An Analysis of Vehicle Emission Factors to Determine Highest Priority Factors

Thomas Jefferson High School for Science and Technology

Rationale

As wildfires plague the drought-ridden American west, and oil spills allow for the captivating viral videos of flames upon the ocean, the continued rise of global temperatures and gaping holes in the ozone layer have caused industries to grapple with their role in the earth's current implosion. Of the contributing factors, the prominence of vehicles is certainly to blame: their negative effects have been extensively documented. They are a necessity to a functioning global society, and yet, are largely responsible for the deterioration of the environment it inhabits. Due to both federal regulation and the competitive desire for a free market, vehicle companies have begun to take initiative to limit their influence and reduce the particle emissions of their products. To ensure the effective impact of their actions, these companies look to studies to evaluate the most influential measures: this is an extension of previous studies to further inform the vehicle industry, and, hopefully, increased security in the future of society. To assess the factors of vehicular emissions, it is necessary to establish the significance of the emission that is attributed to the vehicles and the implications that these effects have on the health and safety of the human population and environment.

The first study I reviewed looked at the effect of vehicle emission factors by recording the levels of particulate matter, known to originate from vehicles on the lanes of tunnels, during different seasons (Li et al., 2020). The measurements were subsequently compared to the null levels of particulate matter to find the contribution attributed to vehicle emissions. The findings of this study suggest that the levels of emissions vary with larger amounts in the winter and fewer emissions in the summer. However, even when accounting for the variability observed by this study, they still concluded that on average 37% of CO (carbon monoxide) emissions and 39% of NOx emissions are caused by vehicles. This study provides two variables of significance caused by vehicles that should be included in a multivariable regression analysis of the characteristics that contribute to these emissions. Additionally, it provides an explanation for possible errors in multiple regression analysis and during a discussion is an important factor to consider when addressing the limitations of my study.

Extending upon the study by Li et al., an evaluation of the significance of these emissions is critical to understanding the importance of actions to prevent them. That is to say, even if vehicles cause 39% of carbon monoxide, what impact does this have on our environment and the posterity of humanity? Utilizing years of life lost, chronic bronchitis, and restricted activity days as indicators of the rapid carbonization of society, a research team in Poland evaluated the effect of fine particle emission of 2.5 mm or less (Wywra et al., 20). This study provides many parallels to the emission of vehicles because the particles measured and those from emission are essentially the same, thus the findings of this study can be used as evidence to support reforms to vehicles. They found that in Poland during 2015, 218,415 years of life, 73,487 years of restricted daily activity, and 24,907 new cases of chronic bronchitis were caused by particle emissions, this equated to 16.15 billion euros of lost earnings. This suggests that an investigation into the limitation of vehicle emission is a worthwhile endeavor because of the potential positive impact that it can have on the quality of life and net profit of a country.

A legislative approach was investigated by Zavala et al. and aimed to understand the regulatory power of limitations to vehicle emissions (2017). They study a variety of trucks in Mexico City with varying levels of exhaust regulators by recording the emitted levels of CO and NOx among other factors. They concluded that although these regulations were effective at limiting the emissions, off-road testing was not entirely representative of the conditions faced by these drivers on commutes. As a result, all data taken from off-road measurements is less accurate of a vehicle's performance, because of factors such as the change in elevation, and as previously mentioned, the season of the year. However, these findings were disputed by Quiros et al. published only a year prior in 2016 which study trucks along the California corridor of long haul truck routes. They monitored seven trucks with varying levels of emission reduction as they ran along six heavily used routes with varying characteristics to test the efficacy of the particulate emission factors and ensure they are below the limit allowable by the EPA. To rationalize the varying outcomes of these experiments it is important to recognize the influence of the environment and aftermarket emission regulators used. When read together, they provide a compelling view that a lot of the variation is dependent on the actions of the driver as opposed to the manufacturer. This is not the most reasonable approach because it allows these production facilities to evade blame: instead, it is put on the consumer. To rectify the disparity of onus, an analysis of the contributing characteristics of a vehicle's design is required to understand how to best limit emission on behalf of the manufacturer and not the consumer.

When designing a vehicle though, it is not clear what factors of the design such as the weight or horsepower determine the particle emission, despite the clear implication on our lives as demonstrated by Wyrwa et al. The purpose of this study was to best understand the design elements of a vehicle that most significantly contribute to particulate emissions?

