BIBECHANA

ISSN 2091-0762 (Print), 2382-5340 (Online)

Journal homepage: http://nepjol.info/index.php/BIBECHANA

Publisher: **Department of Physics, Mahendra Morang A.M. Campus, TU, Biratnagar, Nepal**

Real-Time Traffic Prediction Using Monte Carlo Simulation: A Case Study of Kantipath Road, Kathmandu, Nepal

Niraj Kumar Shah[1], Prena Chaudhary[1], and Gopi Chandra Kaphle[1,*]

[1]Central Department of Physics, Tribhuvan University, Kathmandu, Nepal

Email: gck223@gmail.com; gopi.kaphle@cdp.tu.edu.np

Article Information:
Received: August 18, 2021
Accepted: December 30, 2021

Keywords:
Traffic congestion/jam
Distribution characteristics
Composition of vehicles
Traffic prediction
Monte Carlo Simulation

ABSTRACT

This paper deals with the model development for the analysis of traffic data to improve efficiency and sustainability of the existing transportation system of a section of the Kantipath road, Kaiser Library to Jamal. The traffic data of the two extreme locations of the section of the road were taken using digital video recorders. The data were extracted manually and statistically analyzed to test the different statistical distribution of traffic flow using Kolmogorov–Smirnov and Shapiro–Wilk test, and to find out different categories of vehicles along with their flow and peak hour. The Kolmogorov–Smirnov test reveals the traffic flow distribution is the same as Cauchy, Normal, Inverse Gaussian, Logistic, Gamma, and Beta distributions. The dataset of Kaiser Library was used to predict the real-time traffic flow of Jamal employing Monte Carlo Simulation in accordance to normal distribution. The forecasted and observed traffic flows show better agreement.

DOI: https://doi.org/10.3126/bibechana.v19i1-2.46394

1. Introduction

Traffic congestion is the increasing queue of vehicles on a road, characterized by the slow speed of the traffic flow, and set in when the number of vehicles on the road approaches the capacity of the road. When the vehicles are completely stopped for a certain amount of time, the situation is said to be a traffic jam. The problem of traffic congestion is faced in almost all the major cities due to increase in population and vehicles, and the limited availability of land for construction and expansion of roads. The traffic density, often taken as constant, may vary depending on the season, day of a week and even hour of a day. No traffic research has developed any means yet to accurately predict the condition under which a traffic jam may suddenly occur. The individual incidents such as accidents or sudden braking or slowing down amid smooth flowing traffic may cause a ripple effect leading to a sustained traffic jam [1].

Due to the shortage of residential areas in the main parts of the cities, people are often found living in one part of the city while working in the other. Mostly, the public vehicles are overcrowded and grimy inside, so in order to have a pleasurable commute from home to work and back, people prefer private vehicles. According to a 2011 estimate provided by the United States Census Bureau, a total of 132.3 million people commutes between their job and residential locations on a daily basis in the United States [2]. The problem of traffic congestion is severe in developing countries like Nepal, as the influx of people in the cities is alarmingly high. According to the population census of 2011, the total residential population of the Kathmandu valley is approximately 2.51 million [3], distributed in 899 km^2 area, and is estimated to be 3.68 million by 2021 with an annual population growth rate of 3.9% [4]. The public transportation of the valley is mostly owned by the private companies, 97% of the registered vehicles [5], and there is no fixed schedule for the public transportation. The vehicles are mostly random and uncertain with no guided access to the roads. This leads to severe traffic congestion to the roads leading to market areas. Furthermore, according to the Economic Survey 2015/16, the vehicles operating in the valley alone utilize 60% of imported petroleum [6]. It is estimated, from wasted time and additional fuel cost, that traffic jams in Kathmandu cost the country nearly 165.0 million US Dollar a year [7, 8]. This is a huge loss for a developing country like Nepal. Therefore, proper study of the traffic complexity of the valley is required for the improvement of efficiency and sustainability of the roads.

