Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 91 (2025) 536–543

Transportation Research Procedia

www.elsevier.com/locate/procedia

The Science and Development of Transport - TRANSCODE 2025

Evaluating the Impact of Carpool Lanes on Highway Congestion Using Empirical Traffic Data.

Dillon Motilal[a], Patrick Hosein[b]

[a]*Department of Computing and Information Technology, University of the West Indies, St. Augustine, Trinidad and Tobago*
[b]*Department of Electrical and Computer Engineering, University of the West Indies, St. Augustine, Trinidad and Tobago*

Abstract

Traffic congestion remains a serious challenge in many developing countries, namely Small Island Developing States (SIDS), causing long travel times during peak hours, increased economic pressure, and a decline in quality of life. Limited infrastructure and a lack of publicly available traffic data make it difficult to assess the potential High Occupancy Vehicle (HOV) lanes can provide in reducing congestion. To address this, traffic data from California's Performance Measurement System (PeMS) was used to construct an analytical model grounded in macroscopic, statistical relationships between vehicle density, speed, and travel time. LOWESS (Locally Weighted Scatterplot Smoothing) techniques were applied to fit the nonlinear nature of these traffic flow relationships. Using Trinidad and Tobago as the SIDS, Waze was used to identify congestion zones and provide real-time traffic metrics for comparison. A hypothetical 30% reduction in vehicle density was modeled to estimate the potential impact of HOV lane implementation, while their estimated speed and travel time were interpolated from the LOWESS plots. Under this scenario, travel times were meaningfully reduced, bringing average speeds close to low congestion conditions and, in some cases, close to free flow speeds. The analysis covered eight congested segments along several key routes, suggesting that HOV lanes could offer a practical and affordable way to alleviate congestion.

Keywords: Traffic Congestion; High Occupancy Vehicle (HOV) Lanes; Trinidad and Tobago; Macroscopic Model; Vehicle Density; LOWESS Smoothing; PeMS Data; Waze

1. Introduction

In urbanized areas, there is typically an imbalance between vehicle demand and road capacity, which inevitably results in congestion. This, in turn, leads to slower travel speeds, longer trip times, and negative impacts on the

* Dillon Motilal. Tel.: +1-868-743-9807
** Patrick Hosein. Tel.: +1-868-483-4454
E-mail address: dillon.motilal@my.uwi.edu

10.1016/j.trpro.2025.10.069

country's economy. Aftabuzzaman (2007) noted the lack of a universally accepted definition of congestion. However, for this study, the definition adopted is linked to Lomax et al. (1997) and Rosenbloom (1976), where congestion arises when travel demand exceeds the existing road system capacity, resulting in excess delays compared to those of free-flow conditions. As Aftabuzzaman (2007) explained, such conditions lead to increased travel time and costs, which affect the behavior of drivers. This definition is relevant to this study, as it examines highway segments within a developing country that suffers from traffic congestion and its effects during peak hours.

Traffic congestion imposes economic inefficiency, longer commute times, and reduced productivity. These conditions are exacerbated by limited land availability, an abundance of privately owned vehicles, and fiscal limits, all of which are characteristics to SIDS. In such cases, evaluating methods to alleviate congestion not only benefits its citizens financially but also enhances their socioeconomic well-being.

2. Related Work

HOV lanes are designated lanes for vehicles carrying two or more occupants, which reduces the use of single-passenger cars. This implementation incentivizes carpooling to alleviate congestion during peak travel periods (Naseri 2024, 1). Added, the success of HOV lanes largely depends on their design features, such as occupancy requirements along with enforcement. Two-person minimum configurations have been the most successful when compared to three-person minimums, mainly because they are easier for commuters to coordinate, Martin et al. (2004, 13). HOV lanes are considered feasible, as they utilize existing infrastructure. They can be introduced on existing roadways, with enforcement laws and additional incentives to assist in the public adoption of this policy. These characteristics allowed HOV lanes to be experimented with.

