p–ISSN: 2723 - 6609 e-ISSN: 2745-5254
Vol. 5, No. 10, October 2025 http://jist.publikasiindonesia.id/

Linear Regression-Based Traffic Flow Simulation: Vehicle Density and Speed Analysis on Buah Batu Road

Adam Ichwanul Ichsan[1], Putu Harry Gunawan[2]*
Universitas Telkom, Bandung, Indonesia
Email: aichwanulichsan@student.telkomuniversity.ac.id[1], phgunawan@telkomuniversity.ac.id[2*]
*Correspondence

Keywords: linear regression, traffic flow, velocity-density function, simulation.

ABSTRACT

Traffic congestion has become an increasingly severe problem in many major cities around the world, including in the city of Bandung. Population growth and increased vehicle use exacerbate congestion. Jalan Buah Batu, one of the main roads in the city of Bandung, often experiences congestion due to high density. This study explains the traffic flow simulation using the Lighthill-Whitham-Richards (LWR) model with a speed-density function obtained from observation data on Jalan Buah Batu, Bandung. The data included the relationship between vehicle density and speed which was then analyzed using the linear regression method. The approximation of the velocity-density function obtained from linear regression is v(ρ)= -6.904+4.302. Traffic flow simulations were carried out with a road length of 60 meters, a total time of 5 minutes, and high resolution with 300 grid points. At the beginning of the simulation, a peak density of 0.70 occurred in a 15-25 meter road segment. Over time, the peak density shifted and decreased: 0.65 at 20-30 meters at 1.25 minutes, 0.60 at 25-35 meters at 2.5 minutes, and 0.50 at 30-50 meters at the end of the simulation (5 minutes). These results show the movement of vehicles that reduce congestion and improve the smooth flow of traffic. In conclusion, linear regression is effective in determining the velocity-density function.

Introduction

Traffic congestion has become an increasingly severe problem in many major cities around the world (Timpal et al., 2018), including in the city of Bandung (Dewi et al., 2020). Population growth and increased vehicle use have exacerbated congestion in the city in recent years. Bandung has a vehicle-to-population ratio of almost 1:1, with about 1.7 million two-wheelers, 500 thousand four-wheelers, and a population of about 2.5 million people (Susanto et al., 2016), thus causing high density on the city's streets (Wijaksana et al., 2020). One of the areas that often experience congestion is Jalan Buah Batu, one of the main roads in the city of Bandung that connects several important roads such as Jalan Siswa Pejuang 45, Jalan BKR, Jalan Soekarno Hatta, and Jalan Terusan Buah Batu.

Jalan Buah Batu, with its characteristics of being a secondary collector's road with a width of about 13 meters and a length of about 1.70 kilometers (Ayuni & Fitrianah, 2019)Penerapan metode Regresi Linear untuk prediksi penjualan properti pada PT XYZ, is a vital artery that is very important for mobility in the city of Bandung. Its strategic location, close to toll roads, entertainment centers, office areas, and educational institutions, makes it the main choice for people to reach various important locations. (Yermadona & Meilisa, 2020). However, with its various strategic functions, Jalan Buah Batu also faces significant challenges in maintaining smooth traffic and avoiding congestion that can harm mobility and the city's economy. The congestion that often occurs on this road not only interferes with the daily activities of residents but also hurts economic efficiency and quality of life (Gora et al., 2020).

Traffic flow models have been developed by many researchers to understand and address congestion problems. Traffic problems, such as congestion can be explained by traffic flow models. There are two main models in traffic flow: microscopic models that describe the individual behavior of cars such as position, speed, and acceleration, and macroscopic models that use partial differential equations to discuss traffic variables such as flow, speed, and density, also known as the Lighthill-Whitham-Richards (LWR) model. (Gunawan & Rizmaldi, 2019). Microscopic models tend to be more detailed and can capture the individual behavior of vehicles, while macroscopic models focus more on the overall behavior of traffic flows.

This study will explain the simulation of traffic flow using the LWR model. In the LWR model, the traffic flow is governed by the conservation equation, which can be rewritten as the transportation equation. (Muhartini et al., 2021). The velocity variable in the transport equation is represented by the velocity function, which must be defined based on the observation of velocity. This LWR model is very useful in predicting traffic flow patterns and identifying potential congestion points based on changes in vehicle density. (Harmizi et al., 2019).

The purpose of this study is to analyze and simulate the traffic flow model using the velocity-density function obtained from the observation data. The structure of this journal is as follows: The Introduction section introduces the topic. The Methods section discusses the macroscopic model of traffic flow along with linear regression. Finally, the Result and Discussion section presents the conclusions obtained from this study.

