*Article*

**A Decade Long Slowdown in Road Accidents and Inherent Consequences Predicted for South Africa**

**Dimakatso Machetele and Kowiyou Yessoufou\***

Department of Geography, Environmental Management and Energy Studies, APK Campus, PoBox 526, Auckland Park, University of Johannesburg, South Africa

Dimakatso Machetele; dimakatsomachetele@yahoo.com

Kowiyou Yessoufou; kowiyouy@uj.ac.za

\*   Correspondence: <u>kowiyouy@uj.ac.za</u>

**Abstract:** Globally, there are 1.35 million road fatalities every year, which are estimated to cost governments approximately US$ 518 billion, making road fatalities the 8th leading cause of death across all age groups and the leading cause of death of children and young adults. In South Africa, despite tremendous governmental efforts to curb the soaring trajectory of road accidents, the annual number of road fatalities has increased by 26% in recent years. By fitting a structural equation model (SEM) and a GARCH Model (Generalized Auto-Regressive Conditional Heteroskedasticity) to analyze and predict future trend of road accidents (number of road accidents, number of casualties, number of fatal crashes and number of persons killed) in South Africa, we propose and test a complex metamodel that integrates multiple causality relationships. We show an increasing trend of road accidents over time, a trend that is predictable by number of vehicles in the country, the population of the country and the total distance travelled by vehicles. We further show that death rate linked to road accidents is on average 23.14 deaths per 100,000 persons and is predictable following the equation: $y = -0.0114x^2 + 1.2378x - 2.2627$ ($R^2 = 0.76$) with y = death rate and x = year. Finally, in the next decade, the number of road accidents is predicted to be roughly constant at 617,253 accidents but can reach 1 896 667 accidents in the worst-case scenario. The number of casualties was also predicted to be roughly constant at 93 531 over time, although this number may reach 661 531 in the worst-case scenario. However, although the number of fatal crashes may decrease in the next decade, it is forecasted to reach 11 241 within the next 10 years with the worse scenario estimated at 19 034 within the same period. At the same time, the number of persons killed in fatal crashes is also predicted to be roughly constant at 14 739 but may also reach 172 784 in the worse scenario. Overall, the present study reveals perhaps the positive effects of government initiatives to curb road accidents and their consequences; we call for more stronger actions for a drastic reduction in road accident events in South Africa

**Keywords:** Road crashes; fatalities; casualties; persons killed in road accidents; South Africa

## 1. Introduction

Road accidents can be defined as events on the road that involve the collision of either two or more vehicles, or a vehicle and a vulnerable road user (cyclists or pedestrian), or a vehicle and a fixed object, e.g., bridge (RTMC, 2008). According to the leading road safety agency in South Africa, road accidents are classified into four categories of severity, namely: fatal crashes, major crashes, minor crashes, and damage-only crashes (RTMC, 2017). Fatal crashes result to death of one person or more; such crashes may re-

sult in serious and light injuries (RTMC, 2008). Major crashes are defined as crashes in which one person or more people are seriously injured (RTMC, 2017). Minor crashes are crashes in which one or more persons are slightly injured (RTMC, 2017) whereas damage-only crashes are crashes in which no one has been injured or killed but vehicles or property may be damaged (RTMC, 2008).

Globally, there are 1.35 million road fatalities every year, which are estimated to cost approximately US$ 518 billion to governments (WHO, 2018). Existing statistics indicate that more than 90% of road accidents are fatal in both low- and middle-income countries, and these fatal crashes often involve more than 50% of unregistered vehicles (WHO, 2009). The Global Status Report on Road Safety reported road fatalities to be the 8th leading cause of death across all age groups (Donaldson *et al.*, 2009) and are now the leading cause of death of children and young adults aged 5 to 29 years (RTMC, 2019).

In response to these crashes, several initiatives have been taken by different governments across the world; however, the frequency of road accidents remains on the rise with unacceptable consequences (Benlagha and Charfeddine, 2020). For example, existing statistics indicate that more than 90% of road accidents are fatal especially in developing countries, and these fatal crashes often involve more than 50% of unregistered vehicles (WHO, 2009). The increasing ownership of vehicles is one of the major contributory factors to the rise of road fatalities and injuries in developing countries (Nantulya and Reich, 2002).

In South Africa and in the past few years, an increasing trend of road fatalities has been observed (RTMC, 2018), although the country made strides in reducing road fatalities since their peak in 2006 (ITF, 2018). The annual number of road fatalities increased every year between 2013 and 2016 (RTMC, 2017) and, between 1990 and 2017, they increased by 26 % (ITF, 2018). The high number of road accidents and their associated consequences have a significant negative impact on the socioeconomic development of all South Africans (Labuschange *et al.*, 2017). Impacts are measured in terms of loss of lives, grief and suffering as well as the heavy financial burden of road accidents on the country's economy (Verster and Fourie, 2018). The primary underlying factors of road accidents in South Africa have been identified (Verster and Fourie, 2018). These factors are linked to vehicle (7.8%), environment and road (12%) as well as human (80%) (Verster and Fourie, 2018). In 2017 alone, the factor "vehicle" contributed 3% to road accidents in the country; environmental and road conditions 5% and human factors were responsible for 91% of road accidents (RTMC, 2018). However, still in the South African context, we still have limited understanding of the complexity of road accident events as several potential mediators of road accidents are generally not factored in the analysis of road accidents in the country. In the present study, a more complex approach is employed to analyse potential causality relationships between several variables linked to road accident events.

In the face of this alarming road accident statistics in South Africa, existing studies on road safety in the country focus mainly on identifying contributory factors, e.g., human, vehicle, and environmental factors, to road accidents (e.g., Verster and Fourie, 2018). Whilst other studies indicated that road accident events may be more complex than thought (e.g., Eboli and Mazzulla, 2007), we have limited understanding of this complexity in South Africa, owing to the simplistic approach we generally take in analysing road accident events.

The present study proposes and tests a more comprehensive metamodel (Figure 1) formulated by integrating multiple causality relationships among variables previously linked to road accidents. In the proposed theoretical metamodel, road accidents are represented by four main response variables, including total number of crashes (road accidents), total number of casualties, number of fatal crashes and number of persons killed each year. Our first prediction is that the total number of crashes would have a cascading effect on the other three main variables such that, as the total number of crashes increases over time, so too would the total number of casualties, number of fatal crashes and number of persons killed in a cascading manner.

In addition, the metamodel includes six predictor variables. These include total number of vehicles in the country at a given year, number of registered vehicles, number of unroadworthy and unlicensed vehicles, number of driver's license issued, the population of the country and the distance travelled by vehicles each year. Our prediction is that, as the total number of vehicles in the country increases over time, this would increase the likelihood of more crashes, which, as indicated above, would have a cascading effect on the total number of casualties, number of fatal crashes and number of persons killed (positive relationship; Figure 1). Similar to the total number of vehicles, we also predict that, if the number of unroadworthy and unlicensed vehicles increases over time, this would also increase the total number of crashes. This is grounded on the assumption that, not only the drivers of unroadworthy and unlicensed vehicles are more likely to not have a driver's license, but their unroadworthy vehicles are more likely in a defective state conducive for road accidents. To illustrate this, Verster and Fourie (2018) reported that 7.8% of road accidents in South Africa are linked to vehicles, and in 2017 alone, the factor "vehicle" contributed 3% to road accidents in the country (RTMC, 2018). Again, an increase number of unroadworthy unregistered vehicles would thus increase the risk of more road crashes, which would have a cascading effect on the total number of casualties, number of fatal crashes and number of persons killed over time. Furthermore, we predict the increase in human population would also result in the increase in road crashes and thus the increase in total number of casualties, number of fatal crashes and number of persons killed. The rationale of this prediction is that when the population increases, the number of vehicle owners would increase such that more vehicles (both registered and unregistered) would be on the road, thus increasing the risk of road crashes. Also, when more vehicles are on the road, the total distance travelled by vehicles would increase, and this would reduce the quality and strength of the vehicles over time, and thus increasing the risk of road crashes.

