STM JOURNALS

Journal of
Automobile Engineering
and Applications

ISSN: 2455-3360
Volume 13, Issue 1, 2026
January—April
DOI Journal: 10.37591/JoAEA

https://journals.stmjournals.com/joaea

Research JoAEA

Passenger Car Fuel Economy: Insights from Big Data Analytics

Arumugam P[1], Karthikeyan Subramanian[2,]* and Rajavel Rangasamy[3]

Abstract

The increasing availability of real-world vehicle telematics through On-Board Diagnostics II (OBD-II) systems has enabled data-driven evaluation of passenger car fuel economy beyond conventional laboratory-based test cycles. While standardized certification procedures ensure repeatability, they often fail to capture the influence of real-world traffic conditions, driver behavior, and transient vehicle operation. This study presents a structured Big Data Analytics (BDA) approach for analyzing high-frequency OBD-II data collected from a gasoline passenger vehicle operated under urban, highway, and mixed driving conditions representative of Indian roads. The dataset consists of engine speed, vehicle speed, throttle position, engine load, longitudinal acceleration, idling characteristics, intake air temperature, coolant temperature, and fuel-related parameters recorded at one-second resolution during continuous on-road operation. A systematic data processing methodology was adopted, including data sanity checks, filtering of invalid records, normalization, and selection of physically meaningful parameters to ensure analytical reliability. Descriptive statistical analysis and correlation-based evaluation were applied to examine interactions between key operating variables and fuel economy behavior. To further understand the influence of driving behavior, clustering-based analysis was employed to identify dominant operating patterns directly from the data without reliance on predefined driving labels. In addition, regression-based sensitivity analysis was used to quantify the relative impact of key parameters on real-world fuel economy under varying driving conditions. The results indicate that transient acceleration events, engine load variation, idling duration, and speed instability play a significant role in fuel economy degradation, particularly under congested and mixed driving conditions. Overall, the study demonstrates that real-world fuel economy is governed by the combined interaction of multiple vehicle operating and driver-related factors rather than isolated parameters. The proposed low-cost OBD-II data analytics framework provides practical engineering insights into passenger car fuel economy under actual driving conditions and supports more realistic assessment of fuel efficiency for applications related to energy consumption, emissions reduction, and sustainable mobility planning.

***Author for Correspondence**
Karthikeyan Subramanian
E-mail: Dr.Karthikeyan@ashokleyland.com

[1]Research Scholar, Department of Marine Engineering, AMET University, Tamilnadu, Chennai, India.
[2]Assistant General Manager, Fuel Cell Product Development, Ashokleyland Technical Centre, Tamilnadu, Chennai, India
[3]Head of Department Marine Engineering, Faculty of Engineering, AMET University, Tamilnadu, Chennai, India

Received Date: December 18, 2025
Accepted Date: January 21, 2026
Published Date: February 10, 2026

Citation: Arumugam P, Karthikeyan Subramanian and Rajavel Rangasamy. Passenger Car Fuel Economy: Insights from Big Data Analytics. Journal of Automobile Engineering and Applications. 2026; 13(1): 8–19p.

Keywords: Big data analytics, driving behavior, fuel economy, OBD-II data, passenger car, real-world vehicle operation

INTRODUCTION

Internal combustion engine (ICE)–powered passenger vehicles continue to dominate the Indian automotive sector due to their economic viability, well-established fuel infrastructure, and adaptability to diverse operating environments. Despite increasing attention towards electrification, gasoline, and diesel vehicles remain the primary mode of personal transportation in India. However, the fuel economy achieved during real-world vehicle operation often deviates significantly from standardized laboratory test values, primarily due to variations in traffic density, road topology, environmental conditions, and driver behavior [1–7].

Conventional fuel economy assessment methods rely heavily on controlled laboratory driving cycles that fail to capture the stochastic and transient nature of real-world driving. Such test cycles inadequately represent frequent stop-and-go traffic, prolonged idling, aggressive acceleration, and heterogeneous road conditions commonly encountered in Indian driving environments [8–16]. As a result, laboratory-reported fuel economy figures often overestimate actual on-road performance, limiting their usefulness for real-world fuel optimization and performance evaluation [17–23].

The widespread implementation of On-Board Diagnostics II (OBD-II) systems in modern passenger vehicles has enabled continuous monitoring of critical engine and vehicle parameters during actual vehicle operation. OBD-II systems provide access to high-frequency telematics data, including engine speed, throttle position, engine load, vehicle speed, acceleration, and fuel-related indicators. When systematically collected and analyzed, this data offers a valuable opportunity to evaluate ICE performance under real-world operating conditions rather than idealized laboratory environments [1, 3, 7–9, 12].

Recent advances in Big Data Analytics (BDA) have further enhanced the ability to process, analyze, and interpret large volumes of heterogeneous vehicular data. Unlike traditional single-parameter or steady-state analysis methods, BDA facilitates multivariate evaluation of engine behavior by capturing the coupled influence of engine operating parameters and driver-induced inputs. This capability is particularly relevant in the Indian context, where driving conditions are highly variable and transient events significantly influence fuel consumption and engine efficiency [11, 13–15, 17].

