مجلة الغري للعلوم الاقتصادية والادارية
مجلد (22) عدد (1) 2026

# Forecasting Traffic Accident Fatalities in the Kurdistan Region of Iraq using VARMAX Model

Bafrin Hamad Ali Nanakaly
Salahaddin University-Erbil
bafrin.ali@su.edu.krd

Assistant Prof. Delshad Shaker Ismael Botani
College of Administration and Economic
delshd.botani@su.edu.krd

## Abstract:

Traffic accident fatalities are one of the major global concerns, due to multiple factors affecting them and their impact socially and economically. Thus, forecasting the number of fatalities in any region is very significant in reducing this issue, especially when the number of vehicles is notably high and many families possess more than one car. Kurdistan Region of Iraq (KRI) has high fatality rates, as well as a high level of vehicle ownership in each governorate, highlighting the critical need for improved traffic management and safety strategies. This article aims to forecast traffic accident fatalities in KRI using multivariate dynamic time series models, specifically the Vector Autoregressive Moving Average with Exogenous variables (VARMAX) model. All the statistical analyses were performed using the Python programming language based on monthly data from 2015 to 2024. The average temperature and driving speed were included as exogenous factors to provide an extensive perspective of the dynamic regression relationships within these models, while the number of fatalities from three governorates (Erbil, Sulaymaniyah, and Duhok) were treated as endogenous variables. The dataset was divided into two subsets, with 80% used for training the model and 20% reserved for testing the performance. The results revealed that the VARMAX(1,1) model delivers the highest performance among the parameters investigated. Forecasts for traffic accident fatalities in the three governorates reveal a rising trend in 2025 and 2026. These findings underline the urgent need for targeted road safety programs and policy reforms to mitigate the prognosed increase in fatalities.

**Keywords**: *Forecasting, Dynamic Time Series, VARMAX, Traffic accident fatalities, Kurdistan region of Iraq.*

**DOI:** https://doi.org/10.36325/ghjec.v22i1.20796

Al-Ghary Journal of Economic and Administrative Sciences Vol. 22 (No.1) 2026 PP. 243-274

**مجلة الغري للعلوم الاقتصادية والادارية**

**مجلد** (22) **عدد** (1) **2026**

# التنبؤ بوفيات حوادث المرور في اقليم كوردستان العراق بإستخدام نموذج VARMAX

بفرين حمد علي نانەكەلى
جامعة صلاح الدين-أربيل
bafrin.ali@su.edu.krd

أ.م.د. دلشاد شاكر اسماعيل بوتاني
كلية الادارة والاقتصاد
delshd.botani@su.edu.krd

**المستخلص:**

تعد وفيات الحوادث المرورية من إحدى التحديات العالمية البارزة، نتيجة لتعدد العوامل المؤثرة فيها وتزايد آثارها الاجتماعية والاقتصادية. لذلك، التنبؤ بعدد الوفيات في أية منطقة له أهمية كبيرة في الحد من هذه المشكلة وخاصة عندما يكون عدد السيارات مرتفعًا بشكل ملحوظ وامتلاك العديد من العائلات أكثر من سيارة واحدة. يتمتع إقليم كردستان العراق (KRI) بمعدلات وفيات عالية، فضلاً عن ارتفاع مستوى ملكية السيارات في كل محافظة، مما يسلط الضوء على الحاجة الماسة إلى تحسين إدارة المرور واستراتيجيات السلامة. يهدف هذا البحث إلى التنبؤ بعدد وفيات حوادث المرور في إقليم كردستان العراق باستخدام نماذج السلاسل الزمنية الديناميكية متعددة المتغيرات، وتحديدًا نموذج متجه الانحدار الذاتي والاوساط المتحركة بوجود المتغيرات الخارجية (VARMAX). جميع التحليلات الإحصائية أجريت باستخدام لغة برمجة بايثون وإعتمادا على البيانات الشهرية للسنوات 2015 – 2024. تم إدخال معدل درجات الحرارة وسرعة قيادة السيارات كعوامل خارجية لإعطاء نظرة واسعة لعلاقات الانحدار الديناميكي داخل هذه النماذج، بينما تم التعامل مع عدد الوفيات من المحافظات الثلاثة (أربيل والسليمانية ودهوك) كمتغيرات داخلية. تم تقسيم البيانات إلى مجموعتين فرعيتين، استُخدمت 80% منها لتدريب النموذج، و20% لاختبار أدائه. أظهرت النتائج أن نموذج VARAMX(1,1) يحقق أعلى أداء من بين المعلمات التي تم دراستها. تشير توقعات وفيات حوادث المرور في المحافظات الثلاث إلى اتجاه تصاعدي في سنتي 2025 و2026. لذا تبرز هذه النتائج الحاجة الملحة إلى برامج مستهدفة للسلامة المرورية وسياسات إصلاحية للحد من الزيادة المتوقعة في الوفيات.

**الكلمات المفتاحية:** *التنبؤ، السلاسل الزمنية الديناميكية، VARMAX، وفيات الحوادث المرورية، إقليم كوردستان العراق.*

**DOI:** https://doi.org/10.36325/ghjec.v22i1.20796

## 1 Introduction

Traffic accidents, collisions, or crashes occur when vehicles collide with another, a person, an animal, a roadblock, or any stationary obstruction like a tree or a utility pole. Fatalities, injuries, and property loss are all possible outcomes of traffic accidents. Accidents involving motor vehicles can cause death and disability in addition to financial hardship. The loss of life, property damage, and depletion of resources are the main losses caused by traffic accidents.

Fatalities and injuries have risen in recent years. Young people and pedestrians are particularly vulnerable, with numerous occurrences blaming speeding, careless driving, and unsafe pedestrian passageways, and adverse weather conditions. The nonappearance of suitable public transit networks has also been brought about by increased private automobile use, exacerbating traffic congestion, and raising the likelihood of accidents, all of which have been substantiated in data and studies on traffic accidents.

Traffic accidents are getting increasingly significant concern in the Kurdistan Region of Iraq (KRI), which is experiencing rapid urbanization, population growth, and increasing vehicle ownership. With these growing patterns, highways in cities like Erbil, Sulaymaniyah, and Duhok are becoming more and more congested, raising the risk of an accident. Contributing factors include a mix of infrastructure deficiencies, high-speed driving, ineffective traffic law enforcement, at times, a failure to follow safety rules.

Forecasting traffic accidents is a primary priority for governments throughout the globe as they seek to reduce fatalities, property damage, and economic losses. The purpose of forecasting is to know the future rise of

DOI: https://doi.org/10.36325/ghjec.v22i1.20796

road accidents and to judge the feasibility and impact of traffic safety measures equitably. Traffic accident forecasting is the basis and evidence for road traffic security administration, policies, and choices. It may also be used to govern every contributing aspect efficiently and to fulfill the purpose of minimizing traffic accidents. The goal of the traffic accident forecast is to find patterns in the past relevant data and then to anticipate future trends. The associations between external factors and accidents have been used in this way to identify the applicable laws.

The purpose of this article is to identify the best multivariate dynamic models using Vector Autoregressive Moving Average with Exogenous Variables (VARMAX) models for forecasting traffic accident fatalities in the three main governorates of KRI (Erbil, Sulaymaniyah, Duhok). The findings are meant to support the government’s efforts in making better use of its resources, getting better at preparing for emergencies, and improving traffic management planning in the future.