As opposed to the positive relationships predicted above, we predict that the number of registered vehicles as well as the number of issued drivers' license would cause a reduction in the number of road crashes, the total number of casualties, number of fatal crashes and number of persons killed. This is grounded in the assumption that the drivers of registered vehicles are more likely to have a driver license, and in theory, a licenced driver is relatively less likely to cause road crashes than non-licenced drivers. This leads to our last assumption, which predicts that the more driver's licenses issued, the less likely the risk of more road crashes than the other way around. All these hypotheses/assumptions are graphically represented in Figure 1, which represents the metamodel proposed and tested to explain the complexity of road accident events in South Africa.

The aim of the study is to explain the complexity of road accidents in South Africa. The following objectives are set for the study: i) to elucidate the temporal trends of road accidents and related consequences (number of crashes, causalities, fatal crashes, and deaths) in South Africa, ii) to propose and test a model to explain the complexity of road accidents in South Africa, iii) to predict future trends of road accidents and related consequences (number of crashes, casualties, fatal crashes, and road fatalities) in South Africa.

## 2. Materials and Methods

### 2.1. Study area

The study site is South Africa, a country located right at the southern tip of the African continent (Tibane *et al.*, 2016). South Africa stretches from 22°S to 35°S in latitude and 17°E to 33°E in longitude (Tibane *et al.*, 2016). The country is bordered by Namibia, Zimbabwe, Mozambique, and Botswana (Thompson *et al.*, 2019). The Atlantic Ocean and the Indian Ocean are the two coastlines found in the southwest and southeast direction, respectively. The warm temperature conditions in South Africa is due to its location at the subtropical region, moderated by the ocean currents surrounding the country's

coastlines and the altitude of the inferior plateau (Thompson *et al.*, 2019). In 2019, the mid-year population was estimated at 58,78 million and approximately 51,2% (30 million) of the population is female (Stats SA, 2019).

South Africa is ranked 10th in the world in term of the volume of road network estimated at 750 000 km (Tibane *et al.,* 2016). The South African National Road Agency (SANRAL) is responsible for 21 403 km of the national road network (Tibane *et al.,* 2016). Out of the 21 403km, 18 283 km (85%) are non-toll roads and 3 120 km (15%) are toll roads (DoT, 2002). SANRAL's role is to provide effective strategic road infrastructure to assist with the development, accessibility, and finance towards the proclaimed national roads (Tibane *et al.*, 2016). The country's toll-road network consists of approximately 19% (3 120 km) of the national grid and SANRAL manages 1832km of these toll roads (DoT, 2002). In 2017, the number of registered vehicles in South Africa were 12,2 million, and the number of un-roadworthy and unlicensed vehicles had increased from 1,0 million vehicles in 2016 to 1,1 million vehicles in 2017 (RTMC, 2018). The Gauteng province accounted for approximately 40% of the registered vehicle population in the country (RTMC, 2018).

### 2.2. Data Collection
### 2.2.1. Various variables for which data were collected
The variables for which data were collected are defined as follows:

'Total number of vehicles' is the total sum of the number of registered vehicles and both the number of unroadworthy and unlicensed vehicles.

'Number of registered vehicles' is the number of (motorised and towed) vehicles registered on the National Traffic Information System (NaTIS) (RTMC, 2018).

'Unregistered vehicles' are vehicles which are not registered under the licensing department. The owners of unregistered vehicles might have failed to renew the vehicles' licenses or failed to submit the vehicles for compulsory annual roadworthy tests within a certain period (RTMC, 2008).

'The total distance travelled' variable is defined as the total distance travelled by vehicles on the road each year measured as million vehicle kilometres (mvk) (RTMC, 2018).

'Drivers licenses issued' is the annual total number of issued driving licenses obtained by drivers after they have passed their driver's license test (RTMC, 2019).

'Estimated population of South Africa' is the estimate number of all residents of South Africa in a particular year (mid-year point) and based on the latest information (Stats SA, 2017).

'Total number of crashes' is the number of road accidents (fatal, major, or minor crashes) that occurred in each year (RTMC, 2018).

'Fatal crashes' are crashes that result to death of one person or more people in a road accident (RTMC, 2018).

'Number of casualties' is the number of people that have been injured and/or killed in a crash (RTMC, 2019).

'Fatalities' is defined as person or people that are killed during or immediately after a road accident, or death within 30 days after the accident has occurred as a direct result of the road accident (RTMC, 2019).

### 2.2.2. Sources of data analysed in this study
Data analysed in the present study are retrieved from various sources. These sources include the Government's online campaign database known as *Arrive Alive* (www.arrivealive.co.za), various reports of the Road Traffic Management Corporation (RTMC,2006; RTMC, 2007; RTMC, 2011; RTMC, 2012; RTMC, 2013; RTMC, 2014; RTMC, 2015; RTMC, 2016; RTMC, 2017; 2018) as well as reports from *Statistics SA* (Stats SA, 1995; Stats SA, 2007; Stats SA, 2008; Stats SA, 2009; Stats SA, 2010; Stats SA, 2011; Stats SA, 2012; Stats SA, 2013; Stats SA, 2014; Stats SA, 2015; Stats SA, 2016; Stats SA, 2017). Data from *Arrive Alive* (period of 1935 to 2000) was obtained by email request. The data from the

RTMC (period of 1935 to 2017) was obtained from the corporation's website; and the data from *Statistics SA* (period of 1936 to 2017) was retrieved from their website as well as by email request.

### 2.2.3. *How were data collected by various sources?*

*Arrive Alive* is the South African government's campaign to promote road safety and public awareness. They make use of data from both the RTMC and Stats SA and compile their reports for public use or research. The RTMC is the leading agency on road safety and management (RTMC, 2018). They collect their data from the South African Police Service (SAPS), Provisional Traffic Authorities and the Metropolitan Municipalities through Accident Report and Quick Response forms (RTMC, 2015). The Accident Report forms are used to capture data for all road accident types (injuries and damages) while the Quick Response forms are used to record fatal road accidents and are captured by RTMC. Furthermore, the RTMC makes use of other data sources such as the Culpable Homicide Crash: Observation Report (CHoCOR) form, CAS Analyst Report, National Traffic Information System (NaTIS), Statistics South Africa (Stats SA) (RTMC, 2018). The limitation to the methodology the RTMC used to collect their data is that road traffic information on their reports are mainly based on fatal crashes only (RTMC, 2018). While reporting and capturing road accident, data sometimes are duplicated (RTMC, 2013). The process of verifying the data is a lengthy process, which involves comparing the initial data that was received from various police stations at the time of the crash with that received from the central SAPS database (RTMC, 2013). Another limitation is that the consolidated inputs from provinces may not be received on time during the period of data collection for the specific report (RTMC, 2013). There is still a need for the leading road traffic agency to conduct in-depth research to collect scientific based facts to complement their administrative data (RTMC, 2018).

