Forecasting Traffic Accidents in Developing Countries Using Time Series Analysis

Mohammad Ali Khasawneh [1,2], Aslam Ali Al-Omari [3] and Bader Ganam [4]

[1] Associate Professor, Civil Engineering Department, Prince Mohammad Bin Fahd University, Al Azeziya, Eastern Province, Kingdom of Saudi Arabia.
E-Mail: mkhasawneh@pmu.edu.sa. * Corresponding Author.
[2] Professor, Civil Engineering Department, Jordan University of Science and Technology, Irbid, Jordan. E-Mail: mkhasawneh@just.edu.jo
[3] Associate Professor of Civil Engineering, Civil Engineering Department, Jordan University of Science and Technology, Irbid, Jordan. E-Mail: aaalomari3@just.edu.jo
[4] Graduate Student, Civil Engineering Department, Jordan University of Science and Technology, Irbid, Jordan. Email: bader.gannam88@gmail.com

ABSTRACT

The main objective of this study is to develop models that can forecast traffic accidents, injuries and fatalities in developing countries. Consequently, this could set the first step in enhancing traffic safety measures as well as to help planning for the future. Moreover, results of this investigation may assist decision and law makers at transportation and traffic agencies to come up with resolutions for the increasing traffic crashes and consequently loss of lives, properties and the high cost involved. The forecasting models were developed using the well-known time series analysis. Each time series model was decomposed into its components (trend, cyclical, seasonal and irregular components). In conclusion, the forecasting by time series analysis explained a significant proportion of the variability in the dependent variables (accidents, injuries and fatalities) as shown by the relatively acceptable coefficients of determination (R^2).

KEYWORDS: Traffic accidents, Time series analysis, Monthly prediction, Trend, Cyclical, Seasonal, Irregular.

INTRODUCTION

On 31st of August 1869, the world's first road death involving a motor vehicle took place. Irish Scientist Mary Ward died when she fell out of her cousins' steam car and was run over by it (Offaly Historical & Archaeological Society, 2007). After one hundred years, in 1969, various societies were developed around the world as the number of registered vehicles increased dramatically. The number of fatalities and injuries due to traffic accidents increased as a result too. Leeming et al. (1969) and other British road engineers studied the circumstances of deaths from accidents as reported in various states before the widespread introduction of 55 mph (89 km/h) speed limit and drunk-driving laws.

Received on 18/6/2021.
Accepted for Publication on 7/10/2021.

To show the importance of traffic accidents globally, the World Health Organization (WHO), in its global status report on road safety 2009, estimated that the majority of the world's fatalities on roads occur in low-income and middle-income countries, which have only 48% of the world's registered vehicles. It predicts that road traffic injuries will rise to become the 5th leading cause of death by 2030 (WHO, 2009). Additionally, the United Nations Sustainable Development Goals (SDGs) set priorities to prevent road traffic injuries and deaths and to improve air quality, highlighting the need to protect vulnerable populations such as children. SDGs explore the main challenges in these areas, as well as ways of financing actions to achieve these goals (Griggs, 2013).

The health effects of traffic accidents clearly appear to be either psychological or physical. Psychological trauma may affect individuals' ability to work and take on family responsibilities, while physical effects may lead to human as well as economic losses, especially in the form of injuries or fatalities from traffic accidents. The global economic cost of traffic accidents was estimated at $518 billion per year in 2003 and $100 billion of this cost is spent in developing world (WHO, 2004). There are few global estimates of the costs of injury, but some research efforts carried out in 2010 suggested that road traffic accidents cost high-income countries approximately 3% of their gross national product. This cost rises to be 5% in low-income and middle-income countries (United Nations General Assembly in New York, 2003).

The road traffic deaths reached 1.25 million globally and about 90% of fatalities have occurred in the developing world (WHO, 2015). Therefore, the topic of this study is a crucial issue and its findings would be useful for policy makers in the developing countries. Jordan, like many other developing countries, suffers from traffic accidents in the form of both economic and social losses. Jordan is classified as a middle-income country where traffic accidents are considered as one of the major causes of death. Jordan loses approximately 750,000 JD (1,060,000 USD) per day due to traffic accidents or 2,500 JD (3,530 USD) per accident on average according to accident statistics for 2015 (Jordan Traffic Institute, 2015). The increase of population leads to a rapid increase in traffic problems like road accidents among others. Due to this development, traffic accidents rapidly increased during the last few decades.

The number of accidents increased dramatically from 13,567 in 1981 to 111,057 in 2015, meaning a rise of 8.2 times during this period. Regarding the number of injuries, the highest rate was 18,842 injuries in 2000; after that, it slightly decreased to be 16,139 injuries in 2015, but overall, there was an increase during the whole period. Finally, the number of fatalities oscillated between about 450 and about 992 fatalities in the same period. These numbers give an indication about the challenges that Jordan faces regarding the traffic situation.

LITERATURE REVIEW

Al-Omari et al. (2013) studied traffic accident trends and characteristics in Jordan over the period (1998-2010). The paper investigated traffic accidents' different characteristics, such as type of accident, speed limit, age of driver, hour of the day, month, weather condition, pavement condition and level of severity. According to the results of the study, the study found that Friday has the largest severity level as compared to those of other days of the week. Also, the summer months were noticed to have the largest numbers of traffic accidents. The important conclusion was that Jordan witnessed an increase of 223% in general in traffic accidents during that period. The researchers provided a forecast of data up to the year 2030 through the analysis of time series that gives monthly results regarding accidents injuries and fatalities.

A study by Al-Omari et al., published in 2019, aimed to assess traffic accidents in Jordan. It identified key factors that may aid in reducing traffic accidents and provide information to decision-makers to assist in identifying high-risk areas that require remedial programs to reduce the impact of accidents. The information utilized in this study came from a variety of sources and spanned over the years 1981 to 2015. The study employed a method based on the use of a new and simple instrument that can estimate the risks of accidents at various levels; it is called Accident Hazard Scale (AHS) and is used for calculating the risk of an accident. Each domain's elements are ranked from the most dangerous (100 percent age on the hazard scale) to the least dangerous (lowest percentage on the same scale). On this scale, the top governorates in Jordan were Al-Mafraq, Irbid, Amman and Al-Zarqa, in that order. According to the study's findings, the months with the highest percentages of accident dangers were July, August and September, with 100%, 96.4% and 79.2%, respectively. The percentages in the 75% to 100% zone were highest on weekends.

Khasawneh et al. (2018) forecasted traffic accidents in Jordan using regression techniques. Injuries per 1,000 accidents and fatalities per 1,000 accidents were expected to decrease by around 77 injuries and 2 fatalities from 2015 to 2030, which indicates a desirable behavior in traffic safety measures. The accident rate (number of accidents per 10,000 registered cars) was predicted to rise by 232 accidents by 2030, according to a data analysis. This means that by 2030, the severity level will drop by 0.08, to 0.07. There would be a decrease of 594.6 thousand JD per 1,000 accidents throughout the forecasted period. According to regression models, the results of the forecasts for the period (2016-2030) showed improvements compared to those for the period 2001-2015.

Al-Jarrah et al. (2019) investigated key factors influencing the severity of drivers' injuries in car crashes using supervised machine learning techniques. Driver injury severity levels in fatal car collisions were investigated by examining crash risk factors, such as the driver, vehicle, highway and crash features. These approaches included Classification and Regression Trees (CARTs), Artificial Neural Networks (ANNs) and Multinomial Logistic Regression, all of which require supervision. Using the Fatality Analysis Reporting System (FARS) as a source of data, the research was conducted. It was concluded that when it comes to decreasing fatalities and injuries, the usage of seatbelts is crucial. According to the research, older drivers are more likely to suffer fatal and serious injuries than younger drivers. A fatality or a severe injury happened in 94 percent of rollover accidents.

Generally, time series analysis is applied in economic studies and it has been also widely used in different traffic studies during the few recent years, especially traditional analysis, like Autoregressive Moving Average (ARMA) model. To guarantee the simplicity of analysis, the analysis is presented according to a simple procedure shown in the following sections. The main aim of this paper is to apply time series analysis tool on accidents data by separating the monthly series of accidents, injuries and fatalities into their components and then producing their models. Thus, it is possible to set the first step in incorporating study outcomes in the enhancement of traffic safety measures process in developing countries as well as to help in planning for the future. At the beginning, some definitions relevant to time series analysis should be mentioned as follows:

- A *time series* (Y_t) is a sequence of 'n' observations $\{Y_1, Y_2, \ldots, Y_n\}$ on a process at equally spaced points in time.
- *Trend component* (T_t): describes the long-term sweep of the series; it is usually modeled by smooth curves.
- *Cyclical component* (C_t): describes the alternating periods of relative expansion and contraction of more than one-year duration in the series; it consists of cycles that vary in amplitude and duration.
- *Seasonal component* (S_t): describes the pattern of change recurring within periods of one year or less; it consists of a sequence of relatively repetitious cycles the duration of which is one year or less.
- *Irregular component* (I_t): describes the effects of all other factors; it tends to have an irregular, saw-toothed pattern.

