Road traffic accident mortality analysis based on time of occurrence: Evidence from Kerala, India

Vipin N.*, Rahul T.

*Department of Data Science, Prasanna School of Public Health, Manipal Academy of Higher Education, Karnataka, 576 104, India

BInternal Debt Management Department, Reserve Bank of India, Mumbai, Maharashtra, 400 001, India

**ARTICLE INFO**

**ABSTRACT**

Introduction: Road Traffic Accidents (RTA) are a global public health concern imposing huge socio-economic costs on society. Analysis of the incidence of RTA is essential to monitor the effectiveness of different road safety policies implemented to reduce this. The present study analyzes and forecasts time zone-specific RTA mortality reported during 2005–2018 in Kerala State, India.

Methods: An aggregate level secondary time series data of the total Number of Persons Killed (NPK) due to RTA in Kerala during 2005–2018 recorded at consecutive eight time zones were considered for the analysis. The exploratory data analysis and time series regression were mainly used for analyzing the data.

Results: The analysis confirms seasonality and upward trend in the RTA deaths in Kerala (549.5 (266,5660)) during 2005–2018. The minimum number of RTA deaths (129 (109,171)) happened from 00.00 a.m. to 3.00 a.m. and the maximum number of deaths occurred from 18.00 p.m. to 21.00 p.m. (810.5 (760,876)). RTA deaths in Kerala during evening and night-time were relatively higher than other time zones. RTA deaths that occurred immediately before dawn from 06.00 a.m. to 9.00 a.m. (504.5 (469,566)) or after dusk from 15.00 p.m. to 18.00 p.m. (719 (663,766)) were also relatively high in Kerala. Since 2016, in the time zones 00.00 a.m.–03.00 a.m., 09.00 a.m.–12.00 a.m., 18.00 p.m.–21.00 p.m. and 21.00 p.m.–24.00 p.m., there is an increasing trend in the RTA deaths in which during 21.00 p.m.–24.00 p.m. the NPK records a steady increase. Time series regression analysis indicates, for every 3hrs rise in time over the year; approximately four persons were killed (P-value < 0.001) due to RTA in Kerala during 2005–2018. Moreover, on average, more than 673 (P-value < 0.001) people were killed during the 18:00 p.m. - 21:00 p.m. due to RTA in Kerala than that of 00:00 a.m. - 03:00 a.m. The forecast of expected RTA deaths in Kerala based on the fitted TSR model with SAR(1) errors for the years 2019–2021 indicate most road accident deaths will likely occur during 18.00 p.m. and 21.00 p.m. and the least is between 00.00 a.m. and 03.00 a.m. The forecast for 2020 shows a decrease in Kerala’s RTA deaths for every time zone than 2019 and then increases in 2021.

Conclusions: The findings of the study provide the evidence that there is a significant amount of increase in the RTA deaths in Kerala through time. The study insists on implementing and monitoring the road safety policies concerning specific time zones for lowering RTA death rates.

1. Introduction

In India, road transport is the most dominant mode of transportation in terms of traffic share and its contribution to the national economy. With high population growth rates, increasing mobility, and growing numbers of vehicles, tremendous change has occurred to the road transportation network in India over the years. A negative factor associated with the modernization in the road network is the increase in Road Traffic Accidents (RTA). Many people are killed and injured on the roads every day due to RTA leaving behind shattered families and communities imposing huge socioeconomic costs on society. In a developing country like India, roads are the dominant mode of transport and the road network transports the majority of both passenger and freight traffic. The increased amount of motorized traffic on the roads with high acceleration and speed are the critical factors in the causes of RTA injury. The global status report on road safety 2018 by the World Health Organization reported the RTA as the eighth leading cause of deaths globally. The critical and exciting fact about the deaths

---

* Corresponding author.

E-mail addresses: vipinn.stats@gmail.com (V. N), rahultstat@gmail.com (R. T).

https://doi.org/10.1016/j.cegh.2021.100745

Received 12 January 2021; Received in revised form 8 March 2021; Accepted 10 April 2021

Available online 21 April 2021

2213-3984/© 2021 The Author(s). Published by Elsevier B.V. on behalf of INDIACLEN. This is an open access article under the CC BY-NC-ND license.
and injuries due to RTA is that most of them are preventable.