The article is structured as follows: The second Section provides a brief overview of the literature review about VARMAX models for forecasting traffic accidents. The third section gives the theoretical aspect of the study’s core subject area. The fourth section dealt with the practical aspects of the study, including findings, which are divided into two parts. The first part is concerned with some descriptive statistics, the second contains the results of analyzing the traffic accidents data, identifying exogenous variables, and VARMAX modeling. Section five contains the important conclusions and recommendations.

## 2 Literature Review

Box et al. (1976)has significantly influenced the theory and practice of time series analysis. In fields like business and economics, where data often take the form of a time series, the effect has been particularly noteworthy. Holden-Day published their book "Time Series Analysis: Forecasting and Control," which introduced techniques for evaluating time series data and presented the Autoregressive Moving Average (ARMA) models. Their work made crucial tools for subsequent models, such as VARMA and VARMAX possible.

Hamilton (1994)extended the VARMA framework to include exogenous variables, forming the theoretical basis for VARMAX models. This allowed multivariate dynamic models that incorporate both internal relationships and external predictors, enhancing the ability to capture complex interactions in time series data.

VARMAX-modeling of blast furnace process variable was studied by Ostermark et al. (1996). This study used critical process data from a blast furnace to provide initial evidence for VARMAX modeling of hot metal silicon concentration and pig iron temperature. Given that the dynamics of pig iron manufacture were well understood in the blast furnace example, the study's objective was to provide a helpful framework for realistic modeling and control of the manufacturing process. The temperature and silicon content VARMAX models are given.

Tsirigotis et al. (2011) were published a study under the title "Does Weather Information Affect the Performance of Short-Term Traffic Forecasting Models?" This study employed VARX models with exogenous variables to examine the effects of weather, traffic mix, and speed variation on highway traffic flow. The findings showed that although modeling advancements like vector and Bayesian estimates significantly improve models, adding exogenous variables only marginally improves prediction ability. The conventional univariate and multivariate Autoregressive Integrated Moving Average (ARIMA) models were surpassed by the VARX and Bayesian VARX. Important correlations between lane speeds, volume, traffic mix, and weather were found, despite only slight gains in accuracy. Multivariate models like VARMAX can reliably forecast speeds in various lanes, as seen by the 30–50% improvement in performance compared to the ARIMAX model. The use of meteorological data on time series models for speed forecasting is examined in this study.

Winter Road Surface Condition (RSC) Forecasting was researched by Feng et al. (2015). This study was the first to employ VARMAX to evaluate the winter RSC evolution process utilizing univariate and multivariate ARIMA approaches. Surface temperature and pollution layer depth forecasting errors were rather small. The calibrated models were simple, understandable, and typically consistent with what was known about the physical world. The proposed models allowed for more flexibility in terms of reworking, tweaking, and deployment as compared to existing models.

Parvareh et al., (2018) published their research article about “Assessment and prediction of road accident injuries trend using time-series models in Kurdistan”. A time-series analysis was performed to define and forecast the frequency of road traffic accidents that cause injury in Kurdistan region of Iraq. The injuries were grouped into three main categories which were related to the automobile occupants, motorcyclists and pedestrian road traffic collision injuries. The Box-Jenkins time-series analysis was used to model the injury observations applying ARIMA and SARIMA models from March 2009 to February 2015 to forecast the accidents for 2 years. The study was conducted using R-3.4.2 statistical software tool.

Abdulqader et al., (2020) published their study and used the Box-Jenkins approach, which stands for the Autoregressive Integrated Moving Average ARIMA time series model, to examine and predict the patterns of injuries sustained in Erbil city's road traffic accidents on a monthly basis. They got monthly accident data about wounded persons from 2013 to 2018 in Erbil governorate. The data demonstrated that the series has seasonal characteristics, and that traffic-related injury rates decrease over the year. After developing several viable models, SARIMA $(0,1,1)(1,0,1)^{12}$ was judged to be both sufficient and the best model based on a number of performance metrics. A monthly forecast was prepared using the best model and it indicated that the number of wounded cases due to traffic accidents will continue to decline in Erbil city until the end of December 2020.

The VARIMAX model was used by Peng et al. (2021) to forecast the trend of thick fog. With an emphasis on the development of dense fog, this study investigated the connection between visibility and near-surface meteorological conditions. The research examined the relationship between these variables using VARIMAX and airport AWOS observation data. The findings indicated that although the dew point temperature index hurt vision, temperature had the most effect.

Ahmed et al. (2023) researched “Enhancing the ARIMAX Model by using the Bivariate Wavelet Denoising: Application on Road Traffic Accidents”. In this work, the topic of the possibility of improving the traditional ARIMAX model based on the bivariate wavelet analysis

DOI: https://doi.org/10.36325/ghjec.v22i1.20796

technique was discussed. A suggested model was applied to the monthly data that reveals the number of traffic accidents. Reliant on the chosen BWARIMAX (3,0,0) model, a bit fewer will happen in the number of traffic incidents in 2023 compared to 2022.

Ahmed et al. (2023) published a study about “Building SARIMA Models to Analysis and Forecast Time Series Data of Road Traffic Accidents in the Kurdistan Region of Iraq” and they have been explained that the Traffic accidents inflict injuries and fatalities but also bring material harm to society. The primary purpose of this research was to create an ARIMA time series model to investigate and evaluate the number of traffic accidents and the number of fatalities in traffic accidents in the Iraqi Kurdistan Region according to monthly data (2014-2021) and monthly projection for 2022. Data were gathered from Erbil General Directorate of Traffic. The best model was SARIMA (1,1,1) $(0,1,1)^{12}$ for the number of accidents and SARIMA(0,1,1)$(1,1,2)^{12}$ models for the number of fatalities.

Mahmood et al. (2024) published a study titled "Building a Statistical Model to Forecast Traffic Accidents for Death and Injuries by Using Bivariate Time Series Analysis." Focusing on Erbil city, Iraq, they analyzed a bivariate time series consisting of the accident rates of injuries and mortality. Monthly data collection for the two series was placed between January 2015 and December 2020. They concluded that the time series is stationary and that VARMA (1,0) is the best model to describe the occurrences under investigation. As a result, the research recommended that the Kurdistan Region Traffic Department anticipate future accident occurrences using the proposed methodology.

DOI: https://doi.org/10.36325/ghjec.v22i1.20796

## 3 Methodology

### 3.1 Time Series Models

Time series analysis is a mathematical process that is used to analyze data that has been collected successively over time with an emphasis on a particular variable whose value with time. Because previous values might affect future results, time series model stress the chronological structure, in contrast to standard data analysis, which often disregards the sequence of observations. To enhance forecasts and achieve a better understanding of the underlying dynamics, these models seek to find trends, patterns, and seasonal effects in the data(Ali et al., 2025; Zivot & Wang, 2006). Based on current measurements, these models are used to identify underlying trends and forecast new values. The univariate time series ARIMA model, which is used to forecast time series, is expanded upon by this model. Autoregressive (AR), moving average (MA) processes and mixed ARMA will be briefly discussed in this section. The lag operator L, which shifts the subscript of a time series variable backwards by one period, might sometimes be useful for simplifying the notation (Cryer & Chan, 2008). The ARIMA (p, d, q) model, where p is the autoregressive terms number, d is the differences number, and q is the moving average terms number, could be written as:

$$\boldsymbol{Y_t = \sum_{i=1}^{p} \emptyset_i \, Y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t} \qquad \textbf{(1)}$$

