RESEARCH ARTICLE OPEN ACCESS

Comparative Analysis for Black Spot Identification Methods: Combined Methods in Producing Heat Map for Black Spots within Macarthur Highway, Jose Abad Santos Avenue, and Lazatin Boulevard in the City of San Fernando, Pampanga, Philippines

Henri Fred D. Caparas*, Mark Jordan M. Agbuya*, Angelie T. Antonio*, Jhon Lloyd B. Benitez*, Kathleen P. Butiu*, Jupiter D. Pallermo*, James G. Malig*, Rommel C. De Mesa**, John Vincent G. Tongol**

**Student, Department of Civil Engineering, Don Honorio Ventura State University, Bacolor, Pampanga, Philippines*

Email: fredcaparas@gmail.com

***Instructor, Department of Civil Engineering, Don Honorio Ventura State University, Bacolor, Pampanga, Philippines*

--************************----------------------------------

Abstract:

One of the leading causes of mortality in the world is road accidents. As a matter of fact, it is the 8th leading cause of death worldwide. High number of road accidents in a specific zone of a road or stretch is commonly term as black spot. In the Philippines, this is also an issue, specifically in its cities. The study area was the City of San Fernando, Pampanga of Philippines. In determining the black spot, an analysis was due, but there were a lot of methods used. The researchers conducted a comparative analysis between the models and created heat maps based on the objectives of the researchers. The researchers only selected a few of the models that were possible to perform within the given time frame and data availability. The classification of models was divided into two: Probability model and Severity model. Under Probability model, Poisson Regression and Grey-Verhulst were utilized. Under Severity model, Rate Quality Control and Empirical Bayesian were utilized. These were applied on some portions of the three stretches of roads of City of San Fernando, Pampanga, namely MacArthur Highway, Jose Abad Santos Avenue (JASA), and Lazatin Boulevard. To test the accuracy of probability model, Relative Error and Root Mean Square Error (RMSE) were used. The probability model with lower error was used to combine with severity model allowing the researchers to create two heat maps. It was concluded that in consideration of the data collected, Grey-Verhulst is accurate than Poisson Regression showing an error of 12.93% and 7.49, and 22.25% and 8.26, respectively, under Relative Error and RMSE, respectively. This showed that Grey-Verhulst showed a relatively low of error with a difference of 9.32% and 0.67 to Relative Error and RMSE, respectively of Poisson Regression. As a result, Grey-Verhulst was combined to Rate Quality Control and Empirical Bayesian, to accomplish and create the heat maps.

Keywords **—Black spot, Heat map, Transportation Engineering.**

--************************----------------------------------

I. INTRODUCTION

For the past years, road accidents have become a common scene for road users from narrow roads to main roads. Those unforeseen scenarios cause millions of people to be wounded, severely injured, or die in a worse state. With that, road accidents became one of the reasons why the mortality rate continues to grow globally. According to the World Health Organization [1, para.3], this has become the 8th leading cause of death worldwide. This distressing fact also affects geographic boundaries, economies, and society. This only reflects that road or transportation safety has not been given much attention.

The Philippines is facing an issue concerning road accidents, as the number of reported and unreported accidents increases each year. According to the data released by the Philippine Statistics Authority (PSA), the number of traffic accident deaths had increased by 39% within 10 years, from 7, 938 deaths in 2011 to 11, 096 deaths last 2021 involving mostly persons aged between 15 to 29 years old [1, para 4].

To address this issue, The Department of Transportation (DOTr) together with the World Health Organization (WHO), launched a ‘'Philippine Road Safety Action Plan 2023-2028' on May 31, 2023, in Quezon City. The said action plan is divided into five pillars: road safety management, safer roads, safer vehicles, safer road users, and post-crash response; its main goal is to create a safer road environment [1, para 7 to 13]. On the other hand, problems involving road accidents mostly occur on Highways, as a bigger volume of vehicles is concerned. This is partly because cars are used so frequently and have become so widely owned in recent years [2]. Nowadays, many people own motorcycles than ten years ago, but Metro Manila has seen the biggest growth. It is said that approximately 1 million automobiles were registered in 2000 and 3.2 million were registered in 2009, an increase of 320 percent over ten years [2]. With the increasing number of vehicles, accidents become apparent, with these past researchers identified different methods that are applicable in locating and identifying accident-prone areas or black spots within a certain road.

The Province of Pampanga, specifically the City of San Fernando is not exempted from the pressing issue regarding road accidents. Currently, the metro area population of San Fernando in 2023 is 344, 000 that increases around 1.78 percent from 2022. Due to the increase of population, the traffic volume also increases [3]. This is especially between MacArthur Highway, Jose Abad Santos Avenue, and Lazatin Boulevard which are the busiest highways within the locale as it accommodates many different vehicles particularly trucks carrying heavy loads. As emphasized by DPWH, the Annual Average Daily Traffic in JASA Road is 41, 454 in 2021 [4]. The City of San Fernando Police Station located at Barangay Maimpis, City of San Fernando, Pampanga released a list of reported road accidents containing a total number of Reckless Imprudence Resulting to Homicide, Physical Injury, and Damage to Property in the year 2016 to 2022. In the JASA road, there were 28 recorded Homicides, 1,067 for Physical Injury, and 131 Damaged to Property. In Lazatin Boulevard, 1 recorded Homicide, 265 for Physical Injury, and 58 reported Damaged to Property, while 107, 2,576, and 923 for MacArthur Highway respectively. Given the data, it only shows that the area selected by the researchers is accident prone. Human errors, speeding, obstruction of ambulant vendors within the roads, and disobeying basic traffic ordinances are just common causes of the said event.