In this context, a detailed study of the various roads of the Kathmandu valley was drafted by the Japan International Cooperation Agency. The study encourages people to use public transit and explores the goal of creating a long-term urban transportation network with high mobility, safety, and comfort. The study also predicts that if the current situation, without making some urgent changes, is left on its own, 80% of the roads inside the Ring Road will be immensely congested by 2022 [9]. Similarly, the traffic composition of one day of the Maitighar–Tinkune road section of the Kathmandu valley was determined in their peak hour during morning and evening time [10]. The paper also discusses the loss due to the bottleneck scenario. In 2004, the Ministry of Works and Transport (Road Department) of Gaborone, Botswana published a book, the book facilitates the estimation of the future and present traffic demands for the need-based development of infrastructures by considering the future traffic situation. It also highlights the importance of collection and analysis of traffic data on decision making during planning, design, construction and maintenance period of the road network [11]. In the year 2006, Wach et al. determined the vehicles' velocities and collision locations by the means of the Monte Carlo Simulation method [12]. In 2009, Du et al. developed a model for traffic assignment in local area networks using the Monte Carlo Simulation by considering the turning probabilities at intersections [13]. In 2012, Enright et al. in their paper presented a detailed model of bridge loading for free-flowing traffic by using the Monte Carlo simulation. The results obtained from this model match the measurements of five European highways [14]. In 2013, Shengda et al. formulated a way for intelligent traffic lights based on the sequence of traffic flow in a single intersections exploiting complex mathematics and MATLAB [15]. In the same year, Canaud et al. used probability hypothesis density filtering method for the real-time prediction of traffic, and showed that the probability hypothesis filtering method and the particle filter method are equally efficient [16]. The traffic forecasting techniques that are used by the researchers to predict the short-term traffic have their own advantages and disadvantages. The traffic forecasting is done on the basis of past traffic data. In 2015, Mishra et al. used Monte Carlo Simulation to speculate the short-term traffic of Delhi-Gurgaon Expressway [17].

To the best of our knowledge, the implication of the Monte Carlo Simulation for the traffic prediction has not been performed on any road of Nepal till now. Therefore, the present work has a lot of scope to improve the existing transportation system of different roads of Nepal and serves as a basis for further study.

2. Methodology

2.1 Site selection and data collection

Data collection was done using Sony Handycam HDR-PJ10 placed over two locations, one on the footbridge in Jamal and another near the Kaiser Library on the section of the Kantipath road in Kathmandu, Nepal. The site location is shown in Fig. 1(a). The distance between these two places is 555 m. This section of the road is one of the busiest and a prime road leading to main market areas. The video graphic data of these two places were captured for the whole day from 8:00:00 to 19:00:00. The traffic distribution for every 15 minutes during the observation period is shown in Fig. 1(b). The peak point of traffic flow was found at 9:45:00 and at 18:30:00. The observed traffic volume ranged from 4,700 to 7,500 vehicles per hour. From the traffic composition of vehicles, the percentages of two-wheelers and buses were found to be greatest and least respectively as shown in Fig. 2(a). The contribution of public vehicles in the total vehicle flow was found to be almost constant and insignificant compared to private vehicles as shown in Fig. 2(b).

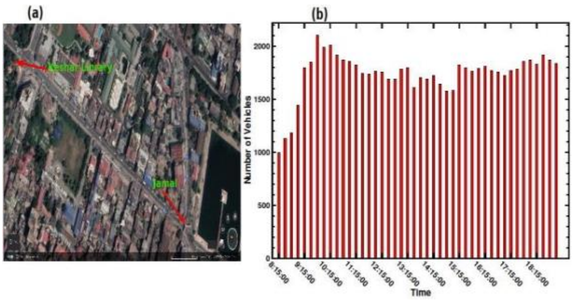

Fig. 1: (a) represents the site location. The red arrows indicate the spot where the video recorders were placed and (b) represents the number of vehicles passing through the section of the road on every 15 minutes interval (June 30, 2019).

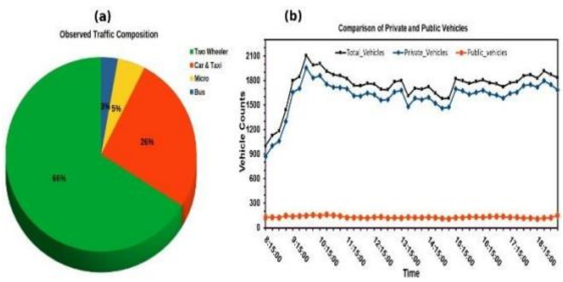

Fig. 2: (a) represents the percentage of observed traffic composition and (b) represents the comparison of private and public vehicles with total vehicles on the section of the road over time. The private vehicles completely dominate over public vehicles.