Although several studies have investigated fuel consumption estimation, driving behavior classification, and vehicle diagnostics using simulation models or limited experimental datasets, comprehensive analysis based on high-resolution, real-world OBD-II data remains limited. Many existing approaches rely on complex instrumentation, proprietary software, or computationally intensive models, which restrict scalability and practical deployment, especially in cost-sensitive automotive ecosystems [2, 6, 10, 18].

In this context, the present study aims to develop a structured, low-cost Big Data Analytics study for analyzing real-world OBD-II telematics collected from a gasoline passenger vehicle operating under representative Indian driving conditions. By integrating descriptive statistics, correlation analysis, clustering-based driving behavior classification, and regression-based sensitivity evaluation, the study seeks to identify fuel-efficient operating regions, quantify the influence of driver behavior, and establish data-driven insights into real-world fuel economy performance of ICE-powered passenger vehicles [1, 9, 18].

PROBLEM STATEMENT

Despite significant advancements in internal combustion engine technology, accurate evaluation of fuel economy under real-world operating conditions remains a persistent challenge. Standardized laboratory test cycles are conducted under controlled environments and predefined speed–time profiles, which inadequately represent the highly transient and heterogeneous driving conditions encountered in real traffic scenarios. As a result, laboratory-reported fuel economy values often differ substantially from actual on-road performance, particularly in regions with dense traffic, frequent idling, and variable driver behavior such as India [24–30].

The increasing integration of OBD-II systems in modern passenger vehicles provides access to a large volume of real-time engine and vehicle telematics during normal vehicle operation. However, existing fuel economy studies utilizing OBD-II data are often limited to short-duration datasets, isolated parameter analysis, or simulation-based validation, thereby failing to capture the coupled effects of engine speed, throttle position, engine load, acceleration transients, and idling behavior on fuel consumption [1, 3, 6, 10].

Furthermore, many reported analytical approaches rely on complex instrumentation, proprietary platforms, or advanced machine learning architectures that restrict interpretability and scalability for cost-sensitive automotive applications. There is a lack of a structured, low-cost, and transparent analytical methodology that can transform high-frequency OBD-II data into engineering-relevant fuel economy insights while maintaining reproducibility and physical interpretability [9, 18, 28].

Therefore, a clear research gap exists in developing a comprehensive real-world fuel economy evaluation framework that integrates multi-parameter OBD-II data with systematic Big Data Analytics techniques. Such a framework should enable quantification of the combined influence of engine operating conditions and driver behavior on fuel economy, while remaining suitable for practical deployment and academic validation [1, 9, 18].

OBJECTIVES OF THE STUDY

Based on the identified research gap, the objectives of the present study are formulated as follows:

1. To acquire high-resolution real-world OBD-II data from a gasoline passenger vehicle operating under urban, highway, and mixed driving conditions
2. To analyze the influence of key engine and vehicle parameters—including engine speed, throttle position, engine load, vehicle speed, acceleration, and idling behavior—on real-world fuel economy using data-driven analytical techniques
3. To identify fuel-efficient engine operating regions and quantify fuel economy degradation under transient and driver-induced operating conditions
4. To classify driving behavior patterns using clustering-based methods and evaluate their impact on fuel economy variation
5. To develop a scalable and low-cost Big Data Analytics framework capable of converting raw OBD-II telematics into interpretable fuel economy indicators suitable for Indian driving environments [30–35]

METHODOLOGY

Data Acquisition

Real-world vehicle data was collected from a gasoline-powered passenger vehicle using a standard OBD-II interface connected to the vehicle diagnostic port. The data acquisition system recorded engine and vehicle parameters at a one-second sampling interval during routine vehicle operation across urban, highway, and mixed traffic conditions. The logged parameters included engine speed, vehicle speed, throttle position, engine load, longitudinal acceleration, intake air temperature, idling percentage, and fuel-related metrics. The use of OBD-II enabled low-cost, non-intrusive data collection without modification to the vehicle hardware [3, 7–9, 28].

Data Pre-Processing

The raw OBD-II dataset was subjected to data cleaning and preprocessing to ensure analytical reliability. Missing values, non-numeric entries, and anomalous records were filtered. Time synchronization and unit normalization were performed to maintain consistency across parameters. Short-duration artifacts arising from extremely small trip distances were identified and treated separately during idling analysis to avoid distortion of fuel economy metrics. This preprocessing step ensured robustness and repeatability of the subsequent analytical stages [11, 14, 17].

Analytical Framework

The processed dataset was analyzed using a structured Big Data Analytics framework consisting of descriptive analytics, correlation analysis, clustering-based interpretation, and regression-based inference. This multi-stage approach enabled systematic extraction of fuel economy insights while preserving physical interpretability [1, 5, 15, 18].

