Forecasting Future Traffic Accidents in Baghdad City

Huda Abdul-Kider Aziz[a], [a]Assist. prof., Department of Geography and Geographic Information Systems, Email: [a]a.prhudaaziz@gmail.com

The present research deals with the future forecast of traffic accidents recorded in the city of Baghdad. Resulting from the large increase in the number of vehicles, traffic accidents are a negative consequence of modern and developed societies. Their impact is most severe when a family member is lost, causing economic, social and psychological damage. The research was based on the use of Box-Genghis models in the future prediction of traffic accidents and drew on the data available for the period (2003-2018), which was obtained from the Directorate General of Traffic. Using the program Gretl, the study showed that the model ARIMA (2.1.1) is the best model the relevant authorities can rely on to predict future traffic accidents.

Key words: *Future forecasting, traffic accidents.*

Introduction

Recent attention has focused on traffic studies, both globally and in the Arab world. This interest is the result of the circumstances and factors of our modern world, where vehicles have become the primary means of transport. Deaths from accidents have increased and, in some countries, have reached levels beyond natural death as well as serious injuries.

This requires the establishment of a large network of roads, which is a fundamental dimension in the planning of modern cities. Whether we accept the present conditions or expect nothing to change, traffic studies and phenomena has emerged as a needed field of inquiry and as at the forefront of topics and issues that deserve study and research.

Research Goal

The present research aims to predict future traffic accidents for the years (2003-2018) in the city of Baghdad, and thereby find the best model for the future forecast of accidents.

Research problem

Traffic accidents are one of the most important problems of urbanization affecting human societies. The seriousness of the damage they cause requires the identification and management of their causes.

