INTERNATIONAL JOURNAL OF APPLIED SCIENCE AND RESEARCH
www.ijasr.org

2018 V0L.-1 ISSU.-2

A Bayesian approach to modelling manual traffic counts

Consul, Juliana Iworikumo
Department of Mathematics and Computer, Niger Delta University, Bayelsa State, Nigeria.
and
Okrinya, Aniayam
Department of Mathematics and Computer, Niger Delta University, Bayelsa State, Nigeria.

ABSTRACT
This research is concerned with the Bayesian modelling of the manual traffic count of vehicles passing through a particular road (Tombia - Amassoma Road) to the Niger Delta University, Bayelsa State of Nigeria. The time intervals between vehicles were recorded in seconds. A Bayesian approach to modelling the occurrence of timed events was used following a Poisson process. The probability density function and distribution function were evaluated using R codes. We investigate the behaviours and patterns of the traffic frequency.

KEYWORDS: Bayesian modelling, vehicle count, traffic, frequency, Poisson process.

INTRODUCTION

Traffic counts provide traffic data from a place to another. It is carried out to determine the volume of vehicles moving on the roads. The traffic count of a day or an hour could vary depending on the different day of the week or different time of a day. Dendrinos (1994) discussed that traffic counts at specific cross sections on urban roads are periodic and accounted through fast Fourier transform and power spectrum density functions of the additional variation beyond rough daily cycle. May (1990) also discussed that traffic flow rates varied with the time of the day and within the hour period.

Road is an important mode of travel for passengers. The flow of traffic or traffic count can be important to the economy of a country. The economy of a country depends on road infrastructure as a medium for conveying the flow of people. The traffic flow can be determined through the frequency of traffic count. It is required that the traffic flow is better predicted in order to avoid traffic jam, thereby taking into account some unforeseen events and react quickly to them. It is also possible to predict the observation of some number of vehicles in the next given time on the road.

In this research, we will use a Bayesian approach which is an explicit way to uncertainty and we combine our prior beliefs (which is in the form of prior probability) and the observed data (which is in the form of a likelihood) to get the form of a posterior distribution. The choice of prior distribution can therefore be important for obtaining useful posterior inferences. We will model the occurrence of timed events (arrival of bus and cars) following a Poisson process. Vayalamkuzhi et al (2016) used a statistical approach where Poisson and negative binomial regression were used to assess the safety performance as occurrence of vehicle crash frequency in India. Most of the studies like Vayalamkuzhi et al (2016), Solomon (1964), Miaou (1993) and Milton et al (1998) have considered modelling vehicle accident frequencies. Joshua et al (1990) used Linear and Poisson regression to model the relationship between highway geometrics and truck crashes in Virginia. Richards et al (2014) used a Bayesian approach to modelling multivariate zero-inflated Poisson model of vehicle crashes.

This research is aimed at modelling the frequency of traffic counts as playing an increasing vital role and a way of understanding journey time reliability. We can also estimate the road usage, trend of the road and possibly plan travel demand and understand the planning implication of traffic network on the road. We will also aim at predicting the observation of some number of vehicles in the next given time on the road.

In Section 2, we will discuss method of vehicle counting and the data collected from the case study. We discuss Bayes' theorem and Bayesian Inference in Section 3, modelling of the frequency of vehicle count data in Section 4 using a Poisson process and an application to the Niger Delta University in Section 5. We will conclude the research and make some recommendations in Section 7.

METHODS

In this section, the Bayesian approach using a poisson process to modelling frequency count is discussed.