SIDS faces limitations in infrastructure, funding, and public access to traffic flow data. In such contexts, simulation provides a cost-effective and scalable approach to transportation planning. This allows researchers to evaluate the feasibility of policies prior to deployment. Su et al. (2024, 2) highlighted that large traffic networks are best represented and analyzed using macroscopic models, which do not require detailed modeling of vehicle-to-vehicle interactions. Among traffic simulations, macroscopic simulations take the least computing time.

Traditional macroscopic simulation tools have been evaluated in studies and have proven to be useful for urban development. However, these tools require data that are not publicly available, which makes it impractical in this case and created a barrier to effectively deploying them for HOV analysis. As a result, this study adopts a macroscopic modeling approach to produce an analytical estimation based on the statistical relationships formed from the empirical traffic data. This method is suited for environments with limited access to data.

3. Methodology

This study utilizes PeMS data from District 3, which encompasses the Sacramento Metropolitan Area and is recognized as an urban region (May et al. 2007, 55). These characteristics make District 3 a suitable point of comparison for SIDS, as it features multilane highways entering densely populated areas and is expected to be affected by peak-hour congestion. The combination of road structure and potential congestion serves as a reference for using this dataset to assess the potential impact of HOV lanes in developing countries. In addition, District 3 has been used in previous studies, such as May et al. (2007), to assess HOV lane performance, lending further support to its relevance in this context.

This dataset contained over 1.38 million hourly records for January 2025, capturing traffic variables such as speed, flow, occupancy, and station length. Roughly 61% of the dataset was complete across these variables. Median imputation was applied to approximately 500 thousand missing values, categorized by unique freeway number and absolute post-mile marker. This ensured that the imputation was localized to the specific station's position along each freeway.

Additionally, there were missing data related to the distance covered by stations and the number of lanes. To address this, columns containing post-mile and lane information were extracted from the metadata file and merged with the main dataset. Only records corresponding to Mainline (ML) lane types were retained for analysis. Finally, speed was converted to meters per second, and the distance between stations was calculated using the differences in their absolute post-mile values.

We computed two important metrics. The **Vehicles per km** gives a measure of traffic density by showing how many vehicles are present on average in each kilometer. This metric is useful for identifying congestion and its relationships, since segments with higher vehicle density typically experience slower speeds and longer delays.

The **Travel Time per km** expresses travel time in seconds per kilometer and compares delay to the congestion within segments. This metric gives a clearer view of how long it takes traffic to move across segments.

To examine how delay behaves with increasing traffic volume, a macroscopic modeling approach was used to analyze the relationships between the traffic variables: vehicle density (veh/km), speed (km/h), and travel time per kilometer (s/km) in the ML lane type configuration. These relationships were first plotted with binned data, as they reduce the computation time, and then fitted with LOWESS smoothing to highlight the overall trends captured within the dataset. Fig. 1a contains a plot of travel time versus vehicle density while Fig. 1b contains a plot of vehicle speed versus vehicle density.

(a) Vehicle Density vs Travel Time

(b) Vehicle Density vs. Speed

Fig 1. LOWESS Smoothed Relationships for 3-Lane Mainline Segments

These plots allowed for the relationship between varying levels of vehicle density and their corresponding changes to travel time and speed to be observed. To apply the model to Trinidad and Tobago, Waze was used to identify congestion segments along busy routes and the estimated speed within those segments. Four routes between cities were studied that typically experience congestion during peak hours. These routes are illustrated in Fig. 2, which shows the morning congestion zones for the Arima to Port of Spain route and the afternoon congestion zones for the return route from Port of Spain to Arima. Similarly, Fig. 3 depicts the morning congestion zones for the San Fernando to Port of Spain route and the afternoon congestion zones for the return route from Port of Spain to San Fernando. The estimated speed values were then matched to the LOWESS curves where the corresponding density and travel time per kilometer values were interpolated from the plots.

To assess the potential impact of HOV lanes, this study manipulated traffic density within the established relationship between vehicle density, speed, and travel time. Specifically, a hypothetical scenario was modeled where vehicle density was reduced by 30%, representing a shift towards carpooling as guided by the modeling strategy described by Naseri et al. (2024). LOWESS curves from 3-lane Mainline segments were used to model this reduction and interpolate the resulting travel time and speed improvements. The 3-lane ML configuration was selected as it is structurally similar to the routes being analyzed.

