May - June 2020
ISSN: 0193-4120 Page No. 3276-3282

Modelling of IoT based Vehicular Emission through Regression Analysis

[1]S. Magesh, [2]Chandar G

[1]Professor, [2]Research Scholar,
[1,2]Department of Computer Science & Engineering
Saveetha School of engineering,
Saveetha Institute of Medical and technical sciences
[1]mageshs.sse@saveetha.com, [2]Chan.it13@gmail.com

Article Info
Volume 83
Page Number: 3276-3282
Publication Issue:
May - June 2020

Article History
***Article Received:** 19 August 2019*
***Revised:** 27 November 2019*
***Accepted:** 29 January 2020*
***Publication:** 12 May 2020*

Abstract
Emissions from vehicles such as cars, motorbikes, trains, aeroplanes account for the major share of rationale towards global warming. The various air pollutants that come from vehicles include CO, CO2, NO, sulphur dioxide and many other particulate substrates. In any vehicle class, the emissions and pollutants primarily depend on engine displacement, engine type, fuel type, vehicle age, vehicle speed, urban metric etc. In this paper, we employed regression analysis to model vehicle emissions based on engine displacement, fuel economy, fuel cost for 6000 miles, fuel cost for 12000 miles and urban metric. Also, based on the analysis using a representative dataset, we have concluded that power model fits the data better than an exponential model for the factor of engine displacement and exponential model fits the dataset better than power model with respect to the factor of fuel economy. The values of the correlation coefficient clearly indicate that the city fuel economy influences the emissions and pollutants than engine displacement. We employed transformation principles to convert non-linear power and exponential models into linear models. We have also suggested that polynomial regression techniques can be employed in future with respect to obtaining more accurate emission models.

Keywords: *Regression, correlation, emissions, factors, power, exponential*

1. Introduction

The excruciating problems of global warming and increased air pollutants due to an exponential rise in the number of vehicles on the road along with the growing economic development require proper attention from the research fraternity. In today's diverse world where people are always on the run towards their goals, less is the importance provided to the mode of transport by them that is possessing seemingly reduced vehicular emissions. The level of emissions from various modes of transport will have a great health hazard in the younger and newer generations in the near future. Through the advent and applications of various cutting-edge technologies, many advanced vehicle configurations are always on the release by their manufacturers. Among the different modes of travel, people in every part of the world feel that owning a car and travel by that car is most convenient and luxurious. Various tech specifications in the car such as the type of engine, engine capacity, gear type, fuel type, mileage, ground clearance, size of the vehicle, the power of the engine, maximum torque etc will greatly influence the customer towards a particular vehicle selection. With societal concern, every customer should give due consideration to the emissions that get out of the vehicle as a result of fuel combustion within the vehicle. A wide range of air pollutants is emitted from cars as a result of fuel combustion and other processes. The pollutants

May - June 2020
ISSN: 0193-4120 Page No. 3276-3282

include Hydrocarbons (HC), carbon mono oxide (CO), Carbon dioxide (CO2), Nitrogen oxides (NOs) and many more[4]. Researchers in various car manufacturers are always on the strive to develop their organization's vehicle to possess excellent features and mileage as compared to others. But, at the same time, the vehicle's emission factors should be reduced greatly in order to align with the objectives of the global climate change program. Many countries have started to provide information in public domains pertaining to the vehicles and their movements especially related to the factors of emissions so that the researchers can always build new vehicle emission models based on the available datasets. These emission models are developed in order to make both the car manufacturers and their customers, aware of the impact of certain factors that greatly influence vehicular emissions.

The emission models can be developed using many strategies and techniques, of which the widely and popularly used techniques are statistical techniques. Regression analysis, a widely used statistical technique in many engineering and scientific applications is considered by many advanced countries like the USA, Japan and Germany towards advanced pollution control technology. With these challenges facing the automobile and transport industry, and the availability of advanced techniques to control emissions from the vehicles, this paper is carefully organized to provide a better insight into the development of emission models based on various regression techniques for indicative conclusions on the influence of engine displacement and fuel economy of the vehicles manufactured in the year 2014. Section 2 of the paper will address the overview of various literature that dealt with emission modelling based on a variety of techniques. The various emission based pollutants and factors influencing the same will be dealt with in detail in Section 3(a). Section 3 (b) deals with the influence of IoT on smart emission control of vehicles. Based on the acquired dataset, different models are created using regression analysis and the relation between CO emissions as a target variable and engine displacement and fuel economy as response variables are carefully analysed. This is given in Section 4. Results of the regression model analysis and Conclusion and Future Work are discussed in Section 5.