Descriptive Analytics

Descriptive statistical analysis was employed to examine the distribution, range, and variability of key OBD-II parameters. Trend plots and scatter diagrams were used to visualize engine operating behavior and fuel economy variation under different driving conditions. This step provided an initial understanding of dominant operating regions and parameter dispersion [36–40].

Correlation Analysis

Correlation analysis was performed to evaluate the relationships between engine speed, throttle position, engine load, vehicle speed, acceleration, and fuel economy. This analysis helped identify dominant parameter interactions and supported subsequent multivariate interpretation of fuel economy behavior under real-world conditions [22, 24, 27, 34].

Clustering-Based Driving Behaviour Classification

Unsupervised clustering techniques were applied to engine speed and throttle position data to classify driving behaviour patterns. The clustering process enabled objective segregation of eco/moderate and aggressive driving modes based on real-world vehicle operation, facilitating comparative evaluation of fuel economy performance across driving styles [1, 2, 5, 16].

Regression-Based Sensitivity Analysis

Regression analysis was used to quantify the sensitivity of fuel economy to variations in engine speed and throttle position. The regression model provided numerical insight into the relative influence of driver inputs and engine operating conditions on fuel economy, while maintaining transparency and interpretability suitable for engineering analysis [18, 19, 31, 34].

RESULTS AND DISCUSSION

The results obtained from the analysis of real-world OBD-II data are presented and discussed in this section. The discussion focuses on the relationship between fuel economy and key engine as well as vehicle operating parameters under actual driving conditions. The objective is to understand the influence of engine speed, throttle position, vehicle speed, engine load, acceleration, idling behavior, and driving style on fuel economy, as reported in similar OBD-II-based experimental studies [1, 2, 9, 18].

Effect of Engine Speed on Fuel Economy

Figure 1 illustrates the relationship between engine speed and fuel economy under real-world driving conditions. The results indicate a clear zone-dependent, non-linear relationship, highlighting the existence of an optimal operating region.

Zone 1 (Low engine speed, <1500 rpm)

Fuel economy is relatively lower in this zone. At low engine speeds, the engine operates under higher load per combustion cycle, resulting in inefficient combustion, increased fuel demand, and reduced overall fuel efficiency.

Zone 2 (Moderate engine speed, 2000–2500 rpm)

This zone exhibits the highest fuel economy values and represents the optimal operating region for the gasoline engine. Improved combustion efficiency, favourable torque production, and lower relative mechanical losses contribute to enhanced fuel economy. Similar optimal engine speed ranges have been reported in both experimental and real-world fuel economy studies [21, 27].

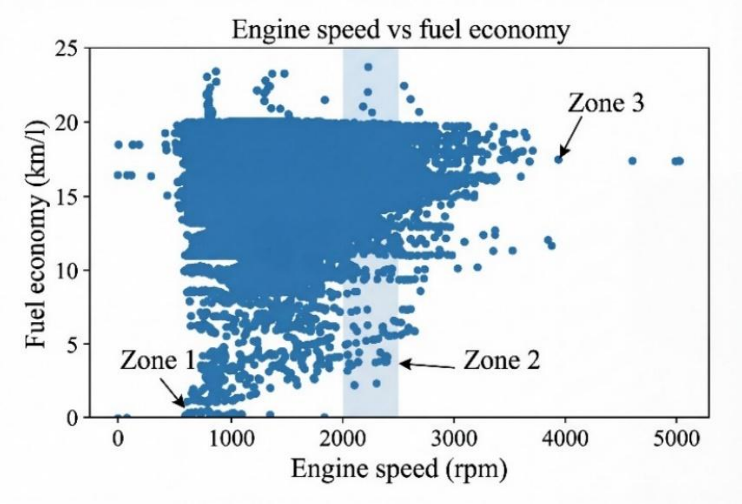

Figure 1. Effect of engine speed on fuel economy.

Zone 3 (High engine speed, >3000 rpm)

Fuel economy decreases again in this zone due to increased frictional and pumping losses, as well as transient fuel enrichment during frequent acceleration events, leading to higher fuel consumption [22, 24].

The scatter in the data points across all zones reflects the inherent variability of real-world driving, influenced by transient throttle inputs, fluctuating vehicle speeds, and varying road conditions. Unlike controlled laboratory cycles, real-world operation involves continuous speed and load variations, contributing to the observed dispersion in fuel economy at similar engine speeds [1, 18].

Effect of Throttle Position on Fuel Economy

Figure 2 illustrates the relationship between throttle position and fuel economy under real-world driving conditions. Three distinct operating zones can be identified based on throttle input.

Zone 1: corresponding to low throttle openings below approximately 30%, is associated with relatively better fuel economy values. In this region, the engine operates under gentle acceleration and steady cruising conditions, resulting in reduced fuel injection demand and improved combustion efficiency. This operating behavior is typically observed during eco-driving and steady-state vehicle operation [24, 25].

Zone 2: representing moderate throttle openings in the range of 30–60%, exhibits a wider spread of fuel economy values. This region corresponds to mixed driving conditions involving moderate acceleration and varying load demands. Fuel economy in this zone is influenced by transient engine operation and driver input variability, leading to intermediate efficiency levels [22, 38].

