Time Series Modeling of Road Traffic Accidents in Amhara Region

Kidane Alemtsega (alemkidtsega@gmail.com)
Bahir-Dar University Institute of Technology, Bahir-Dar City

Research

Keywords: Road Traffic, Accident, Injury, Fatal, Modelling, Time series, ARIMA, Amhara, Bahir dar University

Posted Date: March 16th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-299726/v1

Version of Record: A version of this preprint was published at Journal of Big Data on July 19th, 2021. See the published version at https://doi.org/10.1186/s40537-021-00493-z.

Time Series Modeling of Road Traffic Accidents in Amhara Region

Kidane Alemtsega

Department of statistics, Bahir-Dar University Institute of Technology, Bahir-Dar City, Ethiopia,

Email: almkidtsega@gmail.com

Abstract

Road traffic accidents are commonly encountered incidents that can cause injuries, death, and property damage to members of society. Ethiopia is one of the highest incident rates of road traffic accidents. Report of Transport and Communication from 2012-2014, shows an increment in the number of traffic accidents in Ethiopia. Amhara region accounted for 27.3% of the total road traffic accident-related deaths in Ethiopia during the year 2008/9, which is the highest share among all regions. The current research aims to model the trend of injury, fatal and total road traffic accidents in the Amhara region from September 2013 to May 2017. Monthly reported traffic accidents were obtained from traffic police offices and other concerned governmental organizations at the zonal and regional levels. The most universal class of models for forecasting time series data called Auto-regressive Integrated Moving Averages models are applied to model the trends and patterns of road traffic accident cases in the Amhara region. The average number of observed injury RTA, fatal RTA, and total RTA were 27.2, 14, and 78.2 per month respectively. It was observed that a relatively large number of RTA'S are reported on Tuesday, Thursday, and Saturday relative to other days of the week. The data also reveals that more than 60% of accidents involve drivers between the ages of 18-30 years. ARIMA (2,0,0) (1,0,0) ARIMA (2,0,0) and ARIMA (2,0,0) (1,1,0) were fitted as the best model for total injury accidents, fatal RTA and total RTA data respectively. A 48 months forecast was made based on the fitted models and it can be concluded that road traffic accident cases would continue at the non-decreasing rate in the Amhara region for the predicted periods. Therefore, the findings of this study draw attention to the importance of implementing improved better policies and close monitoring of road trafficking to change the existing trend of road traffic accidents in the region.

Keywords: Road Traffic, Accident, Injury, Fatal, Modelling, Time series, ARIMA, Amhara, Bahir dar University

Background

A road traffic accident is a random event involving a road user that results in property damage, death, or injury[1]. Road traffic accidents cause an estimated 13 million deaths and 20-50 million disabilities worldwide annually, notably, 85% of injury-related deaths occur in developing countries. The burden attributed to road safety is comparable tuberculosis and malaria; approximately it costs 3% of the planet GDP[2]. RTA injuries accounted for 23% of all injury deaths worldwide[3]. Road traffic accidents (RTAs) affect populations all over the world; different local factors influence the causes of RTA in specific regions. The causes of RTAs' among others include human or driver errors, vehicle characteristics, traffic infrastructures including engineering design, road maintenance, and traffic regulation[4]. Driver attitude including road courtesy and behavior[5], driving under the influence of drugs especially alcohol, male sex, use of seat belts, driver age (teenage drivers and elderly drivers) are among the recognized human factors[4]. It is thought that globally 1.4 million people are killed in RTAs and about 20 million to 50 million people are injured or disabled from the effect of RTAs. Unchecked, by the year 2020, RTAs will rank third of all causes of morbidity and mortality globally[6]. Of worry, though is that RTA-related fatalities seem to increase with the gross domestic product (GDP) per capita in lower-income countries and decrease with GDP per capita in wealthy countries[7].

Road traffic accidents are the most frequent causes of injury-related deaths worldwide[8]. According to the world report on road traffic injury prevention[6], traffic accidents account for about 3000 daily fatalities worldwide. Statistical projections show that during the period between 2000 and 2020, fatalities related to traffic accidents will decrease by about 30% in high-income countries. the alternative pattern is predicted in developing countries, where traffic accidents are expected to extend at a quick rate within the years to return. Fatalities due to traffic accidents in Ethiopia are reported to be among the highest in the world. According to global status report on road safety, the road crash fatality rate in Ethiopia was at least 114 deaths per 10,000 vehicles per year, compared to only 10 in the UK and Ireland and 60 across 39 sub-Saharan African countries[9]. Furthermore, it is sad to note

that fatalities due to road traffic accidents are higher among pedestrians in countries like Ethiopia than in developed countries. For instance, 60% of the fatalities in the US account for car drivers, while in Ethiopia only about 5% account for drivers[10]. This is also supported by a recent study where the majority of fatalities were pedestrians (87%) followed by passengers (9%) and drivers (4%), among a total of 25,110 accidents and 3415 fatalities during the period 2000-2009 in Addis Ababa.

Various studies have indicated that Ethiopia has one of the highest fatality rates per vehicle in the world. It is more than 100 fatalities per 10,000 vehicles. This should be compared with Kenya and the United Kingdom, where the figure is about 19 and 2 per 10,000 vehicles, respectively. Road traffic injury is high in Ethiopia, at least 70 people die for every 10,000 vehicle accidents annually[11]. According to this report, 86% of accidents are due to drivers' problems. Car technical problems, pedestrians' mistakes, road quality are among the major causes of road traffic accidents in Ethiopia this day. The other major cause mentioned in the report was driving without a license. Amhara region accounted for 27.3% of the total road traffic accident-related deaths in Ethiopia during the year 2008/9, which is the highest share among all regions[12]. This entails the need to examine the causes of accidents in the region. To date, there have been little types of research on road traffic accidents using pooled data in Amhara National Regional State (ANRS). The aim of the current study is therefore to model the trend of road traffic accidents in the Amhara region using time series modeling and forecast for the future. Road traffic accidents are becoming a major public safety problem and development obstacle in the world. The problem is threatening especially for developing countries like Ethiopia. Besides, the victims are mainly public transport travelers in the working-age group (18–30 years). This alarming statistic underpins the importance of updating and improving accident data records and subsequently the methods of analyzing traffic data as this will help policymakers to formulate evidenced based regulations and road safety measures.

Method

Description of Study Area and Source of Data

Amhara region is one of the regions of Ethiopia, containing the homeland of the Amhara people and its capital is Bahir Dar. Amhara region is bordered by the nation of Sudan to the west, and the Ethiopian regions of Tigray to the north, Afar to the east, Benishangul-Gumuz to the west and southwest, and Oromia to the south. Data used for the analyses was collected from Amhara region police commission traffic departments, transport, and communication bureau at woreda, zonal and regional level for a period of September 2013 to May 2017. The data collected mainly included a monthly recorded number of traffic accidents in the region and other related information about the cause of accidents. The data used in this study was the overall regional data on road traffic accidents. Time series plays significant role in the analysis of road traffic accidents. It is also very useful in forecasting because one of the objectives of time series is to forecast future values. Box-Jenkins (ARIMA) method will be used to derive models for forecasting these data. This method is preferred because of its high accuracy in forecasting data, especially within a short to medium term period[13].

Statistical Analysis

Along with descriptive statistics, time series analysis was conducted for data analysis purposes. The data analysis was carried out using the R-Statistical software package. One of the main objectives of statistics is to forecast the future levels of different processes by studying the behavior of the data in the past. The most important techniques of making inferences about the future based on what has happened in the past are the analysis of time series, which may be defined as a set of observations taking at specified times, usually at equal intervals.

Components of Time Series

In analyzing time series, we may take the observed composite series as a whole for study or study one by one respectively the components in their own right. The components are Secular Trends, Seasonal Trends, Cyclic Variations, and Random or Irregular Variations.

Mathematical Model of Time Series

There are two types of models in time series which are generally accepted as good approximations to the time relationship among the components of the observed data. They are the additive and multiplicative models and are the most commonly assumed relationship between time series and its elements.

Additive Model: This assumes that the value of composite series is the sum of the four components, that is

$$Y=T+S+C+I \quad (1)$$

Where Y- original, T - Values of the secular trend, S - the value of the seasonal component, C - the value of the irregular component, and I - the value of the irregular component

Multiplicative model: assumes that the value of composite series is the product of four component values, that is $E(Y_t)$= constant for all t and $Var(Y_t)$= constant for all t.

Due to the non-stationary nature of most business and economic time series, it is required that stationarity be achieved before building any model. We can differentiate the data.

The $Y= T \times S \times C \times I$ (2)

Generally, the multiplicative model has been considered the standard conventional model for the analysis of time series.

Measurement of Trend

The following are the three methods that are generally used for the study and measurement of the trend component in a time series. That is the Freehand method, the Moving average method, and the Method of least square. In this research work, the method of least-squares was used to estimate the trend value of time series by determining the equation for the best time of fit and also use it to determine whether there will be an increase or decrease in the observed data by predicting for the future occurrence. The study used the Box-Jenkins method to derive models for forecasting these data. This method is preferred because of its high accuracy in forecasting data, especially within a

short to medium-term period. Its model simplicity gives it an advantage of cost and response time because the high cost is required to run and set up complex models (Nihan & Holmesland, 1980).

Model Building

In theory, Auto-regressive Integrated Moving Averages ARIMA Models (Box Jenkins) are the most universal class of models for forecasting time series data. As proposed by Box and Jenkins, that in general, forecasting based on ARIMA models comprises three different steps: Model Identification, Parameter estimation, and Diagnostic checking. Until a desirable model for the data is identified, the three steps have been repeated[14].