Design of the Study

To investigate the emission factors of vehicles, our population of interest is the vehicles themselves and all pertinent measurements about their design. The most comprehensive and reliable dataset available is the EPA's annual set from their National Vehicle and Fuel Emissions Laboratory: it includes all identifying features and variables of interest. As a result of a large amount of available information, I would expect a large amount of design of a vehicle's contribution to its emissions to be captured by a multivariable regression model. All datasets since 1984 are public knowledge and could therefore be used to conduct this analysis. However, because of the fast pace nature of technological innovation, randomly selecting a dataset would limit the predictive power of a model; for this reason, I chose to focus on the most current available dataset from 2022. This choice is supported by the possibility that there is variability between the significant factors from year to year: to best reduce particulate emissions, the current year must be analyzed. Unlike most available datasets, the EPA's collection of vehicles is a census of every road-legal car produced in 2022, not a random sample. This created issues on how to proceed because using the entirety of the dataset would cause significant strain on my computing potential.

Instead, I took a random sample of 100 vehicles as a representative sample of the population of all road-legal vehicles in the US produced in 2022. This will allow me to make inferences about the population and therefore have significant implications. With the creation of the random sample of 100 vehicles, I began by isolating variables with any possible influence and identifying the response variables carbon monoxide and NOx emissions, previously found to be impactful to human health in the literature review. With this list of potential explanatory variables including, rated horsepower, the number of cylinders and gear, transmission type, driving system, and fuel efficiency, I created indicator variables for the categorical variables of transmission type and driving system. I decided to use backward elimination to find the statistically relevant model because without expertise in the possible effects of each of the explanatory variables, running a method allowed the elimination of least significant terms.

Furthermore, by running multiple regression analyses I can compare the relative significance of a variable to affect let's say THC versus CO2.

Using Minitab, I ran the respective regressions with backwards elimination with the null hypothesis that all of the variables would have a slope of zero and the alternative hypothesis that at least one would differ from zero. The general condition shared amongst the two regressions I analyzed was independent observations which was because of the random generation of our sample. However, the data was not linear and therefore a transformation was required before proceeding to an analysis of the data. Based upon the curvature of the normal probability plots, I took the natural log and square root to find the most acquired relationship: natural log for CO and square root for NOx. Upon transformation, the conditions of zero mean, equal variance, and normality were also met.