Method

Traffic Flow model

The macroscopic traffic flow model used in this study is the Lighthill-Whitham-Richards (LWR) model, which is a fundamental model in traffic flow theory. The LWR model is based on the principle of vehicle conservation [9] which states that the number of vehicles remains constant in a road segment over time. The traffic flow model uses the principle of mass conservation. In the context of traffic, this principle states that the flow of vehicles entering and exiting a certain observation point in a certain period remains constant. In this study, the author uses a macroscopic model of traffic flow, specifically the Lighthill-Whitham-Richards (LWR) model. This model describes traffic dynamics using conservation equations:

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho u)}{\partial x} = 0 \quad (1)$$

Here, ρ represents the density of the vehicle, and u is the average speed of the vehicle. The x and t variables indicate position and time, respectively. Observations show that the average speed of vehicles in traffic flow is affected by vehicle density. This relationship can be expressed as:

$$u = v(\rho) \quad (2)$$

In this model, the velocity function v(ρ) is density-dependent. The maximum speed of the vehicle is reached when the density is low or zero, while the speed decreases to zero when the density reaches the maximum level., This can be explained through the following equation:

$$(0) = v_{max} \text{ and } v(\rho_{\text{max}}) = 0 \quad (3)$$

Where Vmax is the maximum velocity and max is the maximum density of vehicles. By incorporating this velocity-density relationship into the conservation equation, we obtain the transport equation:

$$\frac{\partial \rho}{\partial t} + v(\rho)\frac{\partial \rho}{\partial x} = 0 \quad (4)$$

In this study, the velocity function v(ρ) is determined using the linear regression method. This approach was chosen to account for the dependence between velocity and density. By utilizing regression analysis, this study aims to accurately model and investigate traffic flow models.

Linear Regression

Linear regression analysis is an approach method to model the relationship between one dependent variable and one independent variable. In the context of regression, independent variables are used to describe variations in dependent variables. In a simple regression analysis, the relationship between the two variables is linear, which means that a change in variable X will result in a change in variable Y in a fixed way. In contrast, in a non-linear relationship, the change in variable X will not be followed by the variable Y proportionally.

$$v(\rho) = a + b\rho \quad (5)$$

Where:

v = Speed

a = coefficient

b = Intercept

= Density

In this study, the dependent variable is v (Speed), and (Density) is the independent variable.

Methodology

Flowchart System

The first step involves conducting a survey and collecting datasets on the road under study, which is Jalan Buah Batu. The data collection is carried out on different days, at various times, and in random intervals, recording traffic conditions in the form of video files. Once the datasets are obtained, they undergo preprocessing, where the video data is converted into numerical data and then entered into Excel to facilitate data management, eventually being saved in a .csv format.

The next stage is the development of a linear regression model based on the formulas and data previously acquired. Following this, the model undergoes testing and evaluation to determine if it aligns with the real-world graph. If the results from the testing and evaluation do not meet expectations, the process will revert to the model development stage. Once the model meets the expected criteria, the process moves on to the next stage, which involves simulating the validated model.

Observation

The flow of the research system that has been carried out includes: The first step is to survey and take datasets directly on Buah Batu Road with an observation length of 18 meters from one direction. By observing from one direction, it provides a detailed analysis of the movement of vehicles on the road section. For this research, 4 motorcycles are equal to 1 car. The dataset collection was carried out over several days at different and random times, by recording the state of the traffic using a tripod-mounted camera in the form of a video file. Figures 2 and 3 illustrate the planned location and site, depicting the actual conditions at the dataset collection site.