Stationarity

A time series is stationary if there is no systematic change in mean (no trend), variances and strictly periodic variations have been removed. In other words, a time series is said to be stationary if:

$Cov(Y_t, Y_{t+k})$= constant for all t

differenced data will contain one less point than the original. For non-Constance variance, taking the logarithm or square root will stabilize the variance. For non-seasonal data, first-order differencing is usually sufficient to attain apparently $(x_1), \ldots\ldots, (x_n)$ by

$$Y_t = x_{t+1} - x_t = \Delta x_{t+1} \quad (3)$$

Occasionally, second-order differencing is required using the operator Δ^2 where

$$\Delta^2 = \Delta x_{t+2} - \Delta x_{t+1} - 2x_{t+1} + x_t \quad (4)$$

Hence the number of times that the original series is differenced to achieve stationarity is the order of homogeneity.

Differencing

This is a special type of filtering which is particularly useful moving a trend. This is achieved by subtracting each data in a series from its predecessor. For non-seasonal data, first-order differencing is sufficient to obtain apparent stationarity. The concept of backshift operator helps to understand and express differenced ARIMA models.

Auto-Regressive Integrated Moving Average (ARIMA(p,d,q))

ARIMA (p, d, q) model was first introduced by Box and Jenkin in 1976 [32], it can be used for forecasting the non-seasonal stationary time-series data. An ARIMA model is characterized by 3 terms: p, d, q where p is the order of the Auto-Regression (AR) term, q is the order of the Moving Average (MA) term, d is the order of differencing required to make the time-series stationary. Auto-Regression is nothing but the regression of the variable against itself to forecast the variable of interest. It correlates the pattern of the one-time period to its previous time periods. MA is a regression-like model that uses the errors associated with the forecast at a previous time-step to forecast a variable at a later time-step. The following are the generalized equations of p^{th} order AR model and q^{th} order MA model.

$$y_t = C + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdot\cdot\cdot\cdot\cdot\cdot\cdot + \varphi_p\, yt\text{-}p + E_t$$

$$yt = C + E_t + \theta_1 E_{t-1} + \theta_2 E_{t-2} + \cdot\cdot\cdot\cdot\cdot\cdot\cdot + \theta q E_{t-q}$$

ARIMA models are built upon incorporating the AR model, integration (I) and the MA model. The integration (I) is the reverse process of differencing to generate the forecast. The generalized ARIMA model is mathematically represented as

$$yt = C + \varphi_1 y + \varphi_p y_{t-p} + \cdot\ \cdot\cdot\cdot\cdot\cdot + \varphi_n y_{t-n} + \theta_1 E_{t-1} + \theta q E_{t-q} + E_t$$

Where C is intercept, φ_i (i = 1, 2 . . . p) is auto-regressive model parameters, θ_i(i = 1, 2 . . . p) is moving average model parameters, y_t is current time-series value, y_{t-1}, $y_{t-2}...y_{t-p}$ is past values and E_t is random error or residual term for the t^{th} day and it is given by the following equation:

$$Et = y_t - y_{t-1}$$

Box and Jenkins Method

This method of forecasting implements knowledge of autocorrelation analysis based on autoregressive integrated moving average models. The procedure is of four main stages namely: Identification, Estimation, Diagnostics Checking, and Forecasting.

Identification

The first step in developing an ARIMA model is to determine if the series is stationary. If the model is found to be non-stationary, stationarity could be achieved mostly by differencing the series or going for the dickey fuller test. Stationarity could also be achieved by some mode of transformation like the log transformation. Once stationarity has been achieved, the next step is to determine the orders of the autoregressive (AR) and moving average (MA) terms using the Auto Correlation Function (ACF) and Partial Auto Correlation Function (PACF). stationarity so that the new series $(Y_1, \ldots \ldots, Y_n)$.

Estimation

Once the preliminary model is chosen, the estimation stage begins. The purpose of estimation is to find the parameter estimates that minimize the mean square error. Two approaches are used and these include the nonlinear least squares and maximum likelihood estimates. In this method, the R statistical package was used in the estimation.

Diagnostics Checking

Residuals from the model are examined to ensure that the model is adequate (random). The following diagnostics are made: Time plot of the residuals Plot of the residual ACF Normal Quantile Quantile (QQ) Plot.

Forecasting

When a satisfactory ARIMA model is adequate, then we proceed to forecast or predict for a period or several periods ahead. However, chances of forecast errors are inevitable as the period advances.

Results

The current study applied time series analysis to model the patterns of the total number of road traffic accidents that lead to injury, total fatal road traffic accidents, and the total number of traffic accidents that occurred from September 2013 to May 2017 in the Amhara region, Ethiopia.

Descriptive Statistics

Over the 45 months observation (from Sept. 3013 to May 2017), a total of 3385 road traffic accidents were reported. The summary statistics for counts of Total Injuries, Fatal accidents, and Total number of RTA during the study period are presented below.

Table -1 Summary statistics of road traffic accidents

Data	Min.	Q1	Median	Mean	Q3	Max
Injury	*16*	*24*	*26*	*28.2*	*33*	*43*
Fatal	*5*	*12*	*14*	*14*	*16*	*27*
Total	*43*	*68*	*78*	*75.2*	*84*	*108*

As is shown in table -1 above the average number of total accidents per month was around 76. The maximum was 108 and the minimum was 43 which were recorded in October 2014 and March 2013 respectively. It was also shown that an average of 14 fatal accidents has been observed per month over the study period.

Table -2 Total Number of RTA (September 2013 to May 2017) victims in Amhara region by Age

	Age Group	2013	2014	2015	2016	2017
Drivers	<18	5	5	14	8	8
	18-30	80	122	121	150	200
	31-50	57	67	74	80	82
	>51	7	12	33	37	41
Pedestrians	<7	52	75	83	74	61
	7-13	131	55	64	61	72
	14-17	110	149	117	112	135
	18-30	314	426	218	352	404
	31-50	233	274	414	299	289

	>51	74	128	329	326	337
Passengers	<7	16	19	31	34	36
	7-13	22	40	97	94	106
	14-17	155	182	182	180	166
	18-30	832	1101	725	827	982
	31-50	538	650	689	764	802
	>51	94	116	344	229	157

Table- 2 above cross-tabulated the total number of road traffic accidents based on age groups for drivers, pedestrians and passengers. Among drivers, considerably large number of accidents were caused by drivers between the age group 18-30 years. Regarding pedestrians and passengers, although the number of victims were relatively high for the age group 18-30, the problem is common to individuals of all age group.

Table- 3 Total Number of accidents in Amhara region (September 2013 to May 2017) by days of the week

Day	Year				
	2013	2014	2015	2016	2017
Monday	84	111	96	100	93
Tuesday	127	135	128	130	110
Wednesday	67	77	70	97	90
Thursday	108	161	113	127	126
Friday	74	95	93	89	87
Saturday	106	120	129	144	137
Sunday	75	81	76	90	84

In table-3, it was tried to present the observed road traffic accidents by days of the week during the study period. According to the observed data road traffic accidents are more prevalent on Saturday,

Thursday, and Tuesday relative to other days of the week in the Amhara region. It may be related to the market days which are common in the region. So, attention should be given to the mentioned days of the week so that a significant number of accidents can be minimized.

Exploratory Analysis

Time series plots of the injury cases, fatal RTA cases, and total RTA cases are shown below.

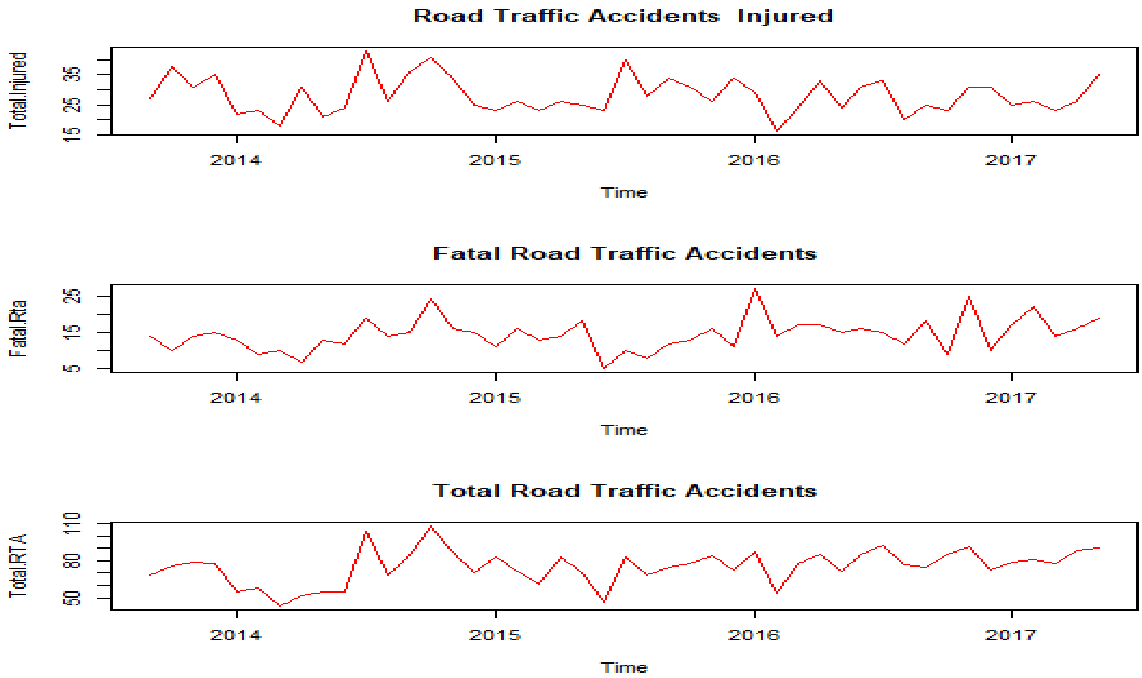

Fig. 1. Time series plot of monthly reported Injury, fatal and total road traffic accidents in Amhara Region (September 2013 to May 2017)

We can't infer from the above time series plot that the trend of the data since the graphs didn't show clear patterns. Therefore, it needs to do some analysis to find out the exact non-stationary and seasonality in the data. As it can be shown from the time series decomposition plots, the observed fatal RTA cases and total RTA cases have an overall increasing trend. On the other hand, injury RTA cases show an overall decreasing trend.