Unfortunately, there is a small number of existing studies on the utilization of time series for accident analysis. Most studies discussed the simplistic and traditional system of models for time series, such as ARIMA model and these studies are summarized herein.

Helfenstein (1990) studied whether or not a reduced speed limit had an effect on accidents by using intervention time analysis. Time series and ARIMA models were applied in the analysis of data. Results showed that the number of accidents had already decreased before the introduction of speed reduction when data was analyzed by seasonal subseries. The study concluded that most methods indicated that the reduction in accidents and injuries lies one quarter ahead of the actual introduction of speed limitation.

Filip et al. (2007) analyzed road risk according to age and gender, based on a time series approach. The methodology was based on classical ARIMA models and state space models that can be considered as dedicated time series models. Results showed that males had higher observed risk than females in the overall trend. Regarding age groups, the riskiest groups were (15-24), (25-44) and (65 or more), respectively for males and were (15-24), (65 or more) and (25-44), respectively for females in risks ordered from most to less dangerous. The study provided a forecast analysis up to 2010 for road safety that concluded that the highest yearly reduction in risk would be found for the oldest and youngest road users in both men and women.

Yu et al. (2015) selected fatalities per 10,000 registered vehicles and gross domestic product (GDP) as dependent and independent variables, respectively. According to analysis, logarithm of fatalities per 10,000 registered vehicles model showed a high fitting degree with $R^2 = 0.98$. Their study provided good forecasting results that could be a good reference for decision-makers.

Based on the literature review conducted, it has been made clear that new approaches such as time series can be utilized to better address the issue of forecasting traffic accidents, injuries and fatalities in developing countries such as Jordan. This work would possibly be extended in the future to utilize other cutting edge techniques for the same. Hence, the main objective of this study is to develop models to forecast and predict traffic crashes and their associated injuries and fatalities from a traffic safety standpoint.

METHODOLOGY

The data used in this research was obtained from various sources and covered the period 1997-2015. Authorities that were the sources of data are Department of Statistics (DOS), Jordan Traffic Institute (JTI), Public Security Directorate (PSD) and Central Traffic Department (CTD).

The forecasting utilizes a traditional tool that is called time series analysis. As the evaluation part shows that monthly numbers of accidents oscillated during a year with accidents in the summer months higher than those in the winter months. To overcome these oscillations, time series forecasting was implemented. The dependent variable (Y) represented monthly numbers of accidents, injuries or fatalities and the independent variable (X) represented months coded from January 1997 (X=1) to December 2014 (X=216) when number of accidents per month is the dependent variable, while Y is coded from January 1997 (X=1) to December 2015 (X=228) when number of injuries or fatalities per month is the dependent variable.

The main formula used in time series analysis is shown in Equation 1 which is called the Classical Multiplicative Model that consists of four terms: trend component which is represented by simple linear regression model (T_t is the dependent variable and X_t is the independent variable representing months) and other terms are cyclical, seasonal and irregular components.

$$Y_t = T_t \,.\, C_t \,.\, S_t \,.\, I_t \qquad t = 1, 2, \dots, n \qquad (1)$$

where:

Y_t: is the response or dependent variable in the t^{th} month;

T_t: is the trend component of series in the t^{th} month represented as a function;

C_t: is the cyclical component of series in the t^{th} month as a percentage;

S_t: is the seasonal component of series in the t^{th} month as a percentage;

I_t: is the irregular component of series in the t^{th} month as a percentage.

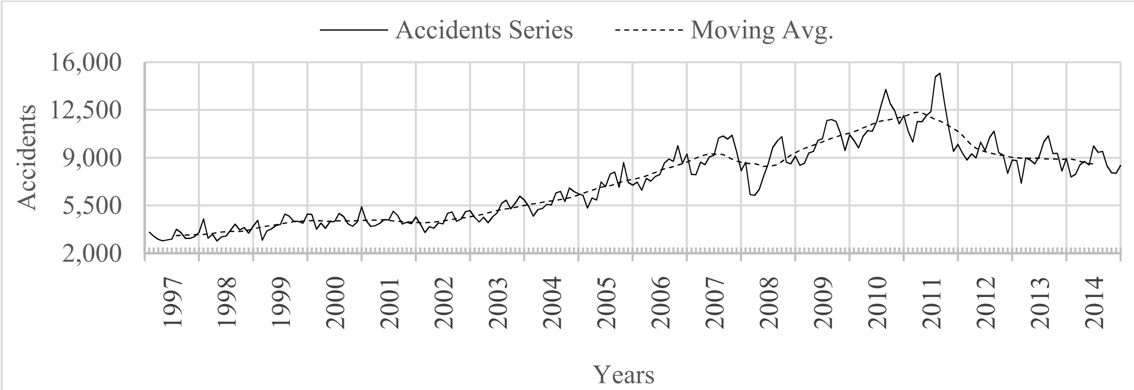
The main aim of this analysis is to predict the monthly rate of accidents, injuries and fatalities. Finally, the residual analysis of implemented models is provided. Essentially, the additive model or what is known as multiple-regression model was avoided here because of the fact that data is highly sensitive to seasonal variation. The final additive model, therefore, will have 12 terms (one term for each month) and is thus very cumbersome.

Procedure of Decomposition of Time Series

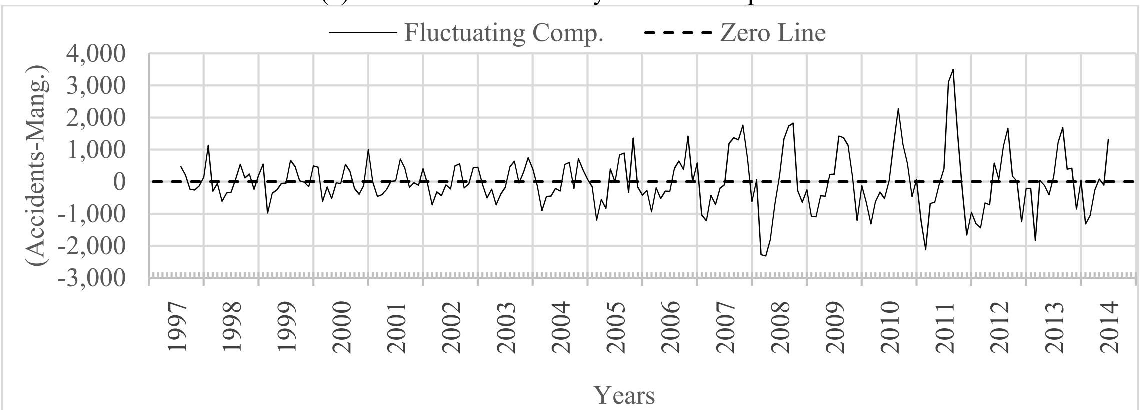
Figure 1(a) shows the accidents series that represents the actual observation of accidents per month over the period 1997-2014 and shows the moving average by a dashed line which is called the systematic component. The moving average depends on the number of observations at each cycle. Here, each year represents one cycle that includes 12 observations for the 12 months. Thus, the k-value of data is equal to 12 and it is an even number.

The weighted average is used to find the moving average according to Equation 2, where p-value represents the trimmed months from the dataset and p here is equal to 6. Therefore, the first and last 6 months were removed from the dataset. If the moving average was estimated by the normal way (mean of 12 observations), the result will be located at 6.5 on the time line of data (6.5 months means 6 months and approximately 15 days). The time line does not have an observation located at 6.5 months (only 6 months or 7 months, not in between). Therefore, the weighted average shifts by 0.5 month to be 7 instead of 6.5. The first and last 6 months of the entire dataset are excluded based on p-value. Therefore, the first moving average was estimated at t=7.

$$M_{avg.@t} = \frac{1}{24} \times (Y_{t-p} + 2Y_{t-p+1} + 2Y_{t-p+2} + \cdots + 2Y_{t-p+10} + 2Y_{t-p+11} + Y_{t-p+12}) \qquad (2)$$



(a) Accidents series and systematic component



(b) Fluctuating component of accidents series

Figure (1): Accidents series and its systematic and fluctuating components

Figure 1(b) is similar to Figure 1(a) in interpretation; the values of fluctuations are made by calculating the number of accidents in a month minus the moving average in that same month. Wider fluctuations indicate an increase in the number of accidents and narrower ones indicate a decrease. From 2003 to 2012, both figures (1 a & b) show a significant increase in accident numbers.