The areas in consideration are located along the highway, wherein it accommodates a large volume of vehicles which lead to the area to be congested. Due to the congestion, road accidents became inevitable. In this context, the researchers wanted to pinpoint those spots that have the higher number of car accidents by utilizing comparative analysis for different black spot identification methods. This allowed the researchers to come up with a combined method for prediction model and severity index that were reliable and accurate which were used by the researchers as a basis in producing heat map. For this purpose, with the aid of a heat map, it was used as an illustration that showed those black spots within a road network, and it gave a clear visual of those hazardous spots which people can interpret easily and fast without further explanation.

1.2 Black Spots

Generally, the first and essential aspect of the accident mitigation process is identifying the specific location of any possible safety problem. The method used to identify whether a location has a safety problem differs depending on the location. Therefore, the most widely used method is based on the road accident history and is also known as the determination of Black Spot Locations [5].

In road safety management, a location where traffic accidents have historically occurred in large numbers is referred to as an accident black spot [6]. Typically, a black spot is a method that is designed to identify the prone areas on a road network in a particular stretch of road section where there has been a high accident rate or events that often result in fatalities or severe injuries. It is used to emphasize the significance of identifying regions and to draw attention to the need for changes in infrastructure, vehicle behavior, and road design to lessen the frequency of incidents in certain areas [7].

1.3 Models

Poisson Regression Model

This model is used to calculate the percentage variation in the risk of death associated with an increase in an interquartile of each pollutant. This model has only one dependent variable (predicted no. of events) and independent variables (or also known as predictors) [7]. It is more suitable than models with multiple linear regressions [8, vol. 11]. It is a generalized linear model used to represent count data, where the response variable is the number of events that occur in a certain period or a given area [7]. In addition, this model will not predict negative values and it is appropriate in counting rare events including road accidents, because of its simplicity and reduced number of parameters, the model is less likely to overfit in certain incidents and is also computationally efficient. On the contrary, the data counts are usually over-dispersed.

Grey-Verhulst Model

The Grey-Verhulst is used in forecasting the incidence of road accidents [9]. In the studies [10][11], the Grey-Verhulst were utilized to forecast their individual subject matter and explored a variety of applications in forecasting their subject, such as forecasting of CO2 emissions [10] and in application to Bitcoin behavior [11]. In China, the model was tested to forecast road accidents and it produced valuable reference for accurate prediction of road accidents [10]. Additionally, improved versions were made and tested, in which the findings showed promising results providing an accurate precision in predicting road accidents with low error [10] [12]. Likewise, simulation and prediction in the grey model is more effective than the traditional GM (1,1) model [13]. With a first-order difference equation, GM (1,1) is a single-variable prediction model that has only one dependent variable and no independent variables [14]. However, two critical problems were found on its function, namely (1) parameter dislocation substitution and (2) unreasonable selection of initial value [13].

Empirical Bayesian Model

The Empirical Bayesian Method combines the two parameters, the observed and the predicted accidents in specific roadways into one statistical model using an equation [15]. Using the empirical Bayes technique, the most accurate estimate of the anticipated number of accidents can be calculated by adding the number of reported accidents to the average number of accidents as indicated by the methodology for predicting accidents [16]. When estimating safety, the Empirical Bayes approach improves estimation precision and accounts for regression-to-mean bias. It is based on the understanding that an accident rate is not the only indicator of the safety of the entity. What is known about the security of similar entities provides another hint. How much weight is assigned to accidents that are likely to occur on similar entities determines the outcome. The power of the EB method lies in its use of a weight that is grounded in reasonable reasoning and actual evidence. This weight will be determined by how strong the accident record is (i.e., how many accidents are expected) and how reliable the SPF is (i.e., how much a particular site's safety may deviate from the average that the SPF represents) [17]. On the other hand, using a Bayesian network to determine the variables influencing the degree of injuries, which were divided into two categories: mildly injured and killed/severely injured. Based on the Bayesian network, the accident type, driver age, lighting, and number of injuries were found to be the factors associated with a fatal or severely injured accident [18]. Furthermore, it can produce accurate and stable estimations, particularly when there is data for each stretch, since it can draw strength from other data points or groups [19]. On the other hand, the challenges of this method is its flexibility and complexity that can make a greater possibility of misconceptions and miscalculations [20].

Rate-Quality-Control Model

When comparing the actual crash rate of the treated location to the average crash rate of similar locations, the rate quality control method makes use of statistical concepts to determine whether the actual crash rate is abnormally high. The treated site is classified as a hazardous location if the actual (observed) crash rate is higher than the critical rate [21]. According to Sardar [22], for every road section, three distinct parameters are calculated using the Rate-Quality-Control Method; these variables are accident rate, accident frequency, and severity index. This method's advantages include cost-effectiveness to apply and ease of use, and as well as it can serve as a guidance to early identify the defects in the production process [23]. However, the disadvantage of this approach is that it only focuses on the quantity of accidents in each segment of road [24].

1.4 Tools

1. RStudio

RStudio is a tool utilizing R, a free open-source statistical programming language, flexible and more interactive for researchers involving statistical analysis [25]. This is the most used tool aside from the known tools such as Python, MATLAB, STATA, and EXCEL. It is a powerful tool for statistical studies and research as RStudio is an integrated development environment (IDE), a software application that helps to code efficiently. It is widely used in the field of research for summarizing, exploring, and visualizing quantitative data [25][26]. In [27], the researchers introduced an introductory class about statistics using R and RStudio in their medical classes, and they were able to find out that there is a positive response from the students. The only restrictions that are present in using R and RStudio, is the mastery of the user of the said tool and programming language.