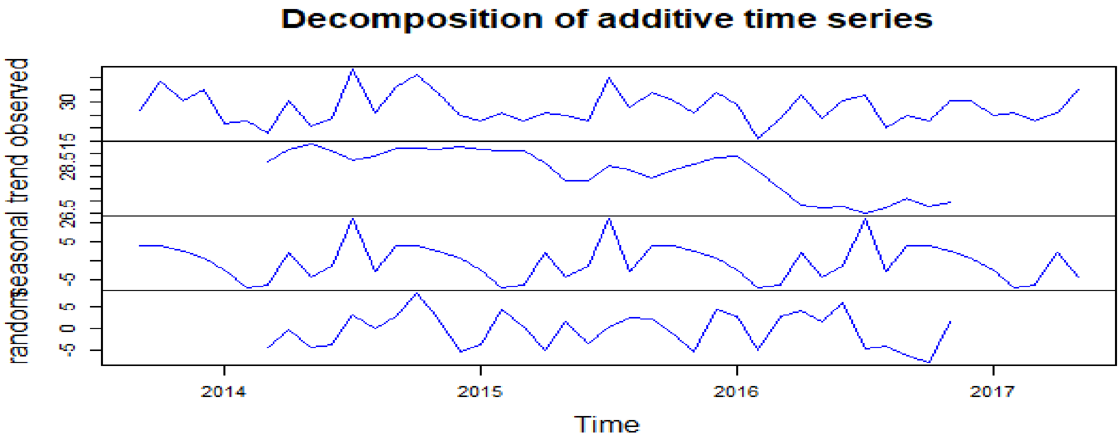

Fig. 2. Time series decomposition of injury RTA data

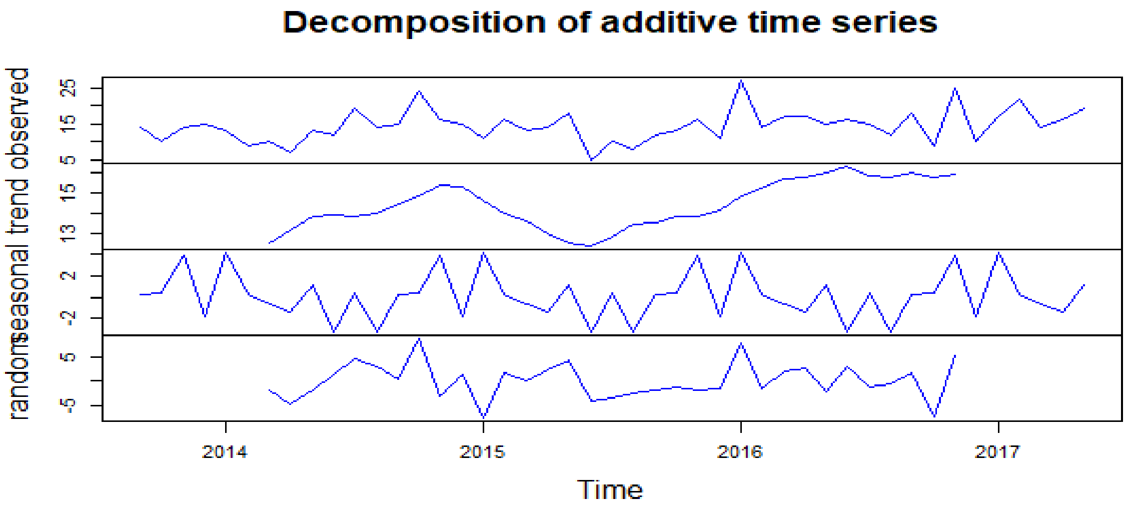

Fig. 3. Time series decomposition of fatal RTA data

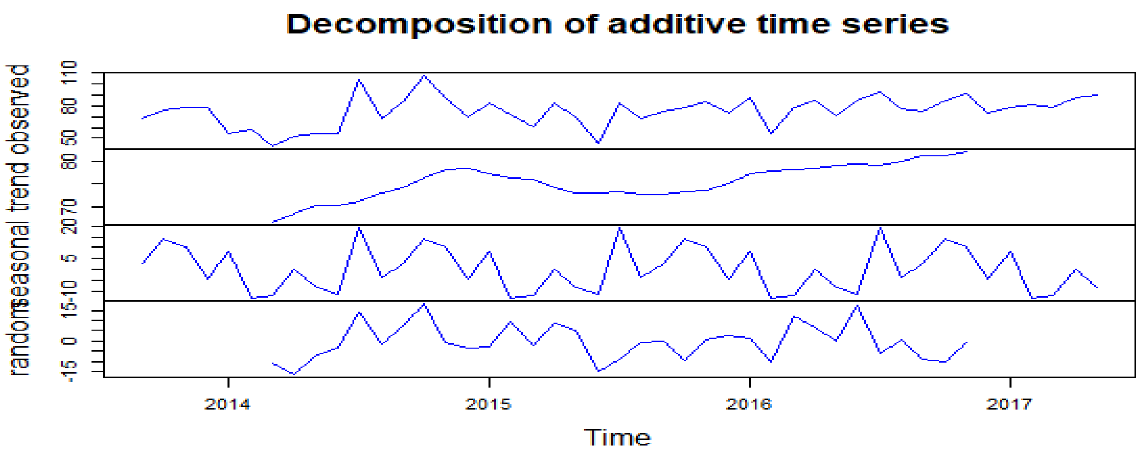

Fig.4. Time series decomposition of Total RTA data

Test of Stationarity

To test the stationarity of the data Augmented Dickey-Fuller test for stationarity was used. According to the ADF-test output in table- 4, fatal RTA and Total RTA data are not stationary. Only the injury RTA data is shown to be stationary since it has a p-value less than 0.05. Therefore, the fatal RTA and Total RTA data should be differenced to achieve stationarity.

Table -4 Summary statistics for ADF test of stationarity

Time Series	Adf-Test Statistic	P-value
Injury RTA	-3.9963	0.01849
Fatal RTA	-2.9403	0.2008
Total RTA	-3.3105	0.08238

Table -5 Summary statistics for ADF test of stationarity after differencing

Time	Adf-	P-value	Differencing(d)
Fatal	-3.843	0.02493	1
Total	-4.921	0.01	2

After applying the first-order differencing on fatal RTA and total RTA cases, it can be shown that both p-values of Adf-Test are small (less than 0.05) which suggests the data is stationary and doesn't need to be differenced further.

Fitting Time Series Models

Once the data is ready and satisfies all the assumptions of modeling, the next step is to select the appropriate ARIMA model i.e. to determine the order of the model to be fitted. This task can be done by examining two important plots called autocorrelation function (ACF) and partial autocorrelation (pacf) plots of the stationary time series. For the current data, the Pacf and ACF plots are shown below.