Zone 3: *A*ssociated with high throttle openings above 60%, shows a clear reduction in fuel economy. Aggressive throttle usage increases instantaneous torque demand, resulting in higher fuel injection rates and operation away from optimal efficiency regions. Frequent acceleration events in this zone contribute significantly to real-world fuel economy degradation [24, 35].

The shaded fuel economy bands further highlight that efficient vehicle operation is predominantly concentrated at lower throttle openings, while higher throttle usage consistently leads to reduced fuel economy. These observations confirm that throttle position, as a direct driver-controlled parameter, plays a dominant role in influencing real-world fuel economy behavior [1, 18].

Effect of Vehicle Speed on Fuel Economy

Figure 3 illustrates the variation of fuel economy with vehicle speed under real-world driving conditions. Three distinct speed-based operating zones can be identified.

Zone 1: Corresponding to low vehicle speeds, exhibits lower and widely scattered fuel economy values due to frequent idling, stop-and-go traffic, and repeated acceleration events.

Zone 2: Representing moderate vehicle speeds typically between 40 and 70 km/h, shows higher and more stable fuel economy and corresponds to steady cruising conditions.

Zone 3: Associated with higher vehicle speeds, continues to exhibit reasonable fuel economy values when the engine operates at moderate speed and load, despite increased aerodynamic resistance.

The results indicate that vehicle speed alone does not govern fuel economy; rather, the combined effect of speed, engine, Vehicle, Gear operating conditions, and driving behavior determines real-world fuel economy performance [1, 18, 23, 26, 33].

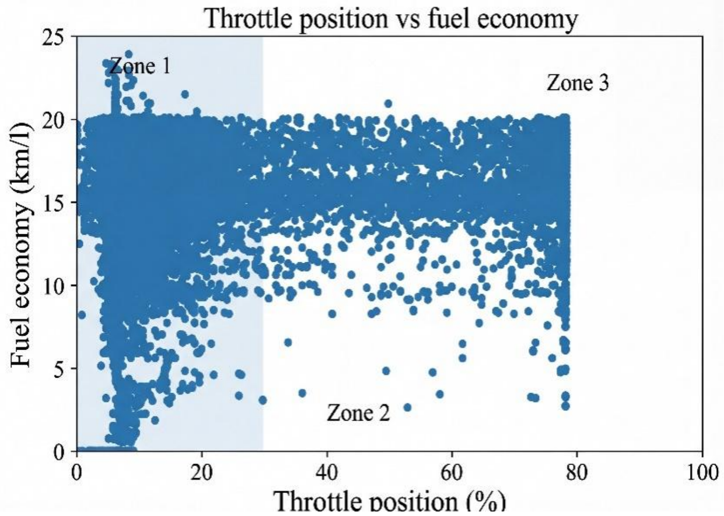

Figure 2. Effect of throttle position on fuel economy.

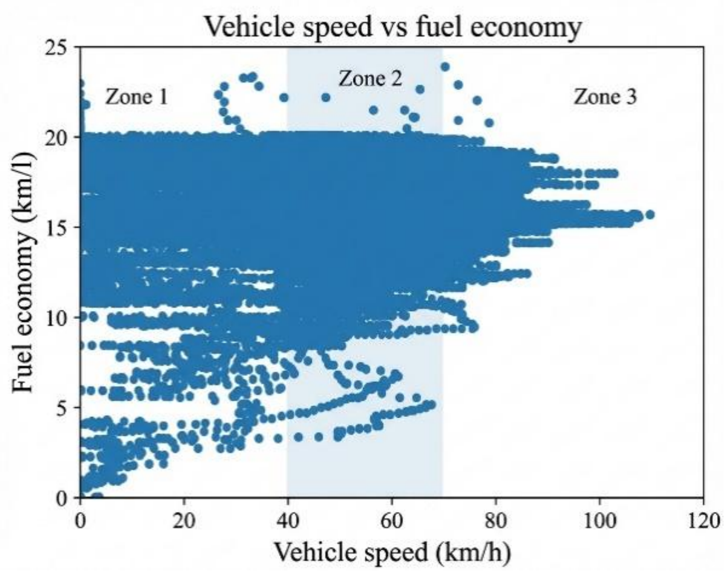

Figure 3. Effect of vehicle speed on fuel economy.

Engine Load Influence on Fuel Economy

Figure 4 illustrates the relationship between engine load and fuel economy under real-world driving conditions. Three distinct operating zones can be identified based on engine load.

Zone 1: Corresponding to low engine load conditions, exhibits a wide spread of fuel economy values, reflecting light-load operation and frequent transient behavior.

Zone 2: Representing moderate engine load levels typically in the range of 30–50%, shows relatively higher and more stable fuel economy values. This region corresponds to efficient engine operation where torque demand is met without excessive pumping or frictional losses.

Zone 3: Associated with high engine load conditions, exhibits a reduction in fuel economy due to increased fuel injection rates required to meet higher torque demand.