2. QGis

Information can be efficiently expressed to a range of individuals through visual presentations. Mapping is one of the most common applications of visual information. There are several mapping programs available, but they are usually overpriced and have a challenging learning curve. A free, open-source program called QGIS (Quantum Geographic Information System) enables users to create, edit, view, analyze, and display geospatial data [28]. According to Ovramenko [29], QGIS provides a user-friendly interface for beginners to generate, modify, and display geographic data. The user will be able to compute line lengths and statistics, and terrain data using fundamental operations. Even more advanced features, like 3D visualization, geo-statistics, and remote sensing analysis, are available in QGIS for experienced users. Users may examine and visualize complex spatial relationships and patterns due to these capabilities. Users can create maps and automate complicated operations.

1.5 Objectives

This study conducts a comparative analysis of different Black Spot Identification (BSI) methods: Poisson Regression Model and Grey-Verhulst Model for the Probability of Road Crash Occurrence, while the Empirical Bayesian Model and Rate-Quality-Control Model for Severity Index, for the researchers to produce a reliable combined method that is accurate in identifying black spots around MacArthur Highway, Jose Abad Santos Avenue, and Lazatin Boulevard of City of San Fernando, Pampanga.

Specific Objectives

The specific objectives of the study are:

1. To determine the black spots within MacArthur Highway, Jose Abad Santos Avenue, and Lazatin Boulevard of City of San Fernando, Pampanga, by utilizing black spot analysis through different methods, namely:

Probability Test
-Poisson Regression Model
-Grey-Verhulst Model

Severity Index
-Empirical Bayesian Model
-Rate-Quality-Control Model

2. To conduct a comparison between these models (under probability test) and identify the accuracy rate (applying relative error and root mean square error ratio) of the stated black spot identification methods for the researchers to recommend the most reliable and applicable method for studying present and future traffic crash accidents.

3. To provide heat mapping for black spots within MacArthur Highway, Jose Abad Santos Avenue, and Lazatin Boulevard of City of San Fernando, Pampanga, that will serve as one of the warning signs for drivers who will cross these roads.

1.6 Conceptual Framework

For the study's implementation, a particular framework model was utilized, aiding in its execution and analysis. The conceptual framework as shown in figure 1 represents how the study was constructed and realized by utilizing the Input-Process-Output (IPO) model.

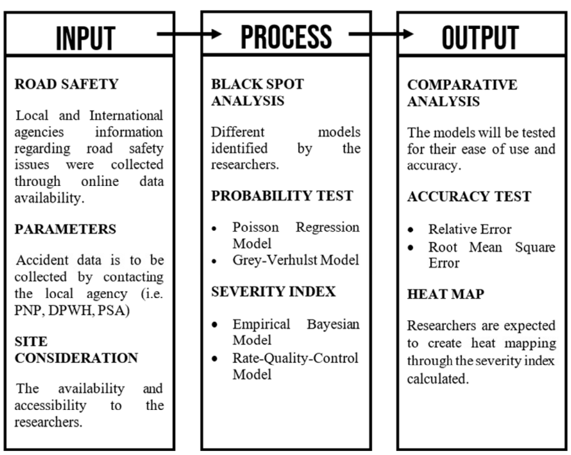

Figure 1. *Conceptual Framework.*

II. Methodology

The following figure is the methodological framework of the study as shown in figure 2.

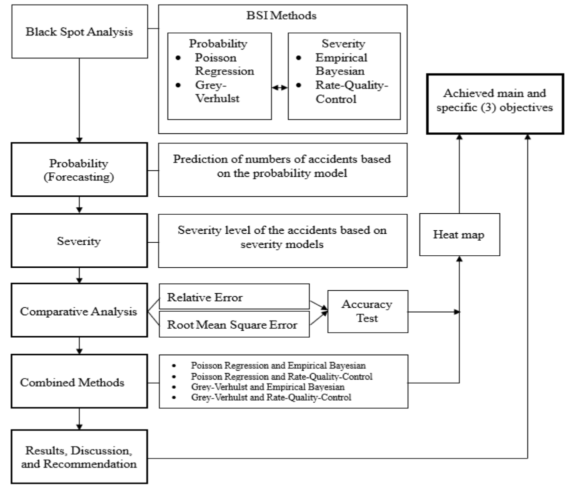

Figure 2. *Methodological Framework*

2.2 Research Procedure

The research procedure of the study was subdivided into 4 stages from performing the models according to their category up to data analysis, results, and discussion. This entails the entire flow of the study, on how the study was conducted in terms of research design and approach. The stages are as follows:

Stage 1

In Stage 1, the researchers performed the models according to their category. In this stage, the researchers used RStudio and R to perform the models under the probability: Poisson regression model and Grey-Verhulst model. Under this category, the results are the predicted numbers of accidents which were accomplished using the gathered data.

POISSON REGRESSION

The Poisson Regression model assumes that the conditional mean of the response variable follows a Poisson Distribution:

$$P(x = k) = \frac{e_{\lambda}^{-\lambda}k}{k!} \quad (Eq.1)$$

k is the number of events occurring in the interval,
λ is the average rate of events occurring in the interval,
e is Euler's number, approximately equal to 2.71828.

And it is specified using a linear predictor function. The model equation can be represented as:

$$log\ log(\mu) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k \quad (Eq.2)$$

μ is the mean of the Poisson distribution (the expected count of events)
$\beta 0$, $\beta 1, \ldots, \beta k$ are the coefficients associated with the independent variables,
$X1, X2, \ldots, Xk$ are the predictor or independent variables.