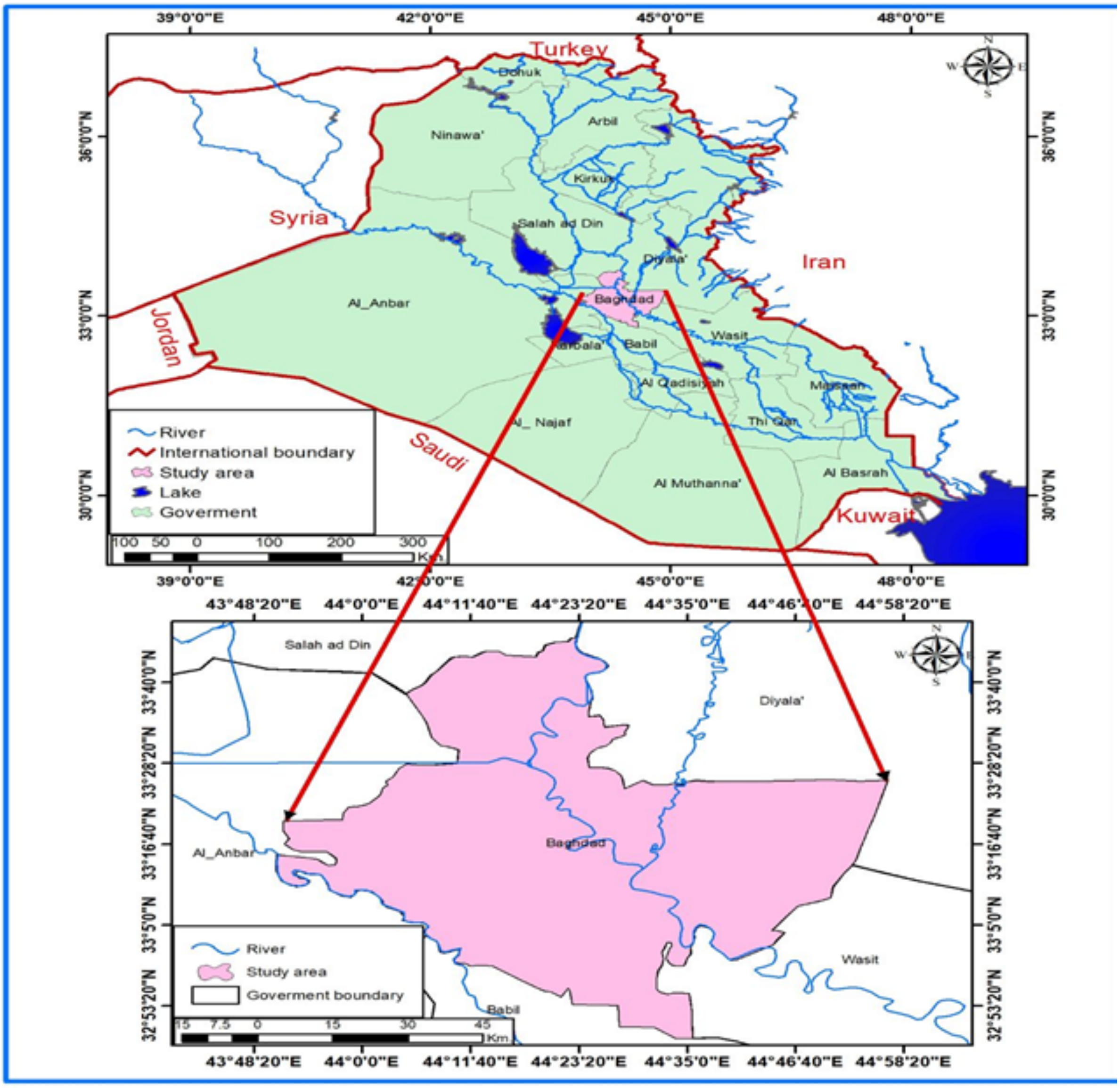
Hypothesis search

Quantitative and statistical analysis methods, including the Box Genghis model, are one way to predict future projections of traffic accidents in Baghdad.

Research Limits

Spatial boundaries: - The study area in the middle section of Iraq was determined between two latitudes (33.23) north, and longitudes (44.30) east, the provinces of Babylon and Karbala on the south and southwest, as well as Anbar on the west map (1).

Map 1. Study area



Source: Republic of Iraq, Ministry of Water Resources, General Survey Authority, Maps Production Section, Digital Unit,2016 ,Scale drawing 1:100000.

Time limits

The time limits for the study were based on the number of traffic accidents recorded for the period between (2003-2018) and available in the Ministry of Interior, Directorate of Traffic in Baghdad.

The concept of traffic accident

A traffic accident is an event that occurs without prior planning by one or more parties on a public or private road. Other vehicles, animals or objects may be involved, and can result in damage and injuries ranging from minor property and vehicles to death and permanent disability (Al-Basrawi, 2010).

Accidents are divided into two parts:
A - fatal accidents: - These accidents result in human death. Whether the person is a passenger of the vehicle, or a pedestrians, there will be material and moral losses.
B- Non-fatal accidents: - They cause serious or minor injuries (Mohammed, 1986).

Elements of traffic accidents

Traffic accidents are the consequence of a person's failure to drive according to traffic rules, and can result in material and human losses that impact society, including permanent disabilities and death. In short, the factors that control the incidents, their forms and seriousness, can be divided into (Al-Attar, 1982) : -

1- The driver: - Is one of the most important elements that can cause the traffic accident. It is the only factor that can contribute to the accident through the mind and senses, which are used to determine the appropriate decision to avoid an accident. Studies have proven that the driver are at fault in 92% of Traffic accidents.
2 **- Pedestrians**: - Are one of the most affected parties in accidents and, as a consequence, may receive major or minor injuries, disability, or death and its consequent moral and material losses.
3 - **Road:** - The surface of a road, as well as the way that its planned and paved is a key element in accidents. Its planning and design, traffic signs and lighting, may contribute to the occurrence of accidents (Saleh, 2007).
4 - **Vehicle**: - The vehicle is an important element for the crisis to occur. A malfunction, a driver's ability or, worse, failure to comply with the safety conditions for driving the vehicle may contribute to the occurance of an accident.