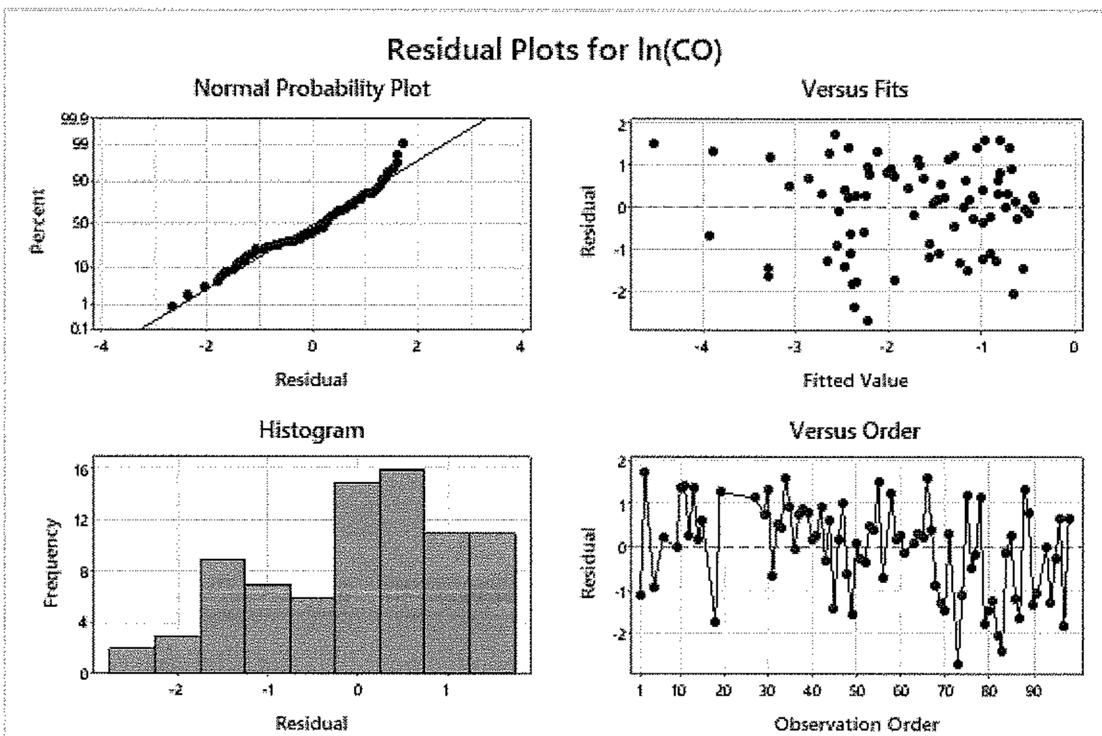

Fig 1. Residual plots of carbon monoxide emissions after a logarithmic transformation. As demonstrated by the visual and the points following the expected distribution of a normal probability plot, are evenly distributed above and below zero.

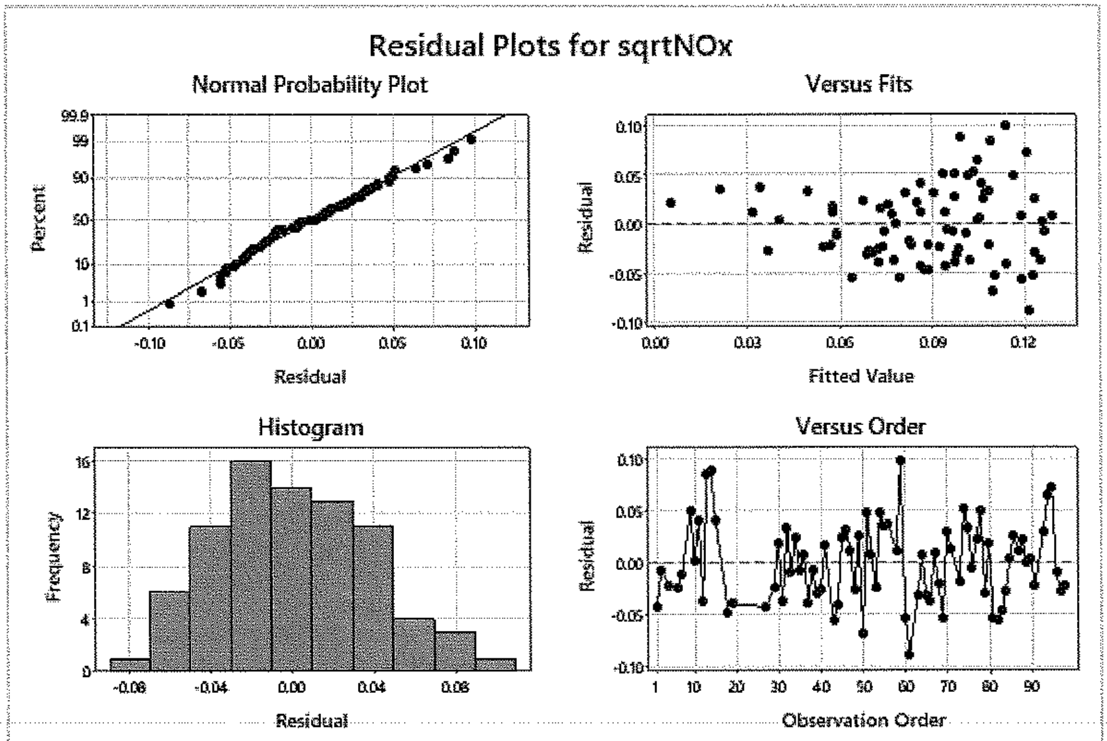

Fig. 2 Residual plots of NOx emissions after a square root transformation. As demonstrated by the visual and the points following the expected distribution of a normal probability plot, are evenly distributed above and below zero.

Results

Based upon the backward elimination multivariable regression analysis of NOx emission, the only statistically significant variable in the fuel efficiency of the vehicle (Fig. 3). It explains 31.07% of the variability in the emission of NOx: this makes intuitive sense as a more efficient motor is likely to produce fewer emissions on average (Fig. 5). From the analysis, with an increase of one mile per hour to the fuel efficiency, a vehicle is likely to produce 0.002 g/mi less of sqrt(NOx): a seemingly small amount. However, given the number of vehicles on the road and the high magnitude of fuel efficiency it is reasonable to reduce NOx emission by increasing fuel efficiency.

