A Grey-Markov Model for the Prediction of Vehicular Accidents along Lokoja-Abuja-Kaduna Express Way in Nigeria

*O.B. Saeed [1], Abubakar U.Y[1], Lawal Adamu[1], Usman Abubakar[2]
[1]Department of Mathematics, Federal University of Technology Minna, Nigeria
[2]Department of Statistics, Federal University of Technology Minna, Nigeria
[1]belaneetech@yahoo.com

Abstract

Providing government of Nigeria with reliable and dependable mathematical model for road safety policy formulation in order to reduce loss of lives and properties along Lokoja-Abuja-Kaduna Express way is the thrust of this research. The World Health Organization (WHO) reported that road traffic accident claims roughly 1.3 million lives annually which make it one of the top causes of death worldwide. A Grey-Markov model that predicts yearly number of Vehicular accidents occurrence has been developed and implemented on Lokoja-Abuja-Kaduna express way in Nigeria. The data used in the research were collected from the archive of federal Road Safety Corps of Nigeria for a period of ten years (2010-2019). The fitted model recorded excellent performance of 93.544% *accuracy; this shows that the model is reliable and dependable. Therefore, results from this model could serves as source of information for road safety policy formulation.*

Keywords: Road, Crashes, Accident, Lokoja-Abuja-Kaduna, Express Way, Markov Chain and Grey-Markov

INTRODUCTION

Lokoja-Abuja-Kaduna express way is one of the major roads that connect the southern and northern parts of Nigeria. As a result of high volume of vehicular movements on this road, hundreds of road accidents that mostly result to thousands of human causalities and loss of properties are recorded every year, most of which results in deaths and disabilities (FRSCN, 2020). Due to complex factors associated with road traffic accident, its occurrence is not easily predictable and its impact has over the years affected economic and social activities of the country. This has consequently reduced Gross Domestic Product (GDP) of the country. In view of this, the research aimed at providing some quantitative information to Government of Nigeria for road safety policy formulation in order to mitigate the loss of lives and properties. Nigeria has roughly 195,000 km of surfaced road, making it the country with largest road network in West Africa and the second largest in south of the Sahara, out of which a proportion of about 32,000 km are federal roads while 31,000km are state roads (F.R.S.N, 2020).

Road traffic crashes have been found to be influenced by many complex factors such as weather, driver factor, over speeding, rear by vehicular density, risk location, driver consciousness, driver fatigue, the nature of the road and so on. Since Road Traffic Crash is generally a random occurrence it is therefore, quite important to select an appropriate forecasting model that will fully capture the behaviour of the system. Grey-Markov stochastic model has been selected for this purpose.

Successful studies of accident occurrence have been reported in the following works: (Li *et.al* 2007) has applied Grey-Markov Model in predicting traffic volume in Manjing City, China. They concluded that the model has high accuracy, reliability and precision in the traffic volume prediction. Presented in (Jian-Yi and Ying, 2014) is a modified Grey-Markov Model. The model was used to forecast the mine safety accident deaths from 1990 to 2010 in China and 2001 to 2014 for coal accidents death were predicted accordingly. The result shows that the new model not only discovered the trend of the mine human error accident death toll but also overcomes the random fluctuation of data affecting precision. It was concluded that the model possesses strong engineering application. Reported in (Xiaoxia, *et.al*, 2018) is a driving risk status prediction algorithm based on Markov Chain. In the study, driving risk states were classified using clustering techniques based on feature variables describing instantaneous risk levels within time windows, where instantaneous risk levels are determined in time-to-collision and time-to-headway two-dimensional plane. Multinomial logistic models with recursive feature variable estimation method were developed to improve the traditional state transition probability estimation, which also takes into account the comprehensive effect of driving behaviour, traffic and road environment factors on the evolution of driving risk status. A "100-car" natural driving data from Virginia technology was employed for the training and validation of the prediction model. The results show that under the 5% false positive rate, the prediction algorithm could have high prediction accuracy for future medium-to-high driving risks and could meet the time line requirement of collision avoidance warning. The algorithm could contribute to timely warning or auxiliary correction to drivers in the approaching-danger state. Successful studies of vehicular accidents occurrence have also been reported in (Bamidele *et.al*, 2006; John *et.al*, 2007; Nyothiri *et.al*, 2018)