Causes of Traffic Accidents

1 - Preoccupation with things outside the perimeter of the road, psychology, and health.
2 - The effect of drinking drugs and alcohol, which leads to loss of concentration.
3 - Element of time and its impact on the incident (Hanna, 1999).
4 - Loose animals, which is one of the most dangerous species that cause the accident.

Means to be Taken in Order to Reduce Accidents

1- Take into account the engineering design of roads and bridges, accommodate the remarkable increase in the numbers of vehicles, as well as develop traffic signals (Al-Adhami, 2010).
2. The driver adheres to traffic standards and rules, and adheres to traffic signs.
3- Media awareness and guidance programs about the severity of traffic accidents.
4 - Provie mobile medical detachments on the roads, which, in order to reduce human casualties, provide immediate care for the wounded, and facilitate their transfer to the hospital.

Practical Part

The data collected on traffic accidents for the years 2003-2018 was used to find the appropriate model for predicting the number of future traffic accidents Table 1.

Table 1: numbers of accidents

No.accidents	years
2063	**2003**
2483	**2004**
2731	**2005**
292	**2006**
201	**2007**
590	**2008**
1422	**2009**
1453	**2010**
1463	**2011**
1750	**2012**
1379	**2013**
1323	**2014**
1248	**2015**
1111	**2016**
1015	**2017**
1023	**2018**

Source: Republic of Iraq, Ministry of Interior, Baghdad Traffic Directorate, Department of Relations and Information, unpublished data for the period (2003-2018).

Data stability

After collecting the data, which is the first phase of the Box Jenkins methodology, we draw the data series for the preparation of incidents from 2003 to 2018 to identify the behaviour of the series as shown in Figure (1).

Figure 1. Draw a series to prepare incidents

In Figure (1), we observe data fluctuation and instability of the time (Abdul Rahman, 2002) series in contrast because the data fluctuates around the average. For more accuracy we draw both the function of s correlation (PACF) respectively after they were calculated according to elf-correlation ACF) and partial self- correlation (PACF) respectively after they were calculated according to table (2).

Table 2: Self-correlation ACF and partial self-correlation functions PACF of the original data

lag	ACF	PACF
1	0.4499*	0.4499*
2	-0.1396	-0.4288*
3	-0.4686*	- 0.2857
4	-0.3109	0.454
5	-0.1533	-0.2694

Significant dat *

Figure 2. String autocorrelation ACF function

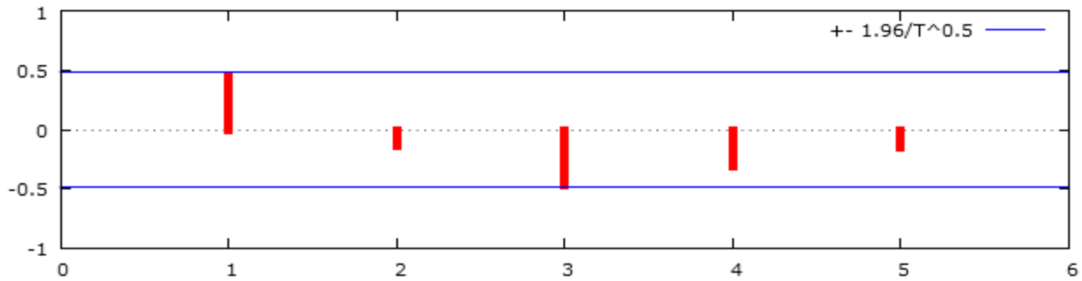
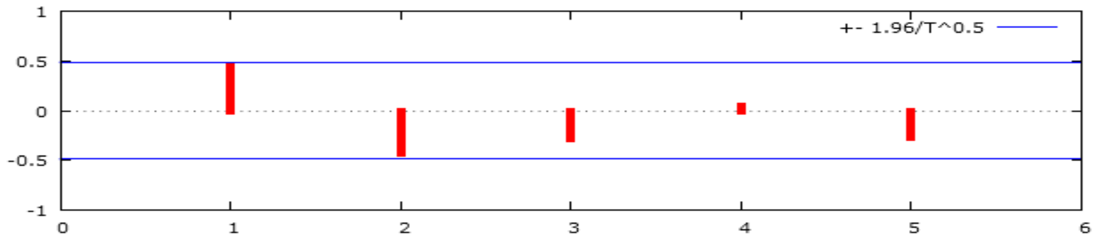


