VEHICLE EMISSION TEST BIG DATA TO INFER ACTIVE VEHICLE OWNERSHIP AS AN INPUT TO TRIP GENERATION MODEL

Manthi Kolambage[1], Thillaiampalam Sivakumar[2], Mavin De Silva[3], Sampath Surendra[4], Amal Kumarage[5]
Department of Transport Management and Logistics
Engineering, University of Moratuwa, Sri Lanka.
[1]*981103mdk@gmail.com* [2]*tsivakumar@uom.lk*
[3]mavinds@uom.lk [4]*sampathsure@gmail.com*
[5]*amalk@uom.lk*

ABSTRACT - With the rapid increment in the demand for transportation, accurately predicting future transport demand has been a critical aspect in developing countries. The trip generation model is the first step of the traditional Travel Demand Forecasting Model which provides the number of trips generated from each zone using socio-economic characteristics as inputs. Vehicle ownership, one of the pivotal inputs to develop the trip generation model, is considered in this study. Vehicle registration data obtained from the Department of Motor Traffic (DMT) may not accurately reflect the number of vehicles actively on the road. Thus, this study considered the Vehicle Emission Test (VET) data as an alternate live stream of data. The objective of this study is to develop a methodology to infer the active vehicle ownership for each Travel Analysis Zone (TAZ) by developing a distance matrix through the QGIS software. The study reveals the number of active vehicles is much less (1/3) than what was registered at DMT in the Western Province in 2019 (the year before the pandemic) highlighting the need for further investigations.

Keywords: Vehicle Emission Testing; Vehicle Ownership; Divisional Secretariat Divisions; Electricity Consumption; Centroids

1. INTRODUCTION

With the evolving dynamics of urban mobility, determining future transport demand is crucial. The travel Demand Forecasting Model (TDFM) is the process of estimating the number of vehicles/travellers that will use individual links or facilities in a transportation network in the future.[1] The Trip generation model, as the first step of TDFM, requires socio-economic and land use characteristics as the inputs for the trip generation step and trip attraction respectively. Vehicle ownership is a pivotal socio-economic input for the trip generation model. [2] Growing urban populations and household incomes lead to a rise in vehicle ownership, creating a greater propensity for transport demand. [3] Traditional methods of obtaining vehicle ownership data such as from household surveys have been more resource intensive, time intensive and money intensive. Therefore, the use of secondary data such as VET data streaming in large amounts. Such big data demands innovative approaches to derive the active vehicle ownership, the number of vehicles actively in use than vehicle ownership obtained from the DMT that may not accurately reflect the number of vehicles actively in use due to factors such as registration renewals, vehicle transfers, and registration lapses. This study aims to close the gap that exists between vehicle ownership information obtained from DMT and true vehicle ownership. The objective of this study is to develop a methodological framework to infer the active vehicle ownership for each Travel Analysis Zone (TAZ). Accurate estimation of active vehicle ownership will become an important input for the trip generation model and improve the accuracy of the outcome of the TDFM and predict future trips more accurately.

2. METHODOLOGY

2.1. Study Area and Data Collection

The comprehensive VET dataset contained important 47 attributes exceeding more than 1.2 million data points in the dataset. Since the dataset comprises a large set of data, the data cleaning process was carried out using R programming language. In the data cleaning process, tests that had been done repeatedly were removed so that the duplicate entries were removed from the dataset. The unique test locations were extracted from the dataset resulting in 34 VET locations from Cleanco and 33 VET locations from Laugfs in the Western Province. The locations were not able to provide the location address or the GPS coordinates, those data were obtained through Cleanco, Laugfs websites and Google Maps.

2.2 Data Cleaning and Analysis

The VET locations obtained as the output of the data cleaning process were fed to the QGIS software along with the DSD boundaries in the western province to visualize VET locations belonging to DSDs. The centroids of each DSD were also fed to the QGIS software and high-populated DSDs such as Colombo and Thimbirigasyaya consisted of more than one centroid. Colombo DSD consisted of 3 centroids while Thimbirigasyaya DSD consisted of 6 centroids. This led to an assumption that each centroid has a similar proportion from the population in their DSD. Considering centroids and VET locations, the distance matrix was generated resulting in finding the nearest VET location to each of the centroids. There were special scenarios where some testing locations belonged to more than one DSD and to sort this, the number of VET tests that had been carried out in that particular testing location were distributed according to the population in DSDs that are related to that testing location as found in the distance matrix. Electricity Consumption data was obtained from Ceylon Electricity Board (CEB) and Lanka Electricity Company (Pvt) Ltd (LECO) and an assumption was made that electricity consumption data represents the population of the western province since electricity has become an essential for the population. [4] There were some specific scenarios where some testing locations had not been the nearest testing location for any of the centroids of DSDs. To deal with this situation, Circular buffers were created assuming that people who live within that buffer area are more likely to approach this testing centre. Then the number of tests that have been done in that particular testing centre were distributed according to the Electricity consumption data points belonging to buffer area and DSDs that are relevant to that testing centre.