METHODOLOGY

The Grey-Markov model consist of GM(1,1) model and Markov chain model. The Grey model deals with an uncertain system of small sample size and poor information involving both known and unknown information. The goal of Grey prediction is to whitening the system and reveals the unknown. However, in Grey model, problem of poor fitting degree and low prediction accuracy may emerge when the change rate of the original data is too large. The Grey –Markov Model (GMM) is an extension of Grey Model (GM) to further reduce the prediction error. The Grey-Markov Model is made up of two components namely Grey and Markov chain model; Markov chain model can handle a situation where the change rate of the original data is too large. Hence, Markov Chain Model makes it possible to solve the problems mentioned in Grey Model. The Grey-Markov Model was established based on the advantage of both methods which adopt GM(1,1) Model to study development regulation of data sequence and uses Markov Models to study vibrating irregularities of data sequence. In general, the combination of the two models have been found to improve the prediction accuracy (Mao and Sun, 2011). As earlier mentioned we shall begin our prediction with GM(1,1) model after which we shall improve the prediction accuracy using Grey-Markov model.

The Grey-System Model GM(1,1)

The raw data series in grey GM(1,1) model is represented by $x^{(0)}(k), k = 1,2,3,\ldots,n, x^{(0)}(k) \geq 0$ and it can also be represented as:

$$X^{(0)}(k) = \left(x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), \ldots, x^{(0)}(n)\right) \tag{1}$$

The accumulated generating sequence is given as:

$$X^{(1)}(k) = \left(x^{(1)}(1), x^{(1)}(2), x^{(1)}(3), \ldots, x^{(1)}(n)\right) \tag{2}$$

where $X^{(0)}(k) = \sum_{i=1}^{k} x^{(0)}(i), k = 1,2,\ldots,n$ (3)

$$x^{(0)}(k) + ax^{(1)}(k) = b \tag{4}$$

Equation (4) represents the original form of the GM(1,1) model, it is a difference equation. The symbol GM(1,1) stands for first order Grey Model in one variable.

Equation (4) is also represented as equation (5)

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b \tag{5}$$

Equation (5) is a differential equation, where a and b are parameters to be identified. a is called developing coefficient and b is grey input.

Equation (6) is the solution of equation (5)

$$\hat{x}^{(1)}(k+1) = \left(x^{(1)}(0) - \frac{b}{a} \right) e^{-ak} + \frac{b}{a} \tag{6}$$

Equation (6) is the time response function while parameters a and b are estimated using Least Square Method as follows:

$$\begin{bmatrix} a \\ b \end{bmatrix} = \left[B^T B \right]^{-1} B^T Y \tag{7}$$

$$\text{Where, } B = \begin{bmatrix} \frac{-[x^{(1)}(1) + x^{(1)}(2)]}{2} & 1 \\ \frac{-[x^{(1)}(2) + x^{(1)}(3)]}{2} & 1 \\ \frac{-[x^{(1)}(3) + x^{(1)}(4)]}{2} & 1 \\ \frac{-[x^{(1)}(4) + x^{(1)}(5)]}{2} & 1 \\ \frac{-[x^{(1)}(5) + x^{(1)}(6)]}{2} & 1 \\ \vdots & \vdots \\ \frac{-[x^{(1)}(n-1) + x^{(1)}(n]}{2} & 1 \end{bmatrix} \tag{8}$$

$$Y = \left[x^{(0)}(2), x^{(0)}(3), x^{(0)}(4), \ldots, x^{(0)}(n) \right]^T \tag{9}$$

The reduction value of equation (6) is given below:

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) = \left(1 - e^{a}\right)\left(x^{(0)}(1) - \frac{b}{a} \right) e^{-ak} \tag{10}$$

Prediction Accuracy Test

To determine the accuracy of our prediction, we shall adopt Mean Absolute Percentage Error (MAPE).

This tool is often used for determining prediction accuracy showing the same characteristics i.e. the smaller the value, the higher the prediction accuracy.

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \tag{11}$$

where;

$\hat{y}_i$ is the Grey Model predicted value.

y_i is the Grey-Model actual value.

n is the number of prediction samples.

Lewis (1982) divided the prediction accuracy of models into four grades and the division of prediction accuracy grades is shown in the table 1 below:

Table 1: Classification of Prediction Accuracy

MAPE	PREDICTION ACCURACY
$<10\%$	High
$10\% - 20\%$	Good
$20\% - 50\%$	Feasible
$>50\%$	Low

The Grey- Markov Model

The Grey-Markov Model (GMM) is an extension of Grey Model (GM) to further reduce prediction errors. In Grey model, the problems of poor fitting degree and low prediction accuracy may emerge when the change range of original data is too large. However, these problems can be well resolved by adopting Markov chain which can narrow down the prediction interval and improve the prediction accuracy. Markov stochastic process improves these limitations of Grey model because it reflects the stochastic volatility impact on elements by determining the transfer law of states (Ducan *et.al*, 1998).