Figure 3. Partial self-correlation function PACF
of the series



From Figure (2), the ACF coefficients are declining exponentially and beyond confidence limits (0.5 ±) at a 95% confidence level. In Figure(3), the behaviour of the PACF partial displacement is observed beyond confidence limits for partial self-correlation coefficients. This is an indication of the lack of stability in the chain.

Located within the boundaries of trust the null hypothesis is H_1:□ EMBED Equat □ on.3□□□ = 0
Located beyond the limits of trust is accepted alternative H_1:□ EMBED Equat □ on.3□□□ ≠ 0

Hypothesis

The null hypothesis, which refers to the equality of coefficients of self-correlation with each other and equal to zero, is rejected. The alternative hypothesis is accepted after taking the logarithm and the first difference becomes the graph of the resulting series as shown in Figure (4). It seems that the series is stable in average and variance.

Figure 4. Plot the data series of records after the logarithm and the first difference

To ensure that the series is stable, both the ACF and PACF functions were plotted after taking the logarithm and the first difference as their coefficients are within confidence limits (at 95% confidence level, i.e. the data is stable. Figures (5) and (6) confirm this.

Table 3: Self-correlation ACF and partial self-correlation
PACF functions of chain stability after taking logarithm and first difference

LAG	ACF	PACF
1	0.1594	0.1594
2	-0.4425	-0.4801*
3	-0.2362	-0.0737
4	0.0409	-0.1405
5	-0.0280	-0.2163

Significant data*

Figure 5. The self-correlation function ACF of the chain after taking the logarithm and the first difference

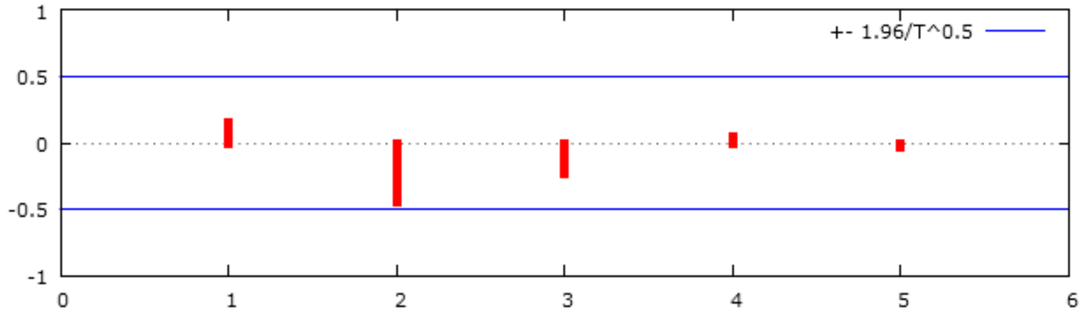
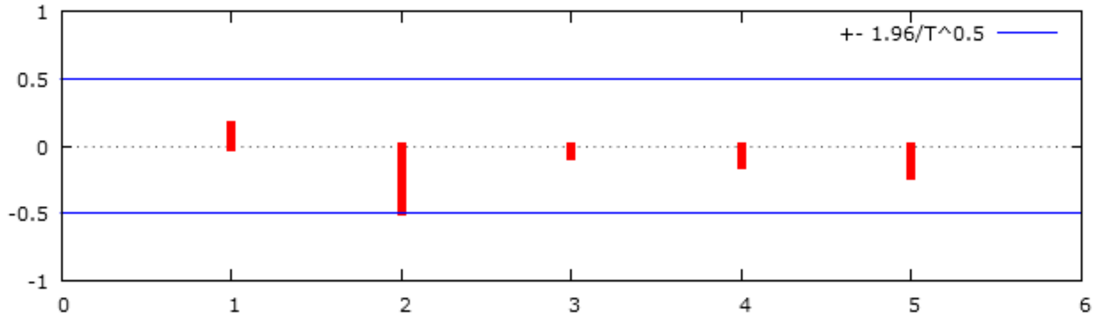


Figure 6. Draw the partial self-correlation function PACF of the series after taking the logarithm and the first difference



Extended Dicky Fuller test (ADF) was carried out to ensure chain stability without a fixed limit and with a fixed limit and with a fixed limit and general direction respectively. The test results were as shown in Table (4).

