World Academy of Science, Engineering and Technology
International Journal of Environmental and Ecological Engineering
Vol:17, No:3, 2023

# Statistical Analysis and Impact Forecasting of Connected and Autonomous Vehicles on the Environment: Case Study in the State of Maryland

Alireza Ansariyar, Safieh Laaly

Open Science Index, Environmental and Ecological Engineering Vol:17, No:3, 2023 publications.waset.org/10012997.pdf

***Abstract*****—**Over the last decades, the vehicle industry has shown increased interest in integrating autonomous, connected, and electrical technologies in vehicle design with the primary hope of improving mobility and road safety while reducing transportation's environmental impact. Using the State of Maryland (M.D.) in the United States as a pilot study, this research investigates Connected and Autonomous Vehicles (CAVs) fuel consumption and air pollutants including Carbon Monoxide (CO), Particulate Matter (PM), and Nitrogen Oxides (NOx) and utilizes meaningful linear regression models to predict CAV's environmental effects. Maryland transportation network was simulated in VISUM software, and data on a set of variables were collected through a comprehensive survey. The number of pollutants and fuel consumption were obtained for the time interval 2010 to 2021 from the macro simulation. Eventually, four linear regression models were proposed to predict the amount of C.O., NOx, PM pollutants, and fuel consumption in the future. The results highlighted that CAVs' pollutants and fuel consumption have a significant correlation with the income, age, and race of the CAV customers. Furthermore, the reliability of four statistical models was compared with the reliability of macro simulation model outputs in the year 2030. The error of three pollutants and fuel consumption was obtained at less than 9% by statistical models in SPSS. This study is expected to assist researchers and policymakers with planning decisions to reduce CAV environmental impacts in M.D.

***Keywords*****—**Connected and autonomous vehicles, statistical model, environmental effects, pollutants and fuel consumption, VISUM, linear regression models.

## I. Introduction

TRAFFIC congestion, air pollution and travel safety have long been among the governments' major concerns. Travel data from 1990 to 2017 indicated that the average travel time in the United States increased from 49.35 to 55.62 minutes [1]. The average commute time to work was estimated at 53.2 minutes according to the 2020 Census data [1]. An annual loss of 99 hours [3] was estimated in 2019 due to traffic congestion caused by increasing number of registered cars in the United States, 276.5 million in 2019 [2]. Nevertheless, new technologies must be implemented in the new generation of light and heavy vehicles, including CAVs, to minimize the impacts of congestion, air pollution, sound pollution, fuel consumption, safety, and public health concerns.

The interactions between CAV technology and the environment have been scrutinized by researchers who utilized either logical estimates [4]-[6], or used mathematical frameworks with the available data to extract numerical results [5] on how CAVs will alter pollutants and energy consumption. The implementation of CAVs, combined with the use of alternative fuels, were promising in the efforts to reduce the amount of GHG emissions from transportation.

This study concentrates on environmental impacts of CAVs and proposes multiple meaningful statistical models for CAVs' pollutants (C.O., NOx, PM), takes into account the correlation of CAVs customers' income level, number of registered CAVs, and other sociodemographic characteristics, and proposes a model to calculate CAV's future fuel consumption in Maryland. A comprehensive dataset for year interval 2010-2021 was collected consisting of the following variables: number of CAVs customers in different counties, number of registered CAVs, vehicle miles traveled (VMT) by CAVs, average fuel consumption, rate of pollutants for each type of CAVs model, race, and the income of CAVs customers.

The remainder of this report is structured as follows: Section II: Data, Section III: Methodology, Section IV: Research questions, Section V: Literature review, Section VI: Modeling and analysis of the collected data for the State of Maryland, Section VII: Models validation, Section VIII: Conclusion.

## II. Data

The real-world data of the State of Maryland were collected and analyzed using a statistical regression model to understand CAVs' current and future environmental impacts. Utilizing statistical modeling, the following variables of the time interval 2010~2021 database were used for analysis: [7]-[13]:

1. Population: number of CAVs customers (consisting of all 23 counties in Maryland)
2. Number of registered CAVs
3. Level of CAV customers' education (bachelor's degree and higher)
4. Age groups of CAV customers' ($\geq$18 years old; people with driving license)
5. Number of VMT
6. Income levels
7. Race of CAVs customers
8. Household size

The purpose of this analysis is to find meaningful regression models to predict the amount of CAV's air pollutants and fuel

Alireza Ansariyar is Ph.D. student, Department of Transportation and Urban Infrastructure Systems (TUIS), Morgan State University, Baltimore, MD, 21251 United States (corresponding author, e-mail: alans2@morgan.edu).