GREY VERHULST

Grey Verhulst model was utilized using RStudio, other researchers shared a code for future researchers a sample of the code that can be used to run a Shiny App. A Shiny app is a web application framework for R, a programming language and environment primarily used for statistical computing and graphics. Shiny allows R users to create interactive web applications directly from R code. It simplifies the process of building web applications by providing a set of functions and tools for creating user interfaces (UI) and connecting them to R code running on a server.

Stage 2

In stage 2, the researchers measure the severity of the predicted numbers of accidents in Stage 1. In this stage the researchers performed the models: Empirical Bayesian and Rate-Quality-Control Method by utilizing MS Excel. After finding the severity of each accident, the researchers will be able to create heat maps but an accuracy test must be conducted first (Stage 3). Thus, allowing the researchers to performed a comparative analysis to know the reliable combined methods in identifying blackspots on Stage 4. The following are the severity models:

RATE QUALITY CONTROL METHOD

The rate quality control method is a statistical method use for identifying black spots. It consists of evaluating three parameters for each road section. The three parameters are:

- Accident Rate
- Accident Frequency
- Severity Index

A critical value was used to compare these parameters. Consequently, if the section where all these 3 indexes are greater than the critical values for each one respectively will be considered a blackspot [30].

Accident Rate, (Cr)

$$Cr = \frac{(A)(1{,}000{,}000)}{(AADT)(365)(L)}) \qquad (Eq.3)$$

Wherein:
Cr = Accident rate of the section in accidents per million vehicle miles of travel,
A = Number of accidents for the study period,
AADT = Average Annual Daily Traffic (AADT) during the study period,
L = Length of the section (in miles).

Accident Frequency, (Ac)

$$Fave = \frac{total\ number\ of\ accidents}{Length\ of\ study\ area}) \quad (Eq.4)$$

$$Ac = F_{ave} + k_a\sqrt{\frac{F_{ave}}{L_j}} - \frac{0.5}{L_j} \qquad (Eq.5)$$

Wherein:
Ac = the critical value for accident frequency (number of accidents)
Fave = the average accident frequency for all road sections
Lj = the length of the road section.
ka = probability factor determined by the level of statistical significance desired for the equation.

$$\alpha = 1\%\ gives\ k\alpha = 2.567$$

$$\alpha = 5\%\ gives\ k\alpha = 1.645$$

$$\alpha = 10\%\ gives\ k\alpha = 1.282$$

kα=1.282 is recommended so that it is better to include more false black spots and thereby get more real black spots.

According to the Black Spot Manual, the RQC Method should be used to determine 5% as the confidence level, though this is up for discussion. To better balance the risk of including false black spots while identifying more genuine ones, some experts advocated for a higher confidence level, such as 10%. Consequently, using a 10% confidence level rather than a 5% level was preferred [31].

Severity Index, (SI)

$$SI = \frac{If\ (9) + Ib\ (3) + Id\ (1)}{A}) \quad (Eq.6)$$

Wherein:
Qj = relative severity value
If = number of fatalities
Ib = number of injured persons
Id = number of damaged to property
A = Number of accidents for the study period

Relative Severity Value, (Qj)

$$Q_j = \frac{SI}{A} \quad (Eq.7)$$

Average Severity Value, (Q_{ave})

$$Q_{ave} = \frac{\sum_{i=1}^{n} SI}{\sum_{i=1}^{n} A} \quad (Eq.8)$$

$$\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Qj - Q_{ave})2 \qquad (Eq.9)$$

Critical Value, (Qc)

The road section is considered to be black spot from severity point of view, if Qj > Qc. Where critical severity value:

$$Q_c = Q_{ave} + k_a\sqrt{\sigma^2} - 0.5 \quad (Eq.12)$$

In order to interpret the results by ranking, researchers generated a table of severity and weights of the respective rank as shown in the table1:

Table 1. *Severity ranks and weights of Rate Quality Control*

RANK	I	II	III	IV	V	VI
VALUES	10	9	8	7	6	5

EMPIRICAL BAYESIAN METHOD

Empirical Bayesian method is a statistical procedure in which the given set of data are used to estimate the prior probability distribution.

$$Y(x) = Y(r) \cdot w + x \cdot (1 - w) \quad (Eq. 13)$$

$$w = (1 + \frac{Y(r)}{k})^{-1} \quad (Eq. 14)$$

$$k = \frac{0.236}{L} \quad (Eq. 15)$$

Wherein;
$Y(x)$ is the expected number of accidents
$Y(r)$ is the average of the predicted number of accidents on the same location
x is the actual or observed number of accidents
w is the weight (formula based)
k is the overdispersion parameter
L is the length of roadway segment in miles

And for identification of severity of the accident, it is define using the SI equation for accident black spot. The SI ratio can be represented as:

$$SI = \frac{PSI}{\sqrt{\sigma_{r|x}^2 + \sigma_r^2}} \quad (Eq. 16)$$

$$PSI = Y(x) - Y(r) \quad (Eq. 17)$$

$$\sigma_{r|x}^2 = (1 - w) \cdot Y(x) \quad (Eq. 18)$$

$$\sigma_r^2 = \frac{Y^2(r)}{kn_0} \quad (Eq. 19)$$

Where $Y(x)$ and $Y(r)$ have the same definition as above. $\sigma_{r|x}^2$ is the variance of $Y(x)$, σ_r^2 is the variance of $Y(r)$, and n_0 is the observed sample size of the model.

To simulate the model, the researchers used statistical tool such as Microsoft Excel respectively to process the given set of data effectively and appropriately. After the simulation of the model, the table below was used as the assessment classification level of SI. It was used to show the level of severity of the accidents and identify the black spots in a road section [32].