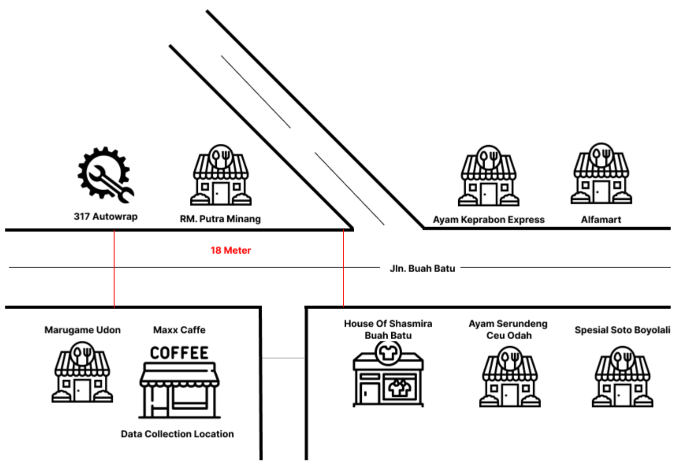

Figure 2
Location Plan of the Observation Area

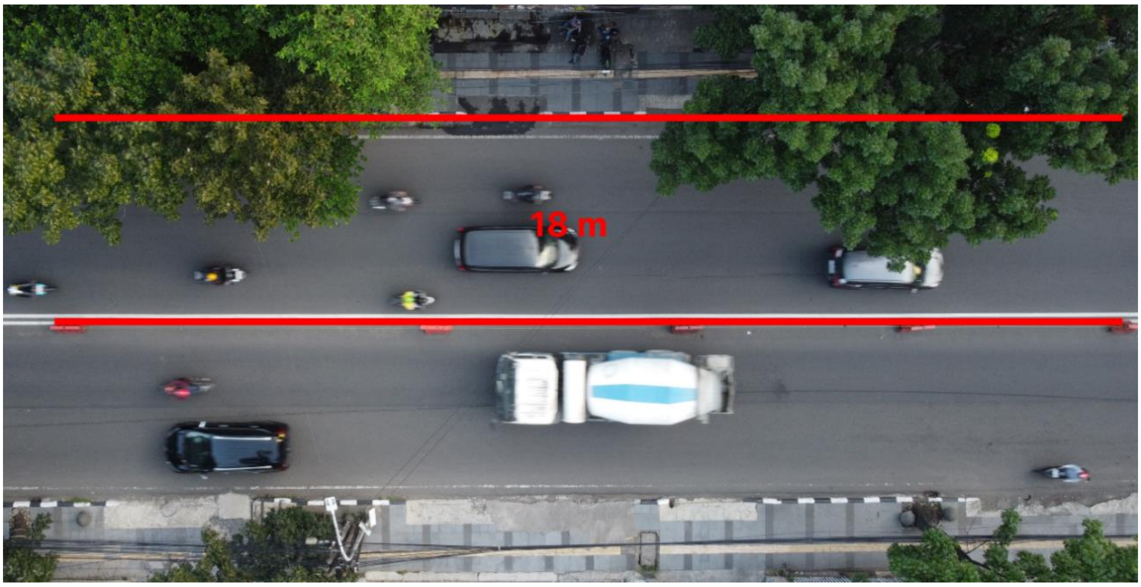

Figure 3
The Real Situation Within the Observation Area

Dataset

The dataset in this study was carried out on Jalan Buah Batu, with an observation length of 18 meters from one direction. Data collection was carried out over several days at different and random times, as well as recording traffic conditions in the form of videos. After obtaining the dataset, the data is pre-processed by converting the video into numerical data which is then entered into Excel format to facilitate data collection and converted into .csv format. An example of the dataset obtained can be seen in Table 1.

Table 1
Example of Dataset of Buah Batu Road

No	T in(mm:ss)	T out(mm:ss)	Speed	Density
1	00:29	00:39	1.80	0.55
2	00:32	00:55	0.78	0.41
3	00:33	00:45	1.50	0.48
4	00:36	01:00	0.75	0.48
5	00:36	00:48	1.50	0.48
…	…	…	…	…
500	07:47	07:50	6.00	0.03

The speed in the dataset was obtained by analyzing videos and applying the formula of the speed $v = L/ \Delta t$ where v is the velocity of the vehicle, L is the length of the road, and Δt is the time interval traveled by the vehicle. Additionally, the density was calculated using the formula of the density = V/Vmax where is the density of the vehicle, V is the traffic volume, and Vmax is the maximum of the traffic volume.

Results and Discussion

Approximation of Velocity Function

The purpose of collecting this observation data is to obtain the velocity function. This is achieved by gathering detailed information, including the average speed and density of vehicles passing through the observation area. The velocity function is obtained as follows:

$$(\rho) = -6.904 + 4.302\rho \quad (6)$$

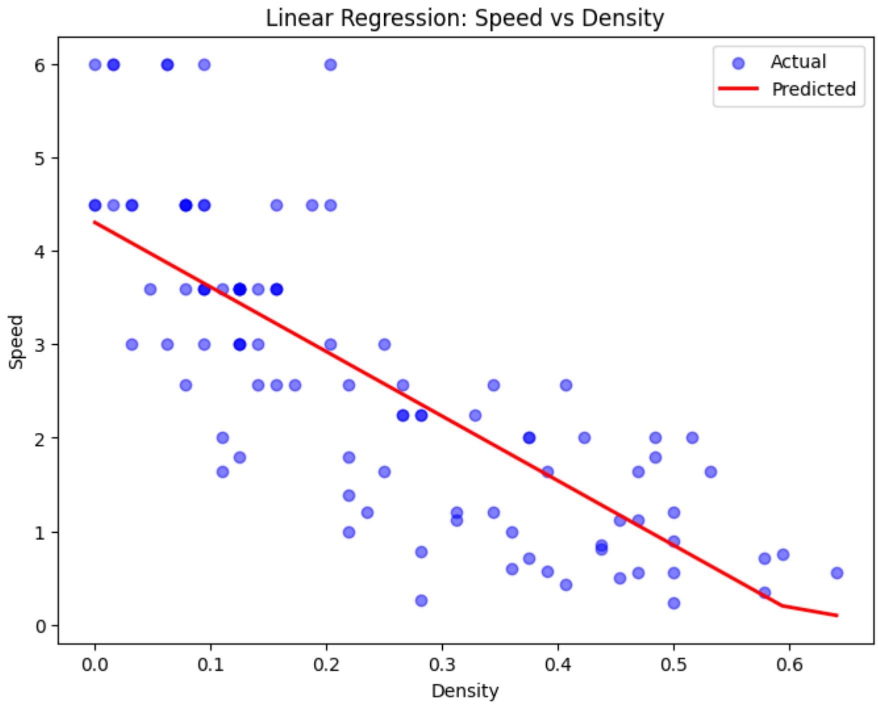

Figure 4
Visualization of linear regression functions vs datasets

This parameter indicates that the traffic speed decreases linearly as the density increases. The negative slope in this graph indicates the inverse relationship between the speed and density of the vehicle. In Figure 3, the velocity function shown provides a fairly accurate representation of the observation data. Thus, in the next section, the derivative linear approach function will be applied as a velocity function in the numerical simulation to be performed. (Rahayu et al., 2020).

Numerical Simulation

As can be seen in Figures 5 and 6, the simulation parameters used include a road length of 60 m(L), with a space grid point of 300 (nx). The total simulation time was set for 5 minutes, $\Delta t = 0.001$ and $\Delta x = L/nx = 0.2$. This approach ensures that the simulation model has high accuracy in representing real traffic conditions throughout the simulation period. At the beginning of the simulation (t = 0.00 minutes), the road segment ranging from 15 to 25 meters showed a high density of vehicles, with the peak density reaching about 0.70. This shows that at the beginning, there was a significant accumulation of vehicles on the road segment.

Figure 5
Initial Condition of Traffic Flow Simulation

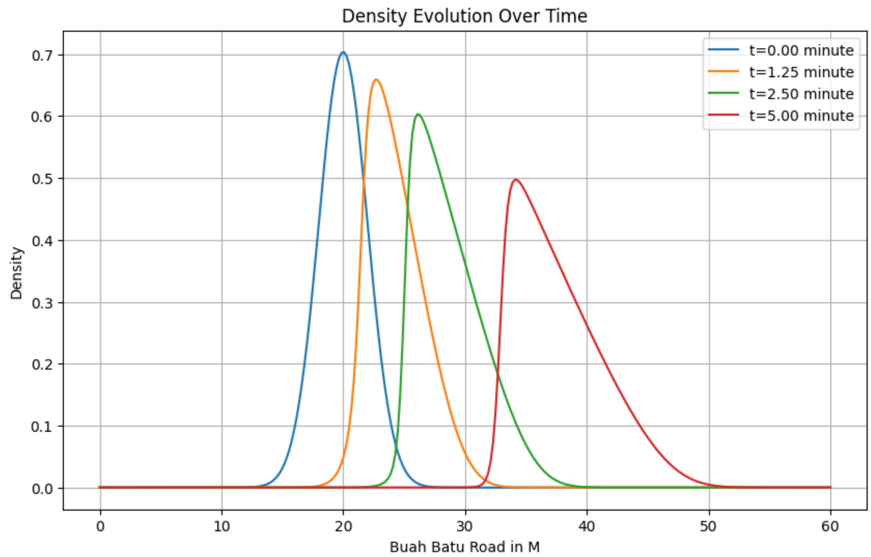

Figure 6 Traffic Flow Simulation

Over time, at t = 1.25 minutes, the peak of density shifts to the right, which is about 20 to 30 meters from the starting point. The maximum density is also slightly reduced to about 0.65. This shows that vehicles are starting to move forward, but there is still congestion in the segment. At t = 2.50 minutes, the density of vehicles continues to shift further to the right, reaching about 25 to 35 meters. The maximum density again decreased to around 0.60. This indicates that the flow of traffic is starting to smooth out and vehicles are slowly spreading along the road. At the end of the simulation (t = 5.00 minutes), the peak density is around 30 to 50 meters from the starting point. The maximum density dropped drastically to around 0.50, indicating that congestion has been significantly reduced and traffic flow has become smoother. The results of this simulation show that over time, vehicles that were initially collected in a particular segment begin to move forward and spread along the road, reducing congestion at the starting point of congestion. The average speed of a vehicle increases as the density decreases, which is reflected in the decrease in the maximum density value over time. Thus, this simulation illustrates the dynamics of traffic flow on Jalan Buah Batu and how congestion can be unraveled over time with increased vehicle movement.