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	0.1563	0.0121	12.97	0.000	
Fuel Eff. (mpg)	-0.002063	0.000341	-6.05	0.000	1.00

Fig 3. Coefficient table of the model of lnNOx

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.0387239	31.95%	31.07%	28.60%

Fig. 4 Model Summary of a multivariable regression of lnNOx

The regression model of the natural logarithm of CO was not so simple. Composed of three different variables, the regression model of ln(CO) is more explanatory of variability, with an R-squared value of 39.96% (Fig. 6). Unlike the model for NOx, there are three significant variables that have a statistically significant effect on the emission rates of the ln(CO). Among them, a continuously variable transmission, Code_CVT, had the largest magnitude of the coefficient (Fig. 5). Despite this large magnitude, the fuel efficiency has the largest potential impact because of the ability to continuously improve the miles per gallon. This suggests that with significant technological advancement ln(CO) emissions could become negligible.

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	2.172	0.764	2.84	0.006	
_ of Cylinders and Rotors	-0.1902	0.0885	-2.15	0.035	1.39
Fuel Eff. (mpg)	-0.0906	0.0125	-7.27	0.000	1.73
Code_CVT	1.321	0.473	2.79	0.007	1.36

Fig. 5 Clioefficient table of the model of sqrtCO

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
1.08667	42.24%	39.96%	35.20%

Fig. 6 Model Summary of a multivariable regression of sqrtCO

Discussion

Based upon the findings of Li et al., it is understandable that much of the emission is explained by external factors that limit the explanatory power of physical properties. However, the results of both of the studies suggest that fuel efficiency is the greatest determinant of particulate emissions. Additionally, it is likely that there is multicollinearity between the factors of the preliminary model before backward elimination: this would suggest that fuel efficiency is explained by some combination of the other factors. To extend upon the conclusion of this study, given the magnitude of fuel efficiency to determine the emissions of CO and NOx, further studies should be conducted to improve fuel efficiency. This research would have far-reaching implications to improve the health and productivity of nations by limiting emissions of particles less than <2.5 mm in width (Wyrwa et al. 2020).

Reflection

Throughout this project, my largest area of growth was the creation of a well-developed and thoroughly researched literature review: this was most important because it helped me to develop a comprehensive understanding of the vehicles' emissions to create a complete research question. This is a beneficial skill to take advantage of in future research studies because it enables a more informative study with substantial results. Additionally, this study taught me to find and modify data sets to allow for significant research to be conducted. In the case of this study, I learned the power of public data sets because they are generally comprehensive and easily accessible, and how to choose a representative sample to optimize the outcomes. Being able to find raw data allows a researcher to verify the results of articles and empowers them to think freely.

References

Clairotte Michaël, Suarez-Bertoa, R., Zardini, A. A., Barouch, G., Jelica, P., Valverde, V., . . . Covadonga, A. (2020). Exhaust emission factors of greenhouse gases (GHGs) from european road vehicles. Environmental Sciences Europe, 32(1) doi:https://doi.org/10.1186/s12302-020-00407-5

Li, X., Dallmann, T. R., May, A. A., & Presto, A. A. (2020). Seasonal and long-term trend of on-road gasoline and diesel vehicle emission factors measured in traffic tunnels. Applied Sciences, 10(7), 2458. doi:https://doi.org/10.3390/app10072458

Quiros, D. C., Arvind, T., Saroj, P., Besch, M., Pragalath, T., Berk, D., . . . Hu, S. (2016). Real-world emissions from modern heavy-duty diesel, natural gas, and hybrid diesel trucks operating along major california freight corridors. Emission Control Science & Technology, 2(3), 156-172. doi:https://doi.org/10.1007/s40825-016-0044-0

Real world vehicle emission factors for black carbon derived from longterm in-situ measurements and inverse modelling. (2021). Atmosphere, 12(1), 31. doi:https://doi.org/10.3390/atmos12010031

Wyrwa, A., Suwała, W., Pluta, M., Olkuski, T., & Raczyński, M. (2020). The impact of decarbonization scenarios on air quality and human health in Poland—Analysis of scenarios up to 2050. Atmosphere, 11(11), 1222. doi:https://doi.org/10.3390/atmos11111222

Zavala, M., Molina, L. T., Yacovitch, T. I., Fortner, E. C., Roscioli, J. R., Floerchinger, C., . . . Jazcilevich, A. (2017). Emission factors of black carbon and co-pollutants from diesel vehicles in mexico city. Atmospheric Chemistry and Physics, 17(24), 15293-15305. doi:https://doi.org/10.5194/acp-17-15293-2017