Bayes' Theorem and Bayesian Inference

Bayes' theorem can be used to show the relationship between two conditional probabilities that are the reverse of each other. Bayes' rule is generally used to change our prior probability distribution which expresses our beliefs about parameters before we see the data to a posterior probability distribution which represents our beliefs about the parameters after we see the given data. In general, Bayes' theorem combines the prior experience (in the form of a prior probability) with observed data (in the form of a likelihood) to interpret these data (in the form of a posterior distribution) in a process known as Bayesian inference. Bayes' theorem is the probabilistic result which plays a central role in Bayesian inference. This is often expressed as

$$\text{Posterior} \propto \text{Prior} \times \text{Likelihood}$$

Bayesian Modelling of the Frequency of Vehicle Count Data

Suppose we have a prior probability density function for a vector $\underline{\theta}$ of parameters, $f_{\theta}(\theta)$ and the probability density function for a vector $\underline{X}$ of observations given θ, $f_{X|\theta}(\underline{x} \mid \underline{\theta})$ is the likelihood. We note that $f_{X|\theta}(\underline{x} \mid \underline{\theta})$ is a function of $\underline{\theta}$ once $\underline{x}$ is observed. Then, the posterior probability density function (pdf) is given by

$$f_{\theta|x}(\underline{\theta} \mid \underline{x}) = \frac{f_{\theta}(\theta) f_{X|\theta}(x|\theta)}{\int f_{\theta}(\theta) f_{X|\theta}(x|\theta) d\theta} \quad (1)$$

We write Equation 1 as

$$f_{\theta|x}(\underline{\theta} \mid \underline{x}) \propto f_{\theta}(\theta) f_{X|\theta}(y|\theta)$$

and we have

$f_{\theta|x}(\underline{\theta} \mid \underline{x}) \propto Prior \times Likelihood$

Again, we note that if the vector of observations, $\underline{Y}$ and (or) the vector of parameters θ is discrete rather than continuous, the probability density function can be replaced by the appropriate probabilities.

Poisson process

A Poisson process can be used in stochastic process for modelling the times at which arrivals enter a system. The Poisson process can be characterized by a sequence of geometrically distributed inter arrival times. We wish to model the occurrence of timed events (buses and cars) which follow a Poisson process with λ (average number per unit time). We suppose that our prior density for the rate of the Poisson process is a gamma (α, β) which is proportional to

$$\lambda^{\alpha-1} e^{-\beta\lambda}$$

We suppose that the events occur at times $t_1, t_2 \ldots\ldots t_n$ where $t_0 = 0$ and we observe the process for τ time units. Then, the likelihood is given by

$$L = \prod_{i=1}^{n} \lambda e^{-\lambda(t_i - t_{i-1})} e^{-\lambda(\tau - t_n)}$$

We can simplify L to

$$L = \lambda^n e^{-\lambda\tau}$$

The probability of n events in $(0, \tau)$ is given as

$$\frac{(\lambda\tau)^n e^{-\lambda\tau}}{n!}$$

$$\propto \lambda^n e^{-\lambda\tau}$$

Hence, the posterior probability density function is proportional to

$$\lambda^{\alpha+n-1} e^{-(\beta+\tau)\lambda} \qquad (2)$$

We have that Equation 2 is a Gamma distribution and so the posterior probability density function is

$$\frac{(\beta+\tau)^{\alpha+n} \lambda^{\alpha+n-1} e^{-(\beta+\tau)\lambda}}{\Gamma(\alpha+n)} \qquad (3)$$

We also have that the posterior mean is

$$\frac{(\alpha+n)}{(\beta+\tau)} \tag{4}$$

and the posterior variance is

$$\frac{(\alpha+n)}{(\beta+\tau)^2}$$

We note that we are not restricted to using a prior of the form

$$\lambda^{\alpha-1}e^{-\beta\lambda} \;.$$

We will again note that we have used this form of prior because it is the conjugate form of prior to the likelihood and it works out neatly. In cases where the form of prior cannot be represented by a conjugate prior, we may have to resort to numerical evaluation of the posterior distribution. Monte Carlo Integration can be used in complex model Berger (1985).

The posterior distribution can sometimes not be represented analytically in realistically complex problems because of the intractability of the normalising constant. This created an obstruction to the implementation of the Bayesian approach until the development of suitable numerical methods (Berger (1985)). Often, it is not feasible to draw independent samples from the posterior distribution since the posterior might not be in a standard form. Hence, sampling is done through a Markov chain which has the required distribution as its stationary distribution by using Markov chain Monte Carlo (MCMC) simulation (Gilks et al, (1996) and Gamerman (1997)). Markov chain Monte Carlo is a generalised and flexible way of simulating a sample from the joint posterior distribution of the unknown parameters. Each sample may depend on the previous one and the sequence of samples follows a Markov chain and so, the past states provide no information about the future state if the present is known.