2.2. Kolmogorov–Smirnov test and Shapiro–Wilk test

A non-parametric Kolmogorov–Smirnov (KS) test is used to verify the goodness-of-fit and to determine whether an underlying probability distribution differs from a hypothesized distribution, or whether there is difference between two distributions [18]. This test is used when a comparison is to be made between observed sample distribution and theoretical distribution. The KS test is used to test whether the given sample has come from some specific distribution or not [19]. The test statistic is given by,

$$D = \text{Maximum } |F_0(X) - F_{data}(X)|$$

Which measures the largest distance between the empirical distribution function F_{data} (X) and the cumulative empirical distribution F_o (X). The null hypothesis is accepted if the test static D is less than the tabulated D value obtained from the KS table at a particular level of significance, otherwise an alternative hypothesis is accepted. The Shapiro–Wilk (SW) test is used to check if a random sample is from a normal distribution or not. Several studies have demonstrated that the KS test, even in its corrected form, is less effecting for checking normality than the SW test [20]. The test statistic is given by,

$$W = \frac{\sum_{t=2}^{n} a_t y_t}{\sum_{t=1}^{n} (x_t - \bar{y})}$$

Where n is number of observations, y_t is value of ordered sample obtained by ordering the disorder sample, x_t and a_t are tabulated coefficients. The null hypothesis is accepted if the test statistic W is greater than the tabulated W value obtained from the SW table at a particular level of significance. In this case, the sample belongs to a normal distribution. The sample does not belong to normal distribution if the alternative hypothesis is accepted.

2.3. Model development for Monte Carlo Simulation

Monte Carlo Simulation is a mathematical tool for simulating risk or uncertainty in a system by generating random variables. This method is widely used in computational physics, quantum mechanics, physical chemistry, and related applied fields, and they have a wide range of applications, ranging from complex quantum chromodynamics calculations to the design of materials that can protect experimentalists while working in radiation-hazardous areas [21-23]. Furthermore, it can also be applied in finance and business. However, in this study, the purpose of developing the Monte Carlo Simulation model is to predict the traffic flow. The section of the Kantipath road chosen for the study purpose is selected in such a way that the road is uniform and there is no on/off ramp or direct access from the adjoining areas. So that the total inflow and outflow of vehicles remain the same.

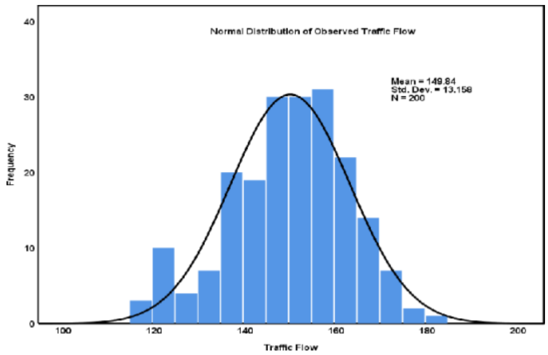

Fig. 3: The normal distribution of the observed data counted on intervals of 1 minute and 20 seconds.

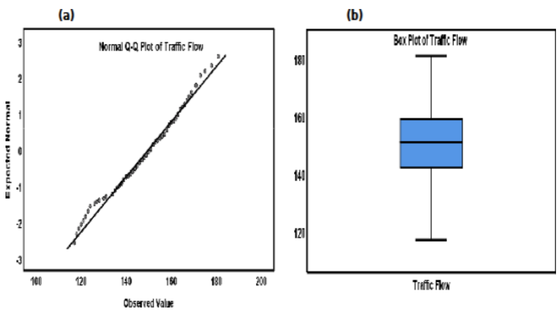

Fig. 4: (a) and (b) represents the normal Q-Q plot and box plot for the test of normal distribution of the sample of traffic data counted on the interval of 1 minute and 20 seconds.