It is necessary to identify the role of the underlying factors that cause the RTA to design and implement policies for reducing RTA. Various factors such as driver’s behavior, vehicle features, road-traffic characteristics and environmental effects act as significant causes for RTA. The use of statistical techniques for analyzing RTA data is well established. There have been many studies reported in the literature regarding the modeling and prediction of RTA using time series and machine learning techniques (see Refs. 7–9 and the references therein). Many researchers analyzed and studied the impact of different factors on the incidence of fatalities and injuries due to RTA. Considered modeling of road traffic fatalities in India. A systematic review of the role of risk factors of deaths related to road traffic crashes in WHO regions is available in Ref. 8. A detailed review of RTA injuries in India reported in Ref. 9 studied the differential patterns, trends and hotspots of road traffic injuries on different road networks in the Vellore district, southern India. Studied the Spatio-temporal analysis of RTA in Indian cities. The studies by 11,12 investigated the role of climatic conditions in RTA. WHO reports more than 90% of road traffic deaths occur in low and middle-income countries. A comprehensive review of RTA in developing countries is available in Ref. 14.

Kerala is a south-western state in India having 2.76% of India’s population according to the 2011 Census; the Government of India. But, the state is three times as densely settled as the rest of India. Moreover, Kerala’s population growth rate is far lower than the national average. In terms of per capita Gross Domestic Product (GDP), economic productivity, and Human Development Index (HDI), Kerala leads many other Indian states. The low GDP and productivity figures juxtaposed with higher development figures than in most Indian states are often dubbed as the “Kerala Phenomenon” or the “Kerala Model” of development by economists, political scientists and sociologists. Although Kerala has remarkable outcomes in many social-economic development indicators, its road safety position is not impressive. Recent reports Kerala has remarkable outcomes in many social-economic development by economists, political scientists and sociologists. Although Kerala has remarkable outcomes in many social-economic development indicators, its road safety position is not impressive. Recent reports indicate that after Maharashtra, Kerala positioned second among different states concerning the RTA rate. RTA indicators, its road safety position is not impressive. Recent reports indicate that after Maharashtra, Kerala positioned second among different states concerning the RTA rate. RTA incidence of total Number of Persons Killed (NPK) due to RTA in Kerala rank at position five.

The present work aims to quantify the uncertainty involved in the incidence of total Number of Persons Killed (NPK) due to RTA in Kerala during 2005–2018. This study’s results may help the policymakers to develop and monitor existing policies and take further actions for the reduction of RTA in the future. The organization of the present work is as follows. In Section 2, we describe the data source and the essential characteristics of the NPK data. Section 3 briefly explains the statistical methodologies used for pursuing this study. The results of the statistical analysis of NPK data are given in Section 4. Finally, the concluding remarks for the present study with our significant findings are given in Section 5.

2. Materials and methods

2.1. Data

The data used for the present study is an aggregate secondary level time series data reported at the official website of Kerala Police (https://old.keralapolice.gov.in/public-information/crime-statistics/road-accident/ accessed on 20-09-2019). The data represent the total number of persons killed due to RTA from 2005 to 2018 in Kerala, India consisting of 112 observations. The data consist of the total number of persons killed in each year according to different eight consecutive time intervals each of 3 h duration such as 00.00 a.m.–03.00 a.m., 03.00 a.m.–06.00 a.m., 06.00 a.m.–09.00 a.m., 09.00 a.m. - 12.00 p.m., 12.00 p.m.–15.00 p.m., 15.00 p.m.–18.00 p.m., 18.00 p.m.–21.00 p.m. and 21.00 p.m.–24.00 p.m. were recorded in the data.

2.2. Time series regression analysis

This section describes the time series analysis methodology used for pursuing this study. A time series is a set of observations denoted by \( X_t \), each one being recorded at a specified time \( t = 1, \ldots, n \). The purpose of time series analysis is two-fold: (i) to model the stochastic (random) mechanism that gives rise to the series of data (ii) to predict (forecast) the values of the series based on the previous history. An observed time series may exhibit time-dependent factors such as trend, seasonality and stochastic time dependence. It is necessary to set up a hypothetical mathematical model to represent the data generating mechanism. Once it has been developed, it may be used in various ways, depending on the particular field of applications. Various parametric approaches are established in the literature for modelling a time series data. Among which, the popular ones are exponential smoothing methods, Box-Jenkins methodology and Time Series Regression (TSR) analysis (see Ref. 15 for a review). The first two methods are more useful if the researcher is more focused on the forecast. The TSR method helps to analyze the impact of the covariate effect and the prediction of future observations. Moreover, from an applied point of view, TSR results are more easily interpretable than the other two.