RESULTS

In this section, we discuss the method of data collection, application to the research and results.

Data Collection

We might want to predict traffic volumes that can be expected on the road during specific period. The traffic variation could have a pattern. We could have hourly patterns of traffic flow which show peaks. The traffic flow may vary throughout the day and night. Robertson (1994) discussed that the size of the data collection depends on the length of the counting period, the type of count being performed, the number of lanes and the volume level of traffic.

Description of Case Study Area

The Niger Delta University was used as a case study in this research. It is situated in Wilberforce Island, Amassoma in Bayelsa State of Nigeria. The Niger Delta University was established in 2000 with its main campus at Wilberforce Island. The University is about 32 kilo meters from the capital city, Yenagoa. The student population was only 1039 when the University started academic activities in the 2001/2002 session (Wikipedia). The only means of transportation was through the water ways. The population increased to 4636 in 2003 and

10294 in 2006. Now, there is a significant increase in the population of students, the number of academic and non-academic staff. Some staff and students now travel from the capital city through the Tombia - Amassoma road to the University since the water ways are no longer being used except for fishing or farming purposes.

Method of Vehicle Counting

The manual counting method was used in this research. Here, we label and organise sheets. A location for observation was selected at Ogobiri Junction along Tombia - Amassoma Road. There was a clear view of traffic from the location and it was away from the edge of the road. The enumerators wore retro refractive dress for safety reasons. All observations were recorded on site.

There are some limitations to the method of counting used in this research. There was the possibility of data error since very few enumerators were used. Due to constraint in manpower, it was impossible to take records for all 24 hours of the day. There are also various factors that could affect the vehicle counting. These include weather conditions, type of road, the method of vehicle counting, traffic flow etc. Bad weather condition can affect the enumerators through interrupted counting process thereby causing data error. The accuracy of traffic counting depends on its duration and variation to traffic flow (Robertson (1994)). Pengjum et al (2012) assessed the accuracy of manual traffic counting by reporting the observation and quantification of errors. The data collection will be more appropriate if the purpose of counting is known.

Data

The following key steps were taken before using the manual vehicle counting. These steps include performing necessary office preparations, selecting proper observer location and labelling data sheets for recording observations.

The data cover the period 7 am to 7 pm, Monday, 23rd July to Sunday, 29th July, 2018 at Ogobiri junction along Tombia - Amassoma Road. The observers have recorded the frequencies of vehicles in opposite direction of travel. The weather conditions throughout the data collection period were ideal and no traffic accidents were recorded on the road throughout the time of the recordings. We have also avoided some variations in the frequency counts of the data collection by avoiding vehicle counting during public holidays, on days preceding public holidays and days with exceptional bad weather conditions. We have also classified cars and buses as passenger cars. Other vehicles such as trucks, wagons, pickup, and motorcycles were also recorded.

Table 1 shows the frequency of traffic count from 7 am to 12 noon on Monday 23rd July, 2018.