Statistics South Africa (Stats SA) is the national statistical service of South Africa (Stats SA, 2019). For their mid-year population estimates of South Africa, they make use of the cohort-component method (Stats SA, 2007). In the cohort-component methodology, a base population is estimated that is consistent with the known demographic characteristics of the country (Stats SA, 2017). The mid-year population estimates are produced by making use of the Spectrum model developed by the Future Group alongside with UNAIDS, WHO and UNICEF (Stats SA, 2013). In 2011, Stats SA made use of census for population estimate. Census is a procedure used to collect basic information on the population and housing statistics of a country for socioeconomic development, creation, and the implementation of policies (Stats SA, 2012). South Africa has conducted three Censuses (1996, 2001 and 2011) (Stats SA, 2012). In addition, Stats SA follows the standards of the International Monetary Fund's (IMF) Special Data Dissemination Standard (SDDS) to publish the mid-year population estimates annually (Stas SA, 2007). The limitation with the methodology Stats SA applies to collect data on the mid-year population estimates for South Africa is that the estimates may change as new data becomes available (Stats SA, 2017). The population estimates are accompanied by a series of revised estimates from the period of 2002-2017 (Stats SA, 2017).

### 2.3. *Data Analysis*

All analyses were done in R (R Development Core Team 2017), and R script used is provided as supplemental information.

Despite the efforts deployed to consult as many sources as possible, there are still some missing values in the dataset. Traditionally, prior to analysis, most statistical packages use listwise deletion to remove entire rows that have missing values; this practice leads to the loss of information sometimes critical for the understanding of the research question at hand or the hypothesis to be tested. To retrieve the information encapsulated in missing values in order to generate a comprehensive dataset, missing values have been imputed. This imputation was done as implemented in the R package *Amelia* (Honaker et al. 2011). Five different imputations were done concurrently on dif-

ferent computers and then combined in a single dataset. The means of those five imputed data were used for the analysis. The complete dataset is presented in Supplemental Table S1. In addition, during the analysis, because variables are not in the same scale (e.g., population in million vs. fatal crashes in hundred), variables were first re-scaled as follows: scaled variable = (observed-mean)/ standard deviation.

To test the adequacy of the proposed metamodel to explain the complexity of road accidents in South Africa, a structural equation model (SEM) was fitted to the imputed data collected. Structural equation modelling (SEM) is a multivariate and powerful technique used to test and evaluate multivariate causal relationships (Fan 2016). The benefits of an SEM is that various causal relationships can be defined and tested simultaneously in one SEM. Each of the relationships or paths in the SEM was translated into GLM models (Generalized Linear Model) with appropriate error structure, depending on the nature of the response variable.

In the present study, four main response variables were defined. These variables include total number of crashes, total number of casualties, number of fatal crashes and number of persons killed. These variables are all 'count data', meaning that the appropriate link function to specify in model fitting is 'negative binomial' (O'Hara and Kotze 2010), given that count data should not be log-transformed (O'Hara and Kotze 2010). Six explanatory variables were defined: total number of vehicles, number of registered vehicles, number of unroadworthy and unlicensed vehicles, number of issued drivers' licenses, population of the country and vehicle distance travelled. Since most of the response variables in the metamodel are 'count data' except the variable 'distance travelled', the negative binomial GLM was fitted. GLM model with quasi-poison error family was appropriate to model 'count data' but to avoid overdispersion the negative binomial GLM was better fitted as demonstrated by O'Hara and Kotze (2010). The GLM model was fitted to model the variable 'distance travelled' rather by specifying the gaussian family of error. All the GLM models were then combined in SEM which was fitted as implemented in the R library *piecewiseSEM* (Lefcheck, 2016). The adequacy of the metamodel was tested based on its overall Goodness-of-fit (C value) and the P value. C values are an indicator of whether an SEM is good enough to explain the data. To tell if an SEM is good enough, the lower the C value the better. Another key parameter of the goodness-of-fit is the P-value of the test of goodness-of-fit. The P-value here tests whether the proposed SEM is different from the best fit to the data at hand. If P<0.05, this means that the proposed SEM is significantly different from the best fit and therefore should be rejected; if P>0.05, this means that the SEM is no different from the best fit to the data and therefore can be used to explain the data (Schermelleh-Engel et al. 2003; Lefcheck 2016).

Prior to all these analyses, multicollinearity among predicting variables are first tested. Unlike in simple regressions, there is no straightforward way to deal with the problem of multicollinearity in SEM. Some authors even argue that multicollinearity is not an issue in SEM (e.g., Maruyama 1998; Malhotra et al. 1999; Verbeke and Bagozzi 2000). This is based on the ground that "if highly correlated variables can be regarded as indicators of a common underlying construct, multicollinearity problems can be avoided" (Grewal et al. 2004). Other authors argued that multicollinearity is a source of unreliable estimates of SEM parameters (e.g., Jagpal 1982, Grapentine 2000; Tarka 2018). There is no doubt that multicollinearity poses a problem in parameter estimates but the problem is particularly complex with SEM for the simple reason that, in SEM, an independent variable in one of the models that form the SEM can be a response variable in another model within the same SEM and *vice versa*.

In the present study, the following is how the multicollinearity problem was dealt with. For each of the models in the SEM, the variance inflation factor (VIF) was calculated in R and predicting variables with VIF>5 (rules of thumb) were considered as highly collinear. Then, using the stepwise selection techniques, collinear variables were removed from the model starting from the variable with the highest VIF. After removing a collinear variable, the model is rerun, and the goodness-of-fit and P values calculated. This stepwise selection was repeated until the model is made up of only variables with VIF<5.

This process was followed for all the nine models that form the SEM. At the end, the SEM was reconstructed using only models with non-collinear variables and then was rerun. The AIC value for this SEM is: AIC=1788.486. The global goodness-of-fit for this SEM in which no model contains collinear variables is as follow: Fisher's C = 1718.486 with P-value = 0 on 56 degrees of freedom. For any SEM, P>0.05 means the SEM is of a good fit for the data while P<0.05 means that the SEM departs significantly from the data analysed (see Lefcheck 2016). In the present case, P=0, meaning that the SEM with only non-collinear variables could not be used to explain the data.

In that case, we started bringing back (into each of the models in the SEM) each of the variables initially removed from the model because of collinearity. In this case, only missing variables that show significant relationships with each response variable in the SEM were brought back into the model, but this was done using one missing variable at a time (stepwise process). Each time, the goodness-of-fit and P value were calculated. It is only when all significant variables were brought back into the SEM that an SEM with P>0.05 was found, indicating that a suitable SEM to explain the data is found.

Furthermore, the $R^2$ for each of the models in the SEM was calculated in scenario 1 where only non-collinear variables were included, and in scenario 2 where the final SEM is constructed with all previously missing but significant variables. This is summarized in Supplemental Table S2. In this table, $R^2$ values do not change substantially in SEM with collinear variables in comparison with SEM without collinear variables, except for variable "driver licence". This means that, although missing variables share similar information with other variables in each of the models in the SEM (that is why they are collinear), they also bear unique or complementary information (that is why it is only when they are added to the SEM that an SEM with P>0.05 is found). These findings support our inclusion of all significant missing variables into the SEM.

The future of road accident in the country was predicted by fitting the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) to the time-series data collected (Table S1). GARCH model was selected to account for the expected high volatility of the variance in the time-series data of road accident (Bollerslev, 1986). The key assumption behind GARCH model is that variance is volatile or is not constant over time in a time series dataset. To test that this assumption is met, the Dickey-Fuller test (Fuller 1976) was run to test the null hypothesis that a unit root is present in an autoregressive model of each of the time series data, and that the process is thus not stationary (non-constant variance over time). The results confirm the null hypothesis, i.e., P>0.05 for each response variable (fatal crashes, P=0.20; persons killed, P=0.09; casualties, P=0.51 and crashes, P=0.99), thus justifying the use of GARCH in this study. GARCH model was fitted as implemented in the R library *rugarch* (Ghalanos, 2020). Different variances of GARCH models were fitted depending on the starting time intervals for the modelling. The best GARCH model was selected using Akaike Information Criteria (AIC). This model was used to forecast future of road accidents in the next 10 years.