An observed time series \( X_t; t = 1, 2, \ldots, n \) may have an additive decomposition

\[
X_t = m_t + S_t + e_t \quad t(2.1)
\]

where the trend component \( m_t \) could be a polynomial of degree \( k \), \( S_t \) may be a periodic function with period \( s \) and \( e_t \) is a sequence of White Noise (WN) process with zero mean and constant variance, \( 0 < \sigma^2_e < \infty \) denoted by \( e_t \sim WN(0, \sigma^2_e) \). The function \( m_t \) may be linear or non-linear in time \( t \). Generally, the trend component \( m_t \) may be a polynomial in \( t \) of the form,

\[
m_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \ldots + \beta_k t^k
\]

In the present study we consider \( k = 1, k = 2 \) and \( k = 3 \) corresponding to a linear, quadratic and cubic trend respectively. The seasonality component \( S_t \) may be modelled using

\[
S_t = \sum_{i=1}^{s} a_i \delta_{i,t} \quad t(2.2)
\]

where, \( a_i \)’s are the regression parameters and \( \delta_{i,t} \) denote indicator functions of time \( t \), defined by

\[
\delta_{i,t} = \begin{cases} 
1 & \text{if } t \text{ corresponds to seasonal period } i \\
0 & \text{otherwise} 
\end{cases}
\]

when \( s = 8 \), we set up 7 dummy variables for each time interval by fixing the time interval 00.00 a.m.–03.00 a.m. as the reference category, i.e., for \( i = 2, \ldots, 8 \), we define,

\[
d_{2,t}, d_{3,t}, d_{4,t}, d_{5,t}, d_{6,t}, d_{7,t} \quad t(2.2)
\]

when \( s = 8 \), we set up 7 dummy variables for each time interval by fixing the time interval 00.00 a.m.–03.00 a.m. as the reference category, i.e., for \( i = 2, \ldots, 8 \), we define,

\[
d_{2,t}, d_{3,t}, d_{4,t}, d_{5,t}, d_{6,t}, d_{7,t} \quad t(2.2)
\]

when \( s = 8 \), we set up 7 dummy variables for each time interval by fixing the time interval 00.00 a.m.–03.00 a.m. as the reference category, i.e., for \( i = 2, \ldots, 8 \), we define,

\[
d_{2,t}, d_{3,t}, d_{4,t}, d_{5,t}, d_{6,t}, d_{7,t} \quad t(2.2)
\]
independently distributed (i.i.d). The Q-statistic for the Ljung-Box test is defined by $Q = n(n + 2) \sum_{k=1}^{h} (n - k)^{-1} r_{k}^{2}$, where $h$ is the maximum lag being considered, $n$ is the number of observations, $r_{k}$ denotes the SACF. If data are WN, $Q$ has $\chi^{2}$ distribution with $(h - k)$ degrees of freedom, where $k$ is the number of parameters in the fitted model.

2.2.2. Regression analysis with time series errors

The serial dependence among the residuals from the fitted regression model can be rectified by specifying a new regression model with time series errors defined by

$$X_{i} = \beta_{0} + \beta_{1} X_{i-1} + \beta_{2} X_{i-2} + \sum_{k=2}^{h} \alpha_{k} \delta_{i-k} + \epsilon_{i}$$

(2.3)

where, $\epsilon_{i}$ is the sequence of autocorrelated residuals from the fitted regression model and admits a SARIMA ($p,d,q$) model can be rectified by specifying a new regression model with time series errors defined by

$$\phi(B)\Phi(B')(1 - B)^{d} \sigma_{i}^{2} \epsilon_{i} = \delta(B)\Theta(B')\epsilon_{i}$$

The multiplicative Seasonal Autoregressive Integrated Moving Average (SARIMA) model is one of the popular time series model existing in the literature for modelling a nonstationary time series data. Here $B$ denotes the back-shift operator. Further we denote, $d$: nonseasonal integer degree of differencing, $D$: seasonal integer degree of differencing, $s$: the periodicity = 8, $\phi(B) = 1 - \phi_{1}B - \phi_{2}B^{2} - \ldots - \phi_{p}B^{p}$: nonseasonal AR polynomial of order $p$, $\theta(B) = 1 + \theta_{1}B + \theta_{2}B^{2} + \ldots + \theta_{q}B^{q}$: nonseasonal MA polynomial of order $q$, $\Phi(B') = 1 - \phi_{1}B' - \phi_{2}B'^{2} - \ldots - \phi_{P}B'^{P}$: seasonal AR polynomial of order $P$ and $\Theta(B') = 1 + \theta_{1}B' + \theta_{2}B'^{2} + \ldots + \theta_{Q}B'^{Q}$: seasonal MA polynomial of order $Q$.