A random sample–taken from the dataset collected at the crossroad by the Kaiser Library, counted for 1 minute and 20 seconds interval each–was analyzed and the normal distribution of the sample was tested employing the KS test and the SW test statistics. The mean and the standard deviation of the fitted normal distribution were used to generate a new dataset with random probabilities, generated by the Monte Carlo Simulation. The validation of the forecasted data was confirmed by comparing it with the dataset collected from the second location. Since the forecasted traffic data was randomly generated, the statistical accuracy of the model was checked with the help of statistical t-test at 5% level of significance and by evaluating the mean absolute percentage error. From the KS statistics, the D value is found to be 0.060 which is less than the critical value 0.074 for the normal distribution and from the SW test, the statistic W value is found to be 0.983 which is greater than the critical value 0.017 and is pretty much close to 1.00. Hence, from both the statistical tests, KS test and SW test, the sample was found to belong to the normal distribution. Moreover, from the visual inspection of histogram (Fig. 3), normal Q-Q plot (Fig. 4(a)) and box plot (Fig. 4(b)) showed that the traffic flow data were approximately normally distributed with a skewness of -0.361 (Standard Error (S. E.) = 0.172) and a kurtosis of -0.209 (S. E. = 0.342).

3. Results and discussion

3.1. Kolmogorov--Smirnov test

The traffic flow distribution obtained from every 15 minutes is compare with eight kinds of

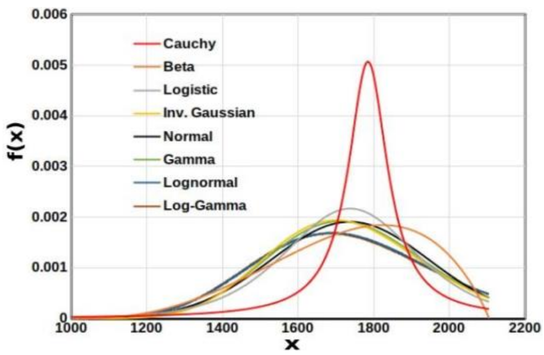

Fig. 5: The probability density functions corresponding to different distributions for the dataset counted for every 15 minutes.

distributions, viz., Cauchy, Normal, Log-Normal, Inv. Gaussian, Logistic, Gamma, Log-Gamma and Beta distributions using Kolmogoro–Smirnov test. The obtained plot for each of the distribution is shown in Fig. 5. The Cauchy distribution looks similar to normal distribution but with much heavier tails. The gamma and Log-Gamma distributions appear same as Normal and Log-Normal distributions respectively.

In KS test, if critical D value 0.2406 at 1% level of significance is higher than calculated D value, then the observed and theoretical distributions are same. From the KS test, it has been found that the traffic flow distribution can be considered as the Cauchy, Normal, Inv. Gaussian, Logistic, Gamma, and Beta distributions. The obtained results are summarized in Table 1.

Table 1: Kolmogorov–Smirnov test distribution result for the data counted for every 15 minutes.

Distribution	Maximum D Value	Remarks
Cauchy	0.08313	Same distribution
Normal	0.22480	Same distribution
Log-Normal	0.26007	Different distribution
Inv. Gaussian	0.22347	Same distribution
Logistic	0.21266	Same distribution
Gamma	0.23957	Same distribution
Log-Gamma	0.26315	Different distribution
Beta	0.21214	Same distribution

3.2 Monte Carlo Simulation

After performing about ten thousand of iterations using Monte Carlo Simulation, a new dataset was generated with random probabilities in accordance with normal distribution for each category of vehicles as shown in Fig. 6. The accuracy of the generated and the observed dataset was compared by computing the mean absolute percentage error (MAPE).

$$\text{MAPE} = \left(\frac{1}{N}\sum_{i=1}^{N}\frac{(x_i - x_i^F)}{x_i}\right) * 100\%$$

Where x_i = observed value of measure; x_i' = predicted value of measure and N = number of observations. The observed and predicted traffic flow of different vehicles during morning, afternoon, and evening hours along with their MAPE are presented in Table 2.