Safieh Laaly is Adjunct Faculty, TUIS, Morgan State University, Baltimore, MD, 21251 (e-mail: Safieh.Laaly@morgan.edu).

consumption in future years and to understand the correlation of pollutants (CO, NOx, PM) and fuel consumption, as dependent variables, with the eight aforementioned independent variables. Numerous regression models were obtained and four highly correlated regression models were identified.

*Methodology*

Maryland transport network was simulated in VISUM macro simulation. The traffic analysis zone or transportation analysis zone (TAZ) and Origin-Destination (OD) demand matrices were obtained from transportation master plans. The network was calibrated based on September morning peak hour for each year, corresponding with the schools' opening with reliable average travel times. The macro model was run, and the output of the macro simulation (pollutants and fuel consumption) was used as the input of statistical analysis in SPSS. A great number of linear regression models were explored and four of the most efficient models were identified. Then, statistical tests for each model were obtained. Version 28.0 of the SPSS software was used for analysis of the data categories.

*Research Questions*

This study aims to answer the following questions:

1. How can the CAV's environmental impact be forecasted for the State of Maryland?
2. What is the correlation of CAVs' pollutants and fuel consumption with the sociodemographic characteristics of its customers?

## III. Literature Review

### *A. CAV Environmental Impact*

The environmental effects of CAVs significantly depend on its market penetration rate (MPR). With high penetration, CAVs' environmental impact is greater. Transportation contributes to approximately one quarter of all energy related greenhouse gas (GHG) emissions in the world [14], and about 29% of total GHG emissions [15] in the U.S. The new generation (personal and shared) of CAVs are reshaping transportation and mobility by replacing human driver and service provider with automation. Improvements in vehicle efficiency and functionality, however, do not necessarily translate to all positive environmental outcomes. It stands to reason that technologies can boost the efficiency of the vehicles, but they may have potential negative impacts which must be recognized. The interactions between CAV technology and the environment have been scrutinized in a number of researches: a) through use of logical estimates of CAVs' fuel consumption, GHG emissions, etc., and b) through mathematical framework using available data to generate numerical results [23].

### *B. GHG Emissions and Energy Consumption*

In 2014, the Society of Automotive Engineers (SAE) classified autonomous vehicles (AVs) in 6 levels of automation from level 0, with no automation, to level 5, full automation [16]. Each type of CAV may have a specific pollution rate. According to the literature [15], major vehicle pollutants are CO, hydrocarbons (HC), NOx, and PM.

The state-of-the-art highlighted the vehicles' size, design, route choice, consumer's choice, penetration levels, road characteristics, and adoption patterns, as "principal factors", with considerable impression on CAVs environmental pollution [23].

Previous studies [4], [17], [23] have identified the general areas where vehicle automation can potentially impact energy and emissions. These studies [4], [18], [23] show that energy consumption and GHG emissions are strongly correlated with the vehicle's acceleration and speed. A particular research [5] utilized the data from previous studies to generate a list of key factors such as vehicle characteristics (i.e., vehicle weight), transportation network, and consumer choice among factors with great impact on CAVs energy consumption and GHG emissions [5], highlighting on the weight of vehicle, its performance, and rightsizing that can reduce energy consumption. An additional factor is the diminished AVs accidents that makes other safety equipment in vehicles unnecessary, therefore, reducing AVs' environmental impact [6] which affects the initial estimates of a range of energy consumption due to AVs' widespread adoption.

Fagnant and Kockelman [19] explored available estimates for potential effects of ten AVs' energy consumption and impacts and discovered that AVs' widespread deployment can lead to dramatic fuel savings. However, the amount of CAVs' fuel consumption and resulting pollutants is impacted also by traffic congestion during peak hours. Previous studies (e.g., [17]-[19]) have also documented that emissions and fuel consumption increase significantly in congestion situations. Moreover, as the MPR of CAVs increases, a considerable decrease in fuel consumption and related pollution should be expected if combined with effective strategies to solve congestions problems.