4. Results

To understand the distribution of key traffic variables after data cleaning, exploratory data analysis (EDA) was conducted using histograms and boxplots. The histograms and boxplots revealed important distribution patterns in the dataset. **Travel Time** exhibited a heavy right skew, with most values concentrated below 1000 seconds and a long tail indicating occurrences of high congestion. **Speed** showed a roughly normal distribution centered around 30 m/s, with low-end outliers indicating slow traffic conditions and high-end outliers mostly falling between 33 and 35

Fig 2. Congestion on the Arima to Port of Spain Route at 7am (top) and Port of Spain to Arima at 5pm (bottom).

m/s, with one instance exceeding 40 m/s. **Total Flow** was right-skewed, showing a median flow rate of nearly 1000 vehicles per hour and a vast number of outliers past 6000 vehicles per hour. The periods of low vehicle flow may either indicate congestion periods or off-peak hours, while the outliers may point to peak hours of travel right before congestion occurs. **Distance** between stations showed a strong right skew, with a few large gaps likely representing missing sensors or long segments. These insights validated the need to retain outliers to preserve meaningful traffic behaviors.

To model the nonlinear relationships between Vehicle Density, Speed and Travel Time, LOWESS was applied. These observed relationships formed the basis for modeling congestion levels along real-world segments. For instance, travel time per kilometer increased dramatically beyond 55 vehicles/km. Once congested zones were identified, Waze-estimated travel speeds were used to extract the estimated vehicle density and travel time values from the LOWESS graphs through interpolation.

Across all eight segments, the modeled time saved from the carpooling scenario ranged from 2.56 to 14.81 minutes, depending on baseline traffic density and segment length. Table. 1 summarizes the time saved in each segment analysis. A column was dedicated to the perceived impact based on the percentage of time saved. The defined thresholds were as follows: moderate impact was assigned to segments with a percentage time saved between 30% and 39%, while high impact was anything above 40%. These results formed the basis for evaluating the potential benefits of HOV lanes on congestion across major urban corridors.

5. Discussion

Combining the traffic flow patterns extracted from California's PeMS dataset alongside the congested segments identified, the model estimated the effects of a 30% shift to carpooling on traffic conditions. As demonstrated in the results section, a reduction in delay across all segments was observed along the Churchill Roosevelt, Uriah Butler, and Solomon Hochoy Highways, with some benefiting more than others.

Fig 3. Congestion on the San Fernando to Port of Spain Route at 7am (left) and Port of Spain to San Fernando at 5pm (right)

Table 1. Summary of Modeled HOV Effects by Segment

Segment	Distance (km)	Density (veh/km)	Time Saved (min)	% Time Saved	Post-HOV Speed (km/h)	Impact Category
Arima to POS 1	9.6	158	14.5	48%	35.68	High Impact
Arima to POS 2	9.3	95	5.67	43%	75.76	High Impact
San Fernando to POS 1	15.9	109	10.47	40%	61.42	High Impact
POS to Arima 1	7.0	132	5.61	37%	44.70	Moderate Impact
POS to Arima 2	4.8	176	10.41	53%	31.26	High Impact
POS to Arima 3	3.7	122	2.56	36%	49.86	Moderate Impact
POS to San Fernando 1	22.8	106	14.81	41%	64.53	High Impact
POS to San Fernando 2	13.9	135	11.89	38%	42.97	Moderate Impact

This chapter interprets the modeling results, evaluates their effectiveness locally, recognizes the limitations and the assumptions, and provides policy recommendations and directions for future work. Through interpreting the findings documented in the results section, the Modeled carpooling scenario improved travel speed and reduced travel time across all eight segments, especially in zones with high congestion. High congestion was defined as any segment with a baseline density exceeding 125 vehicles per kilometer, based on the threshold observed from Fig. 1a.