Furthermore, when there is extraneous information that might influence the time series. The ARIMAX model helps by taking into account external factors (exogenous variables) that could have an impact on the target variable. It is necessary to estimate model parameters, provide forecasts, and include exogenous variables in order to determine the AR, I, and MA

DOI: https://doi.org/10.36325/ghjec.v22i1.20796

components of the model the notation ARIMAX(p, d, q, s) is written as (Cryer & Chan, 2008; Haydier et al., 2023; Tibshirani, 2023):

$$\boldsymbol{Y_t} = \sum_{i=1}^{p} \emptyset_i \boldsymbol{Y_{t-i}} + \sum_{j=1}^{q} \boldsymbol{\theta_j \varepsilon_{t-j}} + \sum_{r=1}^{s} \boldsymbol{B_r X_{t-r}} + \boldsymbol{\varepsilon_t} \qquad \textbf{(2)}$$

where, $Y_t$ is the dependent variable, $\emptyset_1, \emptyset_2, \dots, \emptyset_p$ are the coefficients of the autoregressive terms, $\theta_1, \theta_2 \dots, \theta_q$ are the coefficients of moving average terms, $\varepsilon_t$ is the error term (white noise), $X_{t-r}$ is the exogenous variable lagged by r periods, and $B_r$ is the coefficient for the exogenous variable.

Multivariate time series analysis is a subfield of multivariate statistics that focuses on specifically temporally dependent data and many interrelated time series in studies such as economics, finance, and traffic systems depend on understanding not only the individual behavior of each variable but also their dynamic interactions, an essential aspect for accurate predictions and well-informed decision making. This approach is significantly more complex than univariate time series analysis (Tsay, 2014). Univariate time series data with autoregressive and moving average components are molded as ARMA together. These models can be generalized at accommodate multiple interdependent time series leading to development of Vector Auto-Regressive (VAR) models, which reflect the dynamins interaction between several endogenous variables. To further incorporate the influence of external factors, VAR models have been expanded to Vector Autoregressive with Exogenous variables (VARX) in order to better calculate the impact of external factors. Simultaneously, the Vector Autoregressive Moving Average (VARMA) model has been created concurrently by incorporating the Moving Average (MA) component into the multivariate framework. The VARMA and VARMAX models are more complex to specify individually and estimate accurately because of the additional complications introduced

DOI: https://doi.org/10.36325/ghjec.v22i1.20796

by this generalization that are not present in univariate models, most remarkably, the identifiability problem, where multiple parameterizations may represent the same underlying process. The most comprehensive version is the VARMAX model, which associates exogenous effects with the MA component for a reliable study of multivariate time series data impacted by outside forces (Lütkepohl, 2005; Tsay, 2005).

### 3.2 VARMAX Models

The VARMAX models are multivariate time series models meant to capture the dynamic interrelationships among numerous endogenous variables while simultaneously accounting for the effect of external (exogenous) factors. It develops the classic univariate ARMAX framework by incorporating several interdependent time series, enabling both autoregressive and moving average components across equations. This model is particularly well-suits for complex systems, such as those in economics, engineering, finance, and transportation, where variables do not change over time but also exhibit reciprocal effects. The VARMAX model offers a comprehensive structure for both explanation and forecasting in multivariate time series research by modeling these interactions and the impacts of external factors. (Boddu et al. 2017). The VARMAX (p, q, s) model equation written as (Shumway & Stoffer, 2011; Chatfield, 2001):

$$\boldsymbol{Y_t = C_t + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{r=1}^{s} B_r X_{t-r} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t} \qquad \textbf{(3)}$$

where: $Y_t$ is a $K \times 1$ vector time series variables over time t, where K is the number of variables, $c$ is a $K \times 1$ vector constant (mean values), i.e.,$\mu = [\mu_1, \mu_2, \dots, \mu_k]$, $\phi_i$ is a matrix of coefficients of size $K \times K$ corresponding to the lags of the variables, $\theta_j$ is a matrix of coefficients of size $K \times K$, representing the influence past error term on the current values of variables,

$X_{t-r}$ are Vectors of exogenous variables with lag r, $B_r$ is a coefficient matrix for the exogenous factors, $\varepsilon_t$ is a vector of error terms or residuals at time t.

The VARMAX model comprises autoregressive (AR) terms that account for the past values of dependent variables and moving average (MA) terms that capture previous error impacts. Additionally, the model contains exogenous variables to quantify the influence of external factors on the system. Each equation in the system incorporates lags of all endogenous and exogenous variables, allowing for dynamic interactions. This structure makes the VARMAX model particularly fit for forecasting and evaluating systems impacted by both internal dynamics and external inputs. Data preparations, stationary testing, model identification (lag order selection), model estimate, evaluation, and forecasting are the main steps in creating VARMAX model (Box et al., 2016; Shumway & Stoffer, 2011; Tsay, 2014; Wei, 2019).

### 3.2.1 Model Identification (specification)

Model identification is the first point in producing a model. In this work, the VARMAX model has been selected for its capability to assess multivariate data with interdependencies and exogenous factors. After collecting and preparing the data, the stationary is checked using the Augmented Dicky fuller test (ADF), which is checking for the presence/absence of unit roots in the time series (Chatfield, 1995; Wei, 2005). The ADF test statistic is the t-statistic of the estimated coefficient obtained from least squares regression. However, the ADF test statistic is not essentially t-distributed further down the null hypothesis; instead, it has a fastidious non-standard large-sample distribution under the null hypothesis of unit root. Fortunately, percentage points of this limit (null) distribution

DOI: https://doi.org/10.36325/ghjec.v22i1.20796

have been documented (Cryer & Chan, 2008; Wei, 2005). To address non-stationarity, the series is changed through differencing. The appropriate Lag orders for AR, MA, and exogenous components are found using ACF and PACF plots. This step ensures that the data is properly formatted and suitable for the VARMAX model estimate (Box et al., 2016; Box et al., 2008; Shumway & Stoffer, 2011; Tsay, 2014; Wei, 2019).

### 3.2.2 Model Selection Criteria

Model selection criteria in multivariate time series are statistically used to evaluate and contrast models to identify the most appropriate one for a given set of data. Known metrics for model selection include ( Box et al., 2008; Tsay, 2014):

*a) Akaike Information Criterion (AIC)*

To choose the appropriate model specification, the Akaike information criterion (AIC) is used to specify the best- performing VARMAX model with the lowest possible score, given the trade-off between complexity and goodness of fit, a model's relative quality is determined by its AIC, which is lower for superior models (Shumway & Stoffer, 2011). The AIC criterion is defined as:

$$AIC = ln\left|\hat{\Sigma}_{\varepsilon}\right| + \frac{2}{n}\left(kr + \frac{k(k+1)}{2}\right) \quad (4)$$

*b) Bayesian Information Criterion (BIC)*

The Bayesian Information Criterion (BIC) also known as Schwarz criteria is another model selection Criterion based on Bayesian likelihood maximization. The BIC's definition is (Shumway & Stoffer, 2005).

$$BIC = log\left|\hat{\Sigma}_{\varepsilon}\right| + k^2 p \ln n/n \quad (5)$$

*c) Hannan-Quinn Information Criterion ( HQIC)*

DOI: https://doi.org/10.36325/ghjec.v22i1.20796

It is a statistical metric used for model selection, mainly in time series analysis and econometrics. It stabilizes model fit and complexity, like AIC and BIC, but with a different penalty function (Box et al. 2008).

$$\boldsymbol{HQIC = -2\,ln(L) + 2 * k * ln\,(ln(n))} \quad \textbf{(6)}$$

where: $L$ is the maximized likelihood of the model, $K$ is the number of parameters in the model (including constant, AR, MA, and exogenous terms), $n$ is the number of observations.