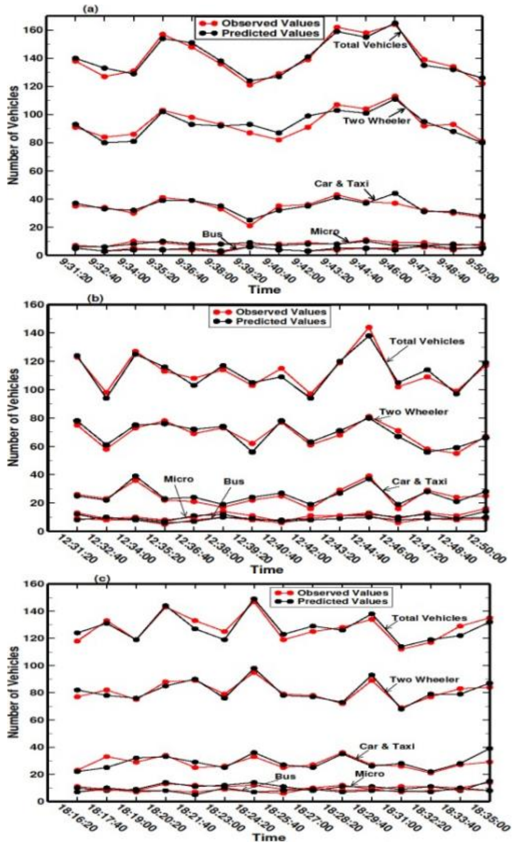

Fig. 6: (a), (b) and (c) represent the observed and predicted number of vehicles of different composition during the morning, afternoon and evening time respectively. The observed and predicted values are in well agreements.

Table 2: Comparison of observed and predicted vehicle flow per 15 minutes

MAPE of observed and predicted flows in the morning hour			
Vehicles	Observed Flow (Vehicles/15mins)	Predicted Flow (Vehicles/15mins)	MAPE (%)
Two Wheeler	1096	1083	1.186
Car & Taxi	458	449	1.965
Micro	92	97	5.435
Bus	50	49	2.000
Total	1696	1678	1.061
MAPE of observed and predicted flows in the afternoon hour			
Two Wheeler	1149	1138	0.957
Car & Taxi	436	452	3.670
Micro	77	80	3.896
Bus	47	49	4.255
Total	1709	1719	0.585
MAPE of observed and predicted flows in the evening hour			
Two Wheeler	1202	1186	1.331
Car & Taxi	480	493	2.708
Micro	79	87	5.063
Bus	50	47	6.000
Total	1811	1813	0.110

4. Conclusion

By analyzing the traffic data, it was found that more than 66% and 90% of the total vehicles running on that section of the road were two-wheelers and private vehicles respectively. The contribution of the public vehicles in the traffic composition on the section of the road seems to be almost constant and insignificant in comparison to private vehicle flows. This indicate the traffic congestion is mainly caused by the private vehicles. Therefore, in order to reduce the traffic congestion, the government of Nepal should increase the number of mass transit vehicles such as well facilitated buses. The traffic flow distribution of the study area, counted on every 15 minutes for the entire day of observation, is obtained to be same as Cauchy, Normal, Inverse Gaussian, Logistic, Gamma, and Beta distributions. From the Monte Carlo Simulation prediction study, following conclusions are drawn:

- The Monte Carlo Simulation turn out to be an excellent tool for the forecast of the real-time traffic flow. The forecasted data obtained by the Monte Carlo Simulation and the real-world observed data shows better agreement.
- This method can be used by local agency for the forecast of traffic flow, as spreadsheet tools and formulae are the basic tools used for the Monte Carlo Simulation.
- The forecasted value of traffic flow can be used for the traffic light design. The traffic light can be adjusted in a manner to reduce the traffic congestion or for the optimum use of the available road networks on the basis of real-time advance forecasted data provided by the use of the Monte Carlo Simulation.
- This method with certain improvements can be used to develop the Intelligent Transport System for the traffic control strategy.

References

[1] G. Orosz and G. Stépán, Subcritical Hopf bifurcations in a car-following model with reaction-time delay, Proc. Royal Soc. A 462 (2006) 2643–2670.

[2] U. S. Department of Commerce, Economics and Statistics Administration, Statistical Abstract of the United States, 131st ed. Maryland USA, U. S. Census Bureau (2012).

