A BAYESIAN FORECAST OF ROAD TRAFFIC FATALITIES IN NIGERIA

***IBEKWE, U. A., AJIJOLA, L. A. & MESIKE, G. C.**
Department of Actuarial Science & Insurance
Faculty of Management Sciences
University of Lagos, Akoka, Lagos, Nigeria.
*****ucheibek@yahoo.comlajijola@unilag.edu.ng; mesikegodson@yahoo.co.uk

Abstract

Road accidents account for a significant proportion of the high mortality rates in Nigeria. Reducing road traffic accidents calls for knowledge of the causative factors and measures for remedial actions. This study therefore attempts to determine whether the targets announced for the U.N. Decade of Action for Road Safety are realizable in order to stabilize and reduce the level of road traffic accidents by about 50 per cent by the year 2020.Projected road traffic fatalities in Nigeria are modelled using an Empirical Bayes approach. The model applied credibility theory to Reported Road accident Cases and Casualties data for the period 2001 to 2016 inclusive to forecast road accidents in subsequent years. Forecast accuracy is assessed within an Exponentially Weighted Moving Average (EWMA) model using Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE).It was discovered that the U.N. targets, though apparently largely achieved by the target year, may not be sustainable into the future in Nigeria if the target date were to be extended by another decade to 2030, unless drastic policy changes based on sound and quantifiable forecasts are introduced and implemented.

Keywords*:* Road traffic accident, Empirical bayes, credibility theory

1. Introduction

Road accidents account for a significant proportion of the high mortality rates in Nigeria and projections show that this could become the second leading cause by the year 2020, next to heart disease (Aggarwal and Oberoi, 2009). These figures appear to be conservative as a result of under-reporting of accidents. Rural roads are often not counted nor are people who die from their injuries later on in hospital (U.N. 2007).

Reducing road traffic accidents calls for knowledge of the causative factors and suggesting remedial actions (Orji, Fadiora, Ogunlola, and Badru, 2002). Such measures will need to be based on accurate data and forecasts. In this study, we set out to achieve this by forecasting the number of road traffic accidents based on the data for seventeen years (2001-2016) and setting this against targets announced by the United Nations regarding road traffic accidents for the years up to 2020. The United Nations General Assembly had proclaimed the period 2011-2020 as the ‘Decade of Action for Road Safety’ with a primary goal of stabilizing and then reducing road traffic fatalities worldwide by about 50% by the year 2020 (www.iris.paho.org). This program of action has been endorsed by more than 100 countries. The World Bank and the World Health Organisation have also endorsed the program.

Providing better data on road traffic injuries and fatalities would be one important tool for tackling road safety problems. Accurate prediction can provide crucially important information for the prevention of road traffic fatalities. Such prediction should help indicate in good time any potential gap between current trends and the targets set by the U.N. The Empirical Bayes Model (EBM) was selected in this study for forecasting traffic road accidents in Nigeria.

2. Literature Review

Several studies have been done on road traffic accidents worldwide. Igboanugo and Ikhuemelo (2007) highlighted the role of multiple policy interventions in reducing road traffic accidents in Nigeria. In a 2001 paper titled “Pattern Recognition for Road Traffic Accident Severity in Korea” Sohn and Shin (2001) found that protective devices have the most impact on mitigating the severity of road traffic accidents. Mesken, Lajunen and Summala (2002) used factor analysis and logistic regression and found that attitude enforcement and better instruction for drivers tended to prevent violations and reduce errors. Aggarwal and Oberoi (2009) found that the most common age group involved were young adults, and that as people got older they tended to be less affected. Males outnumbered females more than ten to one.

According to Eke (2002) and Nwaokoro (2000), the Federal Road Safety Commission came into existence in 1982 and was officially launched in 1995. Operatives of the Commission are known as Road Marshalls (Eke 2001). Forecasting techniques are premised on the assumption that past patterns will be repeated in the future. Some of these techniques include time-series methods, causal forecasting methods, as well as qualitative methods. Various time series models include Song and Chissom (1993), Li and Kozma (2003) Su and Li (2003), Jilani and Burney (2007) and Jilani, Burney and Ardil (2007). In this paper, an actuarial credibility model (EBM) is employed to forecast car road accidents in Nigeria.