The effects of the model varied depending on factors such as the initial congestion level, length, baseline speed, and travel time of the zones being evaluated. Segments with higher baseline densities and lower speeds experienced the most improvement. For instance, Arima to Port of Spain Segment 1, one of the most congested routes, saw a 14.5-minute delay improvement. Likewise, Port of Spain to Arima Segment 2, despite only covering 4.8 kilometers, recorded a 10.41-minute reduction in delay.

Moderately congested segments also benefited from the modeled policy scenario. In the Port of Spain to San Fernando segment 1, the baseline speed recorded was 38 kilometers per hour. However, the length of this segment allowed for the increase in speed to accumulate over its distance, resulting in 14.81 minutes being saved. It is important to note that this method only models the traffic flow conditions in ML lanes. By design, HOV lanes accommodate fewer vehicles due to occupancy restrictions and, as a result, experience lower delays because of a reduced density. This is supported by Fig. 1a which illustrates the inverse relationship between vehicle density and travel time. Evidently, the modeled 30% reduction in traffic flow reflects the impact of carpooling, and HOV lanes may yield greater travel time reductions than general-purpose lanes.

The outcomes identified from the modeling coincide with the principles of traffic flow theory outlined by Gerlough and Huber (1975). As illustrated in Fig. 1a, travel time increases nonlinearly as vehicle density rises, especially after crossing a specific threshold. While the congestion levels analyzed in this study did not reach their maximum capacity, they were relatively high, resulting in long travel times. Therefore, a reduction in vehicle density for segments approaching their maximum threshold would likely yield the greatest reduction in delay, while moderately congested segments would still benefit, though to a lesser extent. This effect was especially evident when comparing the Arima to the Port of Spain segments 1 and 2. Here, they both covered roughly the same distance, but segment 1 had a significantly higher baseline density and saw greater travel time improvements under the HOV scenario.

These estimations support the literature on the usefulness of a macroscopic model for evaluating traffic policies in congested areas and a data-limited environment (Ratrout and Rahman, 2009). While the implementation of HOV lanes is not immediately feasible, this method provided the opportunity to evaluate its potential impact. The use of LOWESS smoothing allowed for the nonlinear trends to be fitted, along with estimating the effects of the carpooling HOV scenario. Notably, the modeled HOV conditions produced from these plots generated travel time estimates that corresponded to low congestion conditions defined by Waze. The alignment between the modeled output and real-time estimates lends credibility to the method used.

Due to the absence of publicly available traffic flow data, this evaluation used PeMS data from California's District 3 to create a macroscopic model of conditions faced by developing countries. Additionally, Waze maps were used to determine the congestion zones along the analyzed routes. This was an essential step to estimate traffic flow conditions in the selected country, even though the reports were dependent on user feedback. The modeled travel time per kilometer was generally consistent with the estimated travel time from Waze, thereby strengthening the decision to use PeMS-derived relationships as an alternative for local traffic behavior. The carpooling adoption rate applied in this study served as a reasonable first estimate, which reflects reluctance to carpool and leaves room for future growth.

This model provides a foundation for policies aimed at reducing congestion. All policies proposed below are built around the idea that, during peak congestion hours, one of the three highway lanes would temporarily be converted into an HOV lane, supported by government enforcement and incentives. For example, during the morning peak, one of the three inbound lanes toward Port of Spain can be designated for HOV use and then later reverted back to general-purpose lanes once this period has passed, with the outbound lanes applying the same reasoning for the afternoon peak.

- **Hybrid HOV-Public Transport Model**: This policy comprises a hybrid approach that merges carpooling with existing public transportation. For instance, commuters from high-traffic areas, such as Arima and San Fernando, can travel to a designated centralized hub. Here, public transport such as PTSC buses may carry passengers to Port of Spain during the morning peak. In the evening, passengers can return to these hubs for transportation. This aims to reduce the number of single-occupancy vehicles entering and leaving Port of Spain as observed in Thorne et al. (2024, 13). Added, this method caters to individuals who are not willing to carpool.
- **Enforcement and Incentives**: Enforcement goes hand in hand with the effectiveness of HOV lanes, as seen from the Jakarta incident documented by Cohen et al. (2022). To avoid a similar issue, giving priority signage for vehicles with two or more passengers, along with adding designated HOV-only entry and exit ramps, is necessary and may be enhanced using a camera monitoring system. Furthermore, incentives should be offered to those willing to use the hybrid HOV public transport policy, such as dedicated parking and reduced parking fees at hubs.