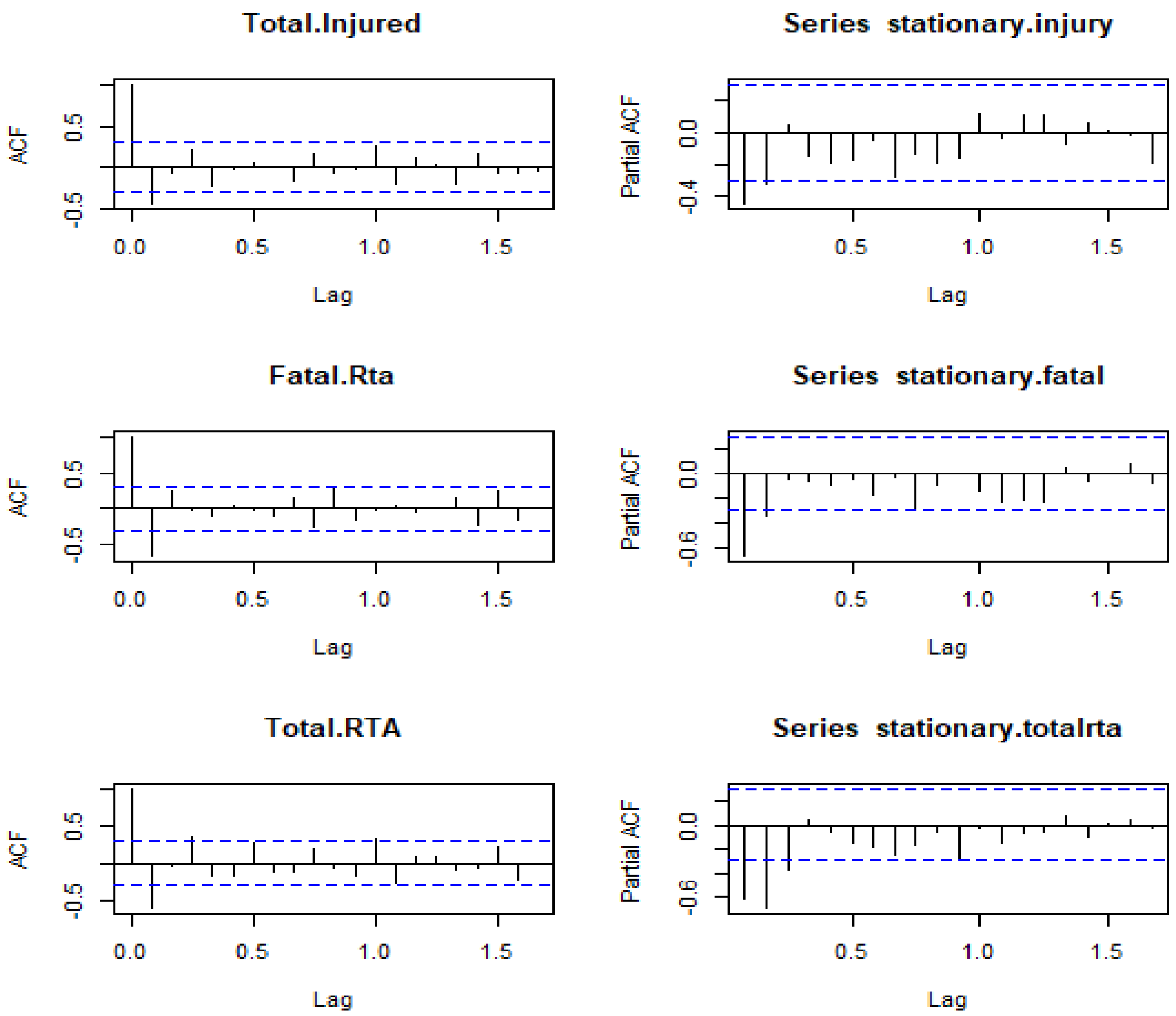

Fig.5 Autocorrelation and Partial autocorrelation function plots of Injury, Fatal and Total RTA

Model Selection

The model selection was done by examining the ACF and PACF plots (fig. 5). The AIC and BIC values are also used to choose the best model. These criteria are closely related and can be interpreted as an estimate of how much information would be lost if a given model is chosen. ARIMA (2,0,0)(1,0,0) [12], ARIMA (2,0,0) and ARIMA (2,0,0) (1,1,0) [12] are chosen as the best models for Total injury, Fatal RTA and Total RTA series respectively. The chosen models above have the lowest AIC and BIC values compared with the other possible models.

Table -6 Parameter estimates & goodness of fit measures of injury, fatal and total RTA

	Parameters	Coefficients	SE	AIC	BIC
Injury RTA	AR1	-0.6309	0.1424	297.21	304.34
	AR2	-0.4299	0.1414		
	SAR1	0.4283	0.1614		
Fatal RTA	AR1	-0.8873	0.1406	264.25	269.61
	AR2	-0.3394	0.1396		
Total RTA	AR1	-1.1453	0.1447	276.14	281.88
	AR2	-0.6523	0.1174		
	SAR1	-0.6643	0.1587		

Model Diagnostics

Now before using the fitted model for forecasting, one should examine the adequacy of the models. This can be commonly checked by examining ACF and PACF plots for model residuals. If model order parameters and structure are correctly specified, we would expect no significant autocorrelations present.

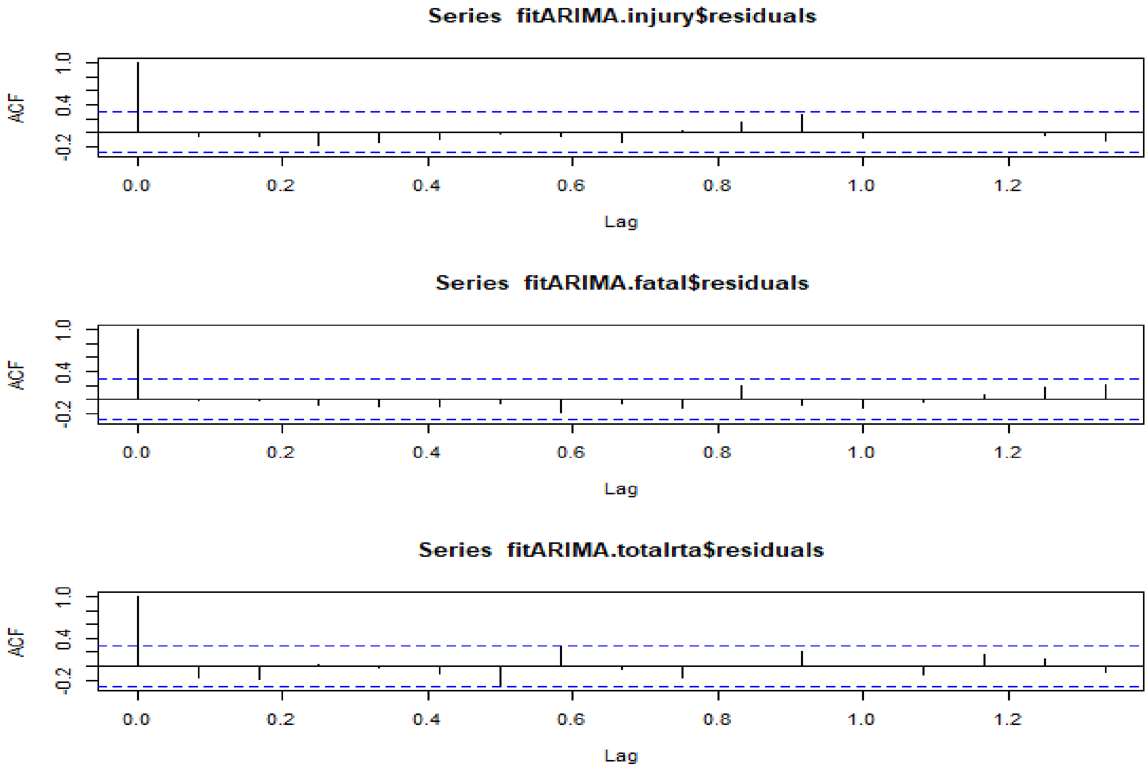

Fig. 4. Diagnostics of residuals for ARIMA of injury, fatal and total RTA data

All standardized residuals plot shows no obvious trend and pattern and looks like an independent and identical distribution. Similarly, all plots of ACF residuals of the diagnostics show no evidence of significant correlation in the residuals. Most of the residuals are located on the straight line except for a few outliers deviating from the normality. In conclusion, the residuals ACF and normal Q-Q plots exhibited a white noise pattern and the goodness of the fitted model.

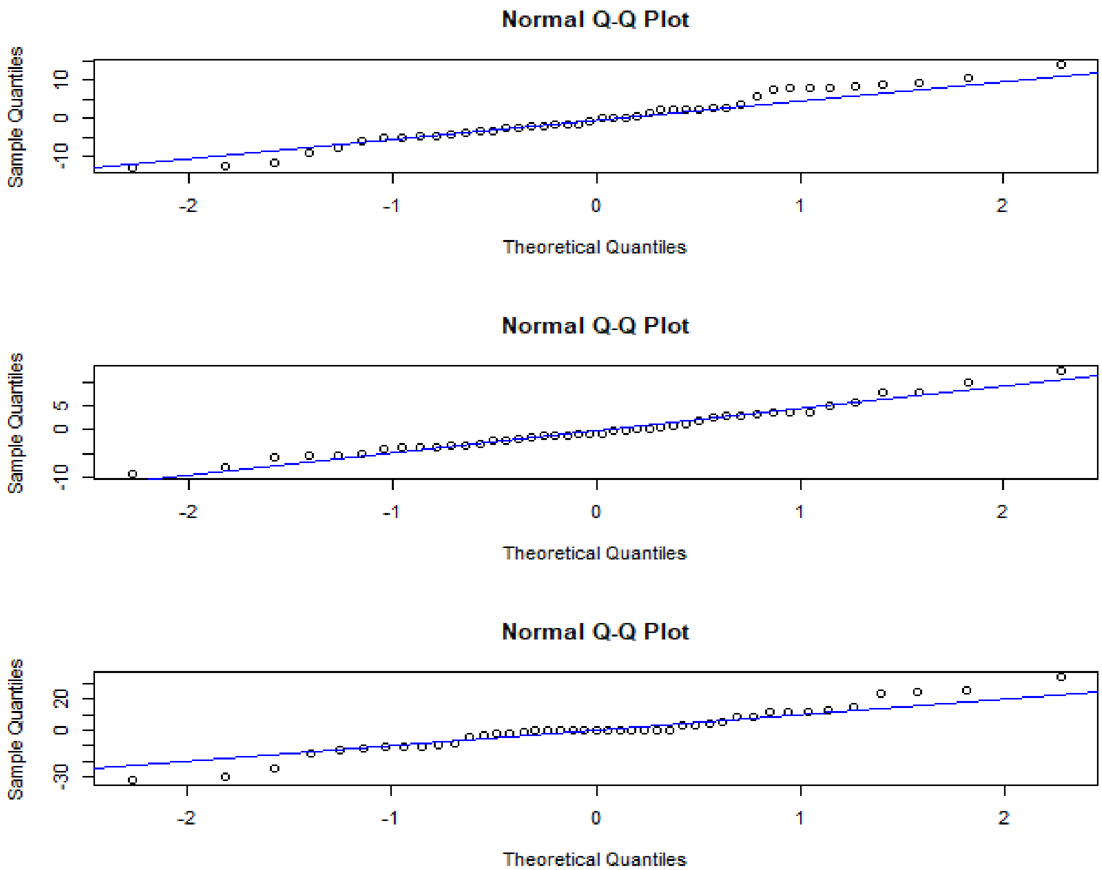

Fig. 5 Normal Q-Q plot of residuals for ARIMA of injury, fatal   and total RTA data

Forecasting

The model was fitted for 48 months period after the diagnostic test. Below is the graph which gives a pictorial view of the observed series, its forecast, and confidence intervals of the forecast.