Now, simple calculations are made to separate the components of accidents series. Firstly, the trend component used in time series analysis is represented by a linear, exponential or Gompertz model. These models are common in time series analysis. When the data has constant percentage change in observations, the exponential model is more adequate. If the data has an upper limit or the rate of observations increased then it decreased, the Gompertz model provides excellent fit, whereas, the linear model is recommended if exponential and Gompertz models do not fit the data well. The other models, like polynomial, are not recommended in forecasting analysis, because predicted values may become negative and this is not logical for accidents. All of these models (linear, exponential and Gompertz) use the total number of accidents per month as a dependent variable and the decoded months as independent variables, where X_t=1 represents the month of January in 1997 and X_t=2 represents February of 1997… and so on, until December 2014 with X_t=216.

Table 1. Results of regression models for accidents series

Model	Equation	R^2	F*
Linear	$T_t = 2888.621 + 39.293X_t$	0.73	570.6
Exponential	$T_t = 3397.771 \times e^{0.006066X_t}$	<u>0.80</u>	<u>836.0</u>
Gompertz	$T_t = 12000e^{-1.717e^{-0.01224X_t}}$	0.77	356.7

Table 1 and Figure 2 summarize the results of fitting these models to determine the best one; the exponential model with $R^2 = 0.80$ and $F^* = 836$ under a confidence level of 95% was found to be the best model.

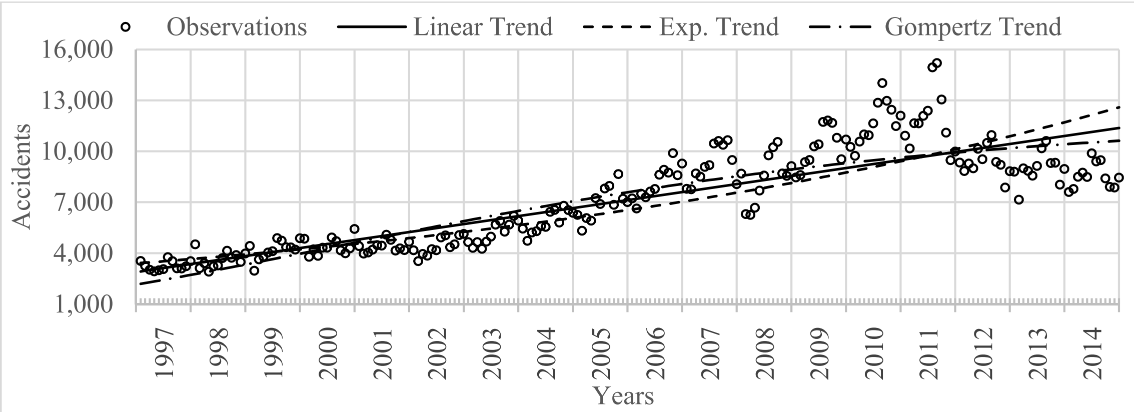


Figure (2): Trends of linear, exponential and Gompertz models for accidents series

The seasonal component is divided into two steps. The first step calculates the seasonal relative values, as shown in Table 2, column 6 by using Equation 3, while the second step arranges seasonal relative values, as shown in Table 3 to find the average for each month over the entire period (see Equation 4). These averages represent the seasonal component recurring in the same pattern every year, as shown in Table 4, column 7.

$$S_t.I_t = \left(Y_t / M_{avg.} \right) \times 100 \quad (3)$$

$$S_t = Averging\ (S_t.I_t) for\ Specific\ month \quad (4)$$

The third component is the cyclical one that is based on finding moving average (Table 2, column 5) and the values are divided then by trend values (obtained from the selected exponential model, as shown in Table 1) for the same month. Results are multiplied by 100 to be reported in percentage. Finally, to find the irregular component, the seasonal relative values (Table 2, column 6) are divided by the seasonal component values (Table 4, column 7). All calculations of accident data decomposition are completed using Microsoft Excel Spreadsheet.

Procedure of Time Series Modeling

Table 4 shows calculations of accidents series forecasting. Columns 3 and 4 represent the observations ordered according to time (column 4) and decoded from 1 to 216 data points (column 3). The moving average and seasonal values shown in columns 5 and 6, respectively, are calculated as in the previous section. The new steps here are the estimated deseasonalized values (column 7), the new trend component (column 8) and the forecasted values (column 9).

At first, deseasonalization that eliminated the seasonal effects from data series was estimated according to Equation 5 that divided observation values by corresponding seasonal values. Thereafter, the trend was found following the same procedure applied in the trend component computation of the previous section, but here the dependent variable is the deseasonalized values, not the observation values, while the independent variable is the same (decoding time X=1 to 216) as shown in Equation 6. The final step calculated the forecast values included in the seasonal effects, as shown in Table 4, column 9, based on Equation 7, by multiplying trend values by the corresponding seasonal values.

$$Deseasonlize = Y_t / S_t \quad (5)$$

$Y_{t.}$: is accidents observation per month at time t;
$S_{t.}$: is seasonal value located at same time t.

$$T_t = Best\ fit\ of (Deseasonalized\ with\ time\ X_t) \quad (6)$$

$T_{t.}$: represented by the best fit modeled according to linear, exponential and Gompertz regression models.

$$Forecast = T_t.S_t \quad (7)$$

$T_{t.}$: is the trend value at time t;
$S_{t.}$: is the seasonal value located at the same time t.

Table 2. Some monthly accidents observations and calculated values for time analysis

1	2		3	4	5	6
Year	**Month**		**t (deCode)**	**Y_t (Accidents)**	**$M_{avg.}$ (Weighted)**	**$S_t.I_t$ (Seasonal Relative)**
1997	1	January	1	3540	-	-
	2	February	2	3244	-	-
	3	March	3	3021	-	-
	4	April	4	2922	-	-
	5	May	5	2992	-	-
	6	June	6	3053	-	-
	7	July	7	3760	3291.2	114.2
	8	August	8	3529	3326.4	106.1
	9	September	9	3100	3336.9	92.9
	10	October	10	3092	3352.4	92.2
	11	November	11	3228	3361.1	96.0
	12	December	12	3524	3379.4	104.3
1998	1	January	13	4518	3386.0	133.4
	2	February	14	3112	3409.2	91.3
	3	March	15	3405	3460.5	98.4
	4	April	16	2909	3519.5	82.7
	..	..	..	..	..	..

Table 3. The seasonal component from average of seasonal relative values

Mon.	**Seasonal Relative (S.I)**																		**S**
	1997	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	
1st	-	133.4	114.1	110.3	99.7	98.1	98.1	97.7	97.6	96.4	88.2	100.7	88.6	94.0	89.9	87.8	97.7	85.2	98.7
2nd	-	91.3	75.2	86.0	89.6	83.1	89.5	84.0	81.6	87.6	86.4	73.5	88.7	88.1	82.7	86.0	79.6	88.1	84.8
3rd	-	98.4	90.9	96.2	90.9	92.5	95.1	91.8	91.6	97.5	95.4	73.0	95.5	94.4	94.5	93.3	100.5	97.0	93.5
4th	-	82.7	93.7	88.0	94.5	89.8	85.5	92.2	87.7	93.3	92.3	78.6	95.4	97.1	94.8	92.6	98.7	101.0	91.6
5th	-	90.2	98.7	99.2	100.3	97.7	92.4	96.4	105.7	96.3	97.8	91.7	102.2	95.4	99.6	106.1	95.4	98.8	97.9
6th	-	91.0	98.9	98.8	100.5	94.8	96.6	95.0	99.9	96.3	99.0	102.2	102.3	100.3	103.5	100.8	101.7	115.4	99.8
7th	114.2	102.6	115.9	112.4	116.1	111.1	108.9	109.2	112.0	105.1	112.9	116.0	113.8	110.0	126.4	111.8	113.8	-	112.5
8th	106.1	115.1	110.9	107.6	109.8	112.4	112.1	110.1	112.7	107.8	114.8	120.4	113.2	119.4	129.9	118.0	119.0	-	114.1
9th	92.9	103.3	100.7	95.2	96.0	95.7	99.3	96.5	95.3	104.6	114.4	120.9	110.8	109.9	113.0	101.9	104.3	-	103.2
10th	92.2	106.6	99.6	91.1	99.2	98.4	105.8	111.8	118.6	116.8	119.8	96.8	101.4	104.8	97.9	100.2	104.8	-	103.9
11th	96.0	93.6	96.4	97.3	97.4	109.4	113.7	106.0	97.7	100.2	108.2	93.0	88.8	96.1	85.1	86.3	90.4	-	97.4
12th	104.3	105.2	111.3	122.8	109.5	109.7	107.2	101.3	94.4	106.7	92.9	97.3	98.9	100.6	91.3	97.7	100.5	-	103.0