- **Scalability**: While this study focused on specific routes, the same methodology may be applied to other highway routes where congestion usually occurs and other developing states. This promotes a scalable, cost-efficient method of congestion management.

This research was faced with several limitations due to data availability and assumptions required for local relevance. First, while the PeMS data set is detailed, it is based on California's infrastructure and driver behavior. Although there were similarities, such as highways passing through densely populated areas, the data may not accurately represent the relationships between driving habits, vehicle types, and drivers' reactions to congestion in Trinidad and Tobago. It also affects how values were estimated in the local context. As sensor data for local roadways was not available, vehicle density and travel time were interpolated from LOWESS plots using an estimated speed from Waze. The use of this dataset and its values assumes that it accurately represents conditions in the country being studied. While internal validation with generated and real-time values was seemingly consistent, the results should be interpreted specifically as approximations. Second, while a 30% reduction in vehicle density was applied uniformly across all segments, carpooling adoption is likely to vary with factors such as the time of day, crime rate, or the route. This assumption was made to allow for a consistent, baseline comparison across segments. Third, the identification of congestion zones relied on Waze, which may introduce potential errors. This is because Waze depends on user reports, and as they do not store historical congestion snapshots, congestion segments had to be captured in real time.

For example, on the day data was collected, Waze estimated the morning travel time from Arima to Port of Spain during peak hours to be 1 hour and 5 minutes. This contradicted comments from drivers during these peak hours, who often reported travel times taking up to 2 hours. Google Maps was referenced, and the estimated times ranged from 50 minutes to 1 hour and 50 minutes. This suggested that on the day data was collected, congestion was not at its peak and was assumed to be the result of the hybridization of work. This assumption aligned with the observation that only two segments had estimated densities exceeding 150 vehicles per kilometer. Referring to the analysis of the results, travel time reduction is expected to scale with congestion, meaning the model could provide even more impactful results under higher congestion. These assumptions and limitations highlight the challenges of evaluating traffic policies in a data-limited environment.

6. Conclusions and Future Work

Driven by the goal of evaluating whether HOV lanes could meaningfully reduce traffic congestion, the limited infrastructure, funding constraints, and lack of publicly available traffic data faced by SIDS posed a unique challenge. To counteract this, a macroscopic model was developed using California's PeMS dataset, paired with LOWESS smoothing and real-time congestion segments identified by Waze.

The model analyzed eight peak-hour congestion segments in Trinidad and Tobago, spanning routes from Arima, San Fernando, and Port of Spain. Applying a 30% reduction in vehicle density, representing potential carpool adoption, produced improvements in travel speed and travel time, which were similar to low congestion conditions identified by Waze. In the highest congestion segment, the Port of Spain to Arima segment 2, speeds nearly doubled from 14.4 km/h to 31.27 km/h, with a reduction in delay by 10.41 minutes in just 4.8 kilometers. Notably, even moderately congested areas showed improvements, indicating that the benefits of HOV lanes are not limited to the highest congested segments.

While the study relied on international data and included several assumptions, the modeling approach provided a solid foundation for testing traffic policies in data-limited environments. The use of Waze for identifying congestion segments, while dependent on user data, served as a reasonable proxy in the absence of sensor data.

Collectively, the study offers a practical framework for evaluating congestion management policies in SIDS. This data-driven approach highlights how congestion can be effectively reduced by simply using existing infrastructure better, making it an affordable and practical method.

There are areas to explore for future research that may strengthen the findings. Firstly, collecting primary data on drivers' willingness to carpool in Trinidad and Tobago through surveys. These would be targeted at drivers in specific locations, so the carpooling adoption rate applied would be specific to the route taken. This would allow for a more realistic scenario based on local behavior rather than relying on international findings. This method could also be applied to any developing state being analyzed. Secondly, this model can be validated by installing road sensors along the examined routes to determine if the vehicle density, speed, and travel times estimated from the LOWESS graphs accurately capture local behavior.