Formulation of Grey-Markov Model

First step in formulating the GMM is to divide the residual errors into q states where each state satisfies the equi-probability principle and is defined as $R_{1,} R_{2,} \cdots, R_q$. Next, the construction of the transition matrix is done by determining the probability from state R_i to state R_j which results in the transition matrix P .

$$P^{(1)} = \begin{bmatrix} p_{(11)}^{(1)} & p_{(12)}^{(1)} & \cdots & p_{(1q)}^{(1)} \\ p_{(21)}^{(1)} & p_{(22)}^{(1)} & \cdots & p_{(2q)}^{(1)} \\ \vdots & \vdots & \cdots & \vdots \\ p_{(q1)}^{(1)} & p_{(q2)}^{(1)} & \cdots & p_{(qq)}^{(1)} \end{bmatrix} \tag{12}$$

$$P^{(m)} = \begin{bmatrix} p_{(11)}^{(m)} & p_{(12)}^{(m)} & \cdots & p_{(1q)}^{(m)} \\ p_{(21)}^{(m)} & p_{(22)}^{(m)} & \cdots & p_{(2q)}^{(m)} \\ \vdots & \vdots & \cdots & \vdots \\ p_{(q1)}^{(m)} & p_{(q2)}^{(m)} & \cdots & p_{(qq)}^{(m)} \end{bmatrix} \tag{13}$$

Where $P_{ij}^{(m)} = \dfrac{M_{ij}^{(m)}}{m_i}, (i, j = 1,2,3,\ldots,L),\ M_{ij}^{(m)}$ stands for the transition from R_i to R_j in m steps and m_i is the number of state R_i .

Next, the residual error must be confirmed.

Let the interval median in $[R_{i-}, R_{i+}]$ be residual error forecasting value as follows:

$$\hat{e} = \tfrac{1}{2}[R_{i-} + R_{i+}] \tag{14}$$

International Journal Industrialist Technology Engineering , Science and Education (IJITESED}Vol 2 No1, June 2021

Hence, the Grey-Markov model is obtained as:

$$\hat{Y}(k+1) = [1+\hat{e}]\hat{x}^{(0)}(k+1)$$

(15)

Where, $\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k)$

Hence:

$$\hat{Y}(k+1) = \left[1 + \tfrac{1}{2}(R_{i-} + R_{i+})\right]\hat{x}^{(0)}(k+1)$$

(16)

RESULTS AND DISCUSSION

Application of Grey System Model for Prediction of Vehicular Accidents along Lokoja-Abuja-Kaduna Expressway

The data used in this research were collected from the archive of Federal Road Safety of Nigeria for the period of ten years (2010-2019).

The summary of the data is presented in table 2 below:

Table 2: Summary of Number of Vehicular Crashes for Ten Years

S/N	Year	Number Of Accident Within The Year
1	2010	877
2	2011	806
3	2012	560
4	2013	1170
5	2014	929
6	2015	672
7	2016	654
8	2017	655
9	2018	780
10	2019	701

Using equation (1) and table (2), we obtain equation 17 below

$$X^{(0)} = (877, 806, 560, 1170, 929, 672, 654, 655, 780, 701) \tag{17}$$

From equation (2) we obtain the accumulated generating sequence as given below:

$$X^{(1)} = (877, 1683, 2243, 3413, 4342, 5014, 5668, 6323, 7103, 7804) \tag{18}$$

Equation (19) below is obtained using equation (7)

$$\hat{a} = \begin{bmatrix} 0.02241 \\ 869.607 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} \tag{19}$$

Substituting for a and b in equation (6), we obtained equation (20) below:

$$\hat{x}(k+1) = 38804.42 - 37927.42e^{-0.02241k} \tag{20}$$

Evaluating equation (20) for $k = 0,1,\ldots,9$ we obtained the following values below:

$$\hat{X}^{(1)} = (877, 1718, 2539, 3343, 4129, 4897, 5649, 6383, 7102, 7805) \tag{21}$$

We compute the simulated value using equation (22) below:

$$\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1) \tag{22}$$

$$\hat{X}^{(0)} = (877, 841, 821, 804, 786, 768, 752, 734, 719, 703) \tag{23}$$

Equation (23) is the simulated values from 2010-2019

Table 3: Comparison of Actual and Grey simulated Value for Vehicular Accident along Lokoja-Abuja-Kaduna Express Way from Year 2010-2019.