2.2.3. Performance accuracy measures

After fitting a time series model, we denote the forecast errors by $e_{n}(l)$, where $n$ is the forecast origin, $l$ is the forecast lead and $L$ denote the number of observations in the data set. One can make use of the following performance evaluation criteria such as Mean Error (ME), Mean Percent Error (MEP), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percent Error (MAPE) for assessing the performance accuracy.

$$ME = \frac{1}{L} \sum_{t=1}^{L} e_{t}(l)$$

$$MPE = \frac{100}{L} \sum_{t=1}^{L} \frac{e_{t}(l)}{X_{t}}$$

$$RMSE = \sqrt{\left\{\frac{1}{L} \sum_{t=1}^{L} e_{t}^{2}(l)\right\}}$$

$$MAE = \frac{1}{L} \sum_{t=1}^{L} |e_{t}(l)|$$

$$MAPE = \frac{100}{L} \sum_{t=1}^{L} \frac{|e_{t}(l)|}{X_{t}}$$

The first two criteria, ME and MPE indicate the bias in forecast whereas the last three RMSE, MAE and MAPE assess forecast accuracy.

3. Results

The present section provides major findings of the study using the methodologies discussed in Section 3. All the statistical analyses were performed two-sided at a 5% level of significance and conducted using R software (version 3.5.3) with the “forecast” (version 8.13) and “astsa” (version 1.12) packages.

3.1. Data analysis

The time plot for the NPK time-series data during 2005–2018 is depicted in Fig. 1. The time plot confirms the seasonal variation in the dataset. The NPK shows an increasing trend in which most RTA deaths occurred in 2012, and the minimum number observed in 2005. The highest (903 during 15.00 p.m.–18.00 p.m.) and lowest (88 during 00.00 a.m.–03.00 a.m.) number of persons who died to RTA was 2012. Table 1 provides the summary statistics of the data. In Table 2 the summary statistics of the data concerning eight different time zones are reported. Table 2, implies that the aggregated count of NPK (549.5 (266.5660)) in each time zone increases over time over the years. The minimum number of deaths due to RTA occurred during 00.00 a.m.–03.00 a.m. The maximum number of deaths occurred during 18.00 p.m.–21.00 p.m. (810.5(760.876)). The incidence of RTA deaths in Kerala during evening and night time are quite high in comparison with other time zones. RTA deaths that occurred immediately before dawn during 06.00 a.m.–9.00 a.m. (504.5(469,566)) or after dusk from 15.00 p.m. to 18.00 p.m. (719(663,766)) were also relatively high. Fig. 2 represents the seasonal plots of the NPK due to RTA in Kerala concerning different time zones through 2005–2018. Fig. 2 indicate generally increasing NPK trend from 00.00 a.m. - 03.00 a.m. to 18.00 p.m. - 21.00 p.m. and then decrease in the 21.00 p.m.–24.00 p.m. time zone every year. Further, Fig. 3 summarizes the information of NPK due to RTA concerning different time zones over 2005–2018. Fig. 3 indicate that comparing with 2016 during the time zones 03.00 a.m.–06.00 a.m., 06.00 a.m.–09.00 a.m., 12.00 a.m.–15.00 p.m. and 15.00 p.m.–18.00 p.m. the NPK exhibits a decreasing nature. Moreover, in the time zones, 00.00 a.
m. – 03.00 a.m., 09.00 a.m. – 12.00 a.m., 18.00 p.m. – 21.00 p.m. and 21.00 p.m. – 24.00 p.m., there is an increasing trend in the total number of persons killed in RTA. Out of which, during 21.00 p.m. – 24.00 p.m. the NPK shows a steady increase.