Conclusion

This study found important results related to the speed function approach using linear regression and traffic flow model simulation. With linear regression, the velocity function is obtained as $v= -6.904+4.302\rho$, indicating the inverse relationship between velocity and density. The simulation was carried out on a 60-meter road for 5 minutes. Initially, the maximum density was at $x = 20$ meters and moved over time reaching $x = 30$ to 50 meters with a density $\rho = 0.50$ at the end of the simulation. The leading vehicle reaches $x = 40$ meters.

The application of linear regression to estimate the effective velocity-density function in traffic flow simulation uses an upwind scheme. These findings demonstrate the importance of linear regression in traffic modeling. Further research can explore non-linear models or machine learning techniques for a more complex relationship between density and velocity. For example, combining real-time traffic data and adaptive traffic signal control systems can improve model accuracy and applicability. Additionally, expanding the model to include variations in driver behavior, road conditions, and weather influences can provide a more thorough understanding of traffic patterns. Alternative numerical methods such as the Lax-Wendroff, Lax-Friedrichs, or MacCormack schemes can also improve accuracy and stability.

Bibliography

Ayuni, G. N., & Fitrianah, D. (2019). Penerapan metode Regresi Linear untuk prediksi penjualan properti pada PT XYZ. *Jurnal Telematika*, *14*(2), 79–86.

Dewi, T., Fajar, Y., Badruzzaman, F., Suhaedi, D., & Harahap, E. (2020). Simulasi Kemacetan Lalu Lintas Pada Lokasi Bundaran Baltos Bandung. *Smart Comp*, *9*(2), 92–95.

Gora, P., Katrakazas, C., Drabicki, A., Islam, F., & Ostaszewski, P. (2020). Microscopic traffic simulation models for connected and automated vehicles (CAVs)–state-of-the-art. *Procedia Computer Science*, *170*, 474–481.

Gunawan, P. H., & Rizmaldi, M. A. (2019). Approximation of velocity-density function for traffic flow model with obstacle problem in Jalan Merdeka Bandung. *2019 7th International Conference on Information and Communication Technology (ICoICT)*, 1–6.

Harmizi, M., Azanuddin, S., & Pd, E. S. (2019). Data Mining Untuk Mengestimasi Tingkat Kecelakaan Lalu Lintas Kab. Serdang Bedagai Menggunakan Algoritma Regresi Linier Berganda. *Jurnal Cyber Tech*, *2*(12).

Muhartini, A. A., Sahroni, O., Rahmawati, S. D., Febrianti, T., & Mahuda, I. (2021). Analisis Peramalan Jumlah Penerimaan Mahasiswa Baru Dengan Menggunakan Metode Regresi Linear Sederhana. *Jurnal Bayesian: Jurnal Ilmiah Statistika Dan Ekonometrika*, *1*(1), 17–23.

Rahayu, W. I., Trigunawan, A., & Andarsyah, R. (2020). *Regresi linier untuk prediksi jumlah penjualan terhadap jumlah permintaan*. Kreatif.

Susanto, I., Studyana, M. D., & Astor, Y. (2016). Analisis Kapasitas Jalan menggunakan Pendekatan Geospasial (Wilayah Studi: Bandung Tengah). *Potensi: Jurnal Sipil Politeknik*, *18*(1).

Timpal, G. S. J., Sendow, T. K., & Rumayar, A. L. E. (2018). Analisa Kapasitas Berdasarkan Pemodelan Greenshield, Greenberg dan Underwood dan Analisa Kinerja Jalan pada Ruas Jalan Sam Ratulangi Manado. *Jurnal Sipil Statik*, *6*(8).

Wijaksana, W. K., Rahayu, W. I., Saputra, M. H. K., & ST, M. T. I. (2020). *Regresi Linear Untuk Prediksi Jumlah Pengunjuing Terhadap Jumlah Petugas Dalam Menentukan Penjadwalan Penjagaan*. CV. Kreatif Industri Nusantara.

Yermadona, H., & Meilisa, M. (2020). Pengaruh Aktivitas Pasar Terhadap Arus Lalu Lintas (Studi Kasus Pasar Baso Kabupaten Agam). *Rang Teknik Journal*, *3*(1), 75–82.