Unit Root Test

Extended Dicky Fuller test (ADF) was carried out to ensure chain stability without a fixed limit and with a fixed limit and with a fixed limit and general direction respectively. The test results were as shown in Table (4).

Table 4: The result test of (Dickey – Fuller)

	estimated value	test statistic	p-value
test without constant	-1.23275	-3.60128	0.0001
test with constant	-1.2476	-3.48496	0.008411
with constant and trend	-1.29532	-3.4917	0.0402

H_0: $\varnothing_1$ = 0
(Having the root of the unit)
H_1 : $\varnothing_1 \neq$ (Lack of root module)

Through Table (4) we find the equation without the constant limit, with the constant constant, and the general trend, the moral value (P-value) has less than the level of significance (0.05) and is located within the unit circle, that is, the chain is stable. The alternative hypothesis is accepted that the chain is stable.

Diagnosing and Estimating the Model and Choosing the Best Model

ARIMA models have been documented and the best model is selected according to the differentiation criteria (Akaike criterion, Schwarz criterion, Hannan-Quinn and the suggested models (Al-Baldawi, 2004) are shown in table (5).

Table 5: Box Genghis models proposed for the series

Models	AIC Akaike	SBC Schwarz	H-Q Hannan-Quinn	Notes
ARIMA(1,1,0)	35.23366	36.64976	35.21857	
ARIMA(0,1,1)	34.31908	35.73518	34.30400	
ARIMA(1,1,1)	35.50067	37.62482	35.47805	
ARIMA(2,1,1)*	32.75583	35.58803	32.72566	Optimal solution
ARIMA(1,1,2)	33.88203	36.71423	33.85186	
ARIMA(2,1,0)	33.84116	35.96531	33.81853	

International Journal of Innovation, Creativity and Change. www.ijicc.net
Volume 10, Issue 9, 2019

ARIMA(0,1,2)	**32.38185**	**34.50600**	**32.35922**	

From Table (5), we can see from the criteria of differentiation and confidence of the selected models that the best of the Box-Jenkins models is the model ARIMA (2,1,1). The values of (Hanan Quinn HQ, Schwartz SBC and AIKI) have the lowest among all the standard values of the proposed models, and the average absolute relative error has reached (4.353) (MAPE = The model parameters were estimated according to the maximum possible control method).
(Exact Maximum Likelihood). Through the Z test, based on the value of p-value, we notice that it is significant because the value of p-value is less than (0.05) as shown in the tables (6).

Table 6: Parameters of the selected model APIMA (2,1.1)

Type	Coefficient	std. error	z	p-value
PHI-1	0.776079	0.208621	3.720	**0.0002*** **
PHI-2	-0.516494	0.205214	2.517-	**0.0118** **
Thita_1	-1.00000	0.221157	-4.522	***$6.14e^{-6}$

The estimated model: (Formula Equation)
$(1-ß\ (1-0.776079\ ß + 0.516494\ ß^2\ )\ Y_1 = (1+ 1.00000)\ \alpha_1$

Model accuracy test

After the diagnosis of the model and determine the degree and estimate must be verified the suitability and efficiency of the model was done through the following tests: -

1- Ljung-Box test :-
$H_0: r_i(u) = r_1(u) = r_2(u) ... = r_K(u) = 0$
$H_1: r_i(u) \neq r_1(u) \neq r_2(u) ... \neq r_K(u) \neq 0$
As shown below (Ljung-Box) test in Table (7)

Table 7: Test results (Ljung-Box)

Lag	12
Chi-Square	7.6
DF	8
P-Value	0.469

Residual Test

The self-correlation and partial self-correlation functions of the residues (errors) of the estimated model were extracted and graphically represented.