[3] CBS, National Population and Housing Census 2011, Central Bureau of Statistics, National Planning Commission Secretariat, Government of Nepal. Kathmandu, Nepal (2012).

[4] E. Muzzini and G. Aparicio, Urban Growth and Spatial Transition in Nepal: An Initial Assessment. Directions in Development, Washington, DC, World Bank (2013). doi:10.1596/978-0- 8213-9659-9. License: Creative Commons Attribution CC BY 3.0

[5] DOTM, Department of Transport Management, Ministry of Physical Infrastructure and Transport (20 July 2017). https://www.dotm.gov.np/en

[6] MOF, Economic Survey 2015/16, Ministry of Finance, Government of Nepal, Kathmandu, Nepal (2016). https://mof.gov.np/uploads/document/file/Final%20Economic%20Survey%202071-72%20English%20(Final)_20150716082638.pdf

[7] B. Budhathoki, Kathmandu should promote walking, cycling and public buses to combat congestion, pollution, Online Khabar (29 March 2021). https://english.onlinekhabar.com/kathmandu-should-promote-walking-cycling-and-public-buses-to-combat-congestion-pollution.html

[8] JICA, Japan International Cooperation Agency, Data Collection Survey on Traffic Improvement in Kathmandu Valley. Ministry of Physical Planning, Works and Transport Management Department of Roads (2012). https://openjicareport.jica.go.jp/pdf/12082459_01.pdf

[9] A. P. Timalsena, A. Marsani, and H. Tiwari, Impact of Traffic Bottleneck on Urban Road: A Case Study of Maitighar–Tinkune Road Section, Proc. IOE Graduate Conference Vol. 5, ISSN: 2350-8914 (Online) (2017) 2350-8906.

[10] B. K. Steinset, Traffic Data Collection and Analysis, Ministry of Works and Transport, Gaborone, Botswana (2004), ISBN 99912-0-417-2.

[11] W. Wach and J. Unarski, Determination of vehicle velocities and collision location by means of Monte Carlo simulation method, SAE Technical Paper (2006) 2006-01-0907. https://doi.org/10.4271/2006-01-0907.

[12] Y. Du, Y. Geng, and L. Sun, Simulation model based on Monte Carlo method for traffic assignment in local area road network, Front. Archit. Civ. Eng. China 3 (2) (2009) 195–203.
[13] B. Enright and E. J. O'Brien, Monte Carlo simulation of extreme traffic loading on short and medium span bridges, Structure and Infrastructure Engineering 9 (12) (2012)1267–1282.
[14] Z. Shengda, W. Lurong, J. Lin, and W. Bizhi, Study on Monte Carlo simulation of intelligent traffic lights based on fuzzy control theory, Sensors and Transducers 156 (9) (2013) 211–216.
[15] M. Canaud, L. Mihaylova, J. Sau, and N. Faouzi, Probability hypothesis density filtering for real-time traffic state estimation and prediction, Am. Inst. Math. Sci. 8 (2013) 825–842.
[16] R. Mishraa, A. P. Singh, A. Sharma, and A. K. Sarkar, Short term traffic prediction using Monte Carlo simulation, 3rd Conference of Transportation Research Group of India (3rd CTRG), 17–20 December (2015), Calcutta.
[17] Y. Dodge, The Concise Encyclopedia of Statistics, Springer (2008).
[18] I. M. Chakravarti, R. G. Laha, and J. Roy, Hand book of Methods of Applied Statistics, Vol. I. Hoboken: John Wiley and Sons (1967).
[19] M. A. Stephens, EDF statistics for goodness of fit and some comparisons, J. Am. Stat. Assoc. 69 (347) (1974) 730–737.
[20] D. W. O. Rogers, Fifty years of Monte Carlo simulations for medical physics, Phys. Med. Biol. 51(13) (2006) R287–R301.
[21] R. Hill, B. Healy, L. Holloway, Z. Kuncic, D. Thwaites, and C. Baldock, Advances in kilovoltage x-ray beam dosimetry, Phys. Med. Biol. 59 (6) (2014) R183–231.
[22] X. Jia, P. Ziegenhein, and S. B. Jiang, GPU-based high-performance computing for radiation therapy, Phys. Med. Biol. 59 (4) (2014) R151–R182.