Table 4. Sample of forecasting number of accidents per month

1	2		3	4	5	6	7	8	9
Year	**Month**		**X_t (deCode)**	**Y_t (Accidents)**	**$T_t.C_t$ ($M_{avg.}$)**	**S_t (Seasonal)**	**Y_t/S_t (Deseas.)**	**T_t**	**$T_t.S_t$ (Forecast)**
1997	1	January	1	3540	-	98.7	3587.1	2952.0	2913.2
	2	February	2	3244	-	84.8	3827.1	2991.0	2535.3
	3	March	3	3021	-	93.5	3232.7	3030.0	2831.6
	4	April	4	2922	-	91.6	3188.6	3069.1	2812.5
	5	May	5	2992	-	97.9	3057.1	3108.1	3041.9
	6	June	6	3053	-	99.8	3058.3	3147.1	3141.7
	7	July	7	3760	3291.2	112.5	3342.6	3186.1	3584.0
	8	August	8	3529	3326.4	114.1	3093.7	3225.1	3678.9
	9	September	9	3100	3336.9	103.2	3003.3	3264.1	3369.2
	10	October	10	3092	3352.4	103.9	2976.4	3303.1	3431.5
	11	November	11	3228	3361.1	97.4	3314.6	3342.2	3254.8
	12	December	12	3524	3379.4	103.0	3420.6	3381.2	3483.4
1998	1	January	13	4518	3386.0	98.7	4578.2	3420.2	3375.2
	…	…	…	…	…	…	…	…	...

Remarks:

Col. 1: number of years.

Col. 2: numbering of months as data points starting from Jan. 1997 with 1 and ending with Dec. 2015 with 228 data series points.

Col. 3: X_t as independent variable. This takes values from 1 to 228.

Col. 4: traffic accidents in each month. Y_t as dependent variable that takes number of accidents in each month.

Col. 5: moving average containing T_t represents trend component and C_t represents cyclical components. M_{avg} calculated by taking the average for every 12 consecutive months.

Col. 6: S_t seasonal effect as percentage. Calculated as shown in Table 3.

Col. 7: de-composition of traffic accidents. Calculated by dividing the accidents by seasonal effect.

Col. 8: T_t the trend values. These come from the regression model.

Col. 9: Tt St represents the values of forecasting by time series analysis. The values are the product of seasonal factors with trend values.

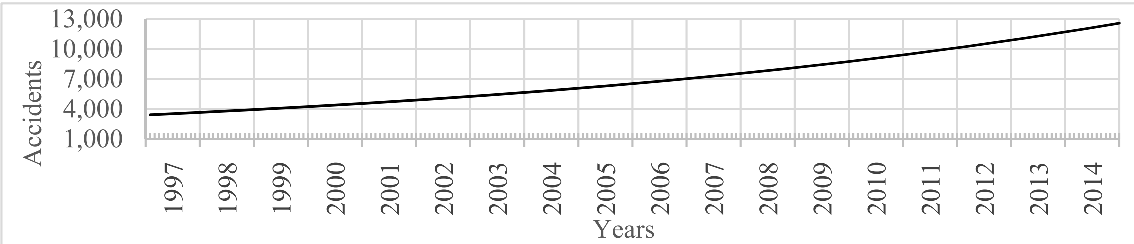
Table 5. Seasonal factor for each month in forecast models

For accidents forecast model												
Month	1	2	3	4	5	6	7	8	9	10	11	12
S_f	0.987	0.848	0.935	0.916	0.979	0.998	1.125	1.141	1.032	1.039	0.974	1.030
For injuries forecast model												
Month	1	2	3	4	5	6	7	8	9	10	11	12
S_f	0.855	0.778	0.930	0.965	1.033	1.041	1.186	1.189	1.070	1.047	0.950	0.949
For fatalities forecast model												
Month	1	2	3	4	5	6	7	8	9	10	11	12
S_f	0.927	0.729	0.915	0.888	0.964	1.052	1.129	1.152	1.106	1.118	0.987	1.034

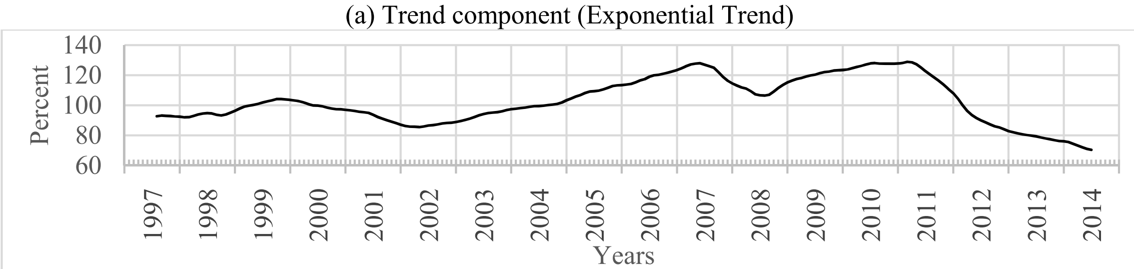
RESULTS AND DISCUSSION

Results of Time Series Components

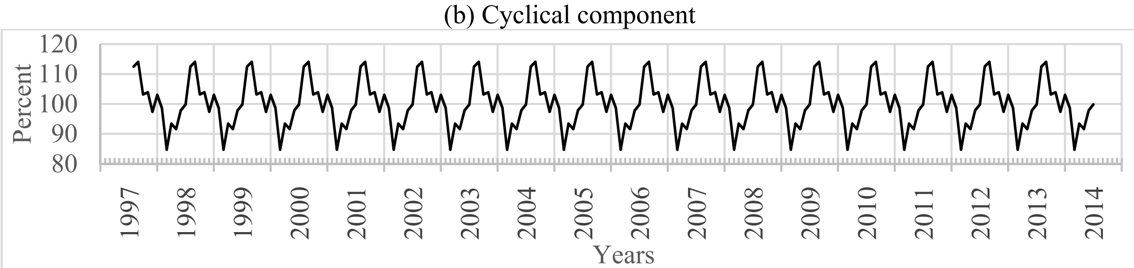
The results of accidents series components are shown and summarized as follows. The trend component shown in Figure 3(a) was represented by an exponential model as was mentioned previously in Table 1. This trend may be natural due to the increase of population and then the increase of drivers' number that lead to increase vehicle number and traffic accidents. The cyclical component shown in Figure 3(b) demonstrates an increase in the number of accidents, with the highest increase (28.9%) in February 2011. More recent years show a decline in accident number per month. The decrement of accidents in recent years happened because the fixed speed camera system was widely applied on major roads. The seasonal component in Figure 3(c) shows seasonal effects with highest and lowest points. These effects are repeated yearly, where more accidents occurred in July, August and September (summer season) and less accidents occurred in February, March and April (winter season). This may be interpreted by the behavior of drivers and people who prefer to stay in their houses and avoid moving around in inclement weather in the winter season. The irregular component in Figure 3(d) represents the other factors. The saw-toothed shape reflects the effects of unique and non-recurring factors, such as new policy law, unusual incidents and others. The irregular pattern is a proof of the random occurrence of accidents according to a number of factors that cannot be controlled.