The two figures (7), (8) show that the values of the correlation coefficients and the values of the partial correlation coefficients for the remainder are shown in the table. It falls within confidence limits (95%), which means that the sequence of residues represents random variables, indicating that the estimated model is good, efficient, and predictable.

Figure 7. ACF self-correlation function for the estimated model residues

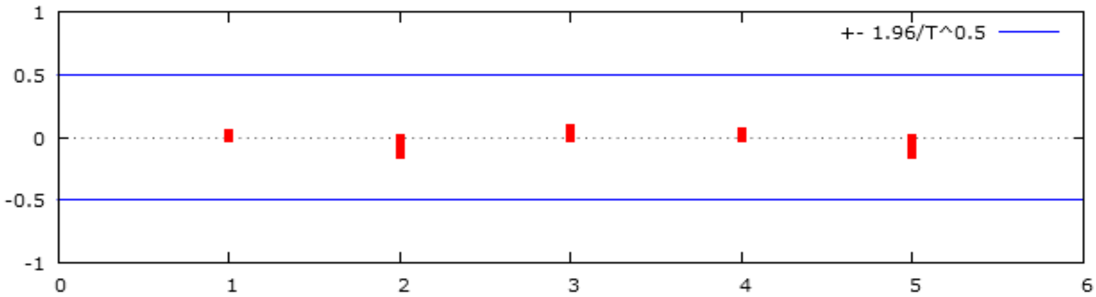
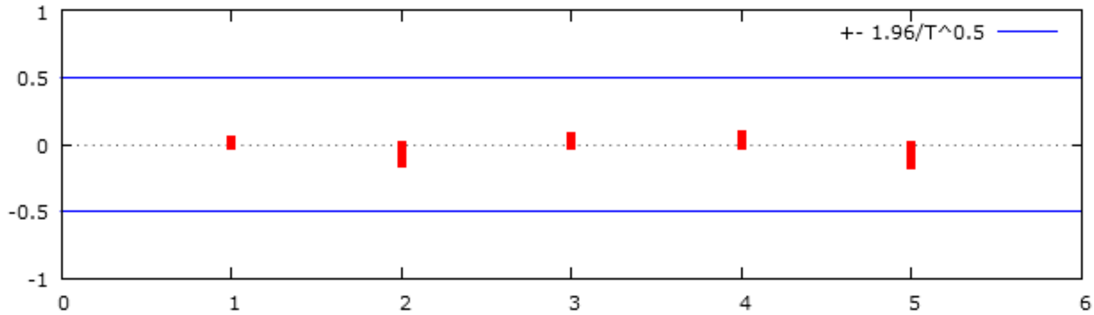


Figure 8. The residues of the estimated model (PACF) draw a partial self-correlation function



Prediction

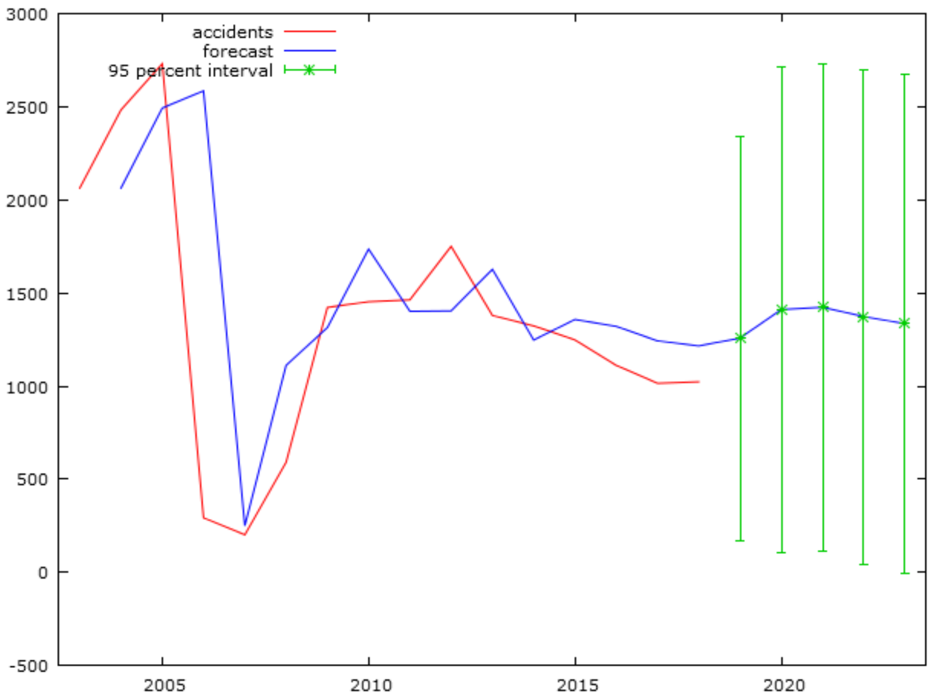
Table (9) shows the predicted values for the period from 2019 to 2023 and the confidence intervals for the predicted values.

Blue indicates predicted values, while green indicates minimum and maximum predicted values at 95% confidence level.

Table 9: Predictive values of the series according to the model ARIMA (2,1,1)

Year	Predictive value	minimum	maximum
2019	1257.50	171.29	2343.71
2020	1410.69	105.46	2715.29
2021	1423.19	116.23	2730.15
2022	1373.02	45.71	2700.33
2023	1334.81	-4.65	2674.27

Figure 9. Draw predictive values and confidence intervals for a series according to ARIMA (2,1,1)



Conclusions

1- The best model of Box Genghis is the ARIM model (2,1,1) since the values of Hanan Quinn are standard.
H-Q, Schwartz SBC and AICI have the least possible combination of standard values for the proposed models.
2. The ARIM model (2,1,1) is the appropriate model to be relied upon in establishing the future forecast of traffic accidents in the city of Baghdad for 2023-2027.
3. The model gave good predictions and close to actual values.

Recommendations

1- Adoption of the model by the General Directorate of Traffic and All Iraq. After we got the number of accidents for the period 2003-2018 for the city of Baghdad, it is possible to update the data and develop the future forecast of accidents there, as well as the general provinces of the country.
2 - Know the number of current and future incidents to identify the factors causing them. Study and analyse their causes and try to address them.
3- Establishing a database of traffic accidents

REFERENCES

Abdul Rahman, A. M. (2002). Statistical Forecasting Methods, King Saud University.

Al-Adhami, D. M. (2010). Scientific methods for determining safety procedures and prevention of vehicle accidents, research submitted to the Ministry of Transport, Baghdad.

Al-Attar, R. (1982). Oyouna and Road Accidents, Ministry of Culture and Information Publications, Dar Al-Zain, Publishing and Distribution, Baghdad.

Al-Baldawi, A. H. A. M. (2004). Applied Statistical Methods, Dar Al-Sharq for Publishing, Jordan, p. 241.

Al-Basrawi, A. K. A. (2010). A study entitled economic and social costs of traffic accidents in Iraq, the Arab Academy in Denmark.

Hanna, D.A. (1999). edition and causes of traffic accidents and future prospects, Master Thesis of the Higher Institute for Urban and Regional Planning, University of Baghdad.

Mohammed, D. T. (1986). A proposed program for traffic safety, unpublished master thesis, Journal of Engineering, University of Baghdad.

Republic of Iraq, (2016). Ministry of Water Resources, General Authority for Survey, Maps Production Section, Digital Unit, Scale, 1/1000000, Iraq.

Republic of Iraq, (2018). Ministry of Interior, Baghdad Traffic Directorate, Department of Relations and Information, unpublished data for (2003-2018).

Saleh B. M. A.-Q. (2007). Technical and Social Dimensions of Traffic Accidents, Center of Studies and Research, Naif Arab University for Security Sciences, Riyadh, Saudi Arabia.