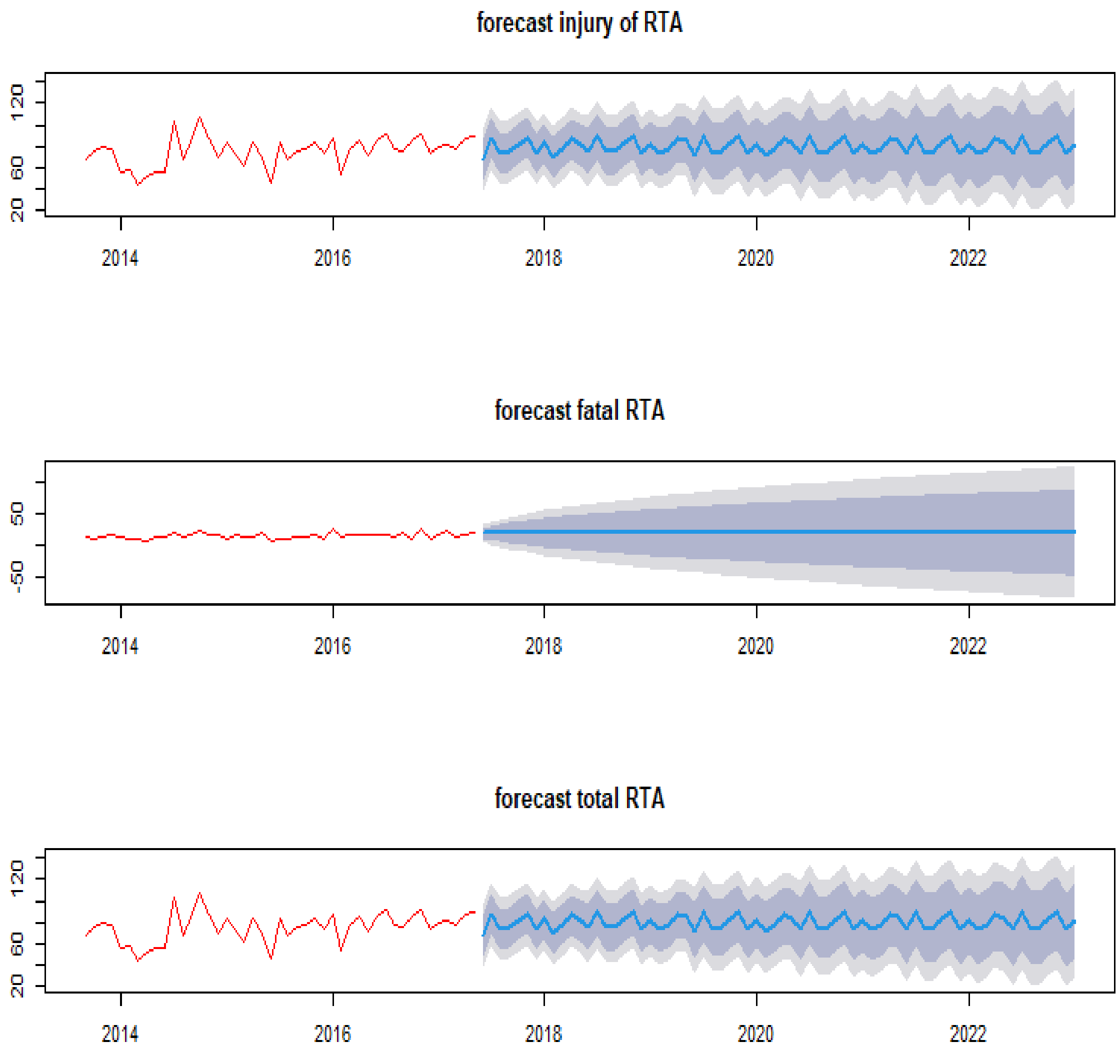

Fig. 3. Graph of injury, fatal and total RTA cases, its forecasts, and confidence intervals

Table 7 Predicted counts of injury and total road traffic accidents from June 2017 to May 2021

Date	***Injury RTA***	***Total RTA***	***Date***	***Injury RTA***	***Total RTA***
Jun-17	29.61	86.76	Jun-19	28.41	84.54
Jul-17	30.35	87.46	Jul-19	28.56	84.63
Aug-17	24.36	81.78	Aug-19	27.29	83.92
Sep-17	26.66	80.99	Sep-19	27.78	83.82
Oct-17	25.74	84.5	Oct-19	27.58	84.26
Nov-17	29.42	86.61	Nov-19	28.37	84.52
Dec-17	29.42	80.27	Dec-19	28.37	83.73
Jan-18	26.66	82.38	Jan-20	27.78	84
Feb-18	27.12	83.09	Feb-20	27.88	84.08
Mar-18	25.74	82.03	Mar-20	27.58	83.95
Apr-18	27.12	85.56	Apr-20	27.88	84.39
May-18	31.27	86.26	May-20	28.76	84.48
Jun-18	28.79	85.12	Jun-20	31	73.56295
Jul-18	29.13	85.37	Jul-20	33	89.36068
Aug-18	26.36	83.36	Aug-20	20	74.36068
Sep-18	27.43	83.09	Sep-20	25	75.00000
Oct-18	27	84.32	Oct-20	23	82.94720
Nov-18	28.7	85.07	Nov-20	31	88.94720
Dec-18	28.7	82.83	Dec-20	31	73.00000
Jan-19	27.43	83.58	Jan-21	25	81.34606
Feb-19	27.64	83.82	Feb-21	26	73.08204
Mar-19	27	83.45	Mar-21	23	78.00000
Apr-19	27.64	84.7	Apr-21	26	87.12023
May-19	29.55	84.94	May-21	35	84.42810

As it can be shown from the forecast plots as well as the table of predicted road traffic accidents, the trends of injury road traffic cases and total traffic cases generally exhibit a similar pattern with the observed data i.e., the situation keeps what it was before. This result signifies the cases will not be decreased for the coming 36 months the same unless new or improved road safety measures are taken. It is in line with what we observe in our day-to-day life that road traffic accidents are common cases here and there.

Discussion

Now a day, the increment of a road traffic accident at a disturbing rate is a major concern in Ethiopia. The current study focuses on the Amhara region which has the highest share of road traffic accidents in the country. The most widely used conventional method of time series is known as Autoregressive Integrated Moving Average (ARIMA) model, also known as the Box-Jenkins method was applied to monthly reported road traffic accident data in four randomly selected zones of the Amhara region from September 2013 to May 2017 to determine patterns of road traffic accident cases in the region. After identifying various tentative models, the appropriate models for the accident cases (Total injury, fatal and total road traffic accidents) are as follows. ARIMA (2,0,0)(1,0,0)[12], ARIMA(2,0,0) and ARIMA(2,0,0)(1,1,0)[12] are chosen models to model the total injury cases, fatal cases and total RTA cases using the data from September 2013 to May 2017. The adequacies of the model were tested by analyzing standard residuals in different forms. 48 months of forecasts were provided for injury cases, fatal cases and total cases would continue in a non-decreasing trend. This study provides reliable and genuine information that could be useful for determining road accident rates in the region supported by [15]. This study would also be used for providing important information to increase the level of awareness among stakeholders

concerning road safety since the problem has become a growing issue in the country as a whole. Most importantly, this study would provide expected benefit to the road users, Road Safety Authority, researchers, and other stakeholders in understanding the rate of the cases of a road accident.

Conclusion

Based on the results of this study, the rate of road accidents is expected to constant at least for the next 5 years. It was found that the incidence of road accidents and injuries as result of accidents in Amhara region can be fitted to ARIMA $(2,0,0)(1,0,0)_{12}$ and ARIMA$(2,0,0)(1,1,0)_{12}$ model. The findings of this study draw attention to the importance of implementing key road safety measures to change the existing treat of road accidents in the Amhara region. Therefore, improved and better policies of the National road safety authority should be introduced with much emphasis on publication and education to ensure a maximum reduction in Road accident crashes.

Abbreviations

PACF: Partial Auto Correlation Function, RTA: Road Traffic Accidents; GDP: Gross Domestic Product; ANRS: Amhara National Regional State; ARIMA: Auto-Regressive Integrated Moving Average; ACF: Auto Correlation Function; AR: Auto-Regressive; MA: Moving Average; AIC: Akaike information criterion; BIC: Bayesian information criterion and SAR: Seasonal Auto-Regressive

Declarations

Ethics approval and consent to participate: Not applicable.

Consent for publication: Not applicable.

Availability of data and materials: Not applicable.

Competing Interests

The authors declare that they have no competing interests.

Funding: No funding was received for this study

Authors' contributions

KAG was responsible for the formulation of the methodology. KAG and MSW made substantial contributions to the conception and design of the study, analyzed and interpreted the data, and was a major contributor in writing the manuscript. KAG and MSW drafted the manuscript. Both authors read and approved the final manuscript.

Acknowledgments

The authors would like to acknowledge the Amhara region police commission traffic departments, transport, and communication bureau to conduct this study and to publish this paper.