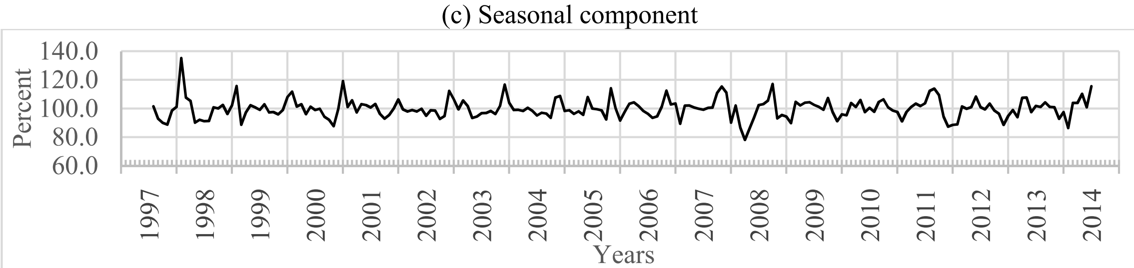
(a) Trend component (Exponential Trend)



(b) Cyclical component



(c) Seasonal component



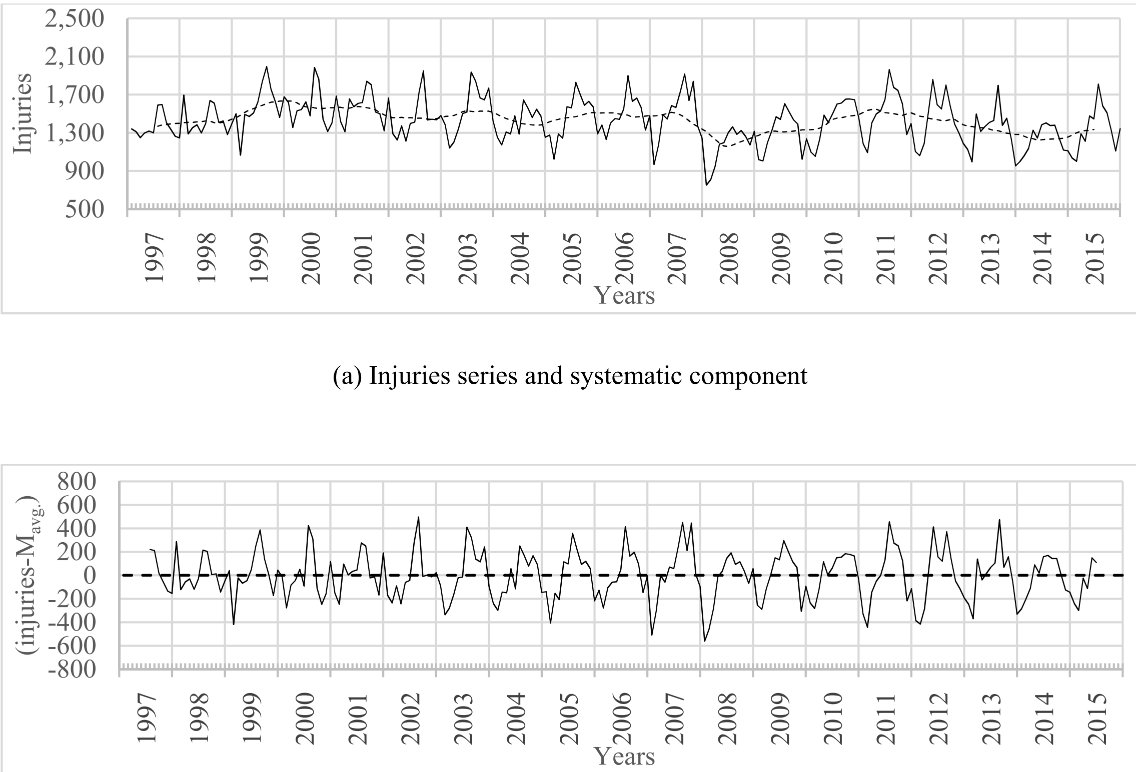
(d) Irregular component

Figure (3): Components of accidents series

Injuries results are summarized in Figures 4 and 5. The injuries series in Figure 4 (a) takes approximately the same pattern (saw-toothed pattern) of the irregular component shown in Figure 5(d). Thus, there is no strong relationship in trend component, see Figure 5(a), which is close to be the average with a slight decrease during the study period. Figure 4(b) is a proof of no-relationship that appears in Figure 4(a), because the observations approximately fluctuated in the same size without contraction or expansion.

Regarding the cyclical component in Figure 5(b), the number of injuries slightly oscillated within ±10% except for the middle of 2008, when it went down by about 20% due to the implementation of new policies. These policies were discussed in a study by Al-Masaeid, H. (2009). The policies can be summarized as follows: (1) Police authorization and execution of traffic law with hardened punishment levels, (2) Associatively with the utilization of the new law, traffic police activities were increased and targeted drivers' mistakes, especially over limit speeds, (3) Police enforcement activities proceeded at a similar level or considerably more up to right now and (4) Beginning from Feb. 2008, the Higher Council for Traffic Safety has met once per month to follow road safety issues. Regarding the seasonal component in Figure 5(c), months from May to October exhibit higher injuries than other months in a given year. The highest increase was by about 19% in the month of August.

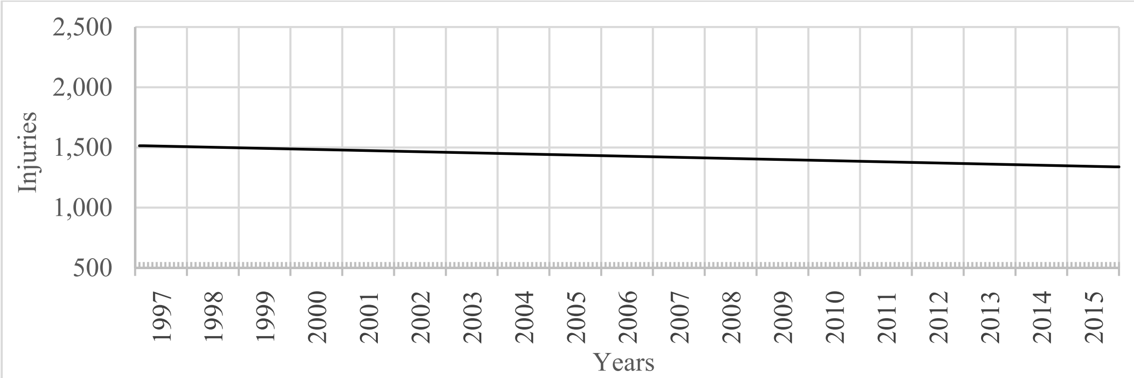
The irregularity of injuries shown in Figure 5(d) has the similar interpretation as that mentioned for accidents series above. Note that the period studied for injuries was 1997-2015 based on data availability.



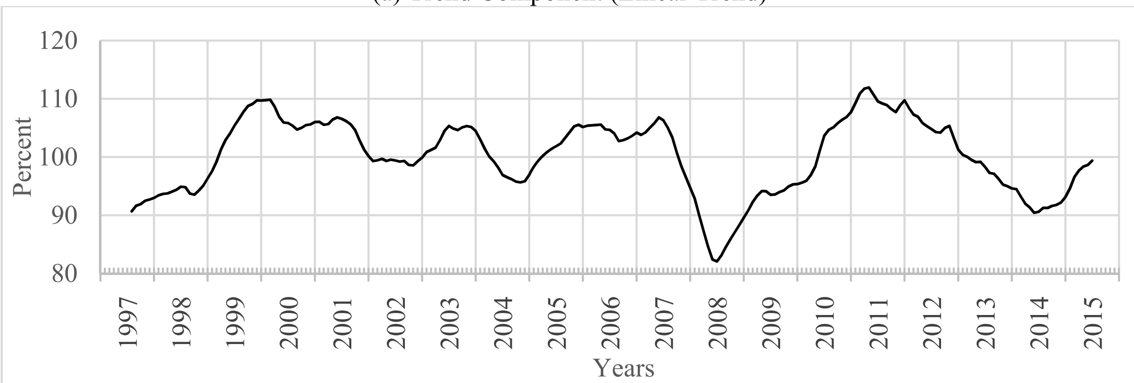
(a) Injuries series and systematic component

(b) Fluctuating component of injuries series

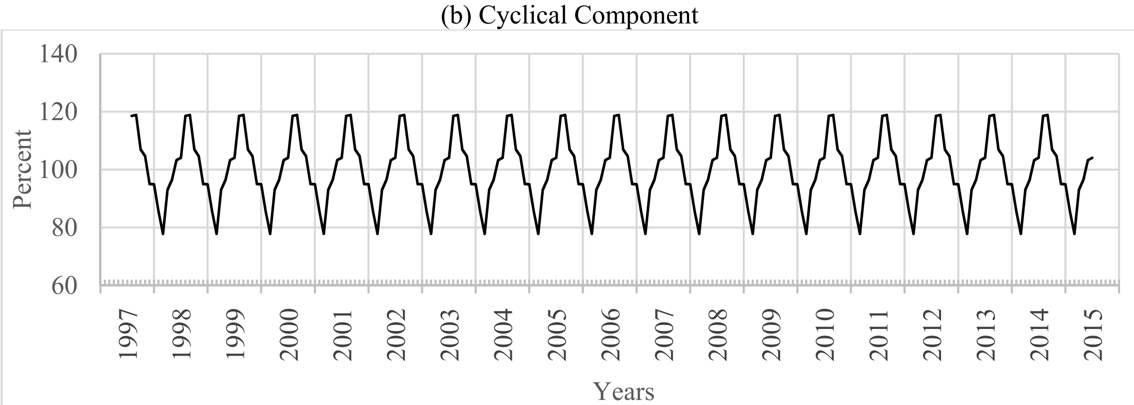
Figure (4): Injuries series and fluctuating component