## 3. Results

### 3.1 Temporal trends of road accidents and related consequences

The data analysed include missing values for some of the variables; these missing values represent 20% of all values across all variables, i.e., 80% of all values are present in the dataset (Figure S1). For further analysis, the missing values have been imputed such that the findings reported below are based on a complete dataset (i.e., without missing values). The analysis shows that the number of crashes (road accidents) increases steadily over time since 1935 but exhibits important variations from the year 2000 (Figure 2). A similar trend is found for the number of casualties, and interestingly, there are always less casualties than crashes (Figure 2): roughly 32% of crashes on average results in casualties (Table S1).

Some of these crashes are fatal, and as can be expected from the patterns in Figure 3, the number of fatal crashes and the number of persons killed in fatal crashes increase too

over time with a sharp increase from the year 2000. Unfortunately, there have always been more deaths (persons killed) than the number of fatal crashes (Figure 3) such that, on average, the number of persons killed represents 1,15 times the number of fatal accidents and 0,03 times the total annual number of road accidents (Table S1).

The death rate follows a polynomial shape (Figure 4) predictable following the equation: $y = -0.0114x^2 + 1.2378x - 2.2627$ ($R^2 = 0.76$) with y = death rate and x = year. This trend can be broken down into four noticeable periods (Figure 4): between 1935 and 1938, the death rate per 100,000 people increased from 6 to 9.5; from 1938 to 1945, there is a decrease from 9.5 to 4.18; the 1945-1972 period witnessed an exponential increase from 4.18 to 37.88 deaths per 100,000 people; this sharp increase has been slowing down since 1972 until 2017 where it decreased from 37.88 to 24.85 and reached the lowest rate of 19.60 in the year 2000 (Figure 4). These trends result in an average death rate of 23.14 deaths per 100,000 people.

### 3.2.    Understand the complexity of road accidents in South Africa
### 3.2.1 Structural equation model of road accidents

The metamodel represented in Figure 1 was tested.  The following parameters of the metamodel suggest that the metamodel is suitable to explain the complexity of the mechanism driving road accidents: Fisher's C = 3.895 with P-value = 0.143 on 2 degrees of freedom.

The coefficients of each of the models in the metamodel are summarized in Table 1. This Table shows that the number of crashes could significantly be explained by total number of vehicles ($\beta = -0.32 \pm 0.06$,   P<0.001), number of registered vehicles ($\beta = 1.76 \pm 0.12$,   P<0.001), number of unregistered vehicles ($\beta = 0.14 \pm 0.04$,   P=0.003) and the population of the country ($\beta = -0.88 \pm 0.13$,   P<0.001). As opposed to expectation, the number of driver licences issued, and total distance travelled by vehicles do not correlate significantly with number of crashes (P>0.05).

Furthermore, the analysis reveals that the number of casualties could be linked significantly to number of registered vehicles ($\beta = 1.44 \pm 0.0.34$, P<0.001) and total distance travelled by vehicles ($\beta = -0.04 \pm 0.02$, P=0.03). None of the other variables tested correlates significantly with number of casualties (P>0.05).

As of the number of fatal crashes, the analysis reveals that the total number of vehicles ($\beta = -0.79 \pm 0.09$,   P<0.001), number of registered ($\beta = 3.73 \pm 0.23$,   P<0.001) and unregistered vehicles ($\beta = 0.22 \pm 0.06$,   P<0.001), the population of the country ($\beta = -3.04 \pm 0.18$,   P<0.001) and the total distance travelled by vehicles ($\beta = 0.20 \pm 0.02$,   P<0.001) correlate significantly with the number of fatal crashes. However, the number of casualties and again the number of driver licence do not seem to determine the number of fatal crashes (P>0.05).

Finally, the findings of the SEM analysis reveal that the total number of vehicles ($\beta = 0.20 \pm 0.06$,   P=0.0017), number of unregistered vehicles ($\beta = -0.10 \pm 0.03$,   P=0.003), number of crashes ($\beta = -0.14 \pm 0.06$,   P=0.04), number of casualties ($\beta = 0.32 \pm 0.04$, P<0.001), distance travelled ($\beta = 0.03 \pm 0.01$,   P=0.009) and obviously the number fatal crashes ($\beta = 0.80 \pm 0.05$,   P<0.001) are significant determinant of the number of persons killed in road accidents (Table 4.1).

### 3.3. Expected trends of road accidents for the future

Using the GARCH model, the future of road accidents was predicted. The number of crashes or accidents is predicted to be roughly constant over time at 617,253 accidents for the next 10 years, with the worse scenario suggesting that this number may reach 1,896,667 (Figure 5A). The number of casualties was also predicted to be roughly constant at 93,531 over time although this number may reach 661,531 in the worst-case scenario (Figure 5B). However, although the number of fatal crashes may decrease over time, it is forecasted to reach 11,241 fatal crashes within the next 10 years with the worse scenario estimated at 19,034 within the same period (Figure 5C). Finally, the number of persons

killed is also predicted to be roughly constant at 14,739 but may also reach 172,784 in the worse scenario (Figure 5D).

### 4. Discussion

#### 4.1. Temporal trajectory of road accidents in South Africa

Over the 82 years (1935-2017) period covered in this study, the findings reveal an increasing trend in number of reported road accidents, casualties, fatal crashes, and deaths. The study also reveals that, unfortunately, 32% of accidents lead to casualties, and when these are fatal, the ratio persons killed: fatal accidents (referred to as accident severity rate in Verster and Fourie, 2018) is 1,15. In their recent study, Verster and Fourie (2018) estimated the severity rate at 1,2-1,3 between 2010 and 2015, a rate similar to what is revealed in the present study, although our present study covers a longer period of time than theirs. In any case, these findings suggest that, overall, in any fatal accident in South Africa, more than 1 person lose their lives. Verster and Fourie's (2018) further showed that the severity rate is not uniform within a week, with the highest severity rates (1,22-1,25) found for Mondays, Fridays, Saturdays, and Sundays (South African Department of Transport, 2015; Verster and Fourie, 2018).

However, the number of road accidents fluctuated more notably from the year 2000 onward, resulting into a sharp increase of fatal crashes and the number of persons killed from the year 2000. This year 2000 corresponds to the period where 75% of road accidents recorded in South Africa were linked to drivers driving under the influence of alcohol (Ncube *et al.*, 2016). In 2000, the cost of road traffic accidents in South Africa was estimated at R13.8 billion (US$2 billion) (Ministry of Transport, 2001). South Africa is the world leading country in drunk driving followed by Canada and the US (McCarhy, 2016). The number of drunk drivers has increased by four times since 1980. It is therefore not surprising that we observe the notable fluctuations of road accident occurrence as well as the increase in fatal crashes and number of persons killed from 2000 onward. Surprisingly, it is in the same year 2000 that the lowest death rate of 19.60 was observed. This implies that the increase in fatal crashes did not result in high death rate after correcting for population size.

Specifically, the death rate increased drastically from 6 death per 100,000 persons in 1935 to 37.88 deaths per 100,000 people in 1972 before it decreased gradually to 24.8. Although the decrease observed in the later stage of the trajectory of death rate seems to be a global trend as this is too reported in several parts of the world, including Australia, Poland, China, etc. (WHO, 2018). Efforts need to be continuously deployed to maintain this decreasing trend because, after a long period of safety improvement, the risky behaviours leading to the re-increase in the number of road accidents may occur again (Bąk *et al.,* 2019) as predicted in the behavioural theory (Wilde, 1988).