Authors' Information

Department of statistics, Bahir-Dar University Institute of Technology, Bahir-Dar City, Ethiopia,

Email: almkidtsega@gmail.com

Reference

1. Farag, S.G., I.H. Hashim, and S.A.J.I.J.T.T.E. El-Hamrawy, *Analysis and assessment of accident characteristics: Case study of Dhofar Governorate, Sultanate of Oman.* 2014. **3**: p. 189-98.
2. Odhiambo, J.N., et al., *Modeling Road Traffic Accident Injuries in Nairobi County: Model Comparison Approach.* 2015. **4**(3): p. 178-184.
3. Violence, W.H.O.D.o., et al., *Global status report on road safety: time for action*. 2009: World Health Organization.
4. Bjerre, J., P.G. Kirkebjerg, and L.B.J.U.f.l. Larsen, *Prevention of traffic deaths in accidents involving motor vehicles.* 2006. **168**(18): p. 1764-1768.
5. Smart, R.G. and R.E.J.C. Mann, *Deaths and injuries from road rage: cases in Canadian newspapers.* 2002. **167**(7): p. 761-762.
6. Peden, M.J.I.j.o.i.c. and s. promotion, *Global collaboration on road traffic injury prevention.* 2005. **12**(2): p. 85-91.
7. Bishai, D., et al., *National road casualties and economic development.* 2006. **15**(1): p. 65-81.
8. Astrom, J., et al., *Signatures of Four Generations of Road Safety Planning in Nairobi City, Kenya In.* 2006. **20**: p. 186-201.
9. Mariam, D.H.J.E.J.o.H.D., *Road traffic accident: A major public health problem in Ethiopia.* 2014. **28**(1): p. 1-2.
10. Persson, A.J.A.i.t.s., *Road traffic accidents in Ethiopia: magnitude, causes and possible interventions.* 2008. **15**: p. 5-16.
11. Mekonnen, F.H. and S.J.E.J.o.H.D. Teshager, *Road traffic accident: the neglected health problem in Amhara National Regional State, Ethiopia.* 2014. **28**(1): p. 3-10.
12. Brief, E.E.A.J.R., *Energy, safety and environment and transport services in Ethiopia.* 2012(2).

13. Johansson, P.J.A.A. and Prevention, *Speed limitation and motorway casualties: a time series count data regression approach.* 1996. **28**(1): p. 73-87.

14. Box, G.E., G.M. Jenkins, and J.F.J.J.o.t.R.S.S.S.C. MacGregor, *Some recent advances in forecasting and control.* 1974. **23**(2): p. 158-179.

15. Al-Ghamdi, A.S.J.J.o.K.S.U.-E.S., *Time series forecasts for traffic accidents, injuries, and fatalities in Saudi Arabia.* 1995. **7**(2): p. 199-217.

Figures

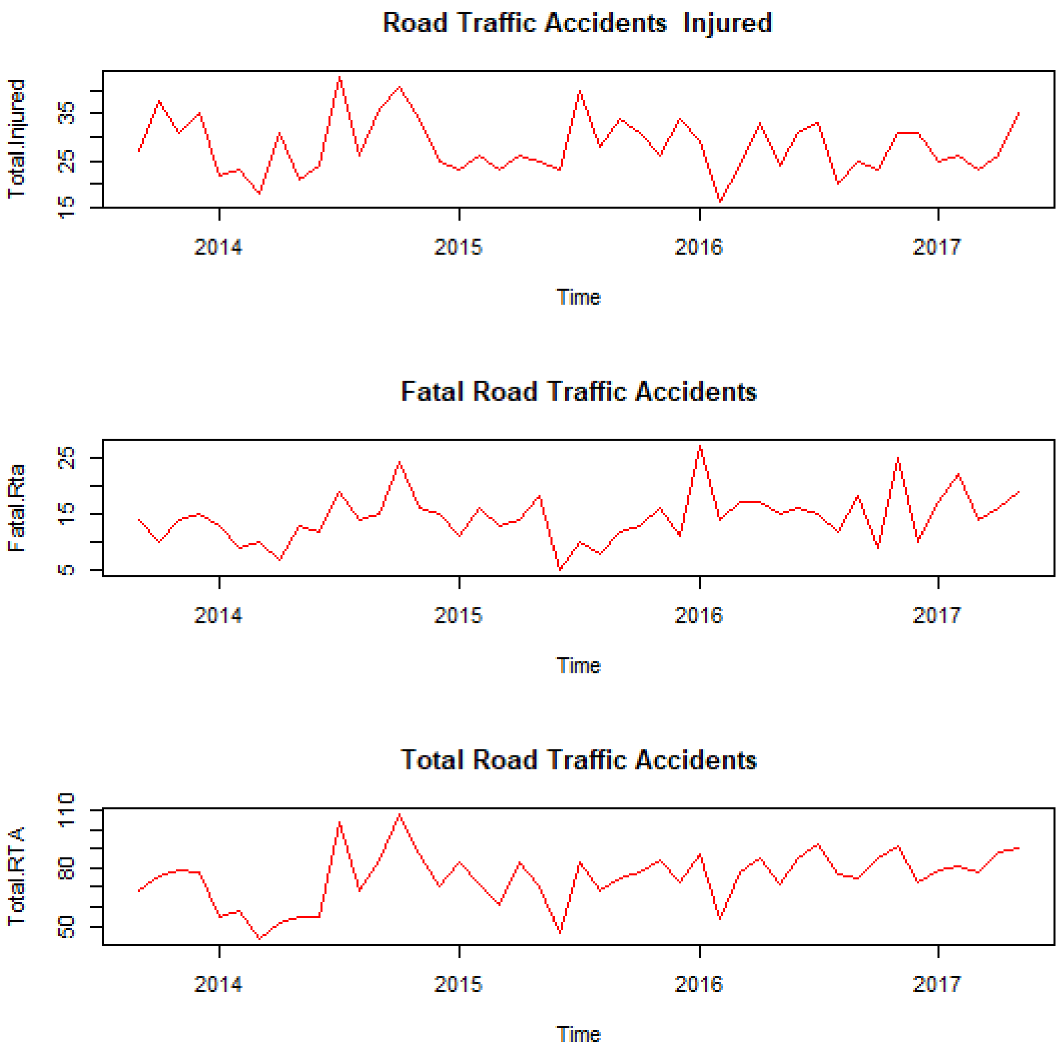

Figure 1

Time series plot of monthly reported Injury, fatal and total road traffic accidents in Amhara Region (September 2013 to May 2017)