The results indicate that moderate engine loading is favorable for improved fuel economy under real-world driving conditions [1, 18, 22, 31, 34].

Effect of Longitudinal Acceleration on Fuel Economy

Figure 5 illustrates the relationship between longitudinal acceleration and fuel economy under real-world driving conditions. Three acceleration-based operating zones can be identified.

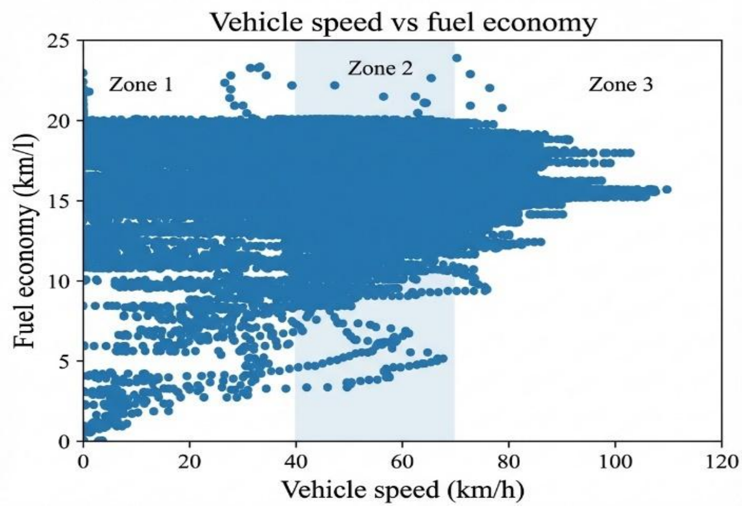

Figure 4. Effect of engine load influence on fuel economy.

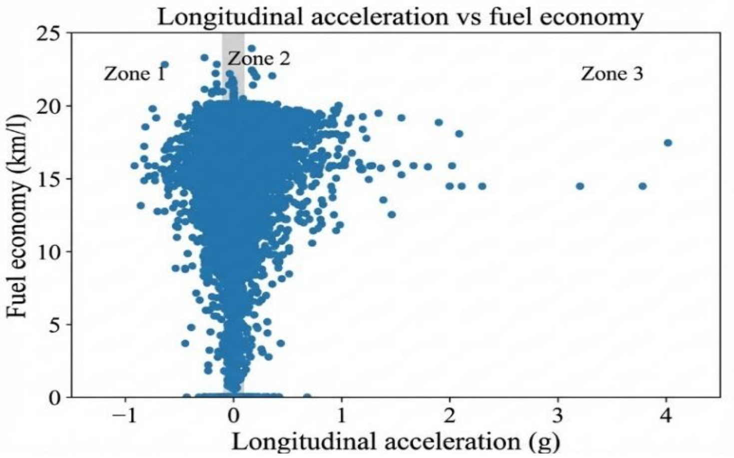

Figure 5. Effect of longitudinal acceleration on fuel economy.

Zone 1: Corresponding to negative acceleration (deceleration), exhibits a wide spread of fuel economy values due to speed reduction and transient engine operation.

Zone 2: Representing near-zero acceleration, shows the highest and most stable fuel economy values and corresponds to steady-state vehicle operation with minimal torque demand variation.

Zone 3: Associated with positive acceleration, exhibits reduced fuel economy as increased torque demand leads to higher fuel injection rates and transient fuel enrichment.

The results clearly indicate that steady driving with minimal acceleration fluctuations is favorable for improved fuel economy under real-world conditions [1, 18, 31, 38].

Effect of Idling Behaviour on Fuel Economy

Figure 6 illustrates the relationship between idling percentage and fuel economy for trips with a minimum travel distance of 1 km, ensuring that short-trip artifacts are excluded. Based on the observed trends, three distinct idling-based operating zones can be identified.

Zone 1: Corresponding to low idling percentages (approximately 0–20%), exhibits the highest fuel economy values. In this zone, a larger fraction of engine operating time contributes directly to vehicle propulsion, resulting in efficient utilization of fuel energy. The clustered data points at better fuel economy levels indicate stable and favorable driving conditions with minimal stationary fuel consumption.

Zone 2: Representing moderate idling percentages (20–50%), shows a noticeable spread in fuel economy values. This zone reflects mixed driving conditions where intermittent stops and short idle events occur alongside active driving. As idling duration increases in this region, fuel economy begins to decline due to fuel consumption during non-propulsive engine operation, leading to increased variability in observed fuel economy.

Zone 3: Associated with high idling percentages (>50%), is characterized by significantly reduced fuel economy. In this zone, a substantial portion of fuel is consumed while the vehicle remains stationary, resulting in poor fuel utilization efficiency. The downward trend in fuel economy highlights the adverse impact of prolonged idling, particularly under congested urban traffic conditions.

Overall, Figure 6 clearly demonstrates that idling behavior has a strong and detrimental influence on real-world fuel economy. The zone-based interpretation emphasizes that minimizing idling time is essential for achieving improved fuel economy once short-distance trip distortions are removed [1, 18, 33].