Table 1: Inter arrival times of arrival of vehicles along Ogobiri Junction

Time	Inter arrival times
7 am to 8 am	10, 100, 150, 200, 100, 200, 100, 200, 165, 200, 248, 600, 650, 300, 300, 10, 30, 20, 80, 55
8 am to 9 am	50, 100, 20, 10, 55, 100, 30, 30, 20, 100, 85, 50, 10, 20, 100, 100, 20, 40, 10, 10, 40, 40, 50, 50, 10, 10, 20, 10, 58, 20, 20, 10, 100, 50, 50, 40, 10, 50, 50, 15, 10, 25, 50, 20, 10, 20, 20, 20, 10, 20, 10, 15, 5, 10, 15, 15, 10, 15, 50, 40, 25, 25, 50, 25, 20, 15, 15, 25, 60, 180, 11, 50, 25, 25, 10, 20, 50, 10, 180, 10, 20, 30, 50, 40, 50, 20, 10, 40, 20, 10, 25, 20, 10, 10, 20, 10, 10, 15, 5, 20, 10, 40, 10, 20, 15, 25, 10, 5, 15, 10, 5, 5, 5, 5, 20, 5, 8, 12, 5, 10, 5, 5, 15, 5, 10, 5, 10, 3, 2, 5
9 am to 10 am	15, 10, 10, 10, 20, 5, 5, 10, 10, 8, 2, 15, 5, 10, 20, 50, 10, 15, 3, 2, 10, 20, 20, 50, 50, 25, 25, 10, 20, 50, 10, 5, 5, 30, 60, 100, 50, 50, 100, 20, 10, 10, 40, 10, 5, 5, 30, 50, 40, 10, 35, 50, 50, 20, 30, 50, 25, 25, 10, 10, 40, 20, 20, 20, 30, 20, 50, 50, 20, 10, 20, 25, 20, 5, 50, 50, 20, 20, 10, 50, 50, 50 25, 25, 20, 10, 20, 50, 20, 10, 20, 50, 50, 20, 10, 20, 100, 50, 50, 100, 200, 30, 200, 100, 100, 50, 50, 300, 100, 60, 75, 50, 30, 20, 20, 20, 60, 60, 75, 100, 80, 100
10 am to 11 am	200, 100, 160, 140, 250, 300, 150, 200, 100, 60, 100, 50, 50, 20, 80, 200, 100, 200, 20, 20, 60, 300, 200, 200, 150, 200
11 am to 12 noon	450, 300, 150, 165, 100, 300, 75, 100, 50, 50, 100, 50, 200, 200, 300, 200, 10, 20, 30, 158, 250, 300

Application: Niger Delta University Road traffic

We record the times of arrivals of motor vehicles passing Ogobiri junction through Tombia - Amassoma road going to the Niger Delta University from 7:00am till 12 noon on Monday 24th July, 2017. We convert the values to time intervals between vehicles in seconds and these are given in Table 1. The values are time interval (in seconds) between vehicles. The first value is the time till the first arrival.

Prior expectation

A Bayesian analysis requires the specification of prior information about the model parameters by expressing beliefs about the parameters in the form of a probability distribution before we look at the observations. The prior distribution should reflect information about the model parameters (James et al (2010)). It is required that reasonable prior specification be used.

We will suppose that our prior expectation for the rate of arrival was 1 vehicle every 20 seconds. That is $\lambda_0 = 0.05$. Again, we suppose that we were almost certain that the rate of arrival would be less than 1 vehicle every 10 seconds. Let us say that $Pr(\lambda < 0.1) = 0.99$. We use a conjugate (gamma) prior.

Then, the mean is

$$\frac{\alpha}{\beta} = 0.05$$

We will require that

$$\int_0^{0.1} f(\lambda)\, d\lambda = 0.99 \qquad (6)$$

where $f(\lambda)$ is the probability density function of a Gamma $(\alpha, \alpha/0.05)$ distribution.

We will evaluate this integral using R codes. We will first set up a vector of values for α and a corresponding vector β. Then, we can evaluate the integral Equation 6 for each pair of values.

We get approximately $\alpha = 8$, $\beta = 160$ with a probability of 0.99. We have that $\tau = 18000$ and $n = 319$. Thus, the posterior probability density function following Equation 2 and 3 is

$$(18000 + 160)^{8+319}\, \lambda^{8+319-1}\, e^{-(160+18000)\lambda}$$

$$= \frac{(18160)^{327} \lambda^{326} e^{-18160\lambda}}{\Gamma(327)}$$

$$= \frac{(18160)^{327} \lambda^{326} e^{-18160\lambda}}{327!}$$

The prior mean was 0.05. The posterior mean following Equation 4 is

$$\frac{8 + 319}{18000 + 160} = \frac{327}{18160} = 0.018.$$

The prior variance following Equation 5 is

$$\frac{327}{18160^2} = 0.000001.$$

The posterior standard deviation is 0.001.

We evaluate the posterior probability density function and distribution function using R codes. We might guess that the posterior distribution is likely to be ± 3 standard deviation of the posterior mean. We have mean ± 3 standard deviation and that is

$$0.018 \pm (0.001) \times 3$$

and we have [0.15, 0.021].

