, No. 2, pp. 1-11, Iraq.
- [7] Akasha, Ali Yousef and Al-Huwaidi, Suleiman Omar (2019) .The Impact of Sulfur Dioxide Emissions from the Mellitah Oil and Gas Complex on Air Quality in Neighboring Cities. Journal of Marine Sciences and Environmental Technologies, Vol. 5, No. 1, Libya.

-
- [8] Ali Khan, Hussein Abdulmutallab (2012) .Factors Affecting Dust Generation in the Industrial Area of Najaf. Al-Thadisiyah Journal of Engineering Sciences, Volume 5, Issue 3, pp. 78-94, Iraq.
- [9] Kadhim, Safaa Karim (2009) .Comparison Between Estimating the Parameters of Multiple Linear Regression Models Using the OLS Method and Linear Object Programming. Journal of Management and Economics, Issue 77, pp. 200-2013, Iraq.
- [10] Mustafa, Suha Walid (2020) .The Environmental Impact of Air Pollution in Basra City with Carbon Monoxide (CO). Arabian Gulf Journal, Volume 48, Issue 3, pp. 361-380, Iraq.
- [11] D. Fogel, Evolutionary Computation: The Fossil Record. Piscataway NJ: IEEE Press, 1998.
- [12] D.E. Goldberg, (1989). Genetic Algorithms in Search. Optimization & Machine Learning. Addison-Wesley. Reading.
- [13] G. Zhang, B. Patuwo, and M. Hue, "Forecasting with artificial neural networks: The state of the art," International Journal of Forecasting, vol.14, 1998.
- [14] S. Warren, (1994)." Neural Networks and Statistical Models", Proceeding of The Nineteenth Annual SAS Users Group International Conference, SAS Institute Inc., Cary, NC, USA.