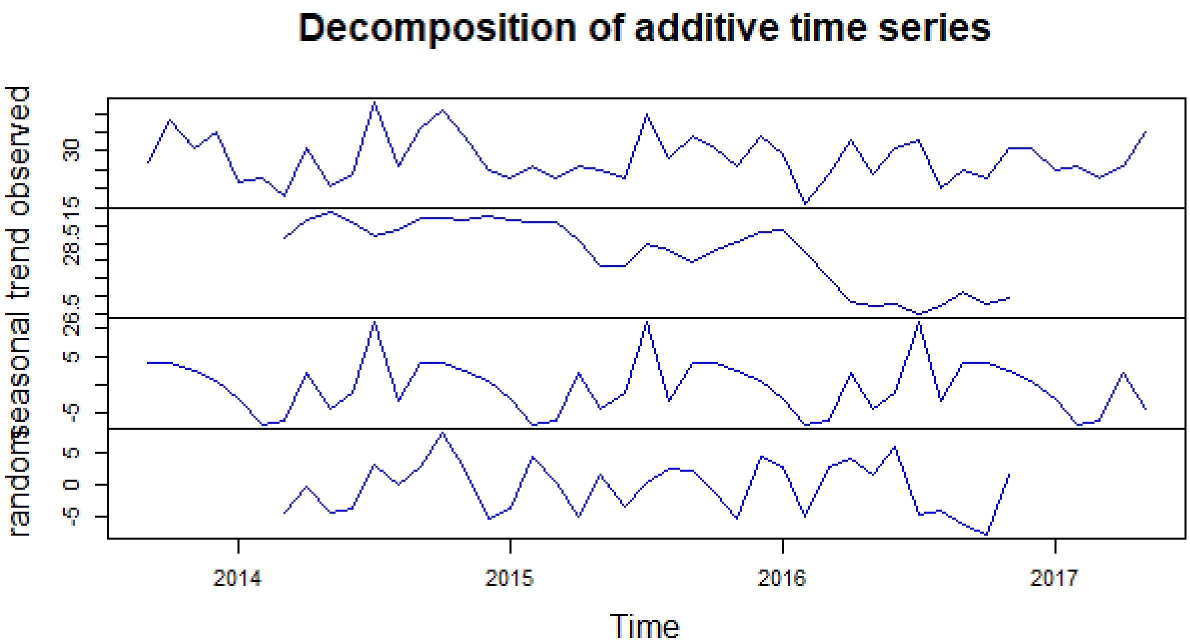

Figure 2

Time series decomposition of injury RTA data

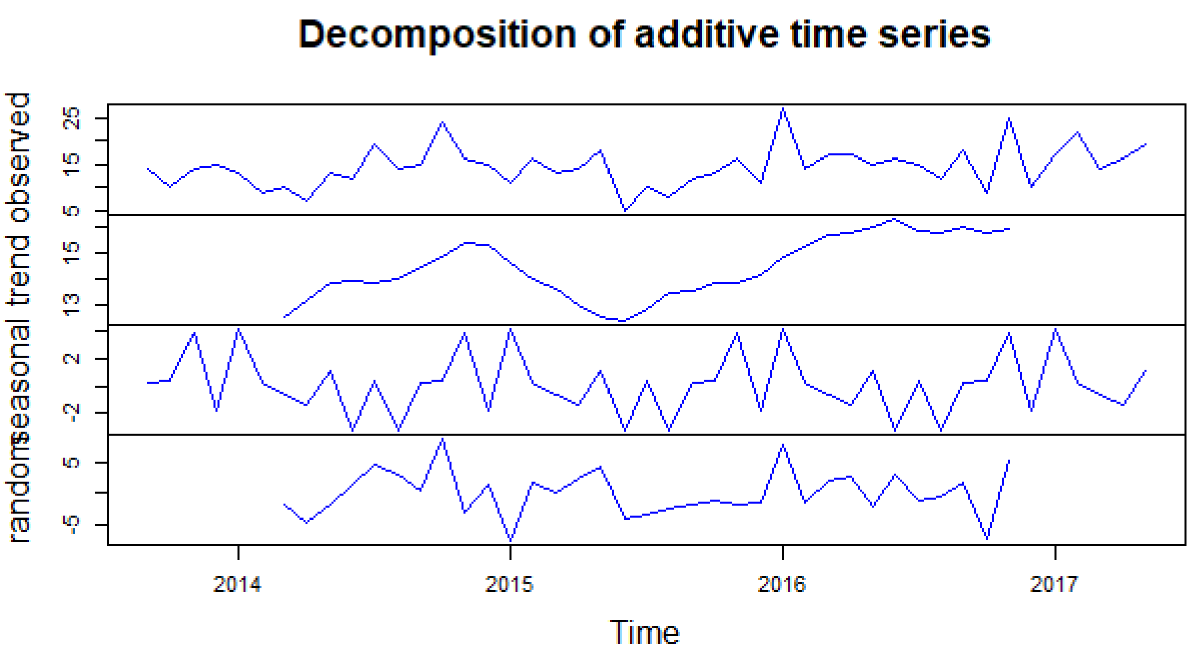

Figure 3

Time series decomposition of fatal RTA data

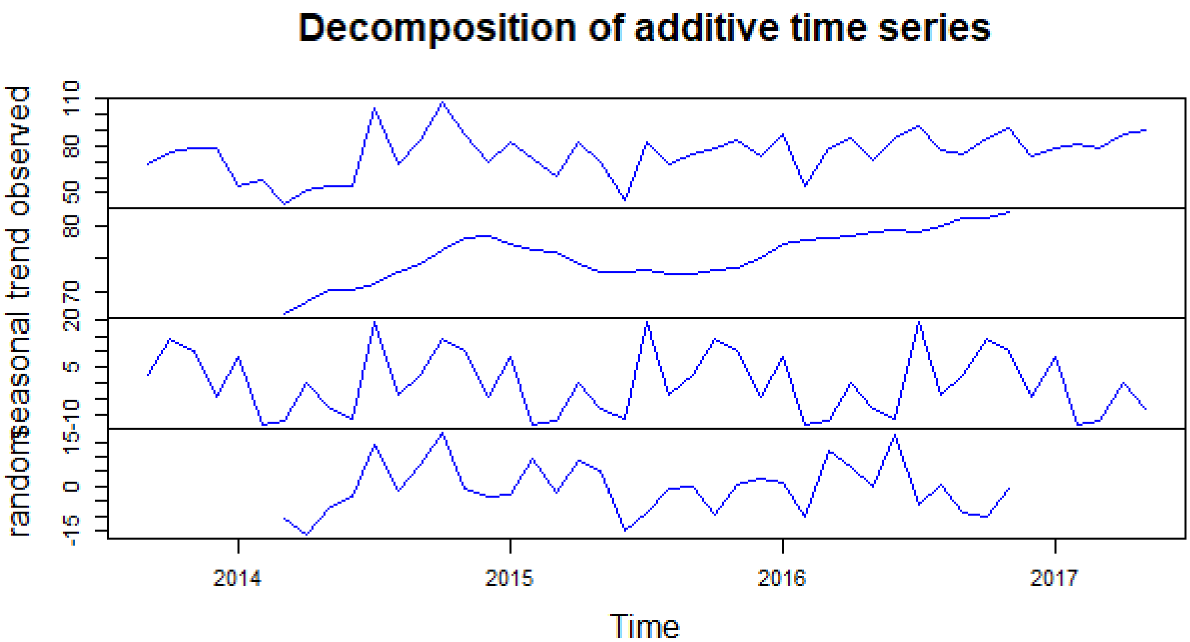

Figure 4

Time series decomposition of Total RTA data

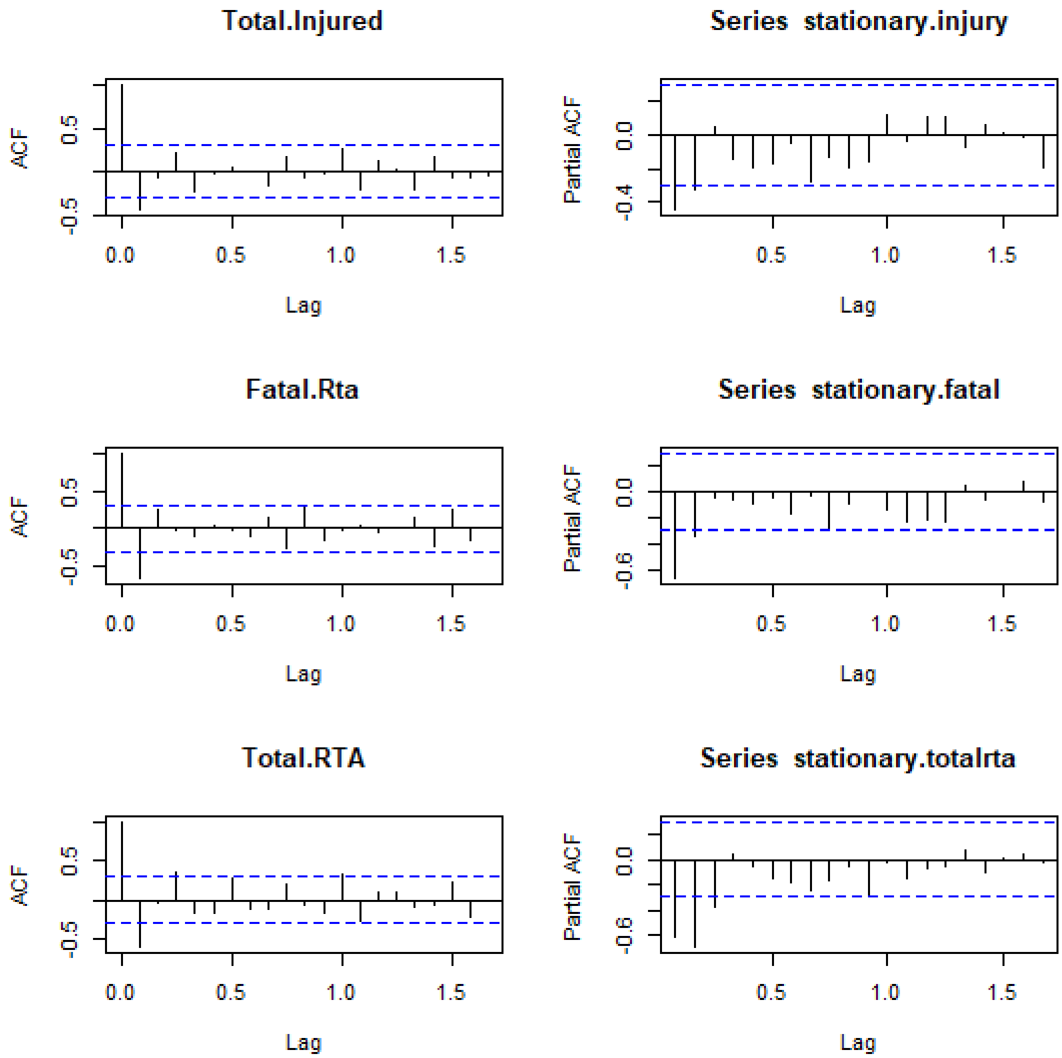

Figure 5

Autocorrelation and Partial autocorrelation function plots of Injury, Fatal and Total RTA