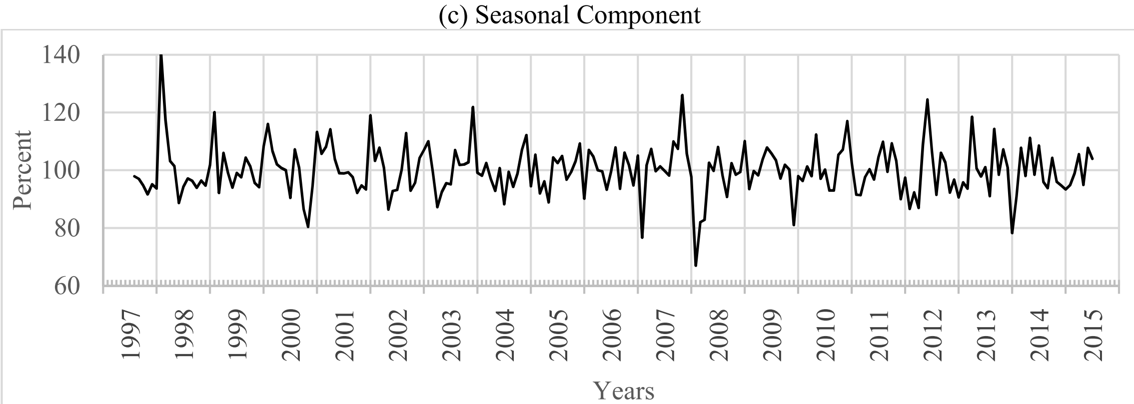
(a) Trend Component (Linear Trend)



(b) Cyclical Component



(c) Seasonal Component

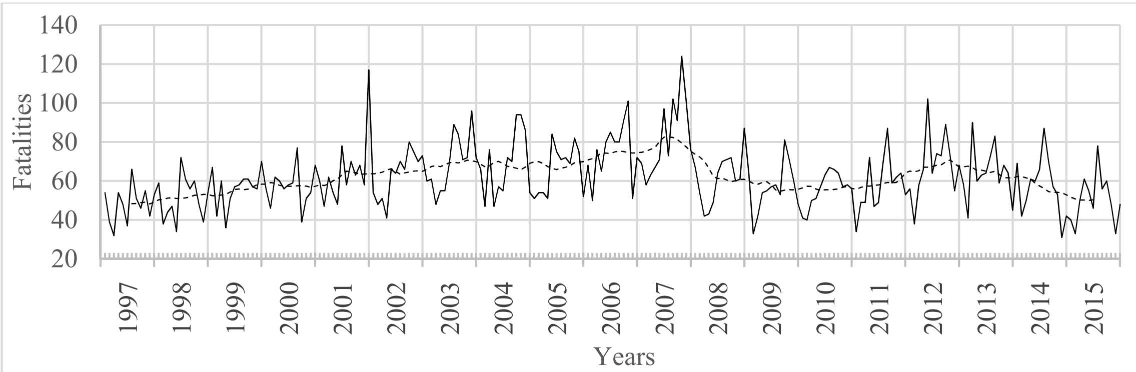


(d) Irregular Component

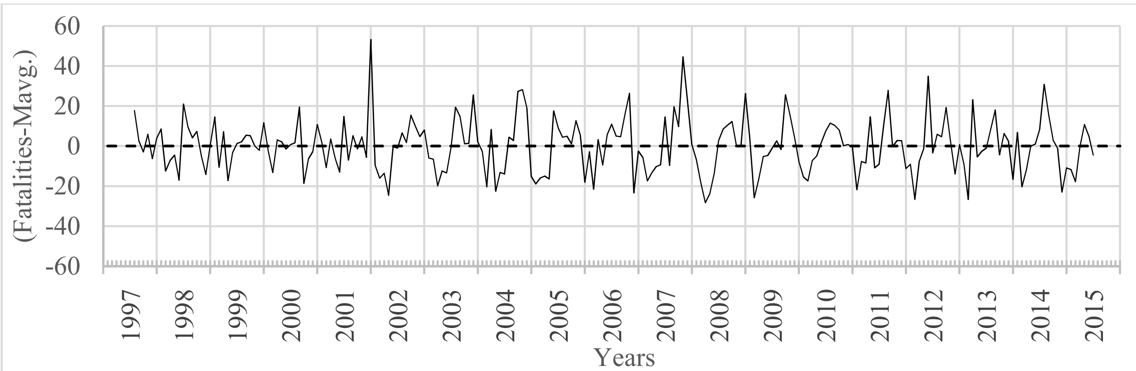
Figure (5): Components of injuries series

For fatalities series, Figures 6 and 7 are provided. Regarding fatalities series and moving average in Figure 6(a), the moving average of fatalities number began with about 50 fatalities in 1997 and ended with approximately the same average.

In the period 2002-2007, the fatalities series slightly increased. In this series, the relationship between number of fatalities and time was not significant, as shown in Figure 7(a). The cyclical component, Figure 7(b), shows three time zones, where the first time zone (1997-2007) contained the highest increase of 33% that occurred in July 2007. The months from June to October with December have higher fatalities than other months every year due to the seasonal component of fatalities series, as in Figure 7(c). This pattern may be attributed to the high accidents that occurred in the same months as mentioned previously in the analysis of accidents series. The irregular component in Figure 7(d) shows the random pattern of road fatalities due to randomness of accidents series.

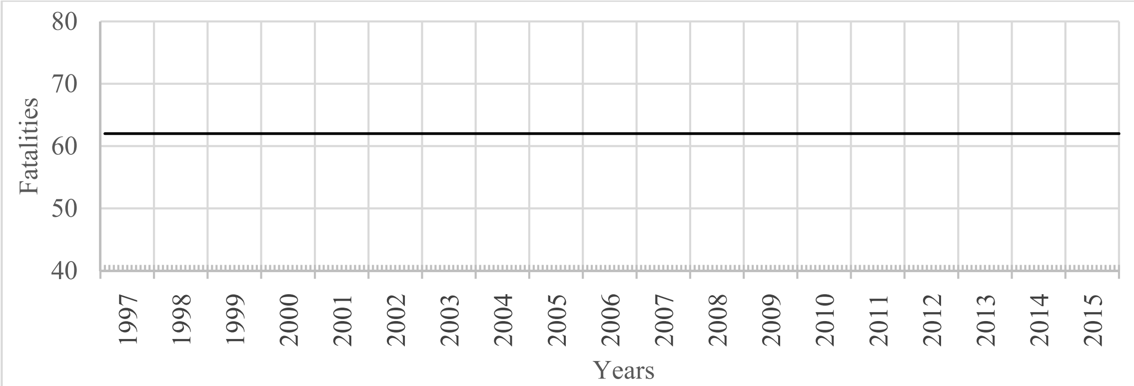


(a) Fatalities series and systematic component



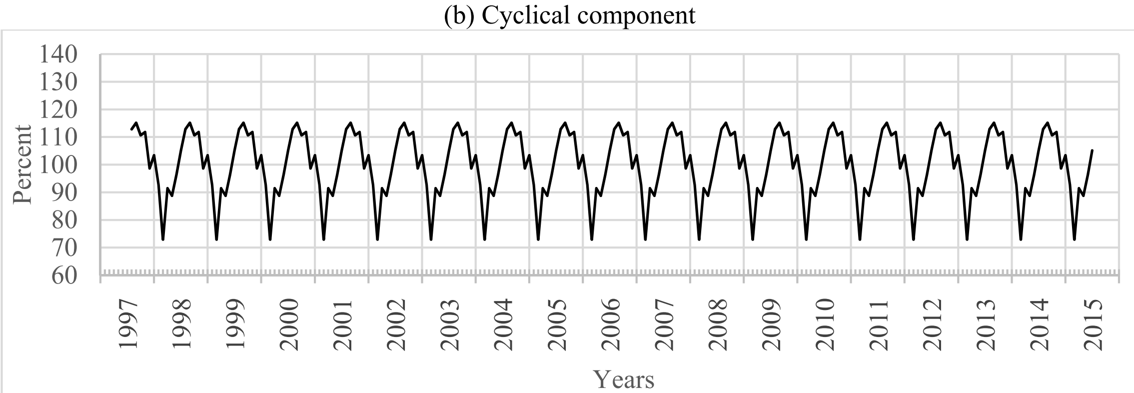
(b) Fluctuating component of fatalities series