S/N	YEAR OF CRASH	ACTUAL NUMBER OF CRASH	GREY MODEL SIMULATED VALUES	RESIDUAL ERROR	RELATIVE ERROR (%)
1	2010	877	877	0	0
2	2011	806	841	-35	-4.3
3	2012	560	821	-261	-46.6
4	2013	1170	804	366	31.3
5	2014	929	786	143	15.4
6	2015	672	768	-96	-14.3
7	2016	654	752	-98	-14.1
8	2017	655	734	-79	-12.1
9	2018	780	719	61	7.8
10	2019	701	703	-2	-0.3

Using equation (11), we observed from table (3) that:

$MAPE = 14.62\%$

$ACCURACY = 100\% - 14.62\% = 85.97\%$

The figure 85.97% indicates that the prediction accuracy is good; however, this prediction accuracy level can be improved using Grey-Markov Model.

Application of Grey-Markov Model for Prediction of Vehicular Accident along Lokoja-Abuja-Kaduna Express Way

The prediction accuracy of the GM(1,1) can be improved by applying the Grey-Markov model. We begin by finding error state of each year. To achieve this, the error is partitioned into states. Due to the small sample size in this study, the error can be divided into three states respectively, using E_1, E_2, E_3 as shown in table 4:

Table 4: Error Partition

State	$E_1(\%)$	$E_2(\%)$	$E_3(\%)$
Error Range	$-46.6 \sim -20.63$	$-20.63 \sim 5.34$	$5.34 \sim 31.31$

From Table (4), we obtained the error states in Table (5)

Table 5: State Division for Actual Vehicular Accident along Lokoja-Abuja-Kaduna Express Way.

S/N	Year Of Crash	Actual Number Of Crash Within The Year	Grey Model Simulated Values	RELATIVE ERROR (%)	States
1	2010	877	877	0	E_2
2	2011	806	841	-4.3	E_2
3	2012	560	821	-46.6	E_1
4	2013	1170	804	31.3	E_3
5	2014	929	786	15.4	E_3
6	2015	672	768	-14.3	E_2
7	2016	654	752	-14.1	E_2
8	2017	655	734	-12.1	E_2
9	2018	780	719	7.8	E_3
10	2019	701	703	-0.3	E_2

From Table (5), we construct transition Count Matrix as contained in equation (24)

$$C = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 3 & 1 \\ 0 & 2 & 1 \end{bmatrix} \tag{24}$$

From equation (24), we construct transition probability matrix using equation (12)

$$P^{(1)} = \begin{bmatrix} 0 & 0 & 1 \\ 0.2 & 0.6 & 0.2 \\ 0 & 0.67 & 0.33 \end{bmatrix} \tag{25}$$

The two steps, three steps and four steps transition probability matrix is calculated using equation (13), hence we have them respectively:

$$P^{(2)} = \begin{bmatrix} 0 & 0.67 & 0.33 \\ 0.120 & 0.494 & 0.386 \\ 0.134 & 0.623 & 0.243 \end{bmatrix} \tag{26}$$

$$P^{(3)} = \begin{bmatrix} 0.134 & 0.623 & 0.243 \\ 0.099 & 0.555 & 0.346 \\ 0.125 & 0.537 & 0.339 \end{bmatrix} \tag{27}$$

$$P^{(4)} = \begin{bmatrix} 0.125 & 0.537 & 0.339 \\ 0.111 & 0.565 & 0.324 \\ 0.107 & 0.549 & 0.344 \end{bmatrix} \tag{28}$$

Evaluating equation (16) for $k = 0,1,2,3,\ldots,9$

$$\hat{Y}_1 = \left[1 + \frac{1}{2}\left(-20.63\% + 5.34\%\right)\right] * 877 = 809.95 \approx 810$$

$$\hat{Y}_2 = \left[1 + \frac{1}{2}\left(-20.63\% + 5.34\%\right)\right] * 841 = 776.70 \approx 777$$

$$\hat{Y}_3 = \left[1 + \frac{1}{2}\left(-46.6\% - 20.63\%\right)\right] * 821 = 545.02 \approx 545$$

$$\hat{Y}_4 = \left[1 + \frac{1}{2}\left(5.34\% + 31.31\%\right)\right] * 804 = 951.333 \approx 951$$

$$\hat{Y}_5 = \left[1 + \frac{1}{2}\left(5.34\% + 31.31\%\right)\right] * 786 = 930.03 \approx 930$$

$$\hat{Y}_6 = \left[1 + \frac{1}{2}\left(-20.63\% + 5.34\%\right)\right] * 768 = 709.28 \approx 709$$

$$\hat{Y}_7 = \left[1 + \frac{1}{2}\left(-20.63\% + 5.34\%\right)\right] * 752 = 694.50 \approx 695$$

$$\hat{Y}_8 = \left[1 + \frac{1}{2}\left(-20.63\% - 5.34\%\right)\right] * 734 = 677.88 \approx 678$$

$$\hat{Y}_9 = \left[1 + \frac{1}{2}\left(5.34\% + 31.31\%\right)\right] * 719 = 850.75 \approx 850$$

$$\hat{Y}_{10} = \left[1 + \frac{1}{2}\left(-20.63\% + 5.34\%\right)\right] * 703 = 649.25 \approx 649$$

We obtained the Grey-Markov Simulated Values presented in Table (6)

Table 6: Comparison of Actual and Grey-Markov Simulated Value for Vehicular Accidents along Lokoja-Abuja-Kaduna Expressway for the Period of 2010-2019.

S/N	Year Of Accidents	Actual Number Of Accidents Within The Year	Grey-Markov Model Simulated Values	Residual Error	Relative Error (%)
1	2010	877	810	67	7.64
2	2011	806	777	29	3.60
3	2012	560	545	15	2.68
4	2013	1170	951	219	18.72
5	2014	929	930	-1	-0.11
6	2015	672	709	-37	-5.51
7	2016	654	695	-41	-6.27
8	2017	655	678	-23	-3.51
9	2018	780	851	-71	-9.10
10	2019	701	649	52	7.42

Using equation (11), we observed from Table (6) that,

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \widehat{y}_i}{y_i}\right| \times 100\% = 6.456\%$$

$ACCURACY = 100\% - 6.456\% = 93.544\%$. This implies that, the prediction accuracy has been improved from 85.97% to 93.544%

The figure 93.544% indicates a high accuracy level.

Table 7: Comparison of Actual Value, Grey Model and Grey-Markov Model Simulated Values for Road Accident along Lokoja-Abuja-Kaduna Express high way

S/N	Year Of Accidents	Actual Number Of Accidents Within The Year	Grey- Model Simulated Values	Grey-Markov Model Simulated Values
1	2010	877	877	810
2	2011	806	841	777
3	2012	560	821	545
4	2013	1170	804	951
5	2014	929	786	930
6	2015	672	768	709
7	2016	654	752	695
8	2017	655	734	678
9	2018	780	719	851
10	2019	701	703	649