Fig. 4(a) indicates seasonality and slowly increasing trend in the data. The nature of the series exhibited in the plot indicates that the data may be stationary. This fact was further confirmed by the Augmented-Dickey-Fuller test. The test rejected the null hypothesis that data is level non-stationary (i.e. there exist a presence of unit-root) with a P-value = 0.01 < 0.05.
3.2. Time series regression analysis

Time Series Regression (TSR) analysis has been applied to assess the relationship between the outcome variable, the NPK due to RTA, \( X_t \) and the variables quantifying the trend and seasonal effects in the different eight time zones. According to the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), the best model is identified. The TSR model’s general structure is given in (2.1) with the seasonality factor as shown in (2.2). The TSR models with different functional choices for the trend component \( m_t \) considered from Fig. 4 were (i) linear trend: \( m_t = \beta_0 + \beta_1 t, \) (AIC = 1249.54, BIC = 1276.73) (ii) quadratic trend: \( m_t = \beta_0 + \beta_1 t + \beta_2 t^2, \) (AIC = 1243.53, BIC = 1273.43) and (iii) cubic trend: \( m_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3, \) (AIC = 1244.48, BIC = 1277.08).

Based on the model selection criteria, the TSR model with quadratic trend component specification was identified as the suitable model. The model diagnostics plots for the fitted TSR model (3.1) is depicted in Fig. 5. Cook’s distance plot in Fig. 5 (a) clearly confirms the presence of outliers. To obtain a better TSR model fit, the outliers in the data have been removed and replaced using linear interpolation method with the `tsclean` function of `forecast` package in R software. The decomposition of the time series data before and after the replacement is displayed in Fig. 4. The TSR model adjusted for outliers with quadratic tend (AIC = 1197.45, BIC = 1224.64) was again fitted to the data. The model fitting information is given in Table 3.

\[
X_t = \hat{\beta}_0 + \hat{\beta}_1 t + \hat{\beta}_2 t^2 + \sum_{i=2}^8 \hat{\alpha}_i \delta_{ti},
\]

3.2.1. Time series regression with SAR(1)(8) errors

The problem of serial dependence among the residuals has been solved by fitting a TSR model with SARIMA (0,0,0) \( \times (1,0,0)^8 \) (also known as SAR(1) of periodicity8) errors. The residuals of the updated regression model was modelled by SARIMA (0,0,0) \( \times (1,0,0)^8 \) Seasonal Autoregressive model of order 1 with periodicity (SAR(1)(8) model) and the new regression model is termed as TSR model with SAR(1)(8) errors. The general form of the fitted model is given in (2.3). The parameter estimates with the corresponding Standard Error (SE) and P-value based on asymptotic z-test are given in Table 3. The estimated white noise variance of the TSR model with SAR(1)(8) errors is \( \hat{\sigma}^2 = 1832 \) with AIC = 1173.09 and BIC = 1203. The residuals of the fitted model are satisfying the white noise assumptions and the Ljung-Box test (\( \chi^2_{1} = 1.44 \), P-value = 0.278) also confirmed this fact.

Fig. 5. Diagnostic plots for checking the influential observations.
The details of the residual analysis are depicted in Fig. 6 (b).

The final fitted TSR model with SAR(1)(8) errors (2.3) provides the following inputs. The regression coefficient of time (4.03) is positive and highly significant (P-value < 0.001), indicating that the numbers of deaths due to Kerala’s RTA are linearly increasing over time. i.e., for every 3hrs increase in the time over the year, approximately four persons were killed due to RTA in Kerala during 2005–2018. Moreover, fixing the time zone 00:00 a.m. - 03:00 a.m. as the reference category, the test for significance of the regression coefficients of dummy variables suggests a highly statistically significant positive time zone effect, indicating that throughout the period 2005–2018, on an average, more than 673 people killed during the 18:00 p.m. - 21:00 p.m. due to RTA in Kerala than that of 00:00 a.m. - 03:00 a.m. Similarly, around 596 more deaths occurred due to RTA in Kerala from 15:00 p.m. - 18:00 p.m. than 00:00 a.m. - 03:00 a.m. There were about 171 more RTA deaths during 2005–2018 from 21:00 p.m. to 24:00 p.m. than 00:00 a.m. - 03:00 a.m.

The minimum number of RTA deaths occurred from 03:00 a.m. to 6:00 a.m., (around 120) compared with 00:00 a.m. - 03:00 a.m. In the forenoon period through 2005–2018, there were more number of RTA deaths occurred from 09:00 a.m. to 12:00 p.m. (around 492) comparing with the time zone 06:00 a.m. to 09:00 a.m. (around 383) than that of 00:00 a.m. - 03:00 a.m.