Table 2.*Classification of blackspot of Empirical Bayesian*

Assessment Level	Non-Black Spots	I	II	III	IV	V
Two-level	[-1,0]	(0, 0.5]	(0.5, 1.0]	-	-	-
Three-level	[-1,0]	(0, 0.33]	(0.33, 0.66]	(0.66, 1.0]	-	-
Four-level	[-1,0]	(0, 0.25]	(0.25, 0.5]	(0.5, 0.75]	(0.75, 1.0]	-
Five-level	[-1,0]	(0, 0.2]	(0.2, 0.4]	(0.4, 0.6]	(0.6, 0.8]	(0.8, 1.0]

Stage 3

In stage 3, the researchers proceed on data analysis with regards to the accuracy of the models and conducted comparative analysis which is further discussed in the data analysis section of this study. Under this stage, the researchers performed the following accuracy test. Below are the following methods for statistical testing that were utilized:

Probability:

The researcher's statistical analysis: Relative Error and Root Mean Squared Error (RMSE).

Stage 4

In this stage, researchers were able to find out which probability model will be used to combined to severity models. As a result, the researchers created two heat maps based on the results from the models. Additionally, 2029 predicted heat map was produced. Furthermore, the researchers systematically draw conclusions which are based on the findings of the study. Additionally, in this stage the researchers did an analytical discussion to fully understand the details of the results. As a result, the researchers drafted several recommendations with accordance to the findings of the study

2.3 Accuracy Test

Probability models were tested under the following accuracy test:

1. Relative Error

$$Relative\ Error\% = \left|\frac{Actual - Predicted}{Actual}\right| \times 100 \quad (Eq. 18)$$

2. Root Mean Squared Error

$$RMSE = \sqrt{\frac{\sum(Y_n - \hat{Y}_n)^2}{n}} \quad (Eq. 19)$$

Where:

RMSE – Root Mean Squared Error
n = number of data points
Y_n = observed value
$\hat{Y}_n$ = predicted value

III. Results and Discussion

3.1 Accuracy Test

Probability

The test utilized and employed is consist of relative error (equation 20) and root mean squared error (equation 21), the results are as follows:

Table 3. *Poisson Regression Accuracy Test*

ROAD	AVERAGE RELATIVE ERROR	ROOT MEAN SQUARE ERROR
DL - JS	11.08 %	8.81
ML - JS	23.02 %	2.59
DL - LZTN	9.35 %	5.21
SJ - LZTN	59.88 %	10.57
DL - MAH	11.57%	10.63
SA - MAH	18.62%	11.12
Avg.	22.25%	8.16

In accordance with table 3, it shows the obtained average relative error and root mean square error along the barangays covered by the JASA, Lazatin and McArthur Highway Road. The yielded value of average relative errors is relatively low which implies that the Poisson regression model is considered effective and a good fit for the given data set. Same way as the value of root means square errors which considered as low value, meaning Poisson regression model accommodates accurate simulation. To come to the point, the said model produces an accurate procedure in a given data set. In line with that, predicting the number of accidents in a given stretch is nearly accurate.

Table 4.*Grey-Verhulst Accuracy Test*

ROAD	AVERAGE RELATIVE ERROR	ROOT MEAN SQUARE ERROR
DL - JS	9.82 %	8.89
ML - JS	18.32%	2.39
DL - LZTN	11.97%	6.88
SJ - LZTN	17.59%	7.7
DL - MAH	9.85%	11.66
SA - MAH	10.02%	7.44
Avg.	12.93%	7.49

Based on table 4, the simulation of the number of accidents along the barangays DOLORES-JS, MAGLIMAN, DOLORES-LZTN, SAN JUAN, DOLORES-MAH and SAN AGUSTIN covered by the JASA, LAZATIN and MAC-Arthur Highway Road Section yielded an average relative error of 9.82%, 18.32%, 11.97%, 17.59%, 9.85% and 10.02%, respectively. The value of relative errors being reasonably low only implies that the Grey-Verhulst model is considered to be efficient and well-suited to the current data set. Similarly, the root means square error values exhibit low figures that are nearly close to zero, indicating that the Grey-Verhulst model provides precise forecasting. In summary, the Grey-Verhulst model produces a necessary fit for the given data sets. Overall precision and prediction of frequency of accidents in the given road stretch is nearly accurate.

3.2 Heat Map

Current and Past Data

Since Grey-Verhulst yielded low errors, in order to create the heat maps, the combination of Grey-Verhulst to Empirical Bayesian and Rate Quality Control was made. The following heat maps are as follows:

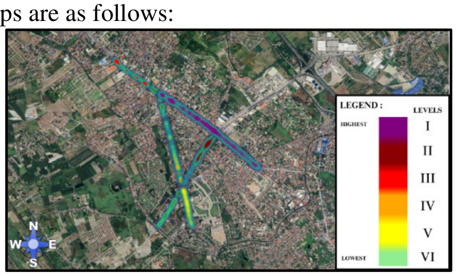

Figure 3. *Heat Map of Grey-Verhulst and Rate Quality Control*

The heat map in figure 3 displays six severity levels, ranging from Level 1 in light green to Level 6 in dark purple. Among the calculated findings, the road in Dolores - MAH exhibits the lowest severity level, followed by Dolores - JASA at the second level, San Agustin - MAH at the third, Dolores - LZTN at the fourth, and San Juan - LZTN at the fifth. The researchers considered by ranking the whole span of years to ensure a thorough assessment. On the other hand, the most severe road condition is observed in Magliman - JASA. These data points have been ranked according to the average of the calculated values, providing a comprehensive overview of road severity across different areas.