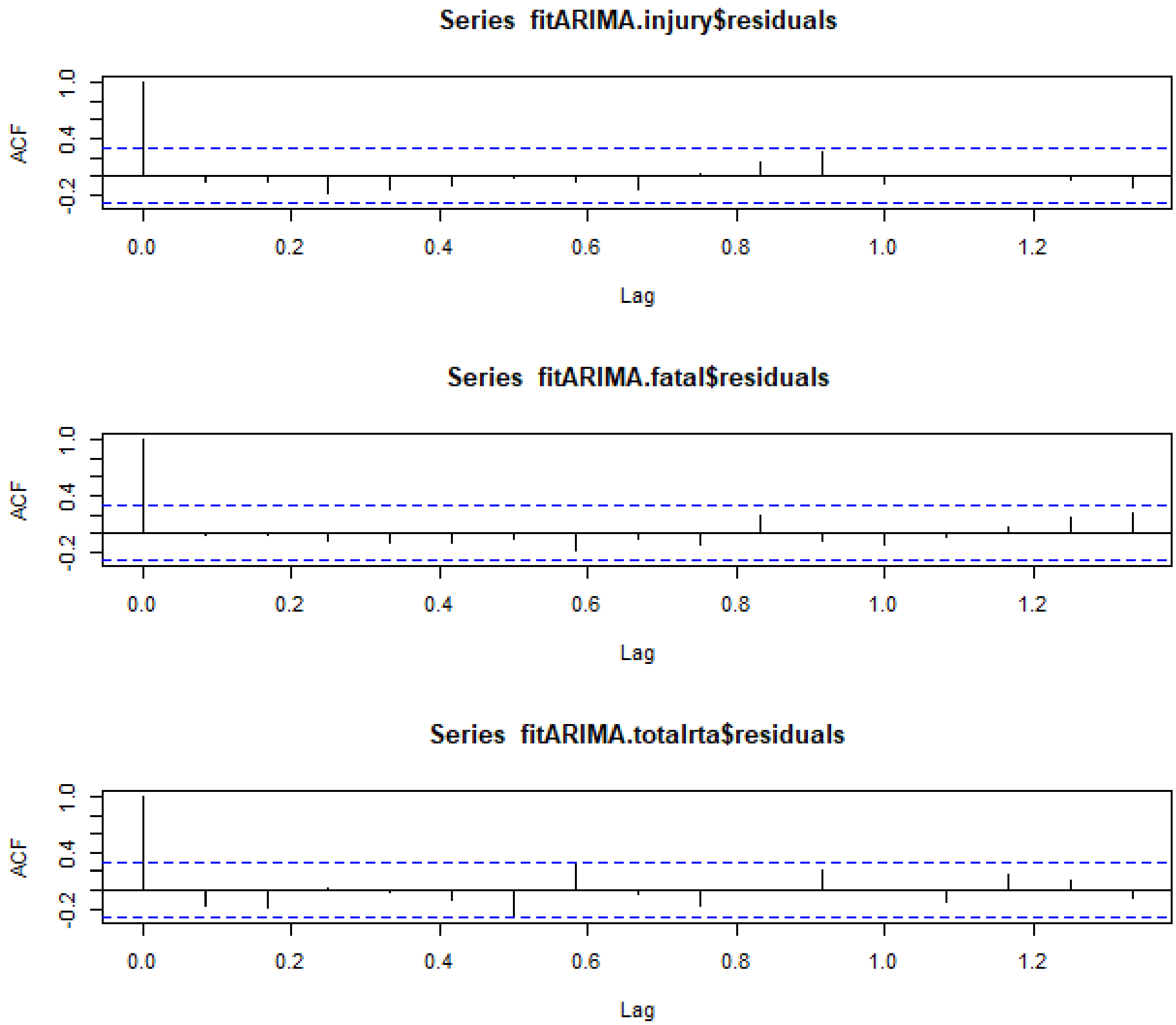

Figure 6

Diagnostics of residuals for ARIMA of injury, fatal and total RTA data

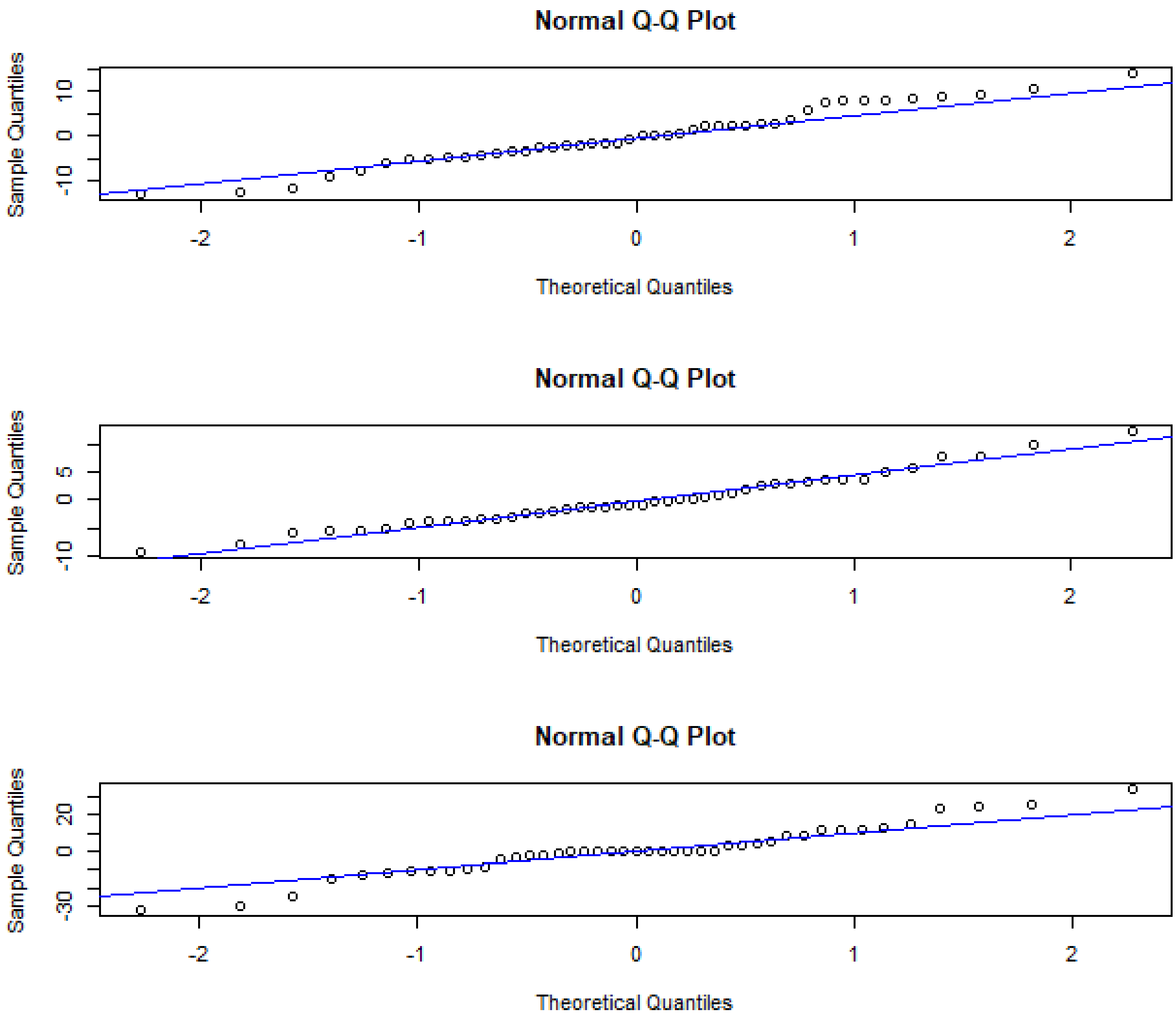

Figure 7

Normal Q-Q plot of residuals for ARIMA of injury, fatal and total RTA data

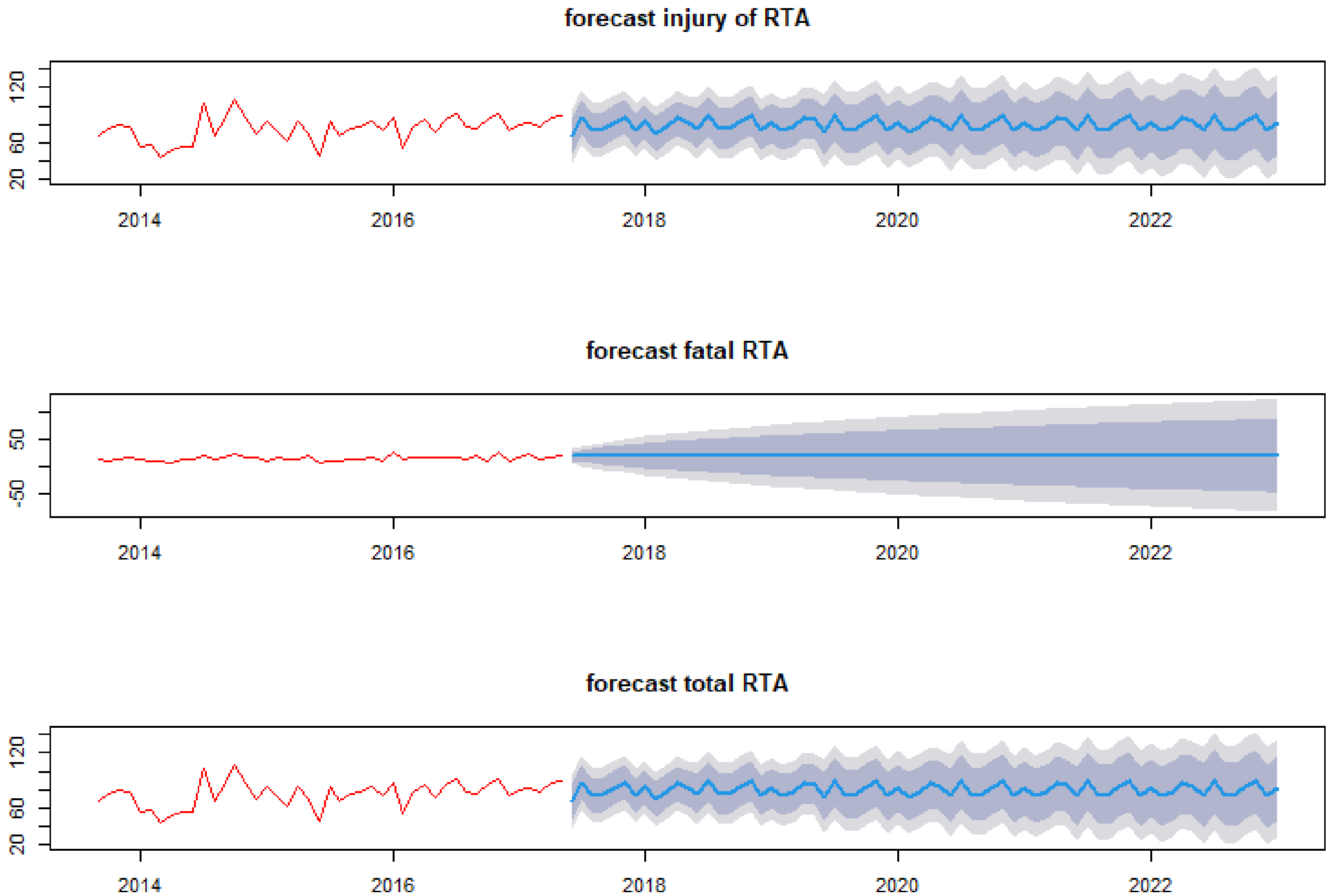

Figure 8

Graph of injury, fatal and total RTA cases, its forecasts, and confidence intervals