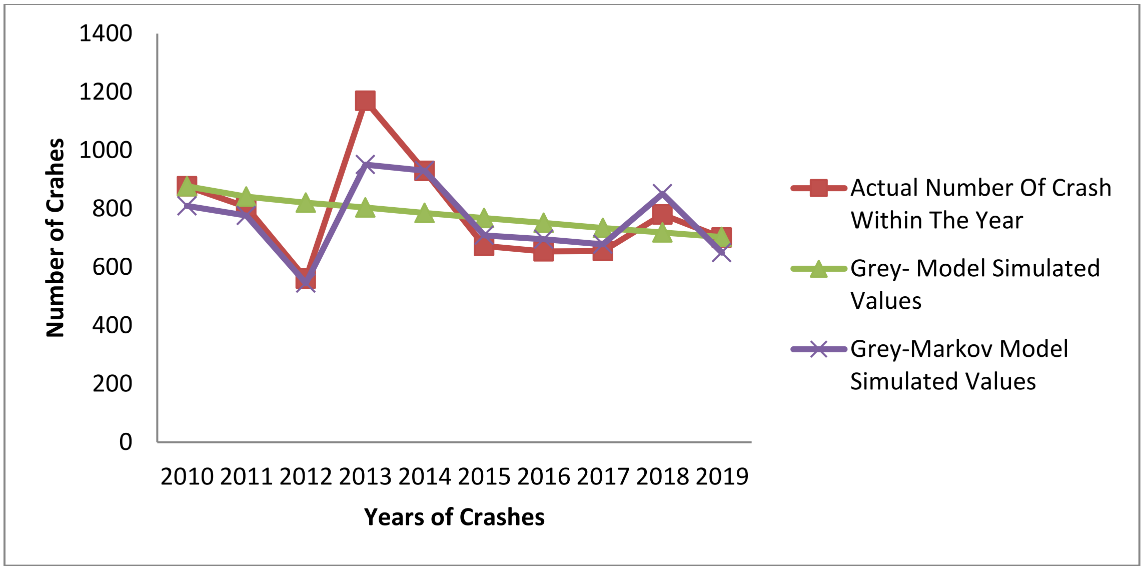

Figure 1: Comparison of Actual Value, Grey Model and Grey-Markov Model Simulated Values for Vehicular Accidents along Lokoja-Abuja-Kaduna Express way.

Grey Forecasting for Vehicular Accidents from Year 2020-2023

Evaluating equation (20) for $k = 9,10,11,12,13$

The predicted values are computed using equation (22) and the results are presented in table 8.

Table 8: Grey-Model Predicted Values from Year 2020-2023

Year Of Accidents	Grey Model Prediction Value
2020	672
2021	657
2022	642
2023	629

Grey-Markov Model Prediction from 2020-2023

To achieve this, we obtained the error state for year 2020 to 2023 by using equations (25) to (27) and the information in table (5) after which we use equation (16) and the error states to make prediction from 2020 to 2023. The results are presented in table 9.

Table 9: Grey-Markov Model prediction value from 2020-2023.

Year Of Crash	Grey- Markov Prediction Value
2020	621
2021	607
2022	593
2023	581

Table 10: Comparison of Grey and Grey-Markov Model predicted values from year 2020-2-23.

Year Of Crash	Grey- Model Prediction Values	Grey-Markov Model Prediction Values
2020	672	621
2021	657	607
2022	642	593
2023	629	581

Conclusion

The G,M(1,1) system model has been used to simulate the number of vehicular accidents along Lokoja-Abuja-Kaduna express way. After which a Grey-Markov model was used to improve the accuracy of the simulated values. The performance of the Grey-Markov model was impressive as it recorded 93.544% accuracy when fitted. This shows that the model is reliable and dependable; therefore the prediction made for the future years should be adopted by Nigeria government as source of information for road safety policy formulation on the express way.

REFERENCE

Bamidele M and Ifeoluwa H. (2016). Modelling of Road Traffic Accidents: A multistate Markov Approach.*Sri Lankan Journal of Applied Statistics Vol. 17(2).*

Duncan, T.E, Hu Y and Pasik-Duncan B. (1998). Stochastic calculus for fractional Brownian motion I, theory. Preprint.

F.R.S.N (2020). Archives of Federal Road Safety of Nigeria, https://www.frsc.gov.ng

Jian-Yi L. and Ying Z. (2014). Application of Grey Markov SCGM $(1,1)_c$ Model to prediction of Accidents Death in Coal Mining. Hindawi Publishing Corporation International Scholarly Research Notices Vol.2014, Article ID 632804, 7 Pages http://dxidoi.org/10.1155/2014/632804

John C., Darwin J., Noime B. and Vensurmar C., (2017). Analysis of Vehicule Crash Injury-Severity in a Super High Way: A Markovian Approach. Industrial Engineering Department Publishing, Adamson University, Manila, Philippines 1000

Lewis J. (1982). "Beta-Blockade after Myocardial infarction View" Imperial Chemical Industries PLC, Pharmaceuticals Division, Macclesfield SK10 4TF

Li Q., Hu Q. and Zangh P., (2007). Application of Grey-Markov Model in Predicting Traffic Volume.Preceeding of 2007 IEEE International Conference on Grey System and Intelligent Services, November 18-2-, 2007, Nianjin, China

Mao Z. and Sun J., (2011). Application of Grey-Markov Model in Forecasting Fire Accidents. The Fifth Conference on Performance-based Fire Protection Engineering, ELSEVIER

Nyothiri A., Weidong Z., Sahraoui D., Yibo A., (2018). Accident Prediction System Based on Hidden Markov Model for Vehicular Ad-Hoc Network in Urban Environments .*www.mdpi.com/Journal /Information 2018, 9,311;doi:10.3390/info9120311*

Xiaoxia X., Long C and Jun L., (2018). Vehicular Driving Risk Prediction Based on Markov Chain Model.Hindawi Discrete Dynamics in Nature and Society, Vol. 2018, Article ID 4954621, 12 Pages http://doi.org/10.1155/2018/4954621.