2. Related Work

According to (Dermot Geraghty et al.) the noise level does not influence the household transport characteristics and that the results recommend that, closeness to transport arteries have more influence towards noise level[6]. The findings suggest that the location of a place ie spatial variable contributes majorly in predicting the urban noise level when compared to an hour of the day, days of the week, the month of the year etc. In addition to it, the contribution towards noise level prediction for seasonal variation is significantly less. The research proposed by (Branimir et al.), suggest that the domestic vehicle that has lower age and mileage, pollutes the environment significantly[3] when compared to other imported vehicles. It also confines that the old age vehicles produce a serious harmful effect which spoils the ecosystem. The regression model (S. D. Oduro et al.) statistically proves that the CO2 emission has linear relationship towards the measure of speed and acceleration of the vehicle[16]. It reasons out that, speed has a greater impact over acceleration as a predictor variable for CO2 emission. The modal emission models (Grace Yi et al.) defines a set of statistical models such as emission maps, regression models and load based model for vehicle emission and fuel consumption[8]. Emission maps are nothing but a matrix containing average acceleration to a combination of speed and acceleration. The findings describe the properties and limitations of the map. Regression model overcomes the limitation of emission maps in terms of sparseness and non-flexibility. The model depicts the existence of a linear relationship of instantaneous speed and acceleration towards vehicle emission. The load-based model figures out the rate of consumption over the engine load.

According to the microscopic fuel consumption and emission model [12], the parameters such as vehicle fuel consumption and emissions are more influential to the level of vehicle acceleration when compared to the vehicle speed. (Grace Yi et al.) proposed least square regression model which incorporates the effect of acceleration and deceleration. The stated model provides a more likely accurate result when compared to CMEM model and POLY model. The author (Lei Yu et al.) suggest that frequent acceleration and deceleration tend to generate more emission. The findings portray the fact that smooth driving, subsequently reduces the rate of emission and also well-planned transportation can change the driving pattern of an individual[14].

3. Emission modelling of Vehicles

3.1 Factors that influence Vehicular Emissions

Anthropogenic contributions and emissions to the atmosphere happen primarily through demographic urbanization and exponential increase in the number of vehicles[12]. Estimating emission data based on different spatial and temporal scale comprises different techniques and is considered to be the primary objective of any emission modelling [10]. Traffic emission information system, to predict air pollution information in real time was discussed in (Liping Xia et al.). The author proposed a Lagrangian model [15]for simulation of the complex road network by creating a traffic information flow database. Also, this database is used to perform comparison tasks between simulated air pollutant observations and real-time air pollution observations. Empirical emission factors of NO2, CO, CO2 were utilized as parameters for the built model. Vehicle network data and road network data are combined to create dynamic models towards achieving integration between TEIS and Traffic GIS. Computer methods for the simulation of air quality process[1] were reviewed from the perspectives of models developed or recommended by governmental agencies for non-regulatory applications. (Aaron Daly et al.). Air pollution model uses the causal relationship between emission, meteorology, atmospheric concentrations and depositions. (Johnson et al. 1980), reviewed the comprehensive long-range transport modelling[11] in the seventies. According to Texas Commission on environmental quality, six categories of emissions data utilized in any emission modelling processes are Point source emissions data, On-road mobile source emissions data, area source emissions data, Non-road and off-road mobile emissions data, biogenic source emissions data and boundary conditions. European and American vehicle emission models such as NAEI(National Atmospheric Emissions Inventory), COPERT(Computer Programme to Calculate Emissions from Road Transport), HBEFA(Hand Book Emission Factors for road transport), ARTEMIS(Assessment and Reliability of Transport Emission Models and Inventory Systems) were reviewed for their applicability in the scenario of New Zealand's traffic[19]. Also, the work by H.Wang identified the scope for improvement in the vehicle emission modelling used by Ministry of Transport in NewZealand based on Vehicle fleet Emission Modelling (VFEM) and Auckland Regional council's Vehicle Emission Prediction Model (VEPM). The work suggests issues and ways to improve the models associated with New Zealand's fleet based emission models. An advanced modelling development model takes into consideration the inclusion of passengers risk attitudes, perceptual conditioning, crowding etc [5].