### *C. The Gap*

Having reviewed numerous related research studies, the following gaps existing in the literature were recognized:

i. Lack of investigative evidence of simultaneous impacts of fuel consumption and environmental pollution changes.
ii. Absence of a "statistical model" based on a real-world data to predict the amount of CAVs' pollution.
iii. No evidence is detected in the previous studies that refer to the interaction of certain variables (i.e., VMT, age, income, race, education, number of registered CAVs, and the exact amount of CAVs' air pollution in a specific time interval).

Therefore, it was recognized that a) the previous studies did not explore a simultaneous impact of the fuel consumption and pollution changes, and b) the interaction of variables was not studied in an integrated statistical model. The principal achievement of this research is developing such statistical models, using real-world data, which has not been explored before.

## IV. Regressions Modeling

### A. Linear and/or Nonlinear Regression

A regression model describes the relationship between a dependent or response variable (Y), and one or more independent or predictor variable (X).

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + \cdots \tag{1}$$

The result of the significant correlations for this study is shown in Table I summarized as: no significant relationship was specified among dependent variables with customers 16 to 20 years old (NDLH1), customers 25 to 29 years old (NDLH3), customers 50 to 59 years old (NDLH6), and VMT. Hereupon, these four variables were not taken into account in the final statistical models. The integrated age group considered at the onset of the modeling showed no significant correlation with dependent variables. The variations and the standard error of an integrated age group were considerable. Hence, the new integrated age groups were re-categorized to determine which group(s) may be most correlated with the dependent variables.

TABLE I
Correlation Results

| Variable | Abbreviations | Amount of CO | Amount of NOx | Amount of PM | Total fuel consumption by CAVs |
|---|---|---|---|---|---|
| Population | P | -.975** | -.935** | -.975** | .940** |
| Total registered cars | TRC | -.960** | -.960** | -.975** | .942** |
| Number of registered CAVs | NRCAVs | -.956** | -.963** | -.956** | .972** |
| Number of educated customers | NoEC | -.984** | -.984** | -.984** | .960** |
| Number of customers_age16to20YearsOld | $NDLH_1$ | | | | |
| Number of customers _age21to24YearsOld | $NDLH_2$ | .870** | .865** | .870** | -.790** |
| Number of customers _age25to29YearsOld | $NDLH_3$ | | | | |
| Number of customers _age30to39YearsOld | $NDLH_4$ | -.947** | -.930** | -.947** | .972** |
| Number of customers _age40to49YearsOld | $NDLH_5$ | .805** | .800** | .805** | -.735** |
| Number of customers _age50to59YearsOld | $NDLH_6$ | | | | |
| Number of customers _age60to69YearsOld | $NDLH_7$ | -.989** | -.950** | -.989** | .967** |
| Number of customers _age70to79YearsOld | $NDLH_8$ | -.989** | -.989** | -.989** | .994** |
| Number of customers _age_greaterthan80YearsOld | $NDLH_9$ | -.984** | -.984** | -.984** | .986** |
| Vehicle Miles Traveled | VMT | | | | |
| Income | Inc | -.934** | -.933** | -.933** | .943** |
| Household size | HHS | -.946** | -.946** | -.947** | .909** |
| Black or African-American customers | BAAC | -.991** | -.965** | -.991** | .972** |
| White customers | WC | .732** | .731** | .731** | -.803** |
| American-Indian customers | AIC | -.993** | -.993** | -.992** | .993** |
| Asian customers | AC | -.986** | -.986** | -.986** | .962** |
| Native Hawaiian customers | NHC | -.992** | -.992** | -.992** | .979** |
| Two or more races customers | TMRC | -.995** | -.995** | -.995** | .979** |

**: Correlation is significant at the 0.01 level

TABLE II
ANOVA Table for CO Pollutant

| Model | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| CO | Regression | 1.408 | 7 | .201 | 5153.240 | .000 |
| | Residual | .000 | 4 | .000 | | |
| | Total | 1.408 | 11 | | | |

Dependent Variable: Amount of CO

### B. Model #1 – Linear Regression - Amount of CO

Factor analysis was applied to highlight the significant coefficients. Factor analysis consists of a number of statistical techniques to simplify a complex set of data and is generally applied to understand the correlation between variables [20]. The number of independent variables should be limited. The factor analysis revealed that the amount of dependent variable, CO, is positively correlated with income, number of registered CAVs, native Hawaiian customers, as well as the number of CAV customers in 21-24 years old, 40-49 years old, 60-69 years old, and 70-79 years old categories. The output of this factor analysis highlighted that a considerable percentage of CO in the State of Maryland is defused by customers in the above four age categories as well as the native Hawaiian race group. The output of CO diffusion linear regression model is summarized in Tables II and III.