3. Methodology

This paper uses a Bayesian credibility model to predict casualty figures in Nigeria for the coming years with a view to providing estimates that could, for instance, help place the U.N. Decade of Action to reduce road traffic accidents on a scientific foundation. Credibility theory will be applied in an empirical Bayesian framework to road accident data in order to predict road accident fatalities in the near future. An Exponentially Weighted Moving Average (EWMA) technique is employed to validate the forecasts using Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and Mean Absolute

Percentage Error (MAPE).

3.1 Credibility Theory

Initially, credibility theory was used in insurance mathematics to produce formulas for experience rating (Klugman, 2012). As explained by Klugman, whereas the sample mean is used in statistics when information on insured persons is available, credibility theory recommends giving some weight to an additional estimator from other sources. While this approach might introduce some bias, the average squared error is likely to be reduced (Norberg, 1980).
Credibility theory thus attempts to use extra (sometimes considerable) information in estimating the parameter of interest (Boland, 2006). In Credibility theory, actuarial and statistical methods are combined to predict future outcomes. This combination tends to improve the estimation process and usually results in more accurate projection of future results (Behan, 2009).
Whitney (1918) in Norberg (2006) suggested the use of a weighted average:

$$\bar{m} = z\hat{m} + (1 - z)\mu \qquad (1)$$

Where $\hat{m}$is the average claim amount and μ is the portfolio mean while z is the credibility factor and$\bar{m}$ is the credibility premium. Mowbray (1914) suggested setting $z = 1$ in (1) for full credibility,

3.2 The Model

In this model, $E[M(\theta)], Var[M(\theta)]$and $E[S^2(\theta)]$are obtained from available data to estimate each risk.
As described in Boland (2006),"the data are presented as $\left\{\left(X_{i_j}\right)_{i=1}^{N}\right\}_{j=1}^{n}$ where X_{i_j} represents the number of accidents in the j^{th} year from the i^{th} risk. An unknown risk parameter θ_i is assumed for the i^{th} risk, which for $i = 1, 2, \ldots., N$ denotes a realization from a random variable θ. We represent the mean and variance of $X_{i_j}|\theta_i$ for $j = 1, \ldots, n$ by $M(\theta_i)$ and $S^2(\theta_i)$ respectively, but do not assume any particular form for the distribution of the random variable''. Boland goes on to assume that ''$X_{i_j}|\theta_i$ for $j = 1, \ldots, n$ are independent and identically distributed for any given risk i, the unconditional random variables $X_{i_j}|\theta_i$ for $j = 1, \ldots, n$ are not necessarily independent even though they are identically distributed''.
The predicted mean $E\left(X_{i,}n + 1|X = x\right)$ for the following year in risk i, is given by

$$\frac{n}{n + E[S^2(\theta)]/Var[m(\theta)]}\bar{x}_i + \left(1 - \frac{n}{n + E[S^2(\theta)]/Var[m(\theta)]}\right)E[m(\theta)] \qquad (2)$$

$E[M(\theta)], E[S^2(\theta)] = \sigma^2$ and $Var[M(\theta)]$ are estimated from actual data.

3.2.1 Unbiased estimators for $E[M(\theta)], E[S^2(\theta)]$ and $Var[M(\theta)]$

Estimator

$$E\left[\widehat{M(\theta)}\right] = \bar{x} \quad (3)$$

$$E\left[\widehat{S^2(\theta)}\right] = \sum_{i=1}^{N}\sum_{j=1}^{N}\left(X_{i_j} - \bar{X}_i\right)^2 / N(n-1) \quad (4)$$

$$Var\left[\widehat{M(\theta)}\right] = \sum_{i=1}^{N}\frac{(\bar{X}_i - \bar{X})^2}{(N-1)} - \sum_{i=1}^{N}\sum_{j=1}^{N}\left(X_{i_j} - \bar{X}_i\right)^2 / [N(n-1)] \quad (5)$$

3.3 The Data

Data on reported road accident cases and casualties obtained from the Federal Road Safety Corps (FRSC) Annual Reports for the period 2001-2016 inclusive. Data on accident cases is subdivided into fatal (if someone is killed), serious (if someone is injured) or minor (nobody injured) as well as casualties (Orji *et al*, 2002). Between 2001 and 2016, the number of accidents decreased 52.78%, injured cases increased by 29.49%, cases of deaths decreased by 49.2%, total casualties increased by 2.9%.

Table1. Reported Road accident Cases and Casualties, 2001-2016

	Cases				Casualties		
Year	Fatal	Serious	Minor	Total	Killed	Injured	Total
2001	6966	8185	5379	20530	9946	23249	33195
2002	4029	7190	3325	14544	7407	22112	29519
2003	3910	7882	2572	14364	6452	18116	24568
2004	3275	6948	4051	14274	5351	16897	22248
2005	2299	4143	2620	9062	4519	15779	20298
2006	2600	5550	964	9114	4944	17390	22334
2007	2162	4812	1503	8477	4673	17794	22467
2008	3024	5671	2646	11341	6661	27980	34641
2009	2460	6024	2370	10854	5693	27270	32963
2010	1178	2819	1333	5330	4065	18095	22160
2011	1764	2485	516	4765	4372	17464	21836
2012	1953	3106	1210	6269	4260	20757	25017
2013	3294	8589	1700	13583	6544	40057	46601
2014	3117	6356	907	10380	5996	32063	38059
2015	2854	6039	841	9734	5440	30478	35918
2016	2638	5633	1423	9694	5053	30105	35158

Source: FRSC Annual Report.

Figure1: Reported Road Accident Cases. 2001-2016

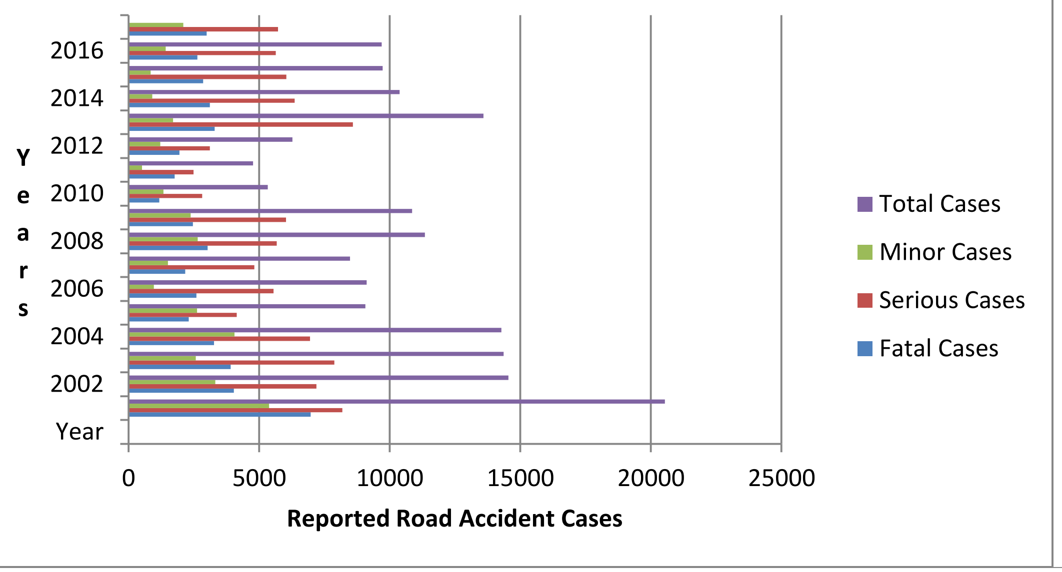
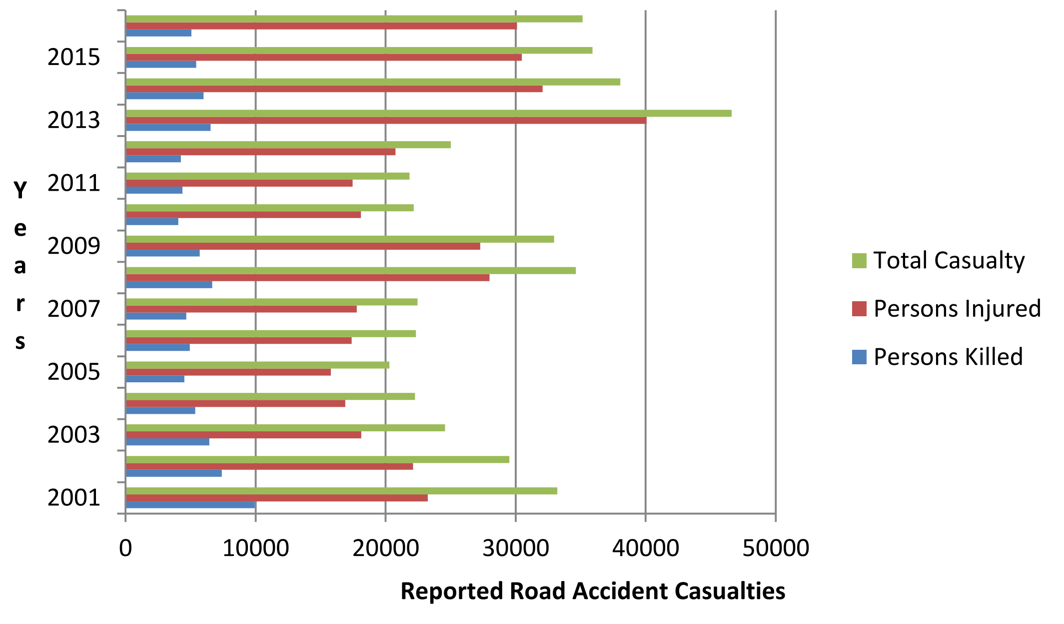