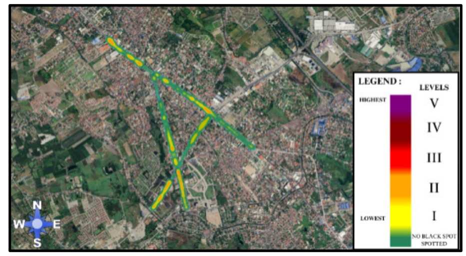

Figure 4. *Heat Map of Grey-Verhulst and Empirical Bayesian*

After combining the Empirical Bayesian and Grey-Verhulst model, the results are tabulated as shown in the figure x. To visualize the tabulated results, the figure 24 provides a heat map of road accidents in 3 different roads in the year 2019. The hotspot for road accidents is represented by warm colors (i.e., yellow, orange and red). In order to acknowledge the model, the year 2019 is the best fit for creating heat map since it recorded the highest number of accidents indicating the black spot in a designated road section.

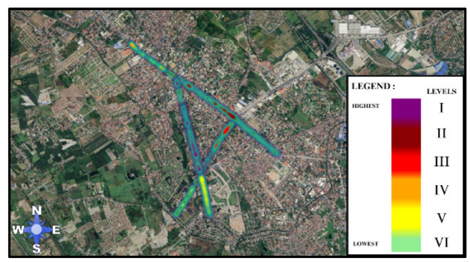

Figure 5. *Heat Map of Grey-Verhulst and Rate Quality Control (2029)*

The figure 5 is the prediction of the number of accidents by year 2029. The models that were utilized here were Grey-Verhulst model and Rate-Quality-Control model.Among the calculated findings, the road in Magliman - JASA still exhibits as lowest severity level while the other Barangays, San Juan - LZTN at the second level, San Agustin - MAH at the third, Dolores - JASA at the fourth, and Dolores - MAH, the most predicted severe road condition is observed in DOLORES - LZTN. This figure indicates the predicted Levels for each Barangays in a stretch that will show in the next 5 years.

IV. CONCLUSION AND RECOMMENDATION

4.1 Conclusions

In collecting data regarding the required specifics, the researchers have only collected the number of accidents and their location (from 2016 to 2022). The number of accidents for each stretch (with a specified length within JASA, LAZATIN boulevard, and Mac Arthur Highway) that were considered were alarming as they are large in numbers, specifically in year 2019. The researchers also tried to collect the specific location of the accident but with no avail as they have only collected the broader area where it happens. In performing the probability models, Poisson Regression perform under lack of variables which resulted on causing minimal errors. On the other hand, Grey Verhulst performs well because it is its strength, being able to perform well relying solely on numbers of accidents. Their errors under relative error and root mean square error are both averaging 22.25% and 8.16, and 12.93% and 7.49, respectively. This shows that both are able to accurately predict the number of accidents but with relatively low difference on their errors. In this case, Grey Verhulst showed a much lower error with a difference of 9.32% in relative error and 0.67 in root mean square error (the values were computed as the difference of their average).

Under the circumstance of lack of variables (predictors that are possible causing road accidents) Grey Verhulst works well in predicting numbers of accidents where data are scarce. With this, researchers used this model to partner it with Empirical Bayesian and Rate Quality Control. The results of the combination (Table 4, 5, 6, and 7) were detailed and made to be easy to comprehend if anyone would look it in that manner. In order to appreciate the results of the combinations, heat maps were produced between these two combinations, showcasing (figure 23 and 24) their difference in classifying black spot. The heat maps are different based on what severity model was used, by utilizing table 1 and 2, the blackspot can be classified. In summary, it is up to individuals or organizations to determine which severity models best suit their situations on investigating particular area.

4.2 Recommendation

The following are the recommendations of the researchers based on findings and conclusions:

1. For organizations (i.e., local government units, agencies, and non-government organizations) planning to adopt the models, it is recommended to conduct a similar study in order to test whether researchers were able to execute properly the models and apply the existing accident data in investigating whether the methods were reliable for locating blackspot.
2. In addition, it is suggested for organizations to take considerations about the road users, as these individuals would benefit from the organizations involved when adopting the study findings and outputs was in success. These will greatly affect the possibility of achieving the researcher's objective which is to produce a heat map in order to provide awareness to all road users so they can easily find where the black spots are located.
3. For reliability, the organizations are advised to perform further optimizations and physical implementations of the heat map to test its efficacy, as the analysis was done in a short period of time while dealing with broad subject matter about the different methods. Additionally, there might be cases where the researchers overlooked those methods, while on the process of executing a model.
4. For future researchers, it is recommended to take note that the models, especially the probability models identified by the researchers, ran on a scenario on only one series of data for considered locations. For that reason, it is advised to conduct research in their efficiency in a new set of series of data. Additionally, since the models researched here are limited it is recommended to consider also other statistical models that are not utilized by the researchers to further enhance the quality of the results concluded in this study, or the other way, future researchers can focus on the combined model provided by the researchers and select multiple study area to strengthen the current objectives.
5. For future researchers who are considering the study as their basis and guide, they must take note that in selecting study area and location for the research focus, they must consider two (2) things. (1) Future researchers are recommended to consider locations that has recorded data about accidents, especially that are predominantly high. Thus, enabling the future researchers to test (performing the models) whether a certain strip within a stretch can be classified either non-black spot or black spot. (2) In collecting data, future researchers are advised, if it possible, to collect data that are detailed, such as exact location (not just the barangay but by kilometer post, or any landmarks), time of occurrence, weather condition, and road condition. Thus, it will help future researchers to strengthen the reliability of Poisson Regression Model, as the variable using this model increases the accuracy of its prediction. Take note that each variable must be tested using Pearson's R correlation or any other statistical tool that tests whether the variable is correlated to the outcome variable.
6. For the models that were used, future researchers must consider the discussion and conclusion part of the study, as the researchers were able to discuss various pros and cons of each model that was used and encountered. Thus, if future researchers were to conduct similar study, they may be come across several things that may be have missed by the researchers.