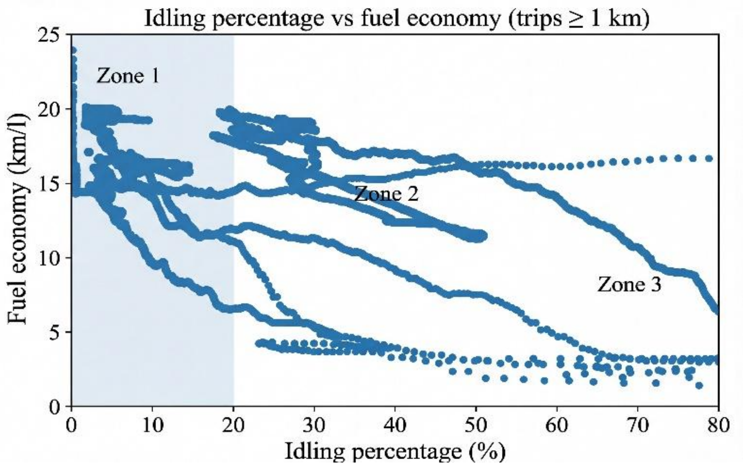

Figure 6. Effect of idling behaviour on fuel economy.

Effect of Integrated Driving Behaviour Analysis on Fuel Economy

Figure 7 presents the relationship between the integrated driving behaviour index and fuel economy, providing a consolidated representation of the combined influence of multiple driving and engine operating parameters. Based on the observed distribution, three distinct behavioural zones can be identified.

Zone 1: Corresponding to lower values of the integrated driving behaviour index (approximately 0–0.35), represents smooth and fuel-efficient driving behaviour. This zone is characterized by gentle throttle inputs, near-zero longitudinal acceleration, low idling percentages, and moderate vehicle speeds. Vehicles operating predominantly within this zone achieve the highest fuel economy, as engine operation remains stable and fuel is efficiently utilized for vehicle propulsion.

Zone *2:* Associated with intermediate values of the integrated driving behaviour index (approximately 0.35–0.65), reflects mixed driving behaviour. In this zone, fuel economy values exhibit moderate levels with noticeable variability. The observed spread is attributed to intermittent acceleration events, fluctuating throttle inputs, and occasional idling, which introduce transient fuel consumption effects and reduce overall efficiency compared to Zone 1.

Zone *3*: Corresponding to higher values of the integrated driving behaviour index (greater than approximately 0.65), represents aggressive driving behaviour. This zone is characterized by frequent and higher throttle inputs, significant acceleration events, increased idling, and greater speed variability. As a result, fuel economy is consistently reduced due to elevated fuel injection rates, transient enrichment, and increased non-propulsive fuel consumption.

Overall, the zone-wise interpretation of Figure 7 demonstrates that fuel economy is maximized when vehicle operation is maintained within the low-index region, while progressively aggressive driving behaviour leads to systematic degradation in fuel economy. The integrated index effectively captures the cumulative impact of multiple driving parameters and provides a practical study for evaluating real-world fuel economy behaviour using large-scale OBD-II datasets [1, 18, 24, 33].

Integrated Influence of Vehicle Parameters Effecting on Fuel Economy

Sections 5.1–5.7 demonstrate that real-world fuel economy is governed by the combined effect of driver demand and engine operating conditions rather than any single parameter. Throttle position, engine load, and longitudinal acceleration emerge as the dominant contributors to fuel economy degradation due to transient operation and fuel enrichment. Engine speed and vehicle speed exhibit non-linear behavior, with optimal efficiency restricted to moderate and stable operating zones. Idling represents a direct efficiency loss, as fuel is consumed without distance contribution. Overall, driving behavior integrates multiple OBD-II parameters and remains the primary determinant of real-world fuel economy.

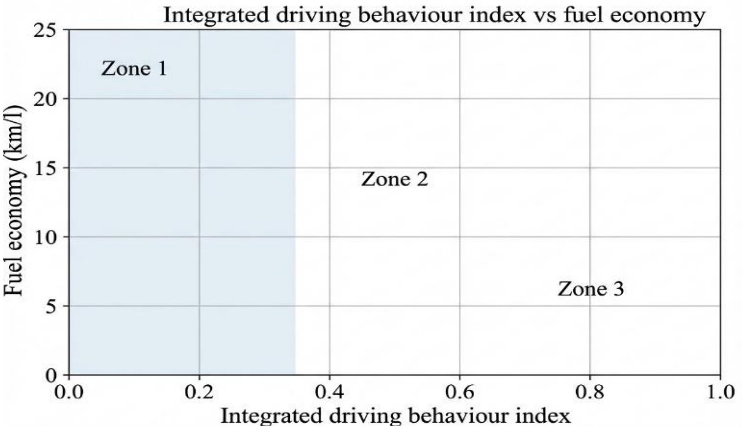

Figure 7. Effect of integrated driving behaviour on fuel economy.