Figure (6): Fatalities series and fluctuating component

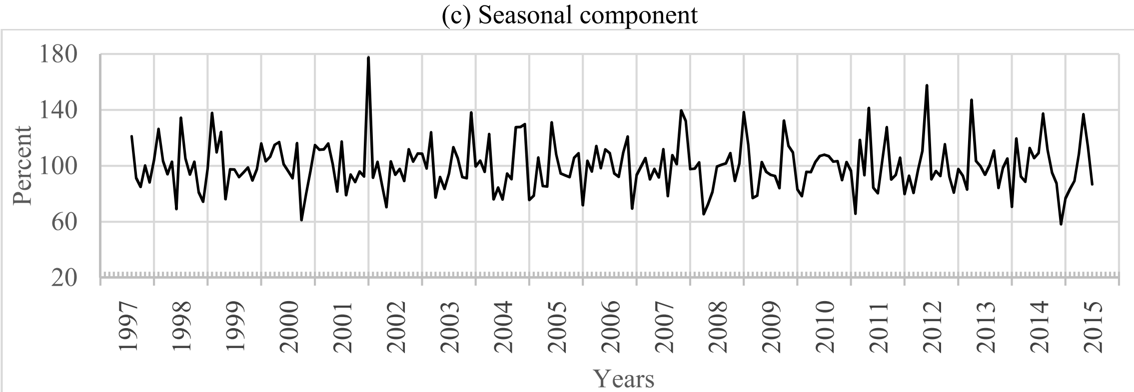


(a) Trend component (constant = average of data = 62.00)

(b) Cyclical component



(c) Seasonal component



(d) Irregular component

Figure (7): Components of fatalities series

Results of Time Series Forecasting

All results of time series forecasting analysis are summarized and provided in the equations and figures below. Each equation consists of two terms: the first term is the value of seasonal factor shown in Table 5 for all equations, while the second is the trend value represented by the regression function. To show the effect of these models, the determination coefficients have been calculated before and after applying time series models.

Regarding accidents' equation, the value of R^2 improved from 0.73 to 0.77 when the time series model was applied (Equation 8). Similarly, injuries and fatalities models (Equations 9 and 10) having 0.05 and 0.00 improved to become 0.54 and 0.22, respectively. This analysis looks like a powerful tool, especially in injuries and fatalities models that managed to convert insignificant relationships into significant ones. Figure 8 shows the applicability of predicted values to actual values with a good match between them.

$$\boldsymbol{FC_{accident.}} = \boldsymbol{S_f} \times (\mathbf{2913.009} + \mathbf{39.013}\boldsymbol{X_t}) \quad (8)$$

R^2 (without time series model) = 0.73	R^2 (with time series model) = 0.77

$$\boldsymbol{FC_{injuries}} = \boldsymbol{S_f} \times (\mathbf{1531.089} - \mathbf{0.900}\boldsymbol{X_t}) \quad (9)$$

R^2 (without time series model) = 0.05	R^2 (with time series model) = 0.54

$$\boldsymbol{FC_{fatalities}} = \boldsymbol{S_f} \times (\mathbf{60.357} + \mathbf{0.01226}\mathbf{X_t}) \quad (10)$$

R^2 (without time series model) = 0.00	R^2 (with time series model) = 0.22

Results of Forecasting for the Year 2030

Figure 9 shows the predicted monthly values in 2030 for numbers of accidents, injuries and fatalities. Regarding Figure 9(a), the highest number of accidents per month was 21,302 in August. The highest numbers of injuries (Figure 9(b)) and fatalities (Figure 9(c)) were 1,388 and 75 in the same month. The lowest values were 15,630 accidents, 913 injuries and 48 fatalities in February.

The percentages of increase of these values with respect to the year 2015 were found to be about 101.3% and 28.9% for accidents and fatalities, respectively. Injuries showed a decrease in the total number for the year by 13.2%.

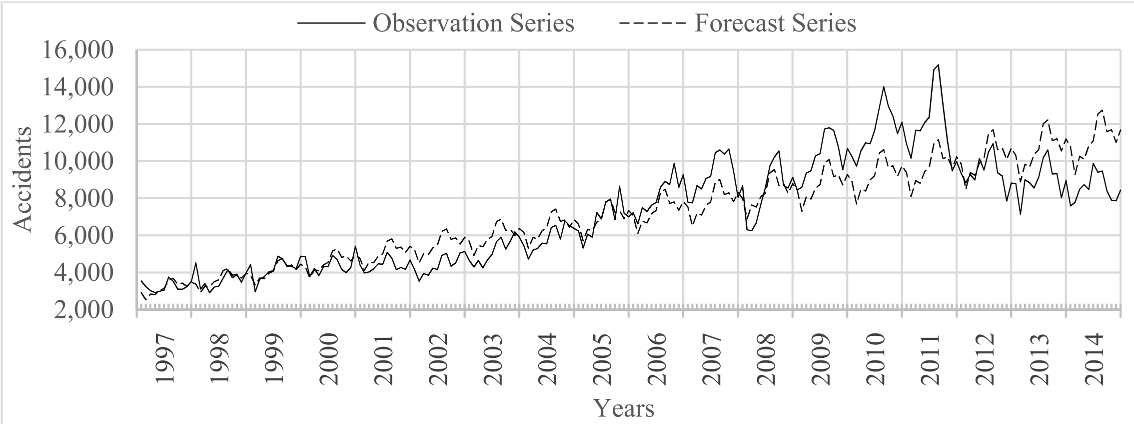
CONCLUSIONS

Time series modeling included forecasting of accidents, where the analysis was based on monthly data of accidents, injuries and fatalities. Each series of these numbers was taken alone and was decomposed into its components (trend, cyclical, seasonal and irregular components). All numbers regarding accidents, injuries and fatalities in summer months were found to be higher than in winter months due to their seasonal components. They had high irregularity according to the analysis of their irregular components. The forecasting by time series analysis in general showed improved coefficients of determination (R^2) for accidents, injuries and fatalities models.

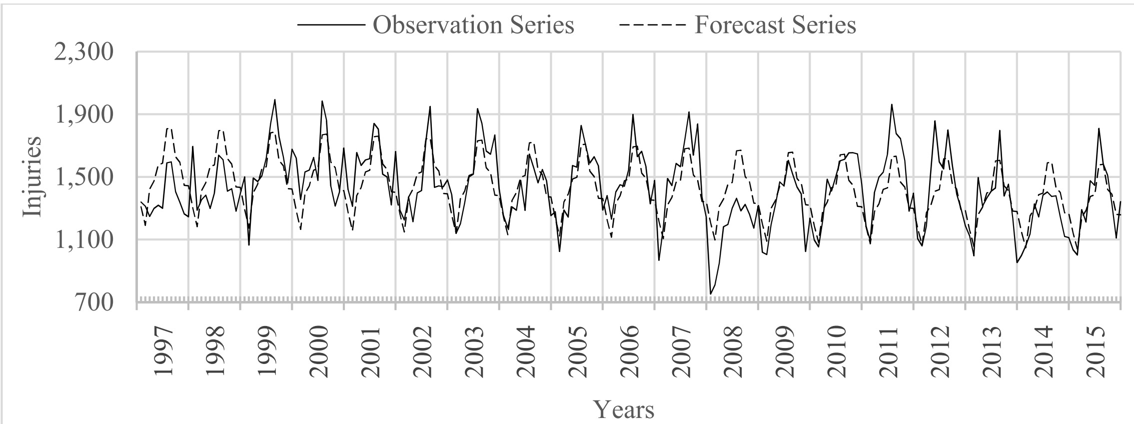
Other conclusions are enumerated as follows:

1. The analysis of time series was a powerful tool in assessing the traffic accidents' situation.
2. The time series modeling showed significant improvement in relationships between time (months) as a single predictor variable and the numbers of accidents, injuries or fatalities as dependent variables.
3. The R^2 value rose slightly from 0.73 to 0.77 for accidents model and significantly from 0.05 to 0.54 for injuries model. For fatalities model, on the other hand, the R^2 value rose from no relationship to slight relationship, but this relationship was not significant with an R^2 value of 0.22.