ACKNOWLEDGEMENT

The researchers would like to extend their deepest gratitude and appreciation to the following:

First and foremost, heartfelt gratitude is expressed to the Almighty God for bestowing strength, wisdom, and perseverance upon the researchers throughout this academic journey. Divine guidance has been an unwavering source of inspiration and solace.

A sincere acknowledgment and gratitude to Engr. Rommel De Mesa, the researcher adviser, for the continuous support and guidance. His expertise and knowledge within the researchers chosen research field surely made an impact and influenced the researchers to make a better output. In addition, his will and drive to help nourished the researchers and boosted the researcher's confidence to accomplish this study.

Furthermore, the researchers like to express their genuine gratitude and appreciation to Engr. John Vincent G. Tongol, the researchers' coordinator, for his constant assistance and guidance all throughout the researcher's journey as the study researchers. His insightful feedback, recommendations and opinions about the study significantly contributed to the success of this research endeavour.

The researchers are profoundly indebted and grateful to the family members, friends, and colleagues who showed an unwavering support, encouragement, and understanding throughout this challenging journey. The faith and trust given to the researchers' abilities and capabilities has become impactful and a constant source of the strength and motivation for the researchers during tough times while conducting the study. Appreciation and gratitude for the financial support given by the abovementioned individuals, as this crucially helped the researchers to finance the research expenses.

Additionally, to Marrielle Lintag, who fuels Henri Fred Caparas' engine, acting as a driving force and pushing him to strive and accomplish this significant part of journey of his life.

Finally, appreciation and gratitude are extended to all individuals who, whether directly or indirectly, contributed to the completion of this research study. The invaluable assistance had been an instrument in order for the researchers to make this study possible.

REFERENCES

[1]. **World Health Organization: WHO.** The Department of Transportation, World Health Organization launched the Philippine Road Safety Action Plan 2023-2028. World Health Organization. May 31, 2023 [Online]. Available: https://www.who.int/philippines/news/detail/31-05-2023-department-of-transportation--world-health-organization-launch-philippine-road-safety-action-plan-2023-2028 [Accessed Nov. 7, 2023].

[2]. **J. R. Regidor,** "TRAFFIC CONGESTION IN METRO MANILA: IS THE UVVRP STILL EFFECTIVE? | Philippine Engineering Journal," Philippine Engineering Journal, 2013. [Online]. Available: https://www.journals.upd.edu.ph/index.php/pej/article/view/3799 [Accessed Nov. 28, 2023].

[3]. **J. A. Añana, K. Calma, C. Flores, and D. Tantoy,** "J-STAGE Home," J-Stage, 2015. [Online]. Available: https://www.jstage.jst.go.jp/?fbclid=IwAR2y9LKcblk090fXZxHiqAx5bex4yEqvxDE5f84zOBRo2dKt6pzrCDCfJKY [Accessed Nov. 28, 2023].

[4]. **"San Fernando, Philippines Metro Area Population 1950-2023,"** MacroTrends, 2023. [Online]. Available: https://www.macrotrends.net/cities/895000030/san-fernando/population [Accessed Nov. 28, 2023].

[5]. **W. Kowtanapanich, Y. Tanaboriboon, and W. Chadbunchachai,** "Applying Public Participation Approach To Black Spot Identification Process," IATSS Research, vol. 30, no. 1, pp. 73–85, Jan. 2006, doi: 10.1016/s0386-1112(14)60158-8.

[6]. **A. Manerikar, D. Buttepatil, and P. Rathod,** "Black Spot Analysis on National Highway - 4," www.academia.edu, vol. Vol. 5, no. Issue 03, Jan. 2016, [Online]. Available: https://www.academia.edu/42710096/IJERT_Black_Spot_Analysis_on_National_Highway_4 [Accessed Nov. 12, 2023].

[7]. **I. Karamanlis et al.,** "Deep Learning-Based Black Spot Identification on Greek road networks," Data, vol. 8, no. 6, p. 110, Jun. 2023, doi: 10.3390/data8060110.

[8]. **M. N. Al-Marafi and K. Somasundaraswaran,** "Review of crash prediction models and their applicability in black spot identification to improve road safety," Indian Journal of Science and Technology, vol. 11, no. 5, pp. 1–7, Feb. 2018, doi: 10.17485/ijst/2018/v11i5/119291.

[9]. **S. Liu & C. Wu,** "Road Traffic Accident Forecast Based on Optimized Grey Verhulst Mode," Proceedings of the 2016 Joint International Information Technology, Mechanical and Electronic Engineering, 2016, doi:10.2991/jimec-16.2016.100

[10]. **M. Tong, H. Duan, & L. He,** "A novel Grey Verhulst model and its application in forecasting CO2 emissions," Environ Sci Pollut Res 28, February 19, 2021, https://doi.org/10.1007/s11356-020-12137-5

[11]. **P. Gatabaz, J.C. Mba, & E. Pindza,** "Grey Verhulst model and its chaotic behaviour with application to Bitcoin adoption," Decisions in Economics and Finance, May 04, 2022, https://doi.org/10.1007/s10203-022-00368-9

[12]. **W. Li, X. Zhao, & S. Liu,** "Traffic Accident Prediction Based on Multivariable Grey Model," ResearchGate, March 2020, doi: 10.3390/info11040184

[13]. **Zeng, B., Tong, M., & Ma, X. (2020)**. A new-structure grey Verhulst model: Development and performance comparison. Applied Mathematical Modelling, 81, 522–537. doi:10.1016/j.apm.2020.01.014

[14]. **B. Zeng, X. Ma, and J. Shi,** "Modeling Method of the Grey GM(1,1) Model with Interval Grey Action Quantity and Its Application," Complexity, vol. 2020, pp. 1–10, Jan. 2020, doi: 10.1155/2020/6514236.