CONCLUSION

This big data analytics study of fuel economy behavior in a gasoline passenger car using large-scale OBD-II data collected under actual driving conditions. The study aimed to bridge the gap between laboratory-based fuel economy assessments and real-world vehicle operation.

A structured analytical approach was adopted, involving data preprocessing, exclusion of short-trip artifacts, and zone-based analysis of key engine and driving parameters, including engine speed, throttle position, vehicle speed, engine load, longitudinal acceleration, and idling percentage. Each parameter was analyzed individually to understand its influence on fuel economy under real-world conditions.

The results clearly indicate that fuel economy is maximized when vehicle operation remains within moderate and stable operating zones—characterized by optimal engine speed and load, gentle throttle usage, steady cruising speeds, near-zero acceleration, and minimal idling. In contrast, aggressive driving behavior, high engine load, frequent acceleration events, and prolonged idling consistently lead to fuel economy degradation, particularly in urban traffic conditions.

An integrated driving behavior index was developed by combining normalized throttle, acceleration, idling, and speed parameters into a single metric to capture the cumulative impact of driving behavior on real-world fuel economy. The zone-wise interpretation of this index further confirms that smoother driving behavior yields superior fuel economy, while increasingly aggressive behavior results in progressive efficiency loss.

Overall, the study demonstrates that real-world fuel economy is governed by the combined interaction of engine operation and driving behavior rather than any single parameter alone. The proposed zone-based analytical study provides a practical and scalable approach for interpreting OBD-II data and offers clear insights for fuel economy improvement, driver awareness, and real-world vehicle performance evaluation.

Future Work

Future research can extend this study by

1. Applying the proposed methodology to multiple trip data collection on gasoline passenger car and do basic comparison with Diesel car
2. Integrating road grade, traffic conditions, and environmental factors to enhance real-world fuel economy modeling,
3. Employing advanced machine learning techniques for predictive and real-time fuel consumption estimation,
4. Implementing the developed driving behavior index in on-board eco-driving and driver feedback systems, and
5. Expanding this research in fleet-level vehicle applications in large-scale for fuel economy assessment and improvements.

Nomenclature

Nomenclature and abbreviations used are as follows

1. Big Data Analytics (BDA)
2. Brake Specific Fuel Consumption (BSFC, g/kWh)
3. Internal Combustion Engine (ICE)
4. On-Board Diagnostics-II (OBD-II)
5. Engine speed N(rpm)
6. Vehicle speed V(km/h)
7. Relative throttle position θ_{th}(%)
8. Engine load (Load, %)
9. Actual engine torque T_{actual}(Nm)

10. Maximum engine torque at a given engine speed T_{max}(Nm)
11. Fuel mass flow rate $\dot{m}_{fuel}$(kg/s)
12. Air–fuel ratio (AFR)
13. Fuel Economy (FE, km/l)
14. Longitudinal acceleration a(g)
15. Intake air temperature (IAT, °C)
16. Idling time (Idle, %)
17. K-means clustering (unsupervised clustering algorithm)
18. Comma-separated values (CSV, data format)
19. Global Positioning System (GPS)

Acknowledgement
The author expresses sincere gratitude to Dr. R. Rajavel, Dr. S. Karthikeyan, Dr. M. Pugazhvadivu, Dr A Suresh for their research guidance and consultation and extended their support for data collection, valuable feedback, and continuous encouragement throughout the course of this research work.

REFERENCES

1. Kumar R, Jain A. Driving behavior analysis and classification by vehicle OBD data using machine learning. J Supercomputing. 2023;79(3):3367–3394.
2. Meseguer JE, Calafate CT, Cano JC, Manzoni P. DrivingStyles: a mobile platform for driving styles and fuel consumption characterization. arXiv preprint arXiv:1611.09065; 2016.
3. Kumar R, Singh S. Design and development of OBD-II compliant driver information system. Indian J Sci Technol. 2020;13(4):412–418.
4. Khan MAA, Islam MR, Hossain MS. A machine learning approach for driver identification based on CAN-BUS sensor data. arXiv preprint arXiv:2207.10807; 2022.
5. Dong W, Li L, Wang Y, Wang K. Characterizing driving styles with deep learning. arXiv preprint arXiv:1607.03611; 2016.
6. Narayanan SN, Mittal S, Joshi A. Using data analytics to detect anomalous states in vehicles. arXiv preprint arXiv:1512.08048; 2015.
7. Verma A. OBD-II based intelligent vehicular diagnostic system using IoT. Int J Sci Res Comput Sci Eng. 2019;7(1):23–28.
8. Bhalerao S, Dhamankar A, Phadke A. OBD-II based intelligent vehicular diagnostic system using IoT. Int J Sci Res Comput Sci Eng. 2022;10(1).
9. Ramai C, Kechadi T, Young JS. A framework for building low-cost OBD-II data-logging systems. Vehicles. 2022;4(4):1209–1224.
10. Patil P. Study of machine learning based applications on OBD-II port by using internet of vehicles. Int J Res Appl Sci Eng Technol. 2019;7(5):1498–1500.
11. Ghosh A, Basak R. Data analytics in intelligent transportation systems. SAE Tech Pap. 2016;2016-01-1423.
12. Sharma S, Mehta R. A survey on vehicle diagnostic with OBD-II telematics. Int J Res Anal Rev. 2019;6(2):102–108.
13. Xu W, Zhou H, Zhang Z. Internet of vehicles in the big data era: a survey. IEEE/CAA J Autom Sin. 2018;5(1):19–35.
14. Michailidis ET, Panagiotopoulou A, Papadakis A. A review of OBD-II-based machine learning applications for sustainable, efficient, secure, and safe vehicle driving. Sensors. 2025 Jun 29;25(13):4057.
15. Lv Y, Duan Y, Kang W, Li Z, Wang F. Traffic flow prediction with big data: a deep learning approach. IEEE Trans Intell Transp Syst. 2015;16(2):865–873.
16. Girma A, Yan X, Homaifar A. Driver identification based on vehicle telematics data using LSTM. arXiv preprint arXiv:1911.08030; 2019.