TABLE III
The Regression Model of CO Pollutant

| Model | | Unstandardized Coefficients B | Unstandardized Coefficients Std. Error | Standardized Coefficients Beta | t | Sig. |
|---|---|---|---|---|---|---|
| CO | (Constant) | 7.766 | .722 | | 10.757 | .000 |
| | Inc | -1.160E-5 | .000 | -.340 | -10.210 | .001 |
| | NRCAVs | -1.243E-6 | .000 | -.416 | -8.082 | .001 |
| | $NDLH_2$ | 6.483E-6 | .000 | .143 | 5.851 | .004 |
| | $NDLH_5$ | -3.734E-6 | .000 | -.338 | -5.366 | .006 |
| | $NDLH_7$ | -3.508E-6 | .000 | -.560 | -7.199 | .002 |
| | $NDLH_8$ | 1.071E-5 | .000 | 1.658 | 9.887 | .001 |
| | NHC | -.001 | .000 | -1.531 | -6.863 | .002 |

Adjusted R square: 0.996, Sig. F change: .000, Durbin-Watson: 2.416

$$CO = 7.77 - 1.160E\text{-}5 * Inc - 1.243E\text{-}6 * NRCAVs + 6.483E\text{-}6 * NDLH_2 - 3.734E\text{-}6 * NDLH_5 - 3.508E\text{-}6 * NDLH_7 + 1.071E\text{-}5 * NDLH_8 - 0.001 * NHC \tag{2}$$

Table III shows the significant errors of all independent variables being less than 5% confidence interval. Durbin Watson (DW) statistic is a test for autocorrelation in the residuals. The DW test was obtained on values from 2 to 4 [21], which indicated a negative autocorrelation between residuals. If DW is between 1.5 and 2.5 then autocorrelation is likely not a cause for concern [22].

### C. Model #2 – Linear Regression - Amount of NOx

A similar methodology to CO was followed to develop linear regression model for NOx. As shown in Table IV, Native Hawaiian customers, NHC, is the only meaningful race group with zero coefficient. Other race groups did not have a significant correlation with dependent variables. The linear regression is shown in (2).

TABLE IV
THE REGRESSION MODEL OF NOX POLLUTANT

| Model | | Unstandardized Coefficients B | Unstandardized Coefficients Std. Error | Standardized Coefficients Beta | t | Sig. |
|---|---|---|---|---|---|---|
| NOx | (Constant) | 84.902 | 15.979 | | 5.313 | .013 |
| | $NDLH_4$ | 4.170E-6 | .000 | .442 | 4.434 | .021 |
| | $NDLH_5$ | -1.636E-5 | .000 | -.992 | -4.688 | .018 |
| | $NDLH_7$ | -1.571E-5 | .000 | -1.678 | -7.827 | .004 |
| | Inc | 4.165E-6 | .000 | .082 | 1.366 | .026 |
| | HHS | -24.137 | 5.064 | -.713 | -4.767 | .018 |
| | $NDLH_9$ | 1.748E-5 | .000 | .411 | 3.158 | .051 |
| | NRCAVs | 9.445E-7 | .000 | .212 | 1.389 | .026 |
| | NHC | .000 | .000 | -.570 | -1.471 | .023 |

Adjusted R square: 0.98, Sig. F change: .000, Durbin-Watson: 2.305

$$NOx = 84.902 + 4.170E\text{-}6*NDLH_4 - 1.636E\text{-}5*NDLH_5 - 1.571E\text{-}5*NDLH_7 + 1.748E\text{-}5*NDLH_9 + 4.165E\text{-}6*Inc - 24.137*HHS + 9.445E\text{-}7*NRCAVs \quad (3)$$

Significant errors of all independent variables are less than 5% confidence interval and DW test (values from 2 to 4) shows negative autocorrelation between residuals. ANOVA table for the NOx regression model is shown in Table V.