### 3.2.3 Maximum likelihood Estimation (MLE)

The parameter estimate is a common first step in defining the linear VARMAX model for time series. Either the maximum likelihood (ML) or the least squares (LS) methods might be used to estimate the VARMAX's parameters (p, q, s). The ML estimates of a VARMAX (p, q, s) model are asymptotically identical to the LS estimates. For the LS techniques under the multivariate normality assumption, which states that the model has a k-dimensional normal distribution (Robert H. Shumway & Stoffer, 2011; Wei, 2019). Maximum likelihood Estimation (MLE) is a commonly used approach in multivariate time series analysis, notably for models like VARMA and VARMAX. Then the ln-likelihood function of the standard normal distribution may be expressed as follows:

$$\boldsymbol{ln\,L(\theta) = -\frac{n}{2}ln(2\pi) - \frac{n}{2}ln|\Sigma| - \frac{1}{2}\sum_{t=1}^{n}(\varepsilon_t'\Sigma^{-1}\varepsilon_t)} \quad \textbf{(7)}$$

and

$$\boldsymbol{S(\Phi, \Theta, \psi) = \sum_{t=1}^{n}(\varepsilon_t'\varepsilon_t)} \quad \textbf{(8)}$$

The least squares approach is used to reduce $S(\Phi, \Theta, \psi)$, while the maximum likelihood method is used to maximize the log-likelihood function (Andayani et al., 2015).

DOI: https://doi.org/10.36325/ghjec.v22i1.20796

### 3.2.4 Residual Autocorrelation Test (Ljung-Box Test)

Diagnosing the presence of autocorrelation in a multivariate time series residuals is conducted using Ljung–Box test which measures the autocorrelation coefficients for residuals of a set of lags in order to quantify the unpredictability of time series residuals. This is how the test hypothesis is written:

$H_0: \rho_1 = \rho_2 = \cdots = \rho_m = 0$ (no serial correlation),

$H_1: \rho_i \neq 0$ for some $i$ satisfying (there is serial correlation)

For some $1 \leq i \geq m$, where $m$ is a positive integer.

$$\boldsymbol{Q_k(m) = T^2 \sum_{\ell=1}^{m} \frac{1}{T-\ell} tr(\hat{\Gamma}'_\ell \hat{\Gamma}_0^{-1} \hat{\Gamma}_\ell \hat{\Gamma}_0^{-1})} \qquad \textbf{(9)}$$

where tr(A) is the trace of the matrix A and T is the sample size.

If the value of the Q is greater than the tabulated value (i.e., p-value <0.05), we would reject the null hypothesis at the significance level α. Which means that there is serial correlation. The Ljung–Box statistics can be used to check the adequacy of the mean equation and can be used to test the VARMAX model (Tsay, 2014).

### 3.3 Forecasting

Forecasting is one of the keys aims in the study of multivariate time series data. forecasting in a VARMAX(p, q, s) model is basically comparable to predicting in a univariate ARIMAX (p,q,s) model. First, the essential notion in the process of forecasting is that the best VARMAX model must be discovered using specific criteria for selection of the best model. Once the model is discovered, it may be utilized for forecasting (Chatfield, 2001; Warsono et al., 2019). After modeling VARMAX, in order to evaluate the Forecasting accuracy, these performance metrics evaluate how well the models fit the data and allow for a qualified analysis of their forecasting powers (Chatfield, 2001; Chukwutoo et al. 2018; Rojas et al., 2017). The accuracy metrics that will be employed are as follows:

- **Mean Absolute Error (MAE)** $\boldsymbol{MAE = \frac{\sum_{t=1}^{n}|\varepsilon_t|}{n}}$ **(10)**
- Mean Squared Error(MSE) $MSE = \frac{\sum_{t=1}^{n}\varepsilon_t^2}{n}$ (11)
- Root Mean Squared Error (RMSE) $RMSE = \sqrt{\frac{\sum_{t=1}^{n}\varepsilon_t^2}{n}}$ (12)
- **Mean Absolute Percentage Error (MAPE)** $\boldsymbol{MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{\varepsilon_t}{Y_t}\right|}$

  **(13)**

## 4 Results and implications

### 4.1 Descriptive Statistics

This study utilizes Traffic and Environmental time series data from KRI, acquired from January 2015 to December 2024 (120 observations). The data consists of the number of accident fatalities in the three main governorates of KRI (Erbil, Sulaymaniyah, and Dohuk), as well as driving speed and temperature which are represented as exogenous variables. The reasons for selecting just these 2 exogenous variables are due to lack of access to recent data, lack of historical data, presence of significant correlation among certain candidate variables, and complexity of the analyses using VARMAX model.

The data was divided into two subsets, namely a training set and a testing set to evaluate the out-of-sample forecasting performance of the VARMAX model. The training set was used to estimate the model parameters, while the testing set was reserved to assess the accuracy of the forecasts on observations not included in the estimation process. This approach helps reduce the risk of overfitting and provides a more objective measure of the model's generalizability. The first subset is called Training dataset – %80 consisting of eight years of data starting from 01/01/2015 to 31/12/2022, this data was used to estimate the ARIMAX, VARX, and VARMAX models for underlying data. The second subset is called Testing data – %20 consisting of two years starting from 01/01/2023 to31/12/2024, this data was used to estimate the forecasting performance of the models. After selecting the best forecast model, the model will be used to forecast future trends in the traffic accident fatalities.

Based on **Table (**1, key descriptive statistics are summarized for the three governorates. Sulaymaniyah recorded the highest monthly average of fatalities with 22.3, compared with 17.7 in Erbil and 9.6 in Duhok. One of the contributing conditions is excessive speed when driving and could also be affected by weather conditions. Remarkably, Sulaymaniyah had the highest average monthly driving speed of 94.3 km/h, twice the average driving speed in both Erbil and Duhok governorates. The measures of skewness for the number of fatality series are positive, it seems more low values than higher value of (0.7-1.0), driving speed series are slightly positively skewed (0.2-1.3), and the Temperature series are close to 0. Shapiro Wilks (S. W.) test is used to check whether the data are normally distributed or not, it is a formal way to the assumption of normality. The significant level (alpha) is set at 0.05, the number of traffic accident fatalities and temperature for all governorates are not normally distributed, while the driving speed series in Sulaymaniyah and Duhok are normal.