So, we create an array of λ values from 0.015 to 0.021 in steps of 0.0001. We plot the probability density function and distribution function in Figures 1 and 2 respectively.

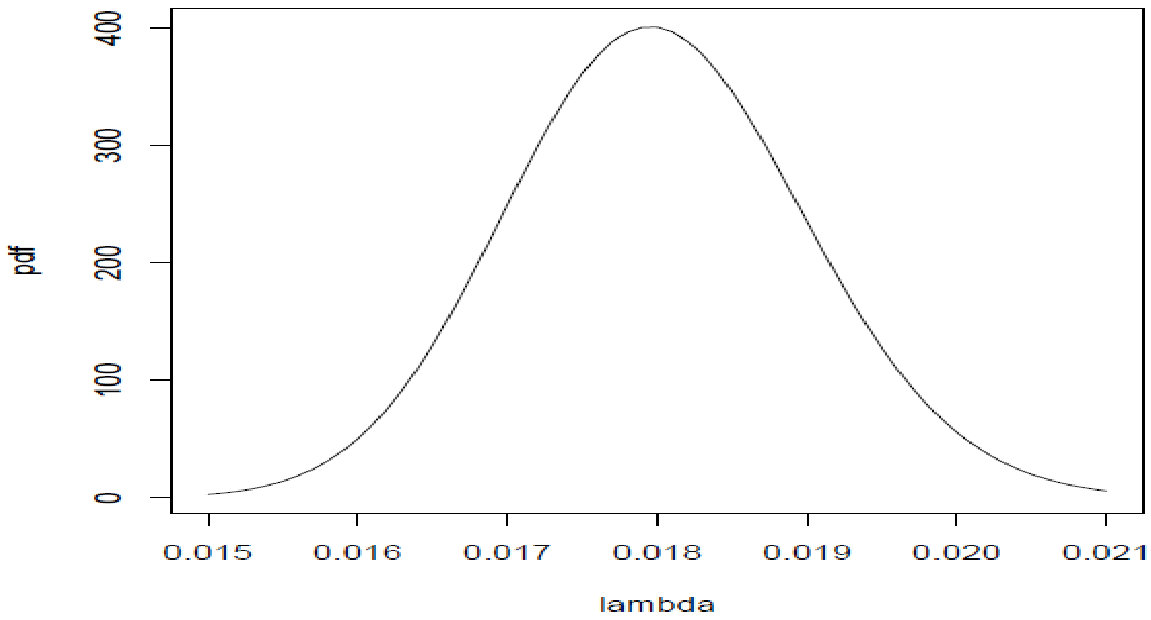

Figure 1: The probability density function of the inter arrival times of arrival of vehicles.