TABLE V
ANOVA TABLE FOR NOX POLLUTANT

| Model | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Regression | 3.141 | 8 | .393 | 1452.995 | .000 |
| Residual | .001 | 3 | .000 | | |
| Total | 3.142 | 11 | | | |

Dependent Variable: Amount of NOx

Table V shows the mean square (0.393) which is defined as the arithmetic mean of the squares of a set of variables in the model. The results indicate that the null hypothesis is rejected. An ANOVA uses null (all group means are equal) and alternative (at least one group mean is different from the rest) hypotheses. The larger the F-statistic, the greater the variation between sample means relative to the variation within the samples. Nevertheless, the larger the F-statistic (in this model: 1452.995), the greater the evidence that there is a difference between the group means [23].

### D. Model #3 – Linear Regression - Amount of PM

The linear regression model for PM pollutants' coefficients is shown in Table VI.

TABLE VI
THE REGRESSION MODEL OF PM POLLUTANT

| Model | | Unstandardized Coefficients B | Unstandardized Coefficients Std. Error | Standardized Coefficients Beta | t | Sig. |
|---|---|---|---|---|---|---|
| PM | (Constant) | 79.556 | 23.860 | | 3.334 | .045 |
| | Inc | -1.236E-5 | .000 | -.156 | -1.827 | .016 |
| | HHS | -20.825 | .002 | -.394 | -2.602 | .038 |
| | NRCAVs | -1.613E-6 | .000 | -.232 | -1.853 | .019 |
| | $NDLH_4$ | 2.301E-6 | .000 | .156 | 2.914 | .032 |
| | $NDLH_5$ | -1.343E-5 | .000 | -.521 | -3.431 | .042 |
| | $NDLH_7$ | -1.568E-5 | .000 | -1.073 | -5.025 | .015 |
| | $NDLH_8$ | 1.607E-5 | .000 | 1.068 | 3.621 | .036 |
| | NHC | -.001 | .000 | -.829 | -2.063 | .031 |

Adjusted R square: 0.991, Sig. F change: .000, Durbin-Watson: 2.223
Dependent Variable: Amount of PM

$$PM = 79.556 - 1.236E\text{-}5*Inc - 20.825*HHS - 1.613E\text{-}6*NRCAVs + 2.301E\text{-}6*NDLH_4 - 1.343E\text{-}5 * NDLH_5 - 1.568E\text{-}5*NDLH_7 + 1.607E\text{-}5*NDLH_8 - .001*NHC \quad (4)$$

As shown in Table VII, the larger the F-statistic (in this model: 1806.871), the greater the evidence that there is a difference between the group means. Also, the amount of p-test (Sig.: .000) specified that the obtained model is reliable.

TABLE VII
ANOVA TABLE FOR PM POLLUTANT

| Model | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Regression | 7.651 | 8 | .956 | 1806.871 | .000 |
| Residual | .002 | 3 | .001 | | |
| Total | 7.653 | 11 | | | |

Dependent Variable: Amount of PM

### E. Model #4 – Fuel Consumption Regression

The input fuel consumption values were specified from macro simulation. The transportation network of the State of Maryland was simulated, then, the VISUM software calculated the amount of fuel consumption from 2010 to 2021. Factor analysis was utilized to mitigate the number of independent variables [23]. Efficient independent variables were identified to obtain the linear regression model of CAVs' fuel consumption.

Table VIII shows the model's coefficients, and Table IX shows the ANOVA table.

$$\text{Fuel consumption} = 3825377850.782 + 211.477*NRCAVs - 47.457*NDLH_4 - 529.996*NDLH_5 - 628.766*NDLH_8 + 2639.858*NDLH_9 + 2744.650*VMT + 1166.517*Inc - 6404.215*HHS - 583.544*WC \quad (5)$$