Many research works have addressed the modelling approaches for emission through transport systems. A framework for providing the interface between urban planning and transport model[9] is discussed in (Gudmundsson et al.) The lacunae of involving human cognitive mechanism in generating travel behaviour and driving behaviour is identified in Stern and Richardson et al. Application of Decision Field theory in creating new frameworks for transport models that will include the effects of travel such as timing, dynamics and type is provided by Stern and Richardson et al. Additional insights into travel behaviour is provided by (Hensher, Rose, et al.) and the potential emission impacts of changes to infrastructure, such as the uptake of electric vehicles (EVs) is discussed[18]. Globally, large-scale energy systems models that provide information about the transport sector internationally is given in[13]. The factors that heavily affect the emissions from compression ignition engine powered vehicles include vehicle class and weight, vehicle age, mileage completed by the vehicle, fuel type, engine displacement, driving in the urban areas, engine exhaust after treatment etc [7]. These factors mostly account for the air pollutants like CO, CO2, NOx, SO2 etc that get emitted out of the vehicles. In this emission modelling research work based on regression analysis, we have taken into account of the factors such as Engine displacement, Fuel Economy, Fuel cost for 6000 and 12000 miles, Urban metric (fuel consumption in Urban conditions in litres per 100 kilometres) for analyzing the CO emissions from the vehicle.

3.2 IoT based Smart Emission Control

IoT has proliferated in all walks of our life and smart emission control is not an exception. Wireless IoT based automotive sensor has been developed as trending technology that measures Nitrogen oxides based emissions in an accurate real-time manner. This is basically accomplished as a response to circumvent U.S emission standards by Volkswagen. IoT based smart emission control helps in providing accurate emission values whenever the environmental agencies required the same. The active RFID and RFID tag will provide the functionalities to collect the emission data from the vehicle exhaust system. Once the data is received by the RFID reader, it is transmitted to the data centre by 4G communication, which has now become a relatively inexpensive mode of communication. For better emissions data collection, Analog to Digital Converter is integrated with the RFID tag. This ADC converts a continuous data stream into a digital value through sampling techniques. ADC integrated with Lambda sensors provides an accurate digital signal from the analog voltage values. With such a sophisticated smart collection of data through IoT for vehicle emissions, it is efficient to analyze the emission data to reduce pollution in huge metropolitan cities around the world.

3.3 Modelling using Regression Analysis

Regression analysis is a statistical technique used in a variety of applications to predict or forecast the target variable based on response variables[2]. Different models can be created in a regression analysis based on the relationship that exists between the target variable and response variables[17]. In the process of building emission models based on the acquired dataset, (Courtesy: Cleaned-up and consolidated car fuel consumptions and emissions data published by the Vehicle Certification Agency (VCA) for years 2000 to 2013[20]. The VCA is an Executive Agency of the United Kingdom Department for Transport. (http://data.okfn.org/data/amercader/car-fuel-and-emissions) the target variable is CO emissions and the response variables are Engine displacement, fuel economy, fuel cost for 6000 miles, fuel cost for 12000 miles and Urban metric.

Exponential Model

$y = ae^{bx}$ .... (1)

can be written as

TEST Engineering & Management

May - June 2020
ISSN: 0193-4120 Page No. 3276-3282

ln(y) = ln(a) + bx .... (2)

Equation (1) represents the non-linear representation of the exponential model which by transformation is converted to linear representation as given in equation (2). In the linear representation, "b" indicates slope and ln(a) indicates intercept and is calculated using equations (3) and (4).

$$b = \frac{n\sum x_i ln(y_i) - \sum x_i \sum ln(y_i)}{n\sum x_i{}^2 - (\sum x_i)^2} \text{.......... (3)}$$

$$ln(a) = \overline{ln(y_i)} - b\,\bar{x} \text{............... (4)}$$

Power Model

$y = ax^b$ ....(5), can be written as

log y = log a + b logx ....(6). Equation (5) and (6) are a non-linear and linear representation of the power model of regression analysis respectively. "b" and "log a" are the slope and intercept of the model respectively.