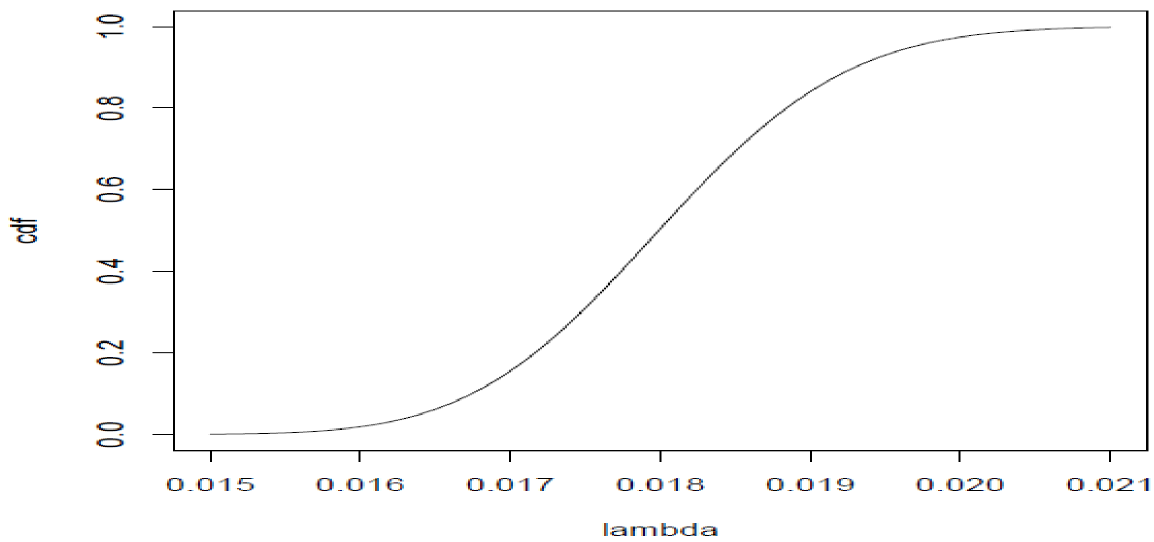

Figure 2: The distribution function of the inter arrival times of arrival of vehicles

We might want to use the distribution function to find the probability

$$Pr(0.017 < \lambda < 0.019) = 0.9 - 0.15 = 0.75.$$

We might want to know the posterior probability of observing j vehicles in the next t seconds. The joint probability density of λ and j is

$$\frac{(18160)^{327}}{326!} \lambda^{326} e^{-18160\lambda} \frac{\lambda^j t^j e^{-\lambda t}}{j!}$$

$$\frac{(18160)^{327}}{326!} \frac{t^j}{j!} \lambda^{326+j} e^{-(18160+t)\lambda}$$

We will integrate out λ and we get the marginal probability of j vehicles. We compare the function with a gamma probability density function and we integrate out that

$$\int_0^\infty \lambda^{326+j} e^{-(18160+t)\lambda}\, d\lambda$$

$$= \frac{\Gamma(327+j)}{(18160)^{326+j}}$$

$$= \frac{(327+j)!}{(18160)^{327+j}}$$

Hence, we have that the probability of observing j vehicles in the next t seconds is

$$\frac{(326+j)!}{326!\, j!} \frac{18160^{327} t^j}{(18160+t)^{327+j}}$$

$$= \binom{326+j}{j} \left(\frac{18160}{18160+t}\right)^{327} \left(\frac{t}{18160+t}\right)^{j}$$

This is a negative binomial distribution. We note that for large n and , this would be approximately a Poisson distribution with mean equal to t times the posterior mean for λ.

DISCUSSION

In this section, we discuss and summarise the results.

A summary of the frequency of vehicle count data

We recall that the data used for this research cover the period 7 am to 7 pm, Monday, 23rd July to Sunday, 29th July, 2018 at Ogobiri junction along Tombia - Amassoma Road. The observers have recorded the frequencies of vehicles in opposite direction of travel. Figure 3 shows the plot of the frequency of vehicle count from Tombia junction to Amassoma against time for the different days of the week. Figure 4 shows the plot of the frequency of vehicle count from Amassoma to Tombia junction against time for the different days of the week.

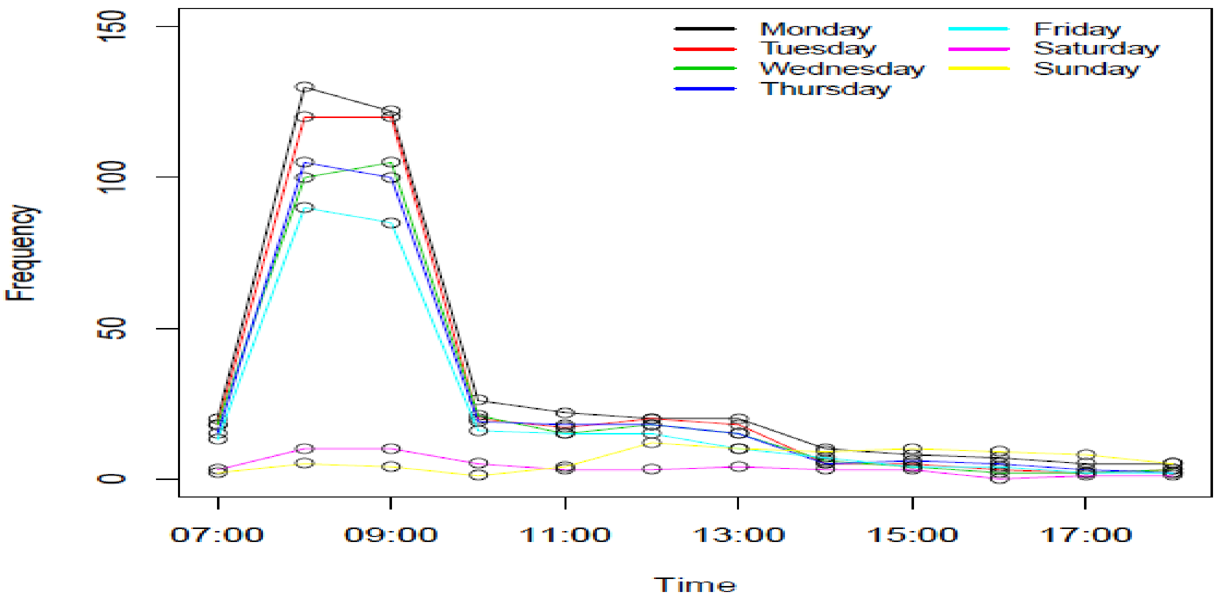

Figure 3: A plot of the frequency of vehicle count from Tombia junction to Amassoma against time for the different days of the week.

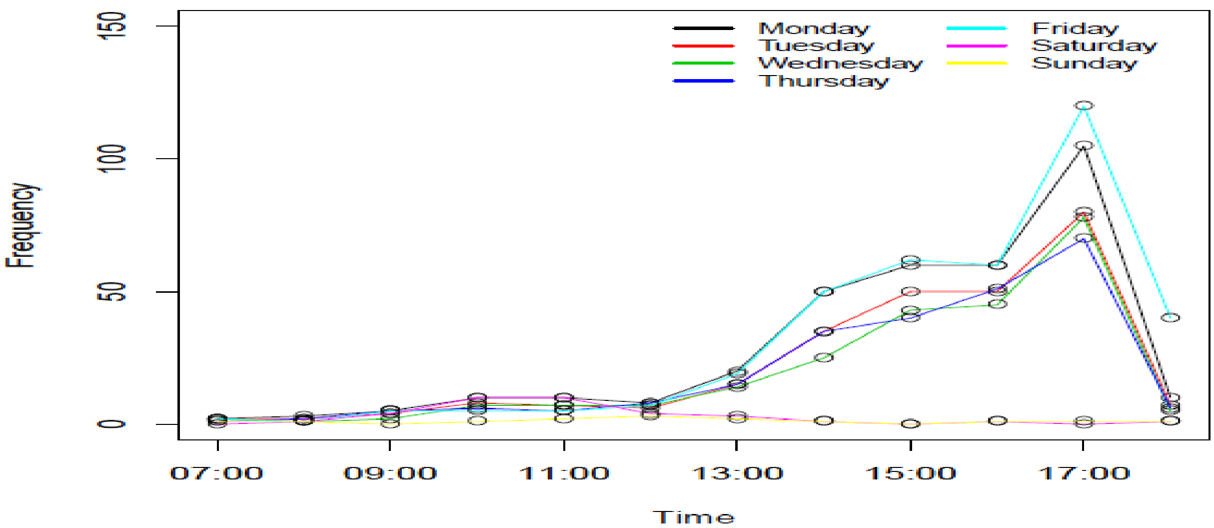

Figure 4: A plot of the frequency of vehicle count from Amassoma to Tombia junction against time for the different days of the week.

The frequency of the traffic count is influenced by the morning peak between 8.00am to 10.00am from Tombia junction to Amassoma. The possible behind the high traffic count frequency could be that most workers and students travel to Amassoma after the weekend. We will also observe that the peak in the morning is followed by a lean flow until night. The traffic during the working days (Monday to Friday) slightly varies but differ during the weekend.

The reverse is the case for the frequency of vehicle count in the opposite direction. The frequency of the traffic count from Amassoma to Tombia junction has a lean flow in the morning in all days. This pattern changes in the evening as workers and students leaving in the capital city (Yenagoa) return home. We observe the peak in the evenings of the working days (Monday to Friday) with highest on Friday since some people travel only at weekends. Again, this pattern varies from Saturday to Sunday.

CONCLUSION

In this research, we have looked into modelling of the frequency of traffic count passing through a particular road (Tombia - Amassoma Road) to the Niger Delta University, Bayelsa State of Nigeria using a Bayesian approach. The manual method of vehicle counting was used to record the time intervals between vehicles in seconds. We have used a Bayesian approach to modelling the occurrence of timed events following a Poisson process. We have evaluated the posterior probability density function and distribution function using R codes. We have also discussed the probability of observing number (say j) vehicles in the next few (say t) seconds.

In this research, the behaviours and patterns of the traffic frequency were investigated. We have observed peaks in the mornings which are followed by a lean flow until night for working days for the frequency of traffic from Tombia to Amassoma Road. This differed during weekends. We have observed a reverse case for the frequency of traffic from Amassoma to Tombia Road. The knowledge of how traffic count or flows could help in setting maintenance programs and evaluating congestion. It is required that we have a better prediction of the traffic count (flow) to avoid traffic jam which can be a cause of economic loss.

RECOMMENDATION

We recommend this research to transportation Engineers for the planning and designing of traffic facilities, forecasting the effect of projected scheme and diagnosing given conditions and finding solutions. The research could also be recommended to the Niger Delta University. The provision of parking spaces could be planned. The need for traffic control devices can be determined. Our finding on the modelling of frequency traffic count could be of help in understanding the decision to building highways for vehicles. It is also recommended that automatic counting method is used as it is not clumsy and based on CCTV video and has reference.

ACKNOWLEDGMENTS

We sincerely acknowledge our students who helped to collect the data used for application.

REFERENCES

[1] Berger J.O., Statistical Decision Theory and Bayesian Analysis (second Edition), Springer-Verlag, New York (1985).
[2] Dimitrios .S. Dendrinos Urban Traffic flows and Fourier transforms, Geographical Analysis, 26, 1994.
[3] Dong C., Clarke D.B and Richards S.H., Differences in passenger car and large truck involved crash frequencies at urban signalized intersections: an exploratory analysis. Accident Analysis & Prevention 62 (1), 87-94 (2014).
[4] Gamerman D., Markov chain Monte Carlo, Chapman & Hall, London (1997)
[5] Gilks W. R., Richardson S., and Spiegelhalter D. J., Markov chain Monte Carlo in practice, Chapman & Hall, London, (1996)
[6] Niger Delta University, "https://en.wikipedia.org/wiki/Niger Delta University
[7] James A. and Choy S. L and Mengersen K., Elicitator: An expert elicitation tool for regression in ecology, Environmental Modelling & Software, 25, (2010)
[8] Joshua, S., Garber, N., 1990. Estimating truck accident rate and involvement using linear and Poisson regression models. Transportation Planning and Technology 15 (1), 41-58 (1990).
[9] May A.D., Traffic flow fundamentals, Englewood Cliff's N.J: Prentice Hall (1990).
[10] Miaou, S.P., Lum, H., Modeling vehicle, accidents and highway geometric design relationships. Accident Analysis & Prevention 25 (6), 689-709 (1993).
[11] Milton, J., Mannering, F., The relationship among highway geometrics, traffic-related elements and motor-vehicle accident frequencies. Transportation 25 (4), 395-413 (1998).
[12] Pengjun Zheng and McDonad Mike, An Investigation on the Manual Traffic Count Accuracy,Procedia - Social and Behavioral Sciences, 43, 226-231, (2012)
[13] Praveen Vayalamkuzhi and Veeraragavan Amirthalingam, Influence of geometric design characteristics on safety under heterogeneous traffic flow, Journal of Traffic Transportation Engineering,3 (6): 559-570 (2016).
[14] Robertson, H. D. 1994. Volume Studies. In Manual of Transportation Engineering Studies, ed. H. D. Robertson, J. E. Hummer, and D. C. Nelson. Englewood Cliffs, N.J.: Prentice Hall, Inc., 6 31.
[15] Solomon, D., Accidents on Main Rural Highways Related to Speed, Driver, and Vehicle. Bureau of Public Roads, Washington DC (1964).