**Table (1): Descriptive statistics summary for the monthly traffic data**

| Statistics | No. of Fatality | | | Driving speed | | | Temperature | | |
|---|---|---|---|---|---|---|---|---|---|
| | Erbil | Suli. | Duhok | Erbil | Suli. | Duhok | Erbil | Suli. | Duhok |
| Min. | 3 | 10 | 2 | 16 | 32 | 7 | 7.6 | 3.4 | 6.2 |
| Max. | 56 | 48 | 26 | 128 | 171 | 65 | 36.6 | 36.1 | 35.2 |
| 1st Qu. | 10 | 17 | 5 | 31 | 70 | 27 | 12.9 | 11.6 | 12.8 |
| Median | 15 | 21 | 8.5 | 45 | 90.5 | 34.5 | 23 | 21 | 21.8 |
| 3rd Qu. | 24 | 26 | 13 | 56 | 113.3 | 44 | 31.8 | 30.7 | 29.8 |
| Mean | 17.7 | 22.3 | 9.6 | 46.8 | 94.3 | 35.7 | 22.5 | 20.6 | 20.9 |
| Variance | 93.5 | 56.5 | 30.7 | 424.5 | 881.8 | 138.1 | 87.7 | 101.9 | 84.5 |
| Skewness | 0.9 | 1.0 | 0.7 | 1.3 | 0.3 | 0.2 | 0.0 | 0.1 | 0.1 |
| Kurtosis | 0.8 | 1.1 | -0.3 | 2.8 | -0.5 | -0.3 | -1.5 | -1.4 | -1.4 |
| S. W. test | 0.9 | 0.9 | 0.9 | 0.9 | 1.0 | 1.0 | 0.9 | 0.9 | 0.9 |
| P-Value | 0.0 | 0.0 | 0.0 | 0.0 | 0.1 | 0.3 | 0.0 | 0.0 | 0.0 |

The correlation matrix was computed using Pearson correlation, the correlation plot for the study data is shown in **Figure (**1, where the correlation coefficient ranges from the light to dark blue (-1 to +1). There is a positive correlation (0.5) between the governorates Erbil and Duhok, a positive correlation (0.4) between Erbil and Sulaymaniyah, and a positive correlation (0.38) between Duhok and Sulaymaniyah. Overall, the correlation matrix shows that environmental factors and local traffic patterns both have a significant impact on road safety

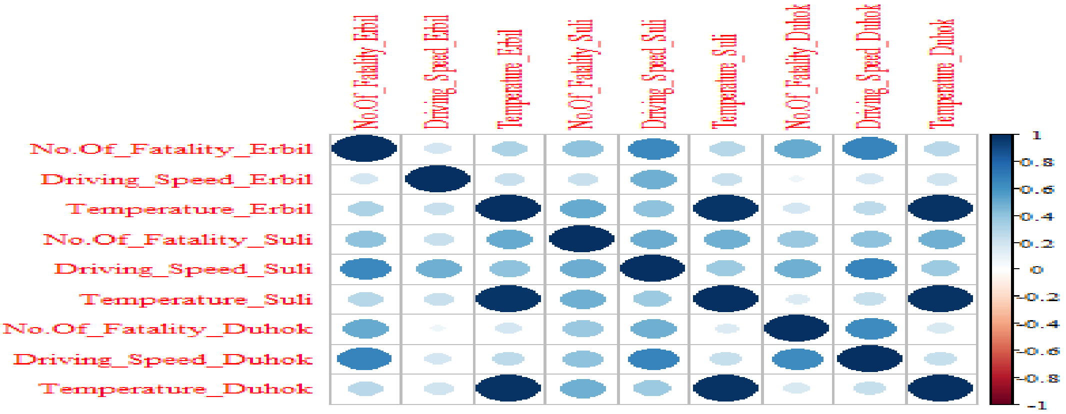


**Figure (1) Pearson's Correlation Matrix Plot of the study data**

### Time Series Analysis (Identification and Stationarity)

Model identification is the initial step in the time series analysis. Choosing the number of accident fatalities in Erbil, Sulaymaniyah, and Duhok governorates as endogenous variables. However, the exogenous variables are driving speed and temperature. **Figure (**2 shows the endogenous variables line plot starting from early 2015 until the end of 2024. On the other hand, **Figure (**3 shows the driving Speed for these accidents in the three governorates for the period 2015 – 2024. The figure demonstrates clearly that Sulaymaniyah (the green line) always maintains the highest driving speed of the three governorates, with frequent spikes well in excess of 150 km/h in previous years (2015–2019), then tapering off in the later years. Erbil (red line) holds moderate driving speeds, typically ranging from 40–90 km/h, with occasional sharp dips beginning after 2020. While, Duhok (blue line) always indicates the lowest driving speed, remaining mostly under 60 km/h, and with relatively smaller variations over the years.

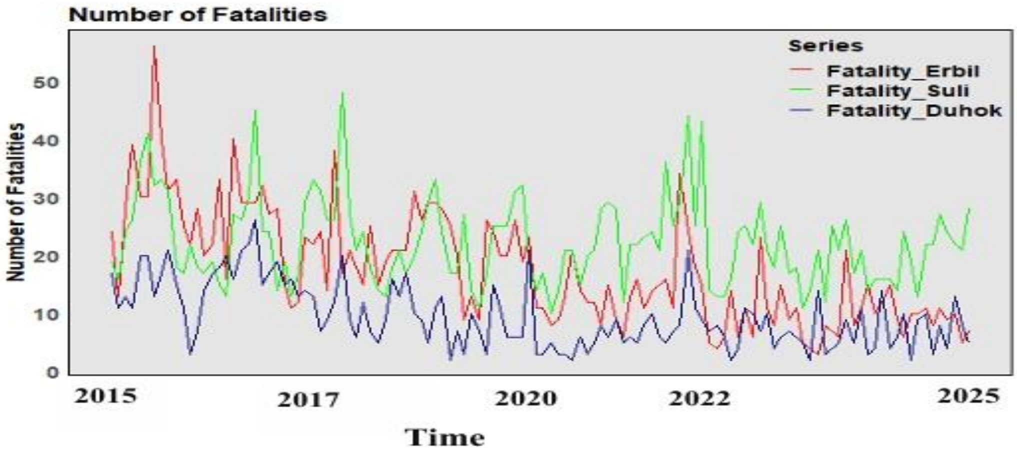


**Figure (2) Number of fatalities trends in KRI governorates (2015 2024)**

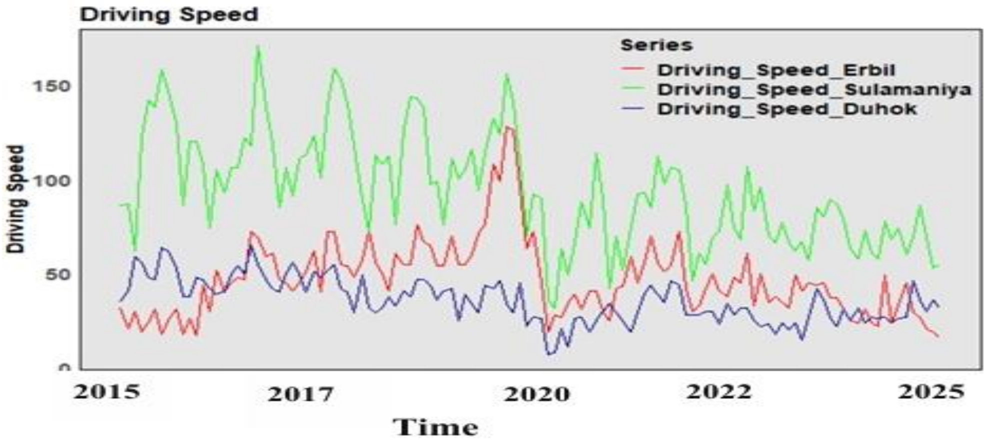


**Figure (3) Driving speeds trends in KRI governorates (2015 – 2024)**

Based on the time series plot for temperature, in general, cloudiness is like the normal climate in KRI appeared, where the high temperature in the summer, moderate in the spring and autumn seasons, and very low in the winter for the three governorates, although there is a slight difference between them.