The overall trend found in the present study is similar to the one reported, also for South Africa, by WHO (2009, 2018), which identified four key points in the trajectory of death rate. They reported that, from 1972 to 1992, the death rate fluctuated within a window of 25-40 deaths/100,000 people. They also indicated that this rate decreased to 20 death/100,000 people from 1992 to 2002, and then increased again to 35 deaths/100,000 people from 2002 to 2007 before it decreases to 21 death/100,000 people from 2007 to 2013 and then a slight increase to the present day. The similarity between the WHO's trend and what is reported in the present study suggests that the data imputation done in the present study did not cause significant bias in the data analysed.

On average, the study reveals a death rate of 23.14 deaths per 100,000 people. This death rate is lower than that reported for low-income countries (27.5 deaths per 100,000 population) and for the African continent (26 deaths/100,000; WHO, 2018) but is similar to what is reported for Brazil in the last 10 years (WHO, 2018). However, it is three times the rate reported (8.3 deaths per 100,000 population) for developed world (WHO, 2018) and 4-times the 6 deaths/100,000 rate reported for Australia (Mooren *et al.*, 2011). In this developed world, the death rate has been slowing down since the 1960s: 27% decrease in

the US and 63% decline in Canada from 1975 to 1988 while an increase was noted in developing countries (an increase by 44% in Malaysia and 243% in China; Kopits and Cropper, 2003). Although we are losing globally 1,35 million people annually owing to road accidents, the youth (5-29 years old; WHO, 2019) is disproportionately the most affected. The consequences of these losses are significant: heavy deterioration of standard of living for family (Silcok, 2003), the loss of 1-2% of the gross national product of developing countries (Peden and Hyder, 2002). To South Africa, it costs 3.4% of its GDP (South African Department of Transport, 2016). This is unaffordable in a country where basic service delivery to community is still a problem, and therefore calls for vigorous actions. These actions may include enforcement of seat belt safety laws, enforcement of speed limits, and strict and unapologetic punishment for drunk driving.

## 4.2. The proposed metamodel for road accidents in South Africa and future trends

The parameters of the metamodel tested show that the metamodel in Figure 1 is suitable to explain the complexity of the mechanism driving road accidents. As predicted in the theoretical framework in Figure 1, the number of crashes is indeed predicted by the total number of vehicles but contrary to prediction, the relationship is negative. This negative relationship means that the increase in the number of vehicles do not lead to the increase in the number of crashes but the later actually decreases. This may imply that, while the number of vehicles is increasing, these vehicles may be in majority good roadworthy cars driven by licenced and careful drivers such that this increase in the number of cars does not result in an increase in crashes. For the next 10 years, the GARCH model employed in this study predicted a roughly constant number of crashes or road accidents (617,253 accidents), but this number may triple, in the worst-case scenario, reaching 1,896,667. This calls for vigorous actions against the worst scenario.

Similarly, the results of this study show a negative relationship between the population of the country and the number of crashes. The initial expectation is that the increase in human population would also result in the increase in road crashes (Al-Reesi *et al.*, 2013). The rationale for this is that when the population increases, the number of vehicle owners would increase such that more vehicles (both registered and unregistered) would be on the road, thus increasing the risk of road crashes. Also, when more vehicles are on the road, the distance travelled by vehicles would increase, and this would reduce the quality and strength of the vehicles over time, and thus increasing the risk of road crashes. However, the total distance travelled by vehicles does not correlate significantly with number of crashes, discounting the importance of this variable. The finding reported in the present study matches what has been found for other countries, such as UK where, in the last 30-years, although the population size has grown by 15%, road fatalities have fallen by 68% (Department for Transport, 2015).

The counterintuitive finding of negative relationships between number of population and number of crashes could mean that the increase in population is followed by an increase in careful and licenced drivers. However, the part of the explanation linked to drivers licence is discounted by the finding that, as opposed to expectation, the number of driver licences issued does not correlate significantly with number of crashes, putting into question the contribution of drivers licence to good driving behaviour on the road. In Botswana, unlicensed drivers, or failure to produce a licence are found to correlate significantly to road fatalities (Mphela, 2011). Interestingly, matching the theoretical prediction, the number of crashes increases when the number of both registered and unregistered vehicles also increases. A potential explanation is that irrespective of whether vehicles are registered or unregistered, too many vehicles on the road predisposes to an increase of the likelihood of accidents. Unregistered vehicles are likely to be unroadworthy and faulty; vehicle roadworthiness was reported as a major contributing factor to road accidents (Bayam *et al.*, 2005; Verster and Fourie, 2018; WHO, 2018). Verster and Fourie (2018) reported that 7.8% of road accidents in South Africa are linked to vehicles,

and in 2017 alone, the factor "vehicle" contributed 3% to road accidents in the country (RTMC, 2018).

Furthermore, as predicted for number of crashes, the expectation is also that the number of casualties would correlate positively and significantly with all six predictors analysed in the study. The findings in this study reveal that the number of casualties indeed correlate positively and significantly but with only the number of registered vehicles and negatively with total distance travelled by vehicles. The positive relationship is an indication that, as the number of registered vehicles increases, more vehicles are on the road, heightening the risk of accidents thus casualties. The negative relationship between casualties and distance travelled is unexpected and means that more vehicles on the road do not increase road casualties.   Again, the GARCH model predict a roughly constant number of casualties for the next 10 years (93,531) but this may increase 6-fold in the worst-case scenario.

As expected, the analysis reveals that the total number of vehicles, number of registered and unregistered vehicles, the population of the country and the total distance travelled by vehicles correlate each significantly with the number of fatal crashes. However, the number of vehicles and the population of the country which were predicted to correlate positively with the number of fatal crashes show instead negative correlations. This is in contrast with the reports from other studies that reported the positive relationships predicted in the theoretical metamodel in the present study (Jacobs and Cutting, 1986; Aderamo, 2012). The contrasting findings in the present study mean that, as the total number of vehicles and the country's population size increase, the number of fatal crashes decreases.

On a global scale, Jacobs and Aeron-Thomas (2000) predicted the road fatalities to reach between 900,000 and 1.1 million in 2010 and between 1 million and 1.3 million in 2020. For the next 10 years, the GARCH model predicted a decreasing trend of the number of fatal crashes but this number may reach 11,241-19,034 in South Africa. Again, the observed decrease in fatal crashes is unexpected and could imply that one or more key variables that decrease the number of fatal crashes while the total number of vehicles and population size are increasing are not taken into consideration in the proposed metamodel in the present study. An example of such variable could be road policy enforcement or improvement, improvement of drivers' consciousness or behaviour that may cause a decline in the frequency of fatal crashes (Shaw and McMartin, 1977; Farmer *et al.,* 1997; Delen *et al.,* 2006; Deffenbacher *et al.,* 2002; Sârbescu and Maracatu, 2019; Papantoniou *et al.,* 2019). Further variables, focused on the drivers' characteristics such as age, gender, driving experience, and education level that are not considered in the present study may be important too as reported in previous studies (Bayam *et al.,* 2005; Eboli and Mazulla 2007; Papantoniou *et al.,* 2019; Papantoniou *et al.,* 2018). Nevertheless, several studies called for road accidents to be regarded as more driven by a failure of a system and not just of a person (the driver) (Reason, 2000). In that system is a complex of interactive factors and an accident is the consequence of a failure of the interactions (Reason, 2000; Bayam *et al.,* 2005; Eboli and Mazulla, 2007; Aderamo, 2012; Hassan and Abdel-Aty, 2013, Bergel-Hayat *et al.,* 2013; Theofilatos, 2017; Lee, *et al.,* 2018; Verster and Fourie 2018; Papantoniou *et al.,* 2019; Papantoniou *et al.,* 2018).

In the present study, six interactive factors/predictors are tested as linked to road accidents and their effects (casualties, fatal crashes, number of persons killed, etc.). Counterintuitively, the number of casualties and the number of driver licences do not correlate significantly with the number of fatal crashes. The lack of significant correlation between number of casualties and number of fatal crashes disproves the cascading effects initially predicted in Figure 1, implying that an increase in road casualties does not necessarily lead to an increase in road fatalities. Again, this could be linked to a potential improvement of drivers' consciousness or behaviour on the road. The lack of correlation between the number of drivers' licence issued and number of fatal crashes dismiss the initial prediction that an increase in number of licences issued should mirror an improvement of driving behaviour or experience (da Silva *et al.,* 2014) and thus limit the

number of crashes/accidents and all inherent consequences (e.g., fatalities). This finding may not be that surprising given the corruption around driver's licence delivery (Arrive Alive, 2020).

Finally, as expected, the total number of vehicles correlate significantly and positively with the number of persons killed: more cars mean high risk of fatalities. One relationship not predicted in Figure 1 is the positive significant correlation found between number of casualties and number of persons killed. In the predicted theoretical metamodel, it was suggested an indirect link between number of casualties and number of persons killed but mediated by number of fatal crashes. The analysis reveals instead a direct link between number of casualties and number of persons killed and reported above there is no significant link between number of casualties and number of fatal crashes. Overall, the finding here means that an increase in number of casualties may heighten the risk of more persons killed without necessarily increasing the number of fatal crashes. Similar positive and significant links were found between distance travelled on one hand and number of fatal crashes on the other with the number of persons killed. The link between distance travelled and number of persons killed were predicted in Figure 1 and means that more cars on the road increases the likelihood of more persons killed in case of road accidents. The positive significant link between number of fatal crashes and number of persons killed goes without saying. Surprisingly, while the number of unregistered vehicles increases, the number of persons killed decreases and similar relationship is found for the link between total crashes and persons killed. As indicated above, there must have been important variables (e.g., improvement in road enforcement, drivers behaviours, etc.) that are not included in the metamodel tested in the present study and which are the basis for the counter-intuitive relationships found in the present study. A constant pattern is predicted for the next 10-years using the GARCH model but could fluctuate between 14,739 and 172,784. In the Mthatha area in South Africa, Meel (2008) predicted 57 road death/100,000; if this rate is generalized to the total population of the country, this would be equivalent to 34,200 road deaths annually, well contained within the 14,739-172,784 range identified in the present study.

Based on the findings of the present study revealing significant correlations between vehicles and various road accident events, we recommend the following:
- Limit the usage of vehicles on the road. This can be done: by promoting a safe public transport system to reduce drastically  the number of vehicles on the roads.
- Vehicles that are not registered should be impounded by traffic officers. Legal provisions must be made towards a harsh punishment of drivers of unregistered vehicles, e.g., jail time or heavy financial penalties.
- Reinforcing road traffic measures such as seat belt law enforcement, severe penalties for reckless and drunk drivers and speed control.
- Old vehicles must be removed from the road. This means that the Traffic Department must set a threshold for kilometrage above which vehicles must not be allowed on the road. This requires a scientific study to identify such threshold.
- Identifying and promoting measures that are working successfully in reducing the number of fatalities in road accident, e.g., promoting seat belt, strict check-up of roadworthiness, severe punishment of drunk drivers. This has to be done by the traffic department.
- Government's *Arrive          Alive*          Road-Safety          Campaign (https://www.arrivealive.co.za/) should be intensified.
- Emergency services should respond quicker to calls of road accidents in order to make it on time at an accident scene to save lives

### 4.3 Potential limitations of the study

There are some debates around road accident data sources and their reliability (Joseph, 2013). A number of studies make use of two data sources (Benlagha and

Charfeddine, 2020). The first data source includes police-recorded data (Manner and Wünsch-Ziegler, 2013; Feng *et al.*, 2016; Lee *et al.,* 2018). The second data source is insurance companies (Krishnan and Carnahan, 1985; Mills and Hambly, 2011). The debates emanate from the mismatch between both sources in term of data provided. For example, 1 out of 5 road deaths is reported by police in the Philippines (WHO, 1996). In Indonesia, police report 40% less road deaths than insurance companies (Lu et al. 1999). Even in China, the Beijing Research Institute of Traffic Engineering (Liren, 1996) estimated that the actual number of people killed in road accidents was over 40% more than the number reported by the police. In term of reliability of data, several studies suggest that insurance companies are more reliable than police-recorded data, given that, not only they are always informed once an accident occurs (fatal accident or not), but also given the professional way (unlike government officials; Joseph 2013) they handle data collection (Cohen, 2005; and González Dan *et al.,* 2017).　In the present study, a multiple-source approach was used as presented in the Methodology to minimize the influence of potential inconsistency in data reported in various sources.

# References

- Aderamo, A,J. 2012: Assessing the Trends in Road Traffic Accident Casualties on Nigerian Roads. *Mediterranean Journal of Social Sciences*, 3, 19-25.

- Al-Reesi, H., Ganguly, S.S., Al-Adawi, S., Laflamme, L., Hasselberg, M. and Al-Maniri, A. 2013: Economic Growth, Motorization, and Road Traffic Injuries in the Sultanate of Oman, 1985–2009. *Traffic Injury Prevention*, 14:322–328.

- Arrive Alive. 2020: Corruption, Traffic Enforcement and Road Safety. Available from: https://www.arrivealive.mobi/corruption-traffic-enforcement-and-road-safety. [Accessed on 02 September 2020].

- Bąk, I., Cheba, K. and Szczecińka, B. 2019: The statistical analysis of road traffic in cities of Poland. *Transportation Research Procedia*, 39, 14–23.

- Bayam, E., Liebowitz, J. and Agresti, W. 2005: Older Drivers and Accidents: A Meta-Analysis and Data Mining Application on Traffic Data. *Expert Systems Application*, 29, 598– 629.

- Benlagha, N. and Charfeddine, L. 2020: Risk Factors of Road Accident Severity and the Development of a New System for Prevention: New Insights from China. *Accident Analysis and Prevention*, 136, 1-15.
- Bergel-Hayat, R., Debbarh, M., Antoniou, C. and Yannis, G. 2013: Explaining the Road Accident Risk: Weather Effects. *Accident Analysis and Prevention*, 60, 456–465.
- Bollerslev, T. 1986: Generalized Autoregressive Conditional Heteroskedasticity 1986. *Journal of Econometrics*, 31, 307–327.
- Cohen, A., 2005: Asymmetric Information and Learning: Evidence from the Automobile Insurance Market. *The Review of Economics and Statistics*, 87, 197–207.
- da Silvaa, F.P., Santos, J.A. and Meireles, A. 2014: Road Accident: Driver Behaviour, Learning and Driving Task. *Procedia - Social and Behavioural Sciences,* 162, 300 – 309.
- Deffenbacher, J.L., Lynch, R., Oetting, E.R. and Swaim, C. 2002: The Driving Anger Expression Inventory: A Measure of how People Express Their Anger on the Road. *Behaviour Research and Therapy*, 40, 717–737.
- Delen, D., Sharda, R. and Bessonov, R. 2006: Identifying Significant Predictors of Injury Severity in Traffic Accidents using a Series of Artificial Neural Networks. *Accident Analysis and Prevention*, 38, 434–444.
- Department for Transport. 2015: Reported Road Casualties, Great Britain: 2015 Annual Report.   Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/556406/rrcgb2015-02.pdf. [Accessed on 2 September 2020].
- Donaldson, L., Brooke, K. and Faux, S. 2009: Orthopaedic Trauma from Road Crashes: Is Enough Being Done? *Australian Health Review*, 33, 72–84.
- Eboli, L. and Mazzulla, G. 2007: A Structural Equation Model for Road Accident Analysis. 4th International SIIV Congress. 12-14 September 2007. Palermo: University of Calabria.
- Fan Y, Chen J, Shirkey G, John R, Wu SR, Park H and Shao C. 2016. Applications of structural equation modeling (SEM) in ecological studies: An updated review. Ecological Processes 5:19.
- Farmer, C.M., Braver, E.R. and Mitter, E.L. 1997: Two-Vehicle Side Impact Crashes: The Relationship of Vehicle and Crash Characteristics to Injury Severity. *Accident Analysis and Prevention*, 29, 399–406.
- Feng, S., Li, Z., Ci, Y. and Zhang, G. 2016: Risk Factors Affecting Fatal Bus Accident Severity: Their Impact on Different Types of Bus Drivers. *Accident Analysis and Prevention*, 86, 29–39.
- Fuller, W. A. (1976). Introduction to Statistical Time Series. New York: John Wiley and Sons. ISBN 0-471-28715-6.
- Ghalanos, A. 2020: Package 'rugarch' Version 1.4-4. Available from: https://bitbucket.org/alexiosg.   [Accessed on 16 September 2020].
- González Dan, J.R., Arnaldos, J. and Darbra, R.M. 2017: Introduction of the Human Factor in the Estimation of Accident Frequencies through Fuzzy Logic. *Safety science Journal*, 97, 134–143.
- Grapentine, Terry. 2000. Path analysis vs. structural equation modeling. *Marketing Res.* 12(3) 12–20.
- Grewal R, Cote JA, and Baumgartner H 2004. Multicollinearity and Measurement Error in Structural Equation Models: Implications for Theory Testing. Marketing Science 23, 519-529.
- Hassan, H.M. and Abdel-Aty, M.A. 2013. Predicting Reduced Visibility Related Crashes on Freeways using Real-Time Traffic Flow Data. *Journal of Safety Research*, 45, 29–36.
- Honaker J, King G, Blackwell M (2011). "Amelia II: A Program for Missing Data." *Journal of Statistical Software*, **45**(7), 1–47.
- ITF (International Transport Forum). 2018: Road Safety Annual Report 2018: South Africa. East Jerusalem.
- Jacobs, G. D. and  Cutting, C.A. 1986. Further Research on Accident Rates in Developing Countries. *Accident Analysis and Prevention*, 18, 119-127.
- Jacobs, G., Aeron-Thomas, A. and Astrop, A. 2000. Estimating Global Road Fatalities. TRL Report 445.

- Jagpal, HS 1982. Multicollinearity in structural equation models with unobservable variables. *J.Marketing Res.* 19, 431–439.

- Joseph, R. 2013: Who are South Africa's Worst Drivers? The Truth doesn't Lie in the Data. Africa Check. Available from: https://africacheck.org/2013/07/04/who-are-south-africas-worst-drivers-the-truthisnt-in-the-data/. [Accessed on 12 January 2018].

- Krishnan, K.S. and Carnahan, J.V. 1985: Analysis of the Effect of Car Size on Accident Injury Probability using Automobile Insurance Data. *Accident Analysis and Prevention*, 17, 171–177.

- Kopits, E. and Cropper, M. 2003: Traffic Fatalities and Economic Growth. The World Bank, Policy Research Working Paper No. 3035. Washington, DC: World Bank.

- Labuschange, F., De Beer, E., Roux, D. and Venter, K. 2017: The Cost of Crashes in South Africa 2016. 36th Southern African Transport Conference (SATC 2017). Pretoria, South Africa: Conference Partners.

- Lee, O., Chae, J., Yoon, T. and Hojin Yang, H. 2018: Traffic Accident Severity Analysis with Rain-Related Factors using Structural Equation Modelling – A Case Study of Seoul City. *Accident Analysis and Prevention*,112, 1–10.

- Lefcheck JS 2016. PIECEWISESEM: Piecewise structural equation modelling in R for ecology, evolution, and systematics. Methods in Ecology and Evolution 7, 573-579.

- Liren, D. 1996: Viewing China Road Traffic Safety and the Counter Measures in Accordance with International Comparison. Second Conference in Asian Road Safety, 28-31 October 1996. Beijing: Beijing Research in Traffic Engineering.

- Lu T, Chou M, Lee M (1999). Regional mortality from motor vehicle traffic injury: relationships among place of occurrence, place of death and place of residence. Accident Analysis and Prevention, 32 (2000) 65-69.

- Malhotra, Naresh K, Mark Peterson, Susan Bardi Kleiser. 1999. Marketing research: A state-of-the-art review and directions for thetwenty-first century. *J.Acad.Marketing Sci.* 27(2) 160–183.

- Manner, H. and Wünsch-Ziegler, L. 2013: Analysing the Severity of Accidents on the German Autobahn. *Accident Analysis and Prevention*, 57, 40–48.

- Maruyama, GM. 1998. *Basics of Structural Equation Modeling*. Sage, Thousand Oaks, CA.

- McCarthy, N. 2016: The Worst Countries in the World for Drunk Driving (Infographic). Available from: https://blogs-images.forbes.com/niallmccarthy/files/2016/08/20160709_Drunk_Driving.jpg. [Accessed on 11 August 2016].

- Meel, B.L. 2008: Fatal Road Traffic Accidents in the Mthatha area of South Africa, 1993 – 2004. *South African Medical Journal*, 98, 716-719.

- Mills, B.N., and Hambly, D. 2011: Analysis of Precipitation-Related Motor Vehicle Collision and Injury Risk using Insurance and Police Record Information for Winnipeg, Canada. *Journal of Safety Research*, 42, 383–390.

- Ministry of Transport. 2001: The Road to Safety 2001-2005: Building the Foundations of a Safe and Secure Road Traffic Environment in South Africa. Available from: http://www.transport.gov.za/projects/index.html. [Accessed on 12 June 2006].

- Mooren, L., Grzebieta, R. and Job, S. 2011: Safe System – Comparisons of this Approach in Australia. In: Australasian College of Road Safety Conference. Melbourne, VIC, Australia.

- Mphela, T. 2011: The Impact of Traffic Law Enforcement on Road Accident Fatalities in Botswana. *Journal of Transport and Supply Chain Management*, 5, 264- 277.

- Nantulya, V.M. and Reich, M.R. 2002: The Neglected Epidemic: Road Traffic Injuries in Developing Countries *British Medical Journal*, 324, 1139-1141.

- Ncube, P. Cheteni, P. and Sindiyandiya, K. 2016: Road Accidents Fatalities Trends and Safety Management in South Africa. P*roblems and Perspectives in Management*, 14, 627-633.

- O'Hara, R.B. and Kotze, D.J. 2010: Do not Log-Transform Count Data. *Methods in Ecology and Evolution*, 1, 118–122.
- Papantoniou, P., Antoniou, C., Yannis, G., and Pavlou, D. 2018: Which factors affect accident probability at unexpected incidents? A structural equation model approach, *Journal of Transportation Safety & Security*, 11, 544-561.
- Papantoniou, P., Yannis, G. and Christofa, E. 2019: Which Factors Lead to Driving Errors? A Structural Equation Model Analysis through a Driving Simulator Experiment. *IATSS Research Journal*, 43, 44–50.
- Peden, M. and Hyder, A.A. 2002: Road Traffic Injuries are a Global Public Health Problem. *British Medical Journal*, 324, 1153.
- R Core Team. 2017. R: A language and environment for computing.
- Reason, J. 2000: Human Error: Models and Management. *British Medical Journal*, 320, 768-770.
- RTMC (Road Traffic Management Corporation). 2006: Year 2005 Traffic Report. Pretoria East: Road Traffic Management Corporation.
- RTMC (Road Traffic Management Corporation). 2007: Interim Road Traffic and Fatal Crash Report for the Year 2006. Pretoria East: Road Traffic Management Corporation.
- RTMC (Road Traffic Management Corporation). 2008: 2007 Road Traffic Report on Number of Registered, Un-Roadworthy and Un-Licenced Vehicles, Driving Licences and PrDPs and Fatal Crashes and Fatalities. Pretoria East: Road Traffic Management Corporation.
- RTMC (Road Traffic Management Corporation). 2011: Road Traffic Report 31 December 2010. Pretoria East: Road Traffic Management Corporation.
- RTMC (Road Traffic Management Corporation). 2012: Traffic Calendar Report 31 December 2011. Pretoria East: Road Traffic Management Corporation.
- RTMC (Road Traffic Management Corporation). 2013: Road Traffic Calendar Report 31 December 2012. Pretoria East: Road Traffic Management Corporation.
- RTMC (Road Traffic Management Corporation). 2014: Road Traffic Calendar Report 31 December 2013. Pretoria East: Road Traffic Management Corporation.
- RTMC (Road Traffic Management Corporation). 2015: Road Traffic Calendar Report 31 December 2014. Pretoria East: Road Traffic Management Corporation.
- RTMC (Road Traffic Management Corporation). 2016: Road Traffic Calendar Report 1 January- 31 December 2015. Pretoria East: Road Traffic Management Corporation.
- RTMC (Road Traffic Management Corporation). 2017: Road Traffic Calendar Report 1 January-31 December 2016. Pretoria East: Road Traffic Management Corporation.
- RTMC (Road Traffic Management Corporation). 2018: State of Road Safety Report: January-December 2017. Pretoria: Department of Transport.
- RTMC (Road Traffic Management Corporation). 2019: Alcohol and its Implications for Road Traffic Crashes in South Africa Phase A- Review. Parow: Violence, Injury and Peace Research Unit SAMRC.
- Sârbescu, P. and Maricuțoiu, L. 2019: Are You a "Bad Driver" All the Time? Insights from a Weekly Diary Study on Personality and Dangerous Driving Behaviour. *Journal of Research in Personality*, 80, 30–37.
- Schermelleh-Engel K., Moosbrugger H, Müller H. 2003. Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures. Methods of Psychological Research 8,23-74Stats SA (Statistics South Africa). 1995: Suid-Afrikaanse Statistieke- South African Statistics. Pretoria: Department of Statistics South Africa.
- Shaw, J.I. and McMartin, J.A. 1977: Personal and Situational Determinants of Attribution of Responsibility for an Accident. *Human Relations*, 30, 95–107.

- Silcok, B.R. 2003: Guidelines for Estimating the Costs of Road Crashes in Developing Countries. London: U.K. Department for International Development.

- Stats SA (Statistics South Africa). 2007: Mid-Year Population Estimates 2007. Pretoria: Department of Statistics South Africa.

- Stats SA (Statistics South Africa). 2008: Mid-Year Population Estimates 2008. Pretoria: Department of Statistics South Africa.

- Stats SA (Statistics South Africa). 2009: South African Statistics, 2009. Pretoria: Department of Statistics South Africa.

- Stats SA (Statistics South Africa). 2010: Mid-Year Population Estimates 2010. Pretoria: Department of Statistics South Africa.

- Stats SA (Statistics South Africa). 2012: Census 2011. Pretoria: Department of Statistics South Africa.

- Stats SA (Statistics South Africa). 2013: Mid-Year Population Estimates 2013. Pretoria: Department of Statistics South Africa.

- Stats SA (Statistics South Africa). 2014: Mid-Year Population Estimates 2014. Pretoria: Department of Statistics South Africa.

- Stats SA (Statistics South Africa). 2015: Mid-Year Population Estimates 2015. Pretoria: Department of Statistics South Africa.

- Stats SA (Statistics South Africa). 2017: Mid-Year Population Estimates 2017. Pretoria: Department of Statistics South Africa.

- Stats SA (Statistics South Africa). 2019: Mid-Year Population Estimates 2019. Pretoria: Department of Statistics South Africa.

- Tarka P 2018. An overview of structural equation modeling: its beginnings, historical development, usefulness and controversies in the social sciences. Qual Quant 52, 313–354.

- Theofilatos, A. 2017: Incorporating Real-Time Traffic and Weather Data to Explore Road Accident Likelihood and Severity in Urban Arterials. *Journal of Safety Research*, 61, 9–21.

- Thompson, L., Bundy, C., Lowe, C., Mabin, A., Vigne, R and Gordon, D. 2019: South Africa. Available from: https://www.britannica.com/place/South-Africa.   [Accessed on 15 September 2019].

- Tibane, E., Lentsoane, N., Honwane, M. and Kraamwinkel, J. (Eds.).   2016: South Africa Yearbook 2015/2016. Pretoria: Government Communications (GCIS).

- Verbeke W., Bagozzi RP. 2000. Sales call anxiety: Exploring what it means when fear rules a sales encounter. *J. Marketing* 64, 88–101.

- Verster, T., and Fourie, E. 2018: The Good, the Bad and the Ugly of South African Fatal Road Accidents. *South African Journal of Science*, 114, 63–69.

- WHO (World Health Organisation). 2009: Global Status Report on Road Safety: Time for Action. Geneva: World Health Organisation.

- World Health Organisation (WHO). 2011:  Global Plan for the Decade of Action for Road Safety 2011-2020. Available from: https://www.who.int/roadsafety/decade_of_action/plan/plan_english.pdf?ua=1.

- WHO (World Health Organisation). 2013: Global Status Report on Road Safety 2013: Supporting a Decade of Action. Geneva: The World Health Organization.

- WHO (World Health Organisation). 2015: Global Status Report on Road Safety 2015. Geneva: The World Health Organisation.

- WHO (World Health Organisation). 2018: Global Status Report on Road Safety 2018. Geneva: The World Health Organisation.

- WHO (World Health Organization). 2019: Road Traffic Injuries. Available from: https://www.who.int/newsroom/fact-sheets/detail/road-traffic-injuries. [Accessed on 21 September 2019].
- Wilde, G.J.S. 1988: Risk Homeostasis Theory and Traffic Accidents: Propositions, Deductions and Discussion of Dissension in Recent Reactions. _Ergonomics_, 31, 441–468.

**Captions**

**TABLES:**

**Table 1.** Coefficients of the structural equation model fit to the dataset analysed in the present study.

**FIGURES:**

**Figure 1.**     Proposed theoretical metamodel to explain road accident events in South Africa.

**Figure 2**. Trends in number of crashes (road accidents) and casualties over the period 1935-2017 covered in this study.

**Figure 3.** Trends in number of fatal crashes and number of persons killed over the period 1935-2017 covered in this study.

**Figure 4.** Trends in the death rate due to road accident. Death rate is estimated over 100,000 people.

**Figure 5.** Prediction of the trends of road accidents and its consequences for the next 10 years in South Africa using GARCH model. A=number of crashes (road accidents), B=number of casualties, C= number of fatal crashes and D=number of persons killed.

**SUPPLEMENTAL INFORMATION**

**Figure S1** Structure of the dataset analysed showing missing and observed values for all variables. Driver_licence=number of driver licence issue per year; dist_travelled=total distance travelled by all registered vehicles every year; casualties=number of casualties reported per accident per year; crashes=number of accidents reported per year; population_SA=human population of South Africa per year; nber_vehicles=total number of vehicles in the country per year; unreg_vehiecles=number of unregistered vehicles estimated per year; fatal_crashes=number of crashes in which at year a person dies; regist_vehicles=number of registered vehicles in the country per year; year=year from 1935 to 2017.

**Table S1.** All data analysed in the present study. These data include raw data collected from various sources as well as the imputed data.

**Table S2.** Comparison of $R^2$ values in SEM with non-collinear versus SEM with collinear variables.

**R script used to analyse data.**