TABLE VIII
THE REGRESSION MODEL OF FUEL CONSUMPTION

| Coefficients | | | | | | |
|---|---|---|---|---|---|---|
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
| | | B | Std. Error | Beta | | |
| Fuel cons. | (Constant) | 3825377850.782 | 841905832.468 | | 4.544 | .045 |
| | NRCAVs | 211.477 | 11.113 | .560 | 19.030 | .003 |
| | $NDLH_4$ | -47.457 | 44.726 | -.059 | -1.061 | .040 |
| | $NDLH_5$ | -529.996 | 91.241 | -.380 | -5.809 | .028 |
| | $NDLH_8$ | -628.766 | 78.050 | -.771 | -8.056 | .015 |
| | VMT | 2744.650 | 140.837 | .162 | 19.488 | .003 |
| | Inc | 1166.517 | 119.578 | .271 | 9.755 | .010 |
| | HHS | -6404.215 | 173.734 | -.224 | -3.686 | .056 |
| | $NDLH_9$ | 2639.858 | 194.886 | .734 | 13.546 | .005 |
| | WC | -583.544 | 91.909 | -.288 | -6.349 | .024 |

Adjusted R square: 0.967, Sig. F change: .000, Durbin-Watson: 2.962
Dependent Variable: Total Fuel Consumption by CAVs

TABLE IX
ANOVA TABLE FOR FUEL CONSUMPTION

| Model | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| Fuel Cons | Regression | 224964489645317122 | 9 | 24996054405503523.5 | 18352.2 | .000 |
| | Residual | 272402647492.072 | 2 | 136201323746.036 | | |
| | Total | 22496721367179204.000 | 11 | | | |

Dependent Variable: Total Fuel Consumption by CAVs

The mean square values in Table IX show that factors (treatments) are significant. The result of the F-test is a positive number which, if it exceeds 1.0, indicates that there are differences between the means. It should be noted that the F-test is one-sided since we are only testing for the possibility that the differences between the means is disproportionately large compared to the variation within the groups.

## V. MODELS VALIDATION

The macro simulation model in VISUM based on the suggested O.D. demand matrices for year 2030 was run. Also, the proposed statistical models in SPSS were validated based on the suggested attributes for 2030 [24]. The reliability of these four models in SPSS was compared with the obtained values from the macro simulation model in VISUM. Table X shows the obtained output of pollutants and fuel consumption for 2030.

TABLE X
VALIDATION OF FOUR PROPOSED STATISTICAL MODELS

| Attribute | Unit | Output from macro simulation model in VISUM | Output from statistical model in SPSS | Percentage of changes (%) |
|---|---|---|---|---|
| CO | kg | 1.122 | 1.051 | 6.3 |
| NOx | kg | 1.966 | 1.795 | 8.7 |
| PM | kg | 2.614 | 2.545 | 2.6 |
| Fuel Consumption | (Liter) | 322642496 | 307546783 | 4.9 |

Table X shows that the values obtained in 2030 by four statistical models are closely related to the values obtained from the macro simulation model in VISUM software (≤ 9% error). Hereupon, the reliability of the statistical models is desirable.

## VI. CONCLUSIONS

This research presented the anticipated environmental impact of CAVs to fill in the gap that exists in the previous findings on how CAVs affects GHG Emissions and fuel consumption. The State of Maryland was selected as a case study and its transportation network was simulated in VISUM software. The real-world demand matrices were implemented to run the macro simulation. Multiple datasets were collected and utilized in linear regression models to obtain significant impact of CAVs on CO, NOx, PM pollutants, and fuel consumption. The values of pollutants and total fuel consumption were obtained from the macro simulation. Other variables specified were: number of registered CAVs, level of education (bachelor's degree and higher), age groups of CAVs customers, VMT, income levels, race groups of CAVs customers, and the household size. Finally, four linear regression models were developed to predict the amount of CO, NOx, PM pollutants, and fuel consumption in the future. The acquired error of three pollutants and fuel consumption by four statistical models was less than 9% in comparison with the obtained values by macro simulation model in VISUM.

It was concluded that CAVs' penetration into the market, and ultimately on the roads, is expected to have a direct and/or indirect impact on the environment. Considering the specified sociodemographic attributes of CAVs customers, the results also displayed significant correlation of pollutants and fuel consumption, with income, segregated age groups of CAV customers, and race (especially the native Hawaiian). Standard error, T-test, F-test, Adjusted R square, Significant F change, and Durbin-Watson test were investigated, and the reliability of the models became evident.

## ACKNOWLEDGMENTS

This study was supported by utilizing the data provided by Urban Mobility & Equity Center (UMEC), a Tier 1 University Transportation Center of the U.S., Department of Transportation and Urban Infrastructure Studies, at Morgan State University.