3. RESULTS AND DISCUSSION

According to the results of the methodology, the number of VETs that have been done at each testing centre are distributed to their respective DSDs. Since each vehicle should be tested once a year, these data imply the active vehicle ownership at each DSD level in the Western Province for the 2019 year.

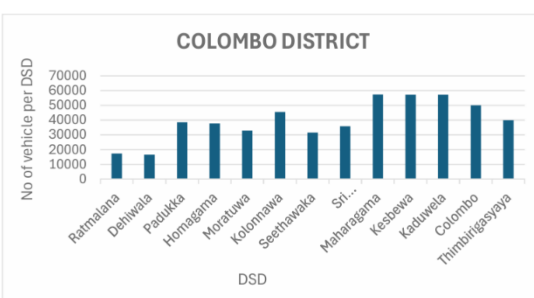

Figure 1 Vehicle Ownership in Colombo District

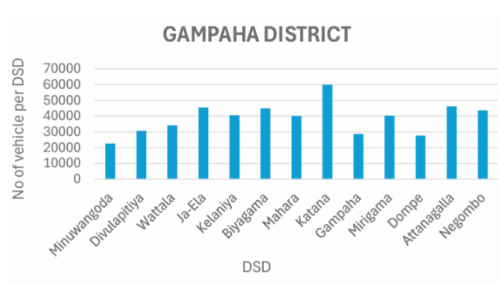

Figure 2 Vehicle Ownership in Gampaha District

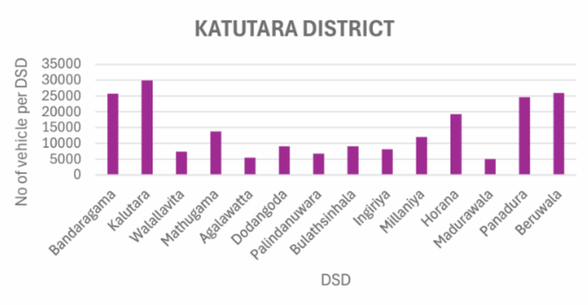

Figure 3 Vehicle Ownership in Kalutara District

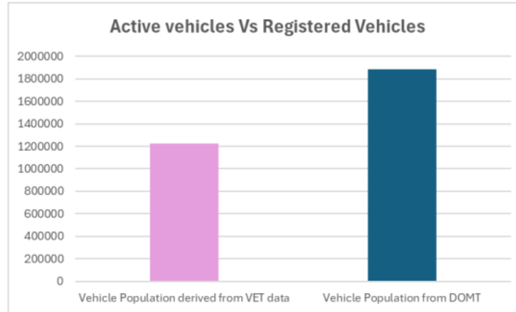

Figure 4 Active Vehicles Vs Registered Vehicles

4. CONCLUSION

This study paves the path to estimate how many vehicles are in use at the DSD level in the Western Province of Sri Lanka. It utilized data from Vehicle Emission Testing (VET) and a GIS-based distance matrix to overcome the difficulty in estimating the active vehicle registration using traditional surveys. The comparison the vehicle population obtained from the DMT and the vehicle population derived from VET data for Western Province in 2019 shows that the actual number of vehicles on the roads is nearly one-third (1/3) less than the vehicle population that DMT represents. A further research is recommended to study the reason for this gap. This study had several limitations including not addressing peoples' preferences and behaviours when choosing a VET centre which should be addressed in future research.

REFERENCES

[1] A. Hakimelahi, K. V. K. Rao, S. L. Dhingra, and F. Peiravian, "Updating of a Travel Demand Model Based on Fuel Consumption Data," in *Transportation Research Procedia*, Elsevier B.V., 2020, pp. 1401–1417. doi: 10.1016/j.trpro.2020.08.171.

[2] A. A. Amavi, J. P. Romero, A. Dominguez, L. dell'Olio, and A. Ibeas, "Advanced Trip Generation/Attraction Models," *Procedia Soc Behav Sci*, vol. 160, pp. 430–439, Dec. 2014, doi: 10.1016/j.sbspro.2014.12.155.

[3] P. Sillaparcharn, "VEHICLE OWNERSHIP AND TRIP GENERATION MODELLING-A Case Study of Thailand " 2007.

[4] H. A. Amarawickrama and L. C. Hunt, "Electricity demand for Sri Lanka: A time series analysis," *Energy*, vol. 33, no. 5, pp. 724–739, 2008, doi: 10.1016/j.energy.2007.12.008.