[15]. **M. Ghadi and Á. Török,** "Comparison different black spot identification methods," Transportation Research Procedia, vol. 27, pp. 1105–1112, Jan. 2017, doi: 10.1016/j.trpro.2017.12.104.

[16]. **M. Nandal,** "An overview of crash prediction models and their suitability in black spot identification," Jun. 01, 2019. [Online]. Available: https://www.jetir.org/view?paper=JETIR1906P56 [Accessed Nov. 12, 2023].

[17]. **E. Hauer, D. Harwood, and M. Griffith,** "Estimating Safety by the Empirical Bayes Method: A Tutorial", Jan. 2002. [Online]. Available: https://doi.org/10.3141/1784-16 [Accessed Nov. 12, 2023].

[18]. **F. Zong, H. Xu, and H. Zhang,** "Prediction for Traffic Accident Severity: Comparing the Bayesian Network and Regression Models", Oct. 30, 2013. [Online]. Available: https://www.hindawi.com/journals/mpe/2013/475194/ [Accessed Nov. 29, 2023].

[19]. **DeepAI,** "Empirical Bayes methods," DeepAI, Jun. 25, 2020. [Online]. Available: https://deepai.org/machine-learning-glossary-and-terms/empirical-bayes-methods?fbclid=IwAR2bTBI3NgmHTrV-3-I-sVUvQS86G6cOucCVLLjgNzZtZKSYaoItS_KLEes [Accessed Nov. 29, 2023].

[20]. **Gupta, S.** Use of Bayesian statistics in drug development: Advantages and challenges. International Journal of Applied and Basic Medical Research, 2(1), 3. doi:10.4103/2229-516x.96789, 2012.

[21]. **J. Weng,** "Improvement of Rate Quality Control Method in Identifying Hazardous Locations", ResearchGate, July 2010. [Online]. Available: https://www.researchgate.net/publication/269139737_Improvement_of_Rate_Quality_Control_Method_in_Identifying_Hazardous_Locations [Accessed Nov. 12, 2023].

[22]. **A. Sardar,** "Black Spot Identification Method Use on Roads," Jan. 29, 2020. [Online]. Available: https://www.linkedin.com/pulse/blackspot-identification-method-use-roads-asif-sardar [Accessed Nov. 12, 2023].

[23]. **"Advantages and disadvantages of quality control,"** Advantages and Disadvantages Table in GCSE Business Studies, Sep. 26, 2018. [Online]. Available: https://getrevising.co.uk/grids/advantages-and-disadvantages-of-quality-control?fbclid=IwAR2F-2s9zj-VMbEewradYRptNSXpKjSpQwnJImSAp6zQi9KLK9kZL2KGdvQ [Accessed Nov. 29, 2023].

[24]. **Wu, L., & Ci, Y.,** Amendment of Quality Control Method in Identifying Black Spots. International Conference on Transportation Engineering 2009. doi:10.1061/41039(345)203, 2009.

[25]. **M. Web.** Using R to analyze your data. Washington University in St. Louis. Available: https://libguides.wustl.edu/R [Accessed Jan. 4, 2024].

[26]. **Expert Research & Data Analysis Help.** R-Studio Data Analysis Services Conducted by Experts. Available: https://www.expertresearch-dataanalysishelp.com/R-software-analysis-services-2699.html [Accessed Jan. 4, 2024].

[27]. **H.A. Silva & A.S. Moura.** Teaching Introductory Statistical Classes in Medical Schools Using RStudio and R Statistical Language: Evaluating Technology Acceptance and Change in Attitude Toward Statistics, Journal of Statistics Education, 28:2, 212-219, DOI: 10.1080/10691898.2020.1773354 [Accessed Jan. 4, 2024]

[28]. **G. Palino and E. Sparks,** "QGIS: An Introduction to an Open-Source Geographic Information System". 2021 [Online]. Available: http://extension.msstate.edu/publications/qgis-introduction-open-source-geographic-information-system [Accessed April 25, 2024]

[29]. **V. Ovramenko,** "Unlock the Power of QGIS: A Step-by-Step Tutorial". September 20, 2023. [Online]. Available: https://www.linkedin.com/pulse/unlock-power-qgis-step-by-step-tutorial-vladimir-ovramenko [Accessed April 25, 2024]

[30]. **B. Seyoum**, "Investigating effects of road geometry on accident Blackspot of Addis Ababa to Gohatsion Trunk Road," Jun. 01, 2023. https://etd.aau.edu.et/items/c15bfd24-b25d-4f8a-9fae-c4ac519ba288/full [Accessed Jan. 04, 2024].

[31]. **K. Sjölinder, H. Ek, and K.-O. Hedman**, "Black Spot Manual," 2001. [Online]. Available: https://www.kgm.gov.tr/SiteCollectionDocuments/KGMdocuments/Eng/Traffic/BlackSpotManual.pdf [Accessed Apr. 17, 2024].

[32]. **W. Yan, W. He, & J. Zhou.** 2021. "Urban Road Accident Black Spot Identification and Classification Approach: A Novel Grey Verhuls–Empirical Bayesian Combination Method" *Sustainability* 13, no. 20: 11198. https://doi.org/10.3390/su132011198