## REFERENCES

[1] Hallman, C., The Average Commute to Work by State and City. 2020.
[2] Hedges, The Number of Cars in the U.S. in 2021/2022: Market Share, Distribution, and Trends, in Financesonline. 2020.
[3] Gardner, G., Americans Lost Nearly 100 Hours to Traffic in 2019, in Forbes. 2020.
[4] Barth, M., K. Boriboonsomsin, and G. Wu, Vehicle automation and its potential impacts on energy and emissions, in Road vehicle automation. 2014, Springer. p. 103-112.
[5] Morrow, W.R., et al., Key factors influencing autonomous vehicles' energy and environmental outcome, in Road vehicle automation. 2014, Springer. p. 127-135.
[6] Brown, A., J. Gonder, and B. Repac, An analysis of possible energy impacts of automated vehicles, in Road vehicle automation. 2014, Springer. p. 137-153.
[7] U.S. Census Bureau, Census 2020 PL94-171 release, Prepared by the Maryland Department of Planning, Projections and Data Analysis/State Data Center, April, 2010-2020.
[8] MVA Vehicle Registration counts summarized by County for F.Y. 2010 to F.Y. 2021, https://opendata.maryland.gov/transportation/mva-vehicle-registration-by-county-fy-2010-to-fy-2/kqkd-4fx8/data
[9] Bachelor's Degree or Higher for Maryland, https://fred.stlouisfed.org/series/GCT1502MD.
[10] Maryland Census 2010-2020 Population & Demographic Changes, https://www.census.gov/library/stories/state-by-state/maryland-population-change-between-census-decade.html.
[11] Annual vehicle miles of travel in millions by functional classification, Maryland department of transportation - state highway administration office of planning and preliminary engineering data services devision.
[12] Median Household Income by the state of Maryland: 1984 to 2020, Income in current and 2020 CPI-U-RS adjusted dollars.
[13] Race and Ethnicity in Maryland in 2010-2020, https://beautifydata.com/united-states-population/race-and-ethnicity/maryland/estimate/2020.
[14] Programme, U.e. We promote sustainable, low-emission transport and work to reduce the sector's contribution to air pollution and climate change. 2021; Available from: https://www.unep.org/explore-topics/energy/what-we-do/transport.
[15] EPA. Fast Facts on Transportation Greenhouse Gas Emissions. 2019; Available from: https://www.epa.gov/greenvehicles/fast-facts-transportation-greenhouse-gas-emissions.
[16] Committee, S.O.-R.A.V.S., Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems. SAE Standard J, 2014. 3016: p. 1-16.
[17] Igliński, H. and M. Babiak, Analysis of the Potential of Autonomous Vehicles in Reducing the Emissions of Greenhouse Gases in Road Transport. Procedia Engineering, 2017. 192: p. 353-358.
[18] Shladover, S.E., D. Su, and X.-Y. Lu, Impacts of cooperative adaptive cruise control on freeway traffic flow. Transportation Research Record, 2012. 2324(1): p. 63-70.
[19] Fagnant, D.J. and K. Kockelman, Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. Transportation Research Part A: Policy and Practice, 2015. 77: p. 167-181.
[20] Kline, P., An easy guide to factor analysis. 2014: Routledge.
[21] Savin, N.E. and K.J. White, The Durbin-Watson test for serial correlation with extreme sample sizes or many regressors. Econometrica: Journal of the Econometric Society, 1977: p. 1989-1996.
[22] Nerlove, M., & Wallis, K. F. (1966). Use of the Durbin-Watson statistic in inappropriate situations. Econometrica: Journal of the Econometric Society, 235-238.
[23] Ansariyar, A., & Laaly, S. (2022, June). Statistical Analysis of Connected and Autonomous Vehicles (CAVs) Effects on the Environment in Terms of Pollutants and Fuel Consumption. In 2022 International Conference on Frontiers of Artificial Intelligence and Machine Learning (FAIML) (pp. 151-156). IEEE.
[24] Jeihani, M., Ansariyar, A., Sadeghvaziri, E., Ardeshiri, A., Kabir, M. M., Haghani, (2022). Investigating the Effect of Connected Vehicles (CV) Route Guidance on Mobility and Equity. https://trid.trb.org/view/1922778