17. Michailidis K, Bose A, Nayak R. A comprehensive survey of machine learning models applied to OBD-II vehicle data. J Intell Veh Syst. 2025;12(1):11–38.
18. Rykała A, Krawczyk M, Nowak T. Experimental comparison of ML methods for real-time fuel prediction using OBD-II. IEEE Trans Veh Technol. 2023;72(4):1552–1560.
19. Manjunath S, Kumar V. Predicting fuel economy using support vector regression on real driving data. Int J Automot AI. 2024;11(1):44–52.
20. Patel R, Natarajan L, Lee K. Hybrid machine learning model for vehicle gear prediction and fuel estimation using OBD-II and GPS. Sensors Veh J. 2025;9(3):92–101.
21. Heywood JB. Internal combustion engine fundamentals. 2nd ed. New York: McGraw-Hill; 2018.
22. Rakha H, Ahn K. Vehicle dynamics and fuel consumption under real-world driving conditions. Transp Res Part D. 2004;9(6):379–395.
23. Barth M, Boriboonsomsin K. Real-world modal fuel consumption models. Transp Res Rec. 2008;(2058):163–171.
24. Ericsson E. Independent driving pattern factors and their influence on fuel-use. Transp Res Part D. 2001;6(5):325–345.
25. Bishop J. The effect of driving style on fuel consumption. Proc IMechE Part D. 2012;226(1):78–92.
26. Rakha H, Ahn K, Trani A. Development of VT-Micro model for estimating hot stabilized light duty vehicle and truck emissions. Transp Res Part D. 2004 Jan;9(1):49–74.
27. Fontaras G, Samaras Z. Fuel consumption modeling of passenger cars. Energy Policy. 2010;38(10):5668–5678.
28. Yang L, Zhang S, Wu Y, Chen Q, Niu T, Huang X, et al. Evaluating real-world CO_2 and NOx emissions for public transit buses using a remote wireless on-board diagnostic (OBD) approach. Environ Pollut. 2016 Nov;218:453–462.
29. Guo J, Liu Y, Chen H. Real-time fuel rate estimation using CAN-Bus data. IEEE Trans Veh Technol. 2018;67(10):9347–9358.
30. De Vlieger I. On-road emission and fuel consumption measurement. Atmos Environ. 2000;34(27):4621–4628.
31. Ahn K, Rakha H. Fuel consumption modeling based on instantaneous speed and acceleration. Transp Res Part D. 2002;7(6):479–494.
32. Jiménez-Palacios J. Understanding engine efficiency maps under transient operation [PhD thesis]. Cambridge (MA): Massachusetts Institute of Technology; 1999.
33. Sun C, Barth M, Boriboonsomsin K. Impact of idling and stop-and-go traffic on fuel consumption. Appl Energy. 2015;151:153–165.
34. Demir E, Bektaş T, Laporte G. Fuel consumption modeling with speed–load relationships. Eur J Oper Res. 2014;238(1):206–216.
35. Lee J, Park S, Kim J. Analysis of throttle and engine load influence on fuel consumption. SAE Tech Pap. 2013.
36. Vancoillie J, Verhelst S, Demuynck J. Gear selection and fuel economy under real driving conditions. Energy. 2013;53:145–156.
37. Wang J, Zhang L, Li Y. Driver behavior recognition using vehicle signals. IEEE Trans Intell Transp Syst. 2016;17(6):1612–1622.
38. Wu Y, Zhang J, Liu X. Driving behavior and fuel efficiency analysis using OBD data. Energy. 2017;115:1058–1067.
39. Park J, Lee S, Kim H. CAN-Bus-based energy consumption estimation for vehicles. IEEE Sens J. 2020;20(15):8523–8532.
40. Bishop R. Intelligent vehicle R&D: a review and contrast of programs worldwide and emerging trends. In: Annales des télécommunications. 2005 Apr;60(3):228–263. Paris: Springer-Verlag.