References

Aftabuzzaman, M., 2007. Measuring traffic congestion-a critical review, in: 30th Australasian transport research forum, ETM GROUP London, UK. pp. –.

Bickel, P., Chen, C., Kwon, J., Rice, J., van Zwet, E., Varaiya, P., 2006. Measuring Traffic. Technical Report. Department of Statistics, University of California, Berkeley. URL: https://www.stat.berkeley.edu/~bickel/traffic.pdf. technical Report.

Boxill, S.A., Yu, L., 2000. An evaluation of traffic simulation models for supporting its. Houston, TX: Development Centre for Transportation Training and Research, Texas Southern University , 25.

Chen, C., Petty, K., Skabardonis, A., Varaiya, P., 2003. Detecting errors and imputing values for loop detector data. Technical Report UCB-ITS-PWP-2003-18. California PATH Program, Institute of Transportation Studies, University of California, Berkeley. URL: https://escholarship.org/content/qt6j93p90t/qt6j93p90t_noSplash_b3faa51086ad3c7d1a28cc92c23b920c.pdf?t=krnrxz. pATH Working Paper.

Cohen, M.C., Jacquillat, A., Ratzon, A., Sasson, R., 2022. The impact of high-occupancy vehicle lanes on carpooling. Transportation research part A: policy and practice 165, 186–206.

Gerlough, D.L., Huber, M.J., 1975. Traffic Flow Theory: A Monograph. Special Report 165. Transportation Research Board, National Research Council. Washington, D.C. URL: https://onlinepubs.trb.org/onlinepubs/sr/sr165/165.pdf. accessed June 2025.

Jain, V., Sharma, A., Subramanian, L., 2012. Road traffic congestion in the developing world, in: Proceedings of the 2nd ACM Symposium on Computing for Development, pp. 1–10.

Lomax, T., Turner, S., Shunk, G., Levinson, H., Pratt, R., Bay, P., Douglas, G., 1997. Quantifying Congestion. Volume 2: User's Guide. Technical Report. Texas Department of Transportation.

Martin, P.T., Perrin, J., Wu, P., Lambert, R., 2004. Evaluation of the effectiveness of high occupancy vehicle lanes. Report to Utah Department of Transportation .

May, A.D., Leiman, L., Billheimer, J., 2007. Determining the effectiveness of hov lanes. UC Berkeley: California Partners for Advanced Transportation Technology .

Naseri, A., Estelaji, F., Samani, A., Omidifar, R., Zahedi, R., Yousefi, H., 2024. Simulating the performance of hov lanes for optimal urban traffic management. Transportation research interdisciplinary perspectives 23, 101010.

Ratrout, N.T., Rahman, S.M., 2009. A comparative analysis of currently used microscopic and macroscopic traffic simulation software. The Arabian Journal for Science and Engineering 34, 121–133.

Remak, R., Rosenbloom, S., 1976. Peak-period Traffic Congestion: Options for Current Programs. Technical Report. The University of Texas.

Sharma, S., Swayne, D.A., Obimbo, C., 2015. Automating the smoothing of time series data. Journal of Environmental & Analytical Toxicology 5, 304.

Su, Y., Ghaderi, H., Dia, H., 2024. The role of traffic simulation in shaping effective and sustainable innovative urban delivery interventions. EURO Journal on Transportation and Logistics , 100130.

Teh, H.Y., Kempa-Liehr, A.W., Wang, K.I.K., 2020. Sensor data quality: A systematic review. Journal of Big Data 7, 11.

Thorne, E., Chong Ling, E., Phillips, W., 2024. Assessment of the economic costs of vehicle traffic congestion in the Caribbean: a case study of Trinidad and Tobago. Technical Report. Naciones Unidas Comisión Económica para América Latina y el Caribe (CEPAL).

United Nations Development Programme, 2024. Small island developing states are on the frontlines of climate change – here's why. URL: https://climatepromise.undp.org/news-and-stories/small-island-developing-states-are-frontlines-climate-change-heres-why. accessed June 14, 2025.