Figure2: Reported Road Accident Casualties. 2001-2016



4. Predictions

The credibility factor for each subsequent year is calculatedusing the sample means and variances to obtain

$(\bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4, \bar{x}_5) = (2970.1875, 5714.5, 2085, 5711, 23475.213)$ and

$variances = (425323.3, 871883.1, 434707.9, 556148.3)$
From which $E[m(\hat{\theta})]$is estimated by $E[m(\theta)] = 4120.172$ and $E[S^2(\theta)] = 2288063$.
Therefore $Var\ [m(\Theta)] = 3359798$.The credibility factor is the same for each case and casualty and is given by:

$$z = \frac{n}{n + E[S^2(\theta)]/Var[m(\theta)]}\bar{x}_i + \left(1 - \frac{n}{n + E[S^2(\theta)]/Var[m(\theta)]}\right)E[m(\theta)] = 0.9565$$

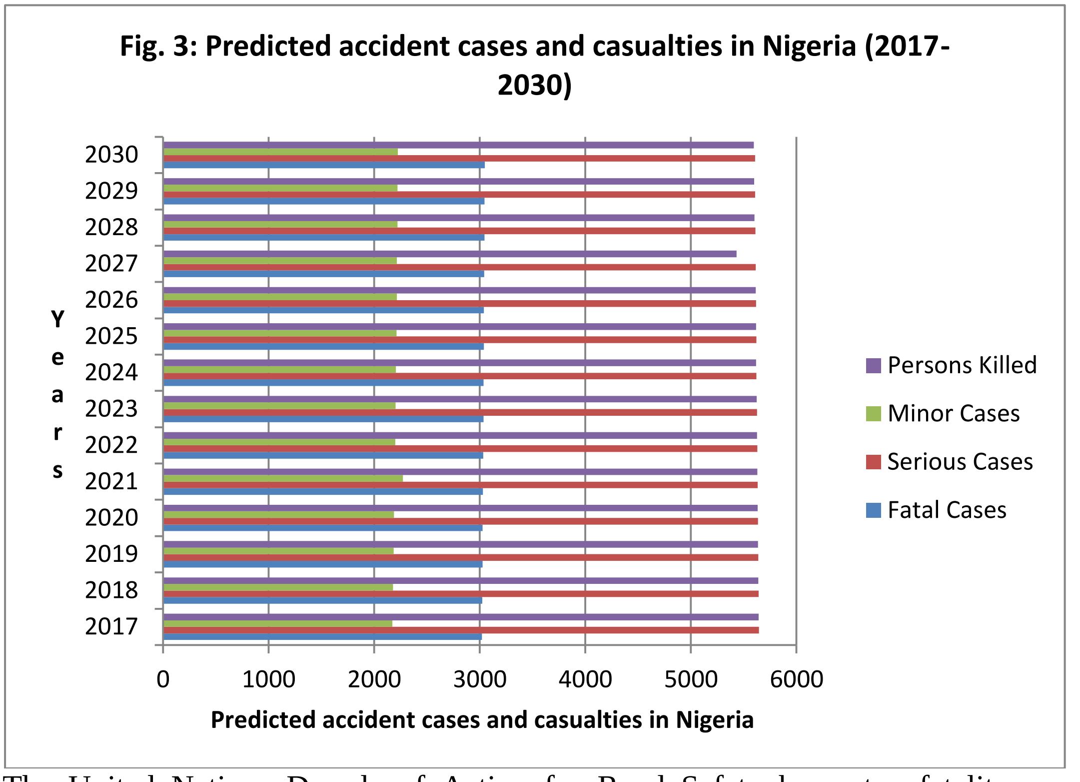
The predicted value for x_1 would be 3020.13. Similarly, the same procedures is applicable for $x_2, x_3,$ andx_4.

5. Discussion and Conclusion

This paper uses an empirical Bayesian model (EBM) from actuarial credibility theory to predict road traffic accidents in Nigeria for the period of 2001-2016, and to forecast the trends in the following years. The forecast results are presented in Table 2.

Table3: Predicted of Road Accident Cases and Casualties

Year	Fatal Cases	Serious Cases	Minor Cases	Persons Killed
2017	3020	5645	2173	5642
2018	3022	5640	2178	5638
2019	3025	5638	2183	5634
2020	3028	5634	2187	5631
2021	3030	5631	2271	5628
2022	3033	5628	2199	5625
2023	3035	5625	2203	5622
2024	3036	5621	2206	5618
2025	3038	5619	2210	5616
2026	3040	5616	2213	5613
2027	3042	5614	2216	5431
2028	3044	5612	2220	5602
2029	3045	5609	2222	5599
2030	3047	5607	2225	5597



The United Nations Decade of Action for Road Safety has set a fatality reduction target of 50% by the year 2020 (www.science.gov). The aim is to first stabilise and then reduce global fatalities by 2020. Our results indicate that the first part of this U.N goal is feasible, and would appear to have been largely achieved, all things being equal. The graph seems to show that the Decade of Action, which took off in 2011 started to bear fruits a few years later, with the number of fatalities stabilizing at about 50% of previous values. However, against the backdrop of decaying road infrastructure in Nigeria, and given the steady decline in the capital expenditure component of the national budget (less than 30%) even this modest first part of the twin goals of the United Nations may not be sustainable into the future in Nigeria. And with the modest first part of the goal compromised, the more ambitious second target of continuing to reduce fatalities might not realistically be within reach. In view of the foregoing, it seems safe to conclude that U.N policy targets of 50% fatality reduction by the year 2020, though largely achieved for now, may not continue into the future in Nigeria unless very drastic policy changes are put in place.

6. Measures of Predictive Accuracy

Since our model will produce an output given any input or set of inputs, we checked these estimated outputs against the actual values that we tried to predict. We call the difference between the actual value and the model's estimate a residual. We can calculate the residual for every point in our data set, and each of these residuals will be of use in assessment. These residuals will play a significant role in judging the usefulness of a model.Forecast accuracy is assessed within an Exponentially Weighted Moving Average (EWMA) model using Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE).

6.1. Root Mean Squared Error ($RMSE$): $RMSE$ is a quadratic scoring rule that also measures the average magnitude of the error. It's the square root of the average of squared differences between prediction and actual observation.$RMSE$ expresses average model prediction error in units of the variable of interest. The metric can range from 0 to ∞ and isindifferent to the

direction of errors. They are negatively-oriented scores, which means lower values are better.The formula for $RMSE$ is:

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(y_j - \hat{y}_j\right)^2} \quad (6)$$

where:

$$y_j = actual\ values\ for\ period\ j$$
$$\hat{y}_i = predicted\ values\ for\ period\ j$$

$j\ goes\ from\ 1\ to\ n, where\ n\ is\ the\ number\ of\ years$

6.2. Mean Absolute Error (MAE):MAE measures the average magnitude of the errors in a set of predictions, without considering their direction. It's the average over the test sample of the absolute differences between prediction and actual observation where all individual differences have equal weight. Effectively, MAE describes the typical magnitude of the residuals. The formula for MAE is:

$$MAE = \frac{1}{n}\sum_{j=1}^{n}\left|y_j - \hat{y}_j\right| \quad (7)$$

where:

$$y_j = actual\ values\ for\ period\ j$$
$$\hat{y}_i = predicted\ values\ for\ period\ j$$

$j\ goes\ from\ 1\ to\ n, where\ n\ is\ the\ number\ of\ years$

A small MAE suggests the model is great at prediction, while a large MAE suggests that your model may have trouble in certain areas.

6.3. Mean Absolute Percentage Error ($MAPE$): The mean absolute percentage error (MAPE) is the percentage equivalent of MAE. The equation looks just like that of MAE but with adjustments to convert everything into percentages**.** The higher the values of $MAPE$, the weaker the prediction performance and vice versa. The formula for $MAPE$ is:

$$MAPE = \frac{100}{n}\sum_{j=1}^{n}\left|\frac{y_j - \hat{y}_j}{y_j}\right| \quad (8)$$

where:

$$y_j = actual\ values\ for\ period\ j$$
$$\hat{y}_i = predicted\ values\ for\ period\ j$$
$j\ goes\ from\ 1\ to\ n, where\ n\ is\ the\ number\ of\ years$

6.4. Forecast Predictive Performance

The table 3 below shows the average$RMSE$, MAE and $MAPE$ statistics calculated from the Exponentially Weighted Moving Average (EWMA) model for each of the variables x_1, x_2, x_3 and x_4. The values have been obtained for $\lambda = 0.97$ (Mina and Xiao, 2001). The Mae measures the average size of the errors in the forecasts. As displayed in the table, the $RMSE$ is everywhere larger than the MAE, showing greater variation in individual errors in the sample, particularly for x_4 and x_2 which show greater difference.
While variablesx_3 and x_1 perform very well using $RMSE$to measure predictive accuracy compared to other two variables, variables x_1 and x_4 performance are great using MAE and $MAPE$ respectively as a measure predictive accuracy.

Table 3: Forecast Performance

Statistic	x_1	x_2	x_3	x_4
$RMSE$	33.61219	66.98928	30.95	1924.399
MAE	4.254718	6.430276	11.91635	7.904794
$MAPE$	0.999752	0.997673	0.999502	0.182666

7. Recommendations.

In order to reduce road traffic accidents appropriate traffic laws and regulations should be enforced. It is also recommended that necessary traffic infrastructural improvement should be put in place by relevant governmental agencies. The media should be also be actively engaged to inform and adequately educate the driving public.

References

Aggarwal, K.K. and Oberoi, S.S.(2009), Distribution of Fatal Road Traffic Accident Cases, *Journal of Punjab Academy of Forensic Medicine & Toxicology*, ISSN 0972-5687.

Annual Abstract of Statistics (2003-2007) Lagos: Federal Office of Statistics. www.oecollaborative.com.

Asogwa, S.E. and Obionu, CN (1985) Urban Pedestrain Accidents and the efficiency of counter measure. *Central Africa Journal of Medicine* 31: 62 -65. www.oecollaborative.com.

Behan, D.F. (2009), Statistical Credibility Theory, Presentation to the South eastern Actuarial Conference, June 2009.

Boland, P.J. (2006). Statistical and Probabilistic Methods in Actuarial Science. Chapman & Hall/CRC Interdisciplinary Statistics Series. Link.springer.com

Eke, N. (2001) Road Traffic Accidents in the Developing World: Who are liable? *Anil Aggrawals Internet Journal of Forensic Medicine and Toxicology* 2, 1. www.oecollaborative.com

Igboanugo A.C. and Ikhuemelo, E.F.(2007), Intervention Analysis of Road Traffic Accidents in Nigeria, Journal of Advanced Materials Research, 18-19, 375.www.safetylit.org.

Jilani, T.A. and Burney, S.M.A (2007), M-factor high-order fuzzy time series forecasting for road accident data, in IEEE-IFSA 2007, World Congress, Cancun, Mexico, June 18- 21.

Jilani, T.A., Burney, S.M.A. and Ardil, C. (2007), Multivariate high order fuzzy time series forecasting for car road accidents, International Journal of computational Intelligence, vol. 4, 1, 15-20. www.dspace.uni.lodz.pl:8080

Li , H. and Kozma, R. (2003), A Dynamic Neural NETWORK Method for Time Series Prediction Using the KIII Model, Proceedings of the 2003 International Joint Conference on Neural Networks, 1, 347-352.www.cyut.edu.tw

Li, S.F.S and Su, H. (2003), Neural Network Based Fusion of Global and Local Information in Predicting Time Series, Proceedings of the 2003 IEEE International Joint Conference on Systems, Man and Cybernetics, 5, 4445-4450.www.cyut.edu.tw

Mesken, J., Lajunen, T. and Summala, H.(2002) Interpersonal violations, speeding violations and their relation to accident involvement in Finland, *Ergonomics*, 45, 7, 469-483. www.tandfonline.com

Mina, J. and Xiao, J. (2001). Return to Risk Metrics. The Evolution of a Standard. Risk Metrics Group.

Mowbray, .H. (1914), How Extensive a Payroll Exposure is necessary to give a dependable Pure premium, Proceedings of the Casualty Actuarial Society, 1, 24-30. www.instruction.bus.wisc.edu

Norberg, R. (1980), Credibility Theory, Department of Statistics, London School of Economics, London WC2A 2AE, United Kingdom. www.isfaserveur.univ-lyon 1.fr

Norberg, R. (1986), Hierarchal Credibility: Analysis of a Random Effect Linear Model with Nested Classification, *Scandinavian Actuarial Journal*, 204-222.www.cambridge.org

Nwokoro, L. O (2005). Anti-Terrorism, Riot, Crowd, Crime Scene and Road Traffic Control, Aba: Ikenigwe Nigeria Ltd.

Orji, E.O., Fadiora, S.O., Ogunlola, I.O. and Badru, O.S. (2002) Road Traffic Accidents in Pregnancy in Southwest Nigeria: a 21-year review, *Journal of Obstetrics and Gynaecology*, 22, 5, 516- 518.

Sohn, S.Y. & Shin, H. (2001), Pattern Recognition for Road Traffic Accident Severity in Korea, *Ergonomics*, 44, 1, 107-117www.easts.info

Song, Q. and Chissom, B.S. (1993), Forecasting Enrolments with Fuzzy Time Series Part 1, *Fuzzy Sets and Systems*, 54, 1-9.www.cyut.edu.tw

United Nations (2007), Nigeria: Driving to Death, Integrated Regional information Networks. www.worldpressreview

Whitney, A.W. (1918), The Theory of Experience Rating, Proceedings of the Casualty *Actuarial Society*, 4, 274-292.www.isfaserveur.univ-lyon 1.fr