Table 4 gives the point forecast with 95% CI for the total number of RTA deaths for the years 2019–2021 according to the fitted the TSR model with SAR(1)(8) errors. The forecast of 2020 indicates a reduction in Kerala’s RTA deaths in every time zone compared with 2019 and 2021. According to the forecast, most road accident deaths will likely occur during 18.00 p.m. and 21.00 p.m. and the least is between 00.00 a.m. and 03.00 a.m. for all three years. The forecast based on the fitted model (TSR Model adjusted for outliers) shows that the number of RTA deaths will continue to increase over the years.

The point forecast for the years 2019–2021 is given in Table 4.

### Table 3
Details of the different TSR model fitting.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>TSR Model 1</th>
<th>TSR Model adjusted for outliers</th>
<th>TSR model with SAR(1)(8) errors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>SE</td>
<td>t-value</td>
</tr>
<tr>
<td>time</td>
<td>3.93***</td>
<td>0.54</td>
<td>7.34</td>
</tr>
<tr>
<td>time²</td>
<td>-0.05***</td>
<td>0.01</td>
<td>-4.81</td>
</tr>
<tr>
<td>00.00 a.m.–03.00 a.m.</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>03.00 a.m.–06.00 a.m.</td>
<td>119.65***</td>
<td>19.78</td>
<td>6.05</td>
</tr>
<tr>
<td>06.00 a.m.–09.00 a.m.</td>
<td>384.83***</td>
<td>19.84</td>
<td>19.40</td>
</tr>
<tr>
<td>09.00 a.m. to 12.00 p.m.</td>
<td>491**</td>
<td>19.88</td>
<td>24.69</td>
</tr>
<tr>
<td>12.00 p.m.–15.00 p.m.</td>
<td>490.79***</td>
<td>19.93</td>
<td>24.63</td>
</tr>
<tr>
<td>15.00 p.m.–18.00 p.m.</td>
<td>600.19***</td>
<td>19.97</td>
<td>30.05</td>
</tr>
<tr>
<td>18.00 p.m.–21.00 p.m.</td>
<td>674.71***</td>
<td>20.01</td>
<td>33.71</td>
</tr>
<tr>
<td>21.00 p.m.–24.00 p.m.</td>
<td>204.71***</td>
<td>20.05</td>
<td>10.21</td>
</tr>
</tbody>
</table>

### Table 4
Point forecast of NPK according to different time zones with 95% CI for the for the years 2019–2021 using fitted TSR model with SAR(1)(8) errors.

<table>
<thead>
<tr>
<th>Time Zone</th>
<th>2019</th>
<th>2020</th>
<th>2021</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Point forecast (95% CI)</td>
<td>Point forecast (95% CI)</td>
<td>Point forecast (95% CI)</td>
</tr>
<tr>
<td>00.00 a.m.–03.00 a.m.</td>
<td>226.29 (98.14,354.44)</td>
<td>121.16 (59.84,182.47)</td>
<td>247.40 (149,459.88)</td>
</tr>
<tr>
<td>03.00 a.m.–06.00 a.m.</td>
<td>233.94 (105.79,362.09)</td>
<td>128.47 (67.11,216.78)</td>
<td>254.39 (141,466.87)</td>
</tr>
<tr>
<td>06.00 a.m.–09.00 a.m.</td>
<td>512.13 (383.97,640.28)</td>
<td>394.41 (333.05,687.62)</td>
<td>508.62 (296,1472.1)</td>
</tr>
<tr>
<td>09.00 a.m. to 12.00 p.m.</td>
<td>614.41 (486.26,742.57)</td>
<td>492.20 (340.83,785.41)</td>
<td>602.10 (389,62814.58)</td>
</tr>
<tr>
<td>12.00 p.m.–15.00 p.m.</td>
<td>544.63 (416.46,767.78)</td>
<td>425.48 (364.12,718.69)</td>
<td>538.32 (325,857,508)</td>
</tr>
<tr>
<td>15.00 p.m.–18.00 p.m.</td>
<td>688.02 (559.87,816.18)</td>
<td>562.56 (450.28,755.87)</td>
<td>669.27 (456,89881.85)</td>
</tr>
<tr>
<td>18.00 p.m.–21.00 p.m.</td>
<td>882.09 (753.93,1010.24)</td>
<td>748.08 (686.72,1041.29)</td>
<td>846.72 (634,24,1059.2)</td>
</tr>
<tr>
<td>21.00 p.m.–24.00 p.m.</td>
<td>569.48 (441.33,697.64)</td>
<td>449.24 (387.88,742.46)</td>
<td>561.04 (348.56,773.52)</td>
</tr>
</tbody>
</table>

*** P-Value < 0.001.

Fig. 6. Residual analysis plots for checking the white noise assumption.
V. N and R. T

Fig. 7. Forecast for the years 2019–2021 with 80% and 95% forecast interval from the TSR model with SAR(1) errors.

TSR model with SAR(1) errors for the years 2019–2021 is graphically represented in Fig. 7.

4. Discussion

Road traffic accidents are the leading causes of injuries, hospitalization, disabilities, and death worldwide, imposing immense socioeconomic costs to society. The available data on RTA deaths shows that it is a significant public health issue in India, which is avoidable to a large extent. Hence an analytical study on the occurrence of RTA is always a relevant research topic with social relevance in the context of a developing country like India with a large, dense population and highly relying on-road transportation for the stability of its economy. The present study focuses on analyzing the number of persons killed due to RTA in Kerala State, India, over 2005–2018 at different time zones of a day. Quantifying the uncertainty in RTA deaths at different time zones and forecasting the same for the forthcoming years were the study’s primary objectives.

Apart from the exploratory data analysis, time series regression analysis was mainly used for analyzing the data. The study explicates that RTA deaths are significantly increasing over the year. The NPK in each time zone of the day also shows an increasing trend over the year, with maximum deaths during evening and nighttime, especially during 18.00 p.m.–21.00 p.m. The RTA deaths that occurred immediately before dawn during 06.00 a.m.–9.00 a.m. and after dusk from 15.00 p.m. to 18.00 p.m. were also relatively high in Kerala. Results from time-series regression analysis confirm that time and time zones have a significant positive effect on RTA and related deaths. The predicted RTA deaths during 2019–2021 also exhibited an upward trend.

This study investigated the intraday seasonality of RTA mortality and forecasted the NPK at different specific time zones of the day in a year. Though the study is the first to analyze the intraday seasonality of RTA mortality in Kerala, it has some limitations. The study suffers from data limitation related to the various factors for RTA in each time zone of the day. The investigation of the relationship between the NPK in different time zones and corresponding risk factors such as driver’s behaviour, vehicle features, road-traffic characteristics and geographic and environmental influences would give much better picture of the scenario. We hope to take this as our future research depending upon the availability of official data. We recommend governments to develop a comprehensive database on road traffic accident details and continuous updating of the same for monitoring and undertaking road safety research.

Being a critical public health and economic concern that adversely affects a society’s central cross-section, road safety needs special attention from the government authorities. The officials have to take necessary actions to improve road safety by implementing and executing all tangible actions such as providing safer roads, proper training of drivers, effective vehicle management and sustained traffic enforcement.

5. Conclusion

There is an imperative need to monitor and predict road traffic accidents to assess the burden of fatalities and non-fatal injuries caused by road traffic crashes. The study’s findings insist on implementing and monitoring the road safety policies concerning specific time zones for lowering RTA death rates. The prevention measures include strengthening road traffic laws, improving roads and vehicles’ safety, and enhancing post-crash care. Despite a better understanding of the problem and its solutions, the political will to carry out the necessary actions is paramount.

Ethics approval and consent to participate

A secondary freely available data released by the Kerala Police was used for the analysis. The data set had no identifiable evidence on the individuals, therefore ethical approval is not required for this work.

Informed consent

Since this study used a data set that is available online in the public domain, there was no need to seek consent to publish the results.

Authors contributions

Dr. Vipin N prepared the data and was a significant contributor in identifying the methodology for analyzing and interpreting the results. Dr. Rahul T contributed to the interpretation of the analysis results. All authors read and approved the final manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or non-profit sectors.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Disclaimer

The views expressed in the article are those of the authors and do not represent the views of the organization in which authors belong.

Declaration of competing interest

The author declare that I have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to express their gratitude to the Editor-in-Chief and the anonymous reviewers for their valuable suggestions which have considerably improved this paper.

References