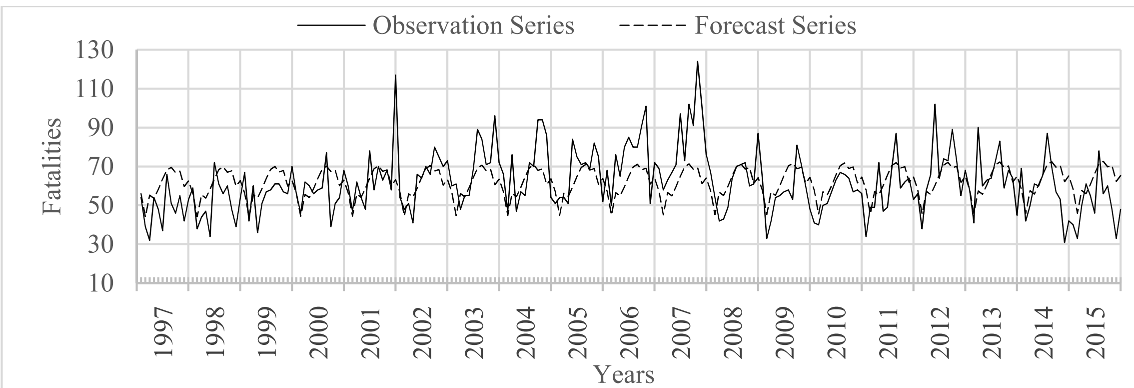
Finally, it is hopeful that the outcomes of this study be treated as the first step towards the enhancement of traffic safety measures in developing countries as well as to help in planning for the future. Moreover, this investigation may assist decision and law makers at different transportation and traffic agencies to come up with better resolutions for the increasing traffic crashes and consequently loss of lives and properties.



(a) Accidents forecast

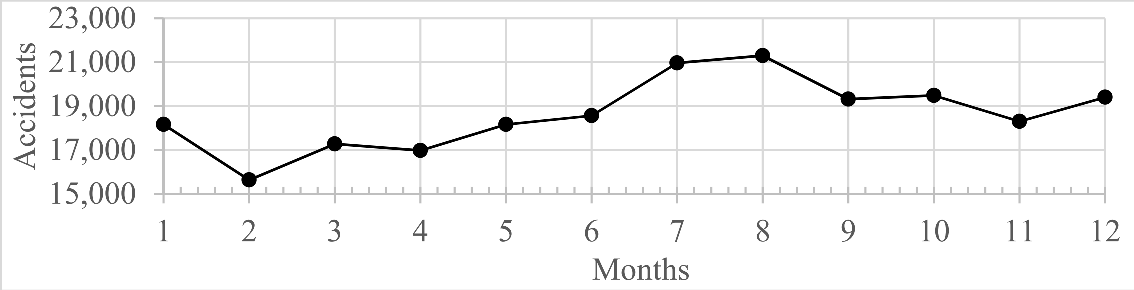


(b) Injuries forecast

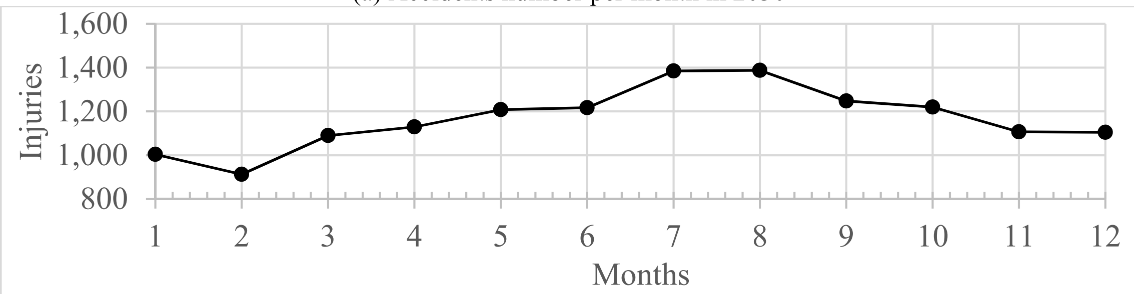


(c) Fatalities forecast

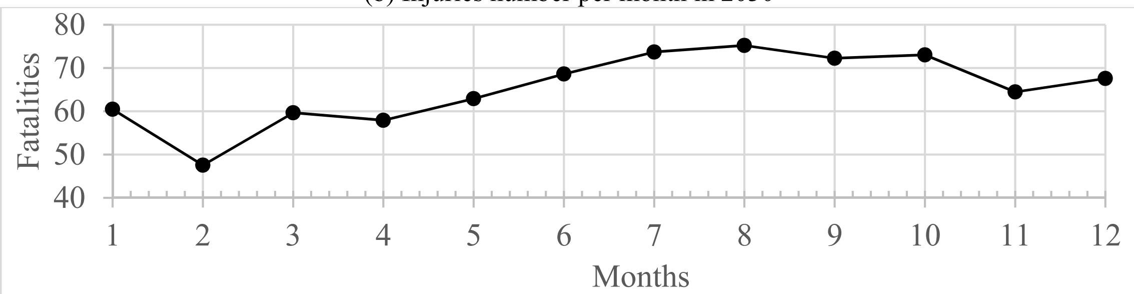
Figure (8): Results of time series forecasting



(a) Accidents number per month in 2030



(b) Injuries number per month in 2030



(c) Fatalities number per month in 2030

Figure (9): Forecasting for the year 2030 by time series models

RECOMMENDATIONS

Traffic safety with the development of cities has become an important factor that must be taken seriously into account. The concept of traffic safety is often associated with a limited budget, so we must determine the places of high risk levels through traffic studies like this study. The cyclical component showed improvements in accidents, injuries and fatalities in the middle of 2008 due to policy laws that were applied. Therefore, decision makers are encouraged to adopt and improve such policies. The implementation of surveillance systems, especially speed limit cameras in major road facilities, is an exceptionally viable method of convincing drivers not to speed and accordingly lessening the quantity of fatalities, injuries and loss of properties. This behavior can be observed in recent years due to accidents cyclical and faulting components as presented in this study. The seasonal components, on the other hand, show many safety concerns in summer months due to increased traffic volume caused by summer vacations and increased tourists' numbers. Therefore, decision makers must give special attention to this period of the year.

REFERENCES

Aljarrah, M.F., Khasawneh, M.A., and Al-Omari, A.A. (2019). "Investigating key factors influencing the severity of driver injuries in car crashes using supervised machine learning techniques". Journal of Engineering Science & Technology Review, 12 (4).

Al-Masaeid, H. (2009). "Traffic accidents in Jordan". Jordan Journal of Civil Engineering, 3 (4), 331-343.

Al-Omari, A., Khasawneh, M., and Ganam, B. (2019). "Evaluation of traffic accidents in Jordan using accident hazard scale". Jordan Journal of Civil Engineering, 13 (1).

Al-Omari, B., Ghuzlan, K., and Hasan, H. (2013). "Traffic accident trends and characteristics in Jordan". International Journal of Civil and Environmental Engineering, 13 (5), 9-16.

Filip, A., Bossche, V., Wests, G., and Brijs, T. (2007). "Analysis of road risk by age and gender category time series approach". Transportation Research Board Journal, 2019, 7-14.

Griggs, D., Stafford-Smith, M., Gaffney, O., Rockström, J., Öhman, M. C., Shyamsundar, P., and Noble, I. (2013). "Policy: Sustainable development goals for people and planet". Nature, 495 (7441), 305.

Helfenstein, U. (1990). "When did a reduced speed limit show an effect? Exploratory identification of an intervention time". Accident Analysis and Prevention Journal, 22 (1), 79-87.

Jordan Traffic Institute (2015). "Yearly report of traffic accidents in Jordan". Ministry of Interior, Public Security Directorate, Amman, Jordan.

Khasawneh, M.A., Al-Omari, A.A., and Ganam, B. (2018). "Forecasting traffic accidents in Jordan using regression techniques". Jordan Journal of Civil Engineering, 12 (4).

Leeming, J.J., Mackay, G.M., Pole, K.F.M., and Fitzgerald, P.J. (1969). "Road accidents: Prevent or punish?" London: Cassell.

Offaly Historical & Archaeological Society. (2007). "Mary Ward 1827-1869: Famous offaly people". (Created on 2 September 2007), 1827-1869. (Accessed on 18 November 2016).

United Nations General Assembly in New York. (2003). "Session 57 verbatim report 86". Page 2. 22 May 2003.

World Health Organization. (2004). "World report on road traffic injury prevention: Chapter 2- The global impact". Page 25. Accessed on 10 November 2016.

World Health Organization. (2009). "Road safety report: Global status report on road safety". Accessed on 18 November 2016. Leading causes of death, 2004 and 2030 compared. Page 9.

World Health Organization. (2015). "Global status report on road safety". Retrieved from: http://www.who.int/violence_injury_prevention/road_safety_status/2015/en/

Yu, Q., Guo, Z., Zhang, Z., and Wang, J. (2015). A forecasting model for setting road safety targets based on time-series analysis". 15th COTA International Conference of Transportation Professionals, July, 24-27, 2015, Beijing, China, ASCE Online Library, 2985-2992.