$$b = \frac{n\sum(logx\_i)(log\,y\_i) - \sum logx\_i \sum log\,y\_i}{n\sum(\log x_i)2 - (\sum \log x_i)2} \text{ .... (7)}$$

$$\log a = \overline{\log y} - b\,\overline{log\ x} \text{ .... (8)}$$

Sum of the squares of the residuals for the Exponential Model and power model is calculated as S_r using equation (9) and equation (10) respectively.

$$S_r = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n}(\ln(y) - \ln(a) - bx)^2_i \text{ .... (9)}$$

$$S_r = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n}(\log(y) - \log(a) - b\log(x))^2_i \text{ .... (10)}$$

After calculating sum of the squares of the residuals error the model that has got the least values fits the data better.

Calculate r the correlation Co-efficient

$$r = \frac{n\sum xiyi - (\sum xi)(\sum yi)}{\sqrt{n\sum x_i^2 - \sqrt{(\sum xi)}}} \text{........(11)}$$

r value lies between -1 and +1. If r has value +1, it means it is positively correlated and if it is 0, it is not correlated and if it has values - 1, it means that it is negatively correlated.

The conversion of exponential and power models of regression (it could be single or multiple) implies that a non-linear relationship is transformed into linear relationships.

4 Results of the Regression Model Analysis

With the representative dataset mentioned in the previous section, we have carried out the calculations to find out the slope and intercept for various factors of Carbon monoxide emissions using MS Excel based on exponential model and power models of regression. The results are tabulated in Table 1 and Table 2. Also, the sum of the squares of the residual error for various factors for exponential and power model was calculated and tabulated in Table 1 and Table 2 respectively.

Table 1: Exponential Model of Regression Analysis

Metric	Slope b	Intercept a	Equation Model	Sum of the squares of the residual error(S_r)
Engine Displacement	$1.0395e^{-05}$	507.571	$y = 507.571e^{1.0395e^{-05}x}$	76.8984
Fuel Cost 12000	-0.08527	1026.315	$y = 266.75e^{0.00028896x}$	746327.78
Fuel Economy	-0.08527	1026.315	$y = 1026.315e^{-0.08527x}$	70.20609
Urban metric	0.03127	335.076	$y = 335.0764e^{0.03127x}$	72.98224

Table 2: Power Model of Regression Analysis

Metric	Slope b	Intercept a	Equation Model	Sum of the squares of the residual error(S_r)
Engine Displacement	0.0233	434.7963	$y = 434.7963x^{0.0233}$	14.505

TEST Engineering & Management

May - June 2020
ISSN: 0193-4120 Page No. 3276-3282

Fuel Cost 12000	-0.001	1426.3387	$y = 1426.3387x^{-0.0001}$	83106532.9161
Fuel Economy	-0.5883	1708.460	$y = 1708.460x^{-0.5883}$	409999864.744
Urban metric	0.5883	113.7825	$y=113.7825x^{0.5883}$	13.454

From the Table. 1 and Table.2, it is vividly clear that Power model suits the best for the factors of engine displacement and urban metric towards carbon mono oxide emissions and that exponential model fits better for the factors of fuel cost for 12000 miles and fuel economy. The graphs (Fig.1, Fig.2, Fig.3, Fig. 4) below indicate the plots for various factors based on the best fit model.

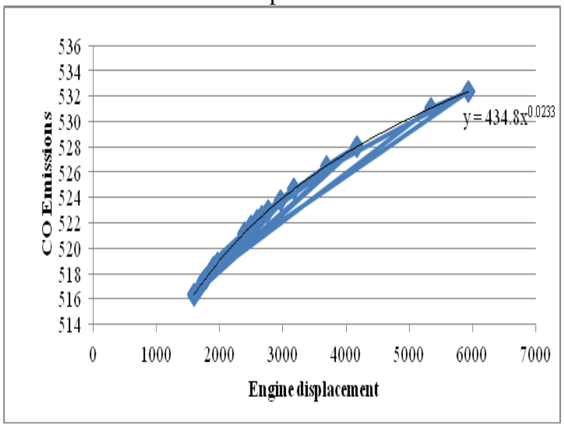

Figure 1: Engine displacement vs. CO Emissions (Power Model)

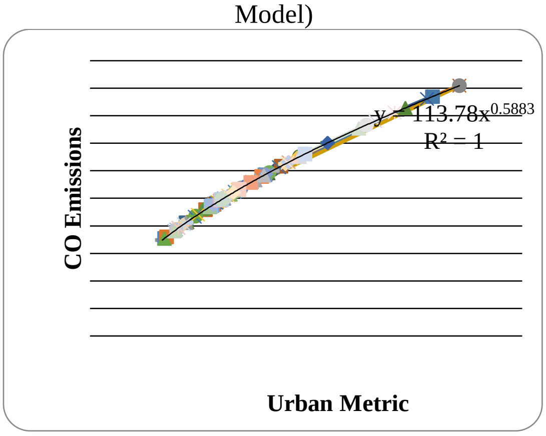

Figure 2: Urban Metric vs. CO Emissions (Power Model)

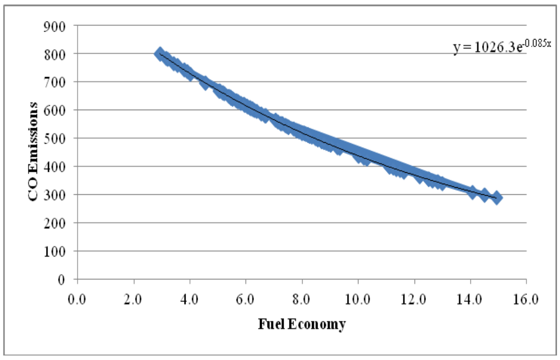

Figure 3: Fuel vs. CO Emissions (Exponential Model)

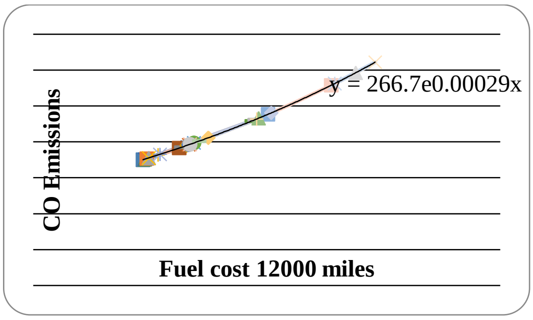

Figure 4: Fuel Cost 12000 miles vs. CO Emissions (Exponential Model)

Table. 3 given below provides the correlation coefficient calculated based on various factors

Table 3. Correlation coefficient

Metric	Correlation Coefficient
Engine Displacement	0.041
Fuel Cost 12000	1
Fuel Economy	-0.344
Urban	0.275

The following observations were obtained out of Table .3:

a. Engine displacement, fuel cost for 12000 miles, urban metric are positively correlated to carbon monoxide emissions.

b. Fuel Economy is negatively correlated with carbon mono oxide emissions

c. Out of the factors that are positively correlated, fuel cost for 12000 miles had the best positive fit.

4. Conclusion and Future Work

Emissions from vehicles contribute a major share in global warming, thereby hindering the sustainable green earth. In this paper, the need for emission modelling is stated with different types of air pollutants and their adverse effects. Modelling of vehicular emissions were carried out in many research works and they were discussed in the literature review. The use of regression analysis for emission modelling along with necessary mathematical background was discussed in this work. The paper concludes with the observations with respect to the emission factors like engine displacement, fuel cost for 12000 miles, fuel economy, an urban metric based on the exponential model and power model. In the future, we plan to undertake polynomial non-linear regression analysis to test the factors for their influence on vehicle emissions.

5. Acknowledgement (For the used datasets)

The authors express their profound thanks to Vehicle certification agency(VCA) for providing with Cleaned-up and consolidated car fuel consumptions and emissions data for years 2000 to 2013. The VCA is an Executive Agency of the United Kingdom Department for Transport. http://data.okfn.org/data/amercader/car-fuel-and-emissions.

References

[1] Aaron Daly., and Paolo Zannetti., "Air Pollution Modeling– An Overview AMBIENT AIR POLLUTION" Published by *The Arab School for Science and Technology*. 2007.

[2] Alessandra Cappiello., Ismail Chabini., Edward K. Nam., Alessandro Luè., and Maya Abou Zeid. "A Statistical Model of Vehicle Emissions and Fuel Consumption". Massachusetts Institute of Technology. Cambridge, MA 02139, USA.

[3] Branimir MilosavljeviT., RadivojePešiT., and PredragDašiT., "Binary Logistic Regression Modeling of Idle CO Emissions in Order to Estimate Predictors Influences in Old Vehicle Park" An ResearchArticle, Hindawi Publishing Corporation, *Mathematical Problems in Engineering*.2015.

[4] Benjamin Afotey., Melanie Sattler., Stephen P. Mattingly ., and Victoria C. P. Chen., "Statistical Model for Estimating Carbon Dioxide Emissions from a Light-Duty Gasoline Vehicle" *J. Environmental Protection*, Vol 4, pp 8-15,2013.

[5] David A. Hensher. , John M. Rose ., Waiyan Leong., Alejandro Tirachini., and Zheng Li.," Choosing Public Transport —Incorporating Richer Behavioural Elements in Modal Choice Models", *Transport Reviews: A Transnational Transdisciplinary Journal*, Vol 3, no.1, pp 92-106,2013.

[6] Dermot Geraghty., and Margaret O'Mahony., "Urban Noise Analysis Using Multinomial Logistic Regression".*J.Transportation Engg.,@ASCE Journal of Transportation Engineering*, 10.1061/(ASCE)TE.1943-5436.0000843, 04016020-1 to 1, 2016

[7] Clark NN., Kem JM., Atkinson CM., Nine RD., "Factors affecting heavy-duty diesel vehicle emissions", *Journal of the Air & Waste Management Association*, pp 84-94, 2002.

[8] Grace Yi., Hualiang Qi., Harry Teng., Lei Yu., "Microscale Emission Models Incorporating Acceleration and Deceleration", *J. Transportation Engineering*, 2004

[9] Gudmundsson *H., Hall R.P., Marsden, G., Zietsman, J., "Introduces into the planning and management of sustainable development in the transportation sector*"Springer Text in Business and Economics,2016

[10] Journard., "Methods of estimation of atmospheric emissions from transport: European scientist network and scientific state-of-the-art action COST 319 final report",1999.

[11] Johnson., "A Review of Long-range Transport Modeling", *Journal of Applied Meteorology*, vol No:19, 1980

[12] Kyoungho Ahn.," Microscopic fuel consumption and emission modelling". Transport emission Model for Line Sources- Methodology,1998

[13] James A. Edmonds., Katherine V Calvin, Leon E Clarke, Anthony C Janetos, Son H Kim, MarshallA Wise, and Haewon C McJeon. 2012. "Integrated Assessment Modeling." Springer, 2012

[14] Lei Yu., Don Small., " Vehicle emissions", Interdisciplinary Lively Applications Project, 2003.

[15] Liping Xia., Yaping Shao., "Modelling of traffic flow and air pollution emission with application to Hong Kong Island", *Environmental Modelling & Software* 20, pp 1175–1188, 2005.

[16] Oduro.S.D., Metia.S, Duc.H., and Ha.P.Q., "CO2 Vehicular Emission Statistical Analysis with Instantaneous Speed and Acceleration as Predictor Variables, ICCAIS, 2013

[17] Steven C. Chapra., Raymond P. Canale.,and Berger Chair., "*Computing and* Engineering: Numerical Methods *for* Engineers". Sixth Edition. Tufts University.

[18] Shafiei, E., Thorkelsson, H., A´ Sgeirsson, E. I., Davidsdottir, B., Raberto, M., & Stefansson, H. "An agent-based modelling approach to predict the evolution of market share of electric vehicles: A case study from Iceland". Technological Forecasting and Social Change, 79, pp 1638–1653.2012.

[19] Wang.H., McGlinchy.I., "Review of vehicle emission modelling and the issues for New Zealand", *Australian Transport Research Forum*,2009

[20] *http://data.okfn.org/data/amercader/car-fuel-and-emissions*)