After collecting and processing the time series traffic accident data, and selecting the endogenous and exogenous variables, the second step to identify the model is checking out the stationarity of the data. This was performed using the ADF test, as seen in **Table (**2 .With a p-value of 0.125, which is higher than the 5% significance level, the ADF test results show that the data for the Duhok governorate is non-stationary. On the other hand, p-values of 0.01 indicated that the data for Erbil and Sulaymaniyah were stationary at level 0.01. First-order differencing (d=1) was used to transform the data series to guarantee consistency across all variables and remove the effects of fluctuating mean and variance over time. The series became stationary after differencing, as indicated by the ADF test findings, which exhibited p-values less than 0.05 for the three-time series.

**Table (2): ADF test for study data before and after differencing (d=1)**

Based on the time series plot for temperature, in general, cloudiness is like the normal climate in KRI appeared, where the high temperature in the summer, moderate in the spring and autumn seasons, and very low in the winter for the three governorates, although there is a slight difference between them.

After collecting and processing the time series traffic accident data, and selecting the endogenous and exogenous variables, the second step to identify the model is checking out the stationarity of the data. This was performed using the ADF test, as seen in **Table (**2 .With a p-value of 0.125, which is higher than the 5% significance level, the ADF test results show that the data for the Duhok governorate is non-stationary. On the other hand, p-values of 0.01 indicated that the data for Erbil and Sulaymaniyah were stationary at level 0.01. First-order differencing (d=1) was used to transform the data series to guarantee consistency across all variables and remove the effects of fluctuating mean and variance over time. The series became stationary after differencing, as indicated by the ADF test findings, which exhibited p-values less than 0.05 for the three-time series.

**Table (3): ADF test for study data before and after differencing (d=1)**

| **Variables** | **Before differencing** | | | **After differencing (d=1)** | | |
|---|---|---|---|---|---|---|
| | **ADF test** | **5% Critical Values** | **p-value** | **ADF test** | **5% Critical Values** | **p-value** |
| No of fatalities Erbil | -5.67 | -2.88 | 0.01 | -6.46 | -2.88 | 0.01 |
| No of fatalities Sulaymaniyah | -5.56 | -2.88 | 0.01 | -6.37 | -2.88 | 0.01 |
| No of fatalities Duhok | -2.87 | -2.88 | 0.125 | -6.37 | -2.88 | 0.01 |

The Null Hypothesis $H_0$: No. of fatalities Series has a unit root (Not Stationary)
The Alternative Hypothesis $H_1$: No. of fatalities Series has not a unit root (Stationary)

DOI: https://doi.org/10.36325/ghjec.v22i1.20796

## 4.3

### 4.3.1

**Table**

| Extended | MA | MA | MA | MA | MA | MA |
|---|---|---|---|---|---|---|
| **AR** | 0 | 0.0401 | 0.0231 | 0.0081 | 0.0069 | 0.035 |
| **AR** | 0.032 | 0.2801 | 0.0416 | 0.4655 | 0.1136 | 0.8271 |
| **AR** | 0.2786 | 0.0799 | 0.2424 | 0.5536 | 0.8918 | 0.9221 |
| **AR** | 0.4212 | 0.1266 | 0.9041 | 0.916 | 0.1629 | 0.9032 |
| **AR** | 0.9155 | 0.5076 | 0.9853 | 0.9571 | 0.8501 | 0.938 |
| **AR** | 0.9991 | 1 | 1 | 0.9997 | 0.9678 | 1 |

**Table**

| Models | **AIC(n)** | **HQIC(n)** | **BIC(n)** |
|---|---|---|---|
| VARMAX | 1889.971 | 1927.287 | 1982.288 |
| VARMAX | 1906.931 | 1953.576 | 2022.327 |
| VARMAX | 1923.544 | 1979.518 | 2062.019 |
| VARMAX | 1946.926 | 2012.229 | 2108.480 |
| VARMAX | 1918.056 | 1955.372 | 2010.373 |
| VARMAX | 1904.228 | 1950.873 | 2019.624 |
| VARMAX | 1909.633 | 1965.607 | 2048.108 |
| VARMAX | 1930.232 | 1995.535 | 2091.786 |
| VARMAX | 1914.211 | 1960.856 | 2029.606 |
| VARMAX | 1917.899 | 1973.873 | 2056.374 |
| VARMAX | 1924.555 | 1989.858 | 2086.109 |
| VARMAX | 1943.096 | 2017.728 | 2127.729 |
| VARMAX | 1911.154 | 1967.128 | 2049.629 |
| VARMAX | 1918.18 | 1983.482 | 2079.734 |
| VARMAX | 1932.108 | 2006.74 | 2116.742 |
| VARMAX | 1946.713 | 2030.674 | 2154.426 |

Depending on $\boldsymbol{Yt}=\boldsymbol{C_t}+\sum_{i=1}^{p}\boldsymbol{\phi_i}\,\boldsymbol{Y_{t-i}}+\sum_{r=1}^{s}\boldsymbol{B_r}\,\boldsymbol{X_{t-r}}+\sum_{i=1}^{q}\boldsymbol{\theta_i}\,\boldsymbol{\varepsilon_{t-i}}+\boldsymbol{\varepsilon_t}$ (3) the VARMAX(1,1) Model could be written as follows:

$$
\begin{aligned}
[\Delta Y_t] = {} & \begin{bmatrix} -3.3 \\ -2.3 \\ 1.5 \end{bmatrix} + \begin{bmatrix} -0.3 & 0.1 & -0.3 \\ 0.3 & -0.02 & -0.3 \\ 0.3 & 0.2 & -0.1 \end{bmatrix} [\Delta Y_{t-1}] \\
& + \begin{bmatrix} -0.2 & 0.05 & 0.1 \\ 0.05 & -0.6 & 0.1 \\ -0.3 & -0.1 & -0.5 \end{bmatrix} [\varepsilon_{t-1}] + [\varepsilon_t] \\
& + \begin{bmatrix} -0.3 & 0.02 & 0.1 & 0.2 & 0.1 & 0.4 \\ -0.2 & -0.03 & 0.2 & 0.1 & -0.2 & 0.6 \\ 0.2 & -0.004 & -0.1 & -0.1 & -0.04 & 0.03 \end{bmatrix} [X_{t-1}]
\end{aligned}
$$

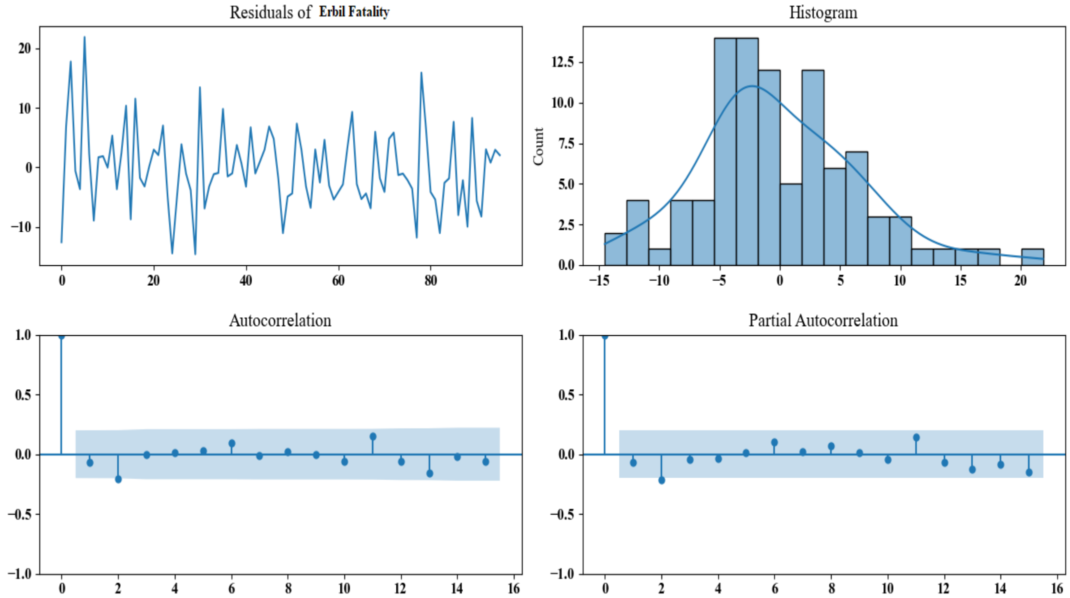
After the VARMAX(1,1) model has been selected and the parameters estimated, the diagnostic checking of the model was tested. For this purpose, the Ljung-Box (Q) Statistics Test was used for checking the autocorrelation in the residuals up to 10 lags. For all three governorates (Erbil, Sulaymaniyah, and Duhok), the P_ values (0.78, 0.32, and 0.43 respectively) were greater than the 0.05 significance level (see **Table (**6). Therefore, the null hypothesis cannot be rejected, meaning that residuals are not serially correlated. This emphasizes that the residuals of the VARMAX(1,1) model adequately capture the temporal dynamic system.

**Table (6): Ljung-Box diagnostic for VARMAX(1,1) model of the traffic accident fatalities in Erbil, Sulaymaniyah, and Duhok.**

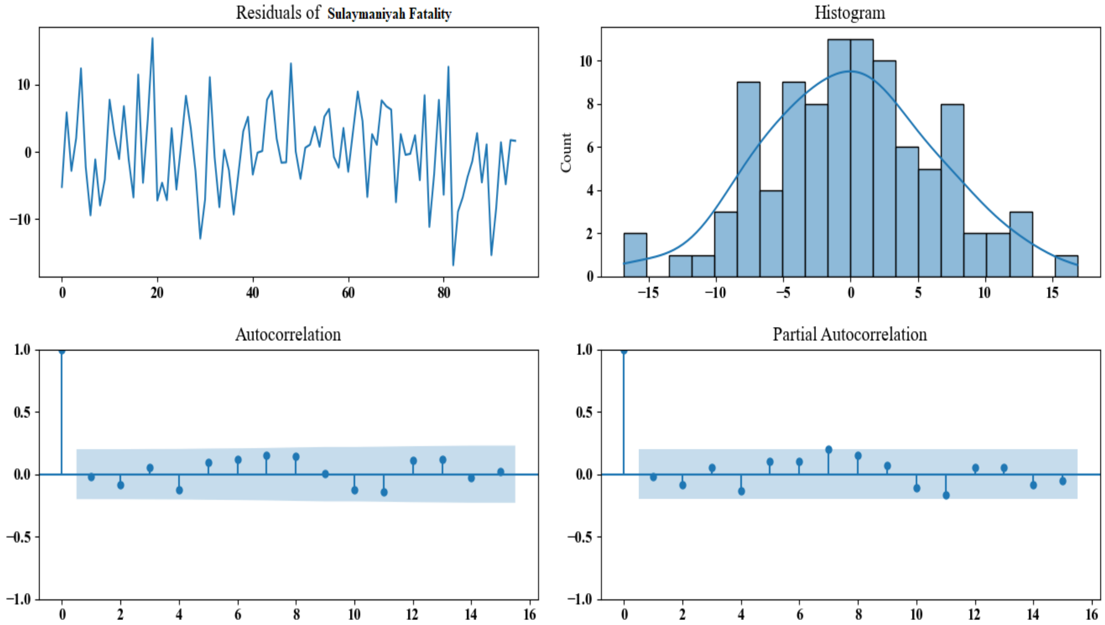
| **VARMAX(1,1) Model** | **Ljung–Box statistics** | | |
|---|---|---|---|
| | d.f | Q* | p_value |
| **Erbil** | 10 | 6.385 | 0.782 |
| **Sulaymaniyah** | 10 | 11.466 | 0.322 |
| **Duhok** | 10 | 10.126 | 0.430 |
| $H_0: \rho_1 = \rho_2 = \cdots = \rho_m = 0$ (There is no serial correlation)<br>$H_1: \rho_1 \neq \rho_2 \neq \cdots \neq \rho_m \neq 0$ (There is serial correlation) | | | |

The results of the diagnosis are a good model fit and confirm that the proposed VARMAX(1,1) model has been successfully capturing monthly traffic fatality dynamics in the three governorates. The residual plots are depicted in **Figure (4** , which have a random pattern around zero with a constant variance.

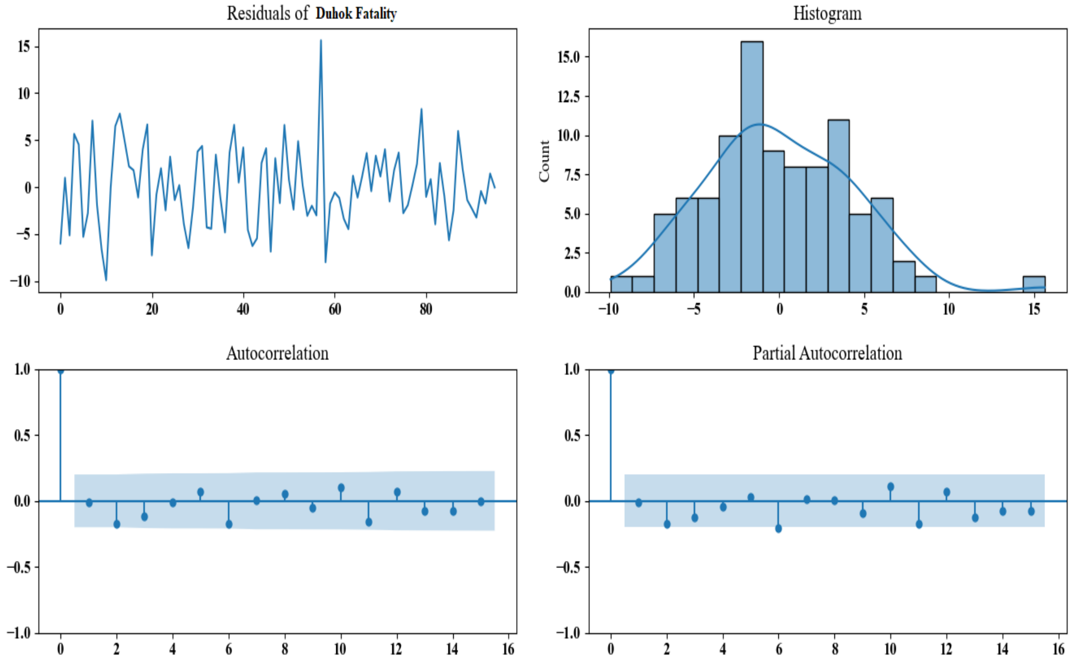
**DOI:** https://doi.org/10.36325/ghjec.v22i1.20796



***a)* Diagnostics of Erbil Fatality**



**B) Diagnostics of Sulaymaniyah Fatality**



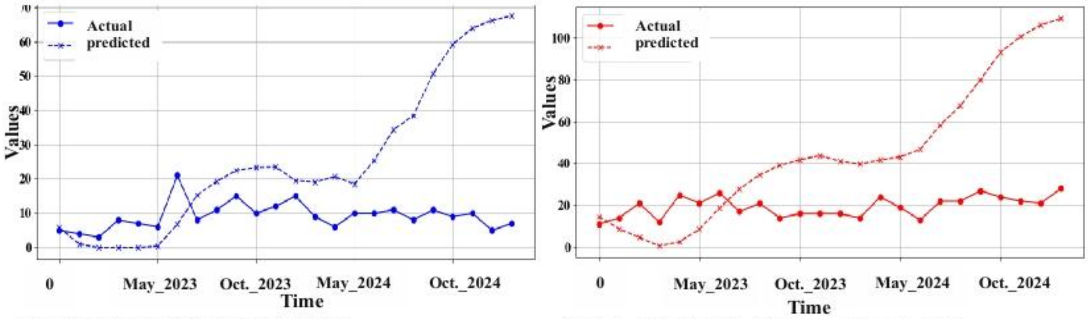
*c)* **Diagnostics of Duhok Fatality**

**Figure (4)** Diagnostic plots related to the normality, residual plots with Ljung–Box test, ACF, and PACF for the VARMAX(1,1) model of monthly traffic fatalities in a) Erbil, b) Sulaymaniyah, and c) Duhok

### 4.3.2.Predicting and Forecasting Traffic Accident Fatalities

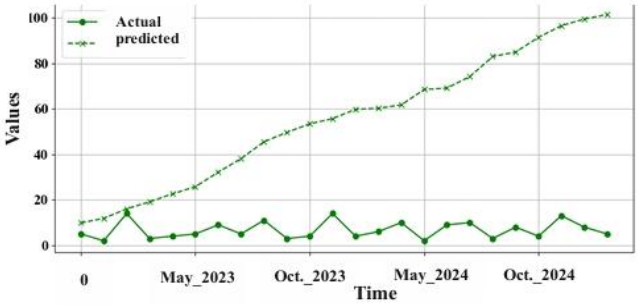
After model diagnostics were applied on the VARMAX (1,1) model using 80% of the dataset, the testing data of the remaining 20% of the dataset is utilized for the prediction process of the 24-month value (2 years) for the number of traffic accident fatalities in KRI. **Figure (**5 ، shows that the VARMAX (1,1) fits very well with real Traffic data. In Sulaymaniyah, accident fatalities constantly display the highest peaks among the three governorates, particularly around December 2024 and August 2024, the high-pitched rise and consequent fall in fatalities to a cyclical pattern, possibly linked to exogenous factors. Erbil’s trend shows fewer sharp increases, indicating qualified stability compared to Sulaymaniyah.

Predicted accident fatalities appear as moderate peaks, match closely with Sulaymaniyah during specific intervals, such as August 2023 and April 2024. Duhok has the lowest predicted accident fatality values all over the period, while the predicted values display minor fluctuations, the general trend is stable with no major spikes.



(a) Actual Vs Predicted Erbil Fatalities

(b) Actual Vs Predicted Sulaymaniyah Fatalities



(c) Actual Vs Predicted Duhok Fatalities

**Figure (5)** Actual vs. Predict plots for monthly traffic accident fatalities for the period 2023 and 2024 in a) Erbil, b) Sulaymaniyah, and c) Duhok

Depending on the VARMAX(1,1) model and using 20% testing data (last 24 months of the traffic accident fatalities), the MAE, MSE, RMSE, and MAPE accuracy metrics of the model are equal to 8.89, 121.59, 10.01, and 76.85% respectively.

The forecast of the traffic accident fatalities in the KRI governorates is presented Table (7. It appears that the accident fatalities in Erbil governorate will increase in the next two years, while in Sulaymaniyah governorate, it will be higher than in Erbil and Duhok. The forecasting data in Duhok governorate compared to the past few years. The historical data reveals a general decline in fatalities, showing a clear rising trend in the number of accident fatalities forecasts across all governorates. Sulaymaniyah, in particular, needs immediate and focused road safety improvements. Erbil also shows modest but persistent growth, although Duhok remains the last impacted. The growing trend calls for early action to avoid future escalation.

**Table (7): Forecasting Number of accident fatalities in KRI's governorates for 24 months using VARMAX(1,1) model**

| 2025 | | | | 2026 | | | |
|---|---|---|---|---|---|---|---|
| **Time period** | **Erbil** | **Sulaymaniyah** | **Duhok** | **Time period** | **Erbil** | **Sulaymaniyah** | **Duhok** |
| Jan. | 7.362 | 27.983 | 6.745 | Jan. | 16.661 | 42.390 | 11.200 |
| Feb. | 7.536 | 27.499 | 6.636 | Feb. | 16.963 | 42.598 | 11.550 |
| Mar. | 8.600 | 28.145 | 6.616 | Mar. | 18.029 | 43.995 | 11.743 |
| Apr. | 9.363 | 29.606 | 6.856 | Apr. | 18.671 | 45.388 | 12.144 |
| May. | 10.198 | 31.058 | 7.328 | May. | 19.666 | 47.432 | 12.567 |
| Jun. | 12.220 | 33.392 | 7.663 | Jun. | 21.465 | 50.012 | 12.895 |
| Jul. | 13.230 | 35.619 | 8.324 | Jul. | 22.583 | 52.618 | 13.485 |
| Aug. | 14.589 | 37.691 | 8.960 | Aug. | 23.692 | 54.679 | 14.030 |
| Sep. | 15.577 | 39.034 | 9.494 | Sep. | 24.664 | 56.200 | 14.486 |
| Oct. | 16.028 | 40.095 | 10.001 | Oct. | 25.026 | 57.514 | 14.978 |
| Nov. | 16.629 | 40.925 | 10.215 | Nov. | 25.698 | 58.910 | 15.337 |
| Dec. | 16.826 | 41.428 | 10.411 | Dec. | 25.874 | 60.163 | 15.792 |

## 5. Conclusions

This article aimed to identify the best multivariate dynamic models using VARMAX models with assisting exogenous variables for forecasting Traffic Accident Fatalities. The best model demonstrated with the highest forecasting accuracy is the VARMAX (1,1). The exogenous variables particularly driving speed and temperature, significantly enhanced the model's predictive power to improve model performance.

Forecasts for 24 months for the periods 2025 and 2026 were generated, revealing a remarkable upward traffic accident fatality across the region. The process of splitting the time series data into training and testing datasets (80% and 20% respectively) proved valuable in identifying overfitting and underfitting of the forecasting results and led to best reliable forecasting model.

These findings are relevant to a variety of stakeholders, such as traffic authorities, urban planners, law enforcement organizations, and policymakers, rely heavily on these estimates. To mitigate the forecast and rise in fatalities, the results provide data-driven evidence in favor of the development of road safety measures, including speed limits and public awareness initiatives.

Future studies could improve robustness by integrating a wider range of explanatory variables, encompassing behavioral, demographic, and meteorological aspects associated with traffic accidents, such as mobile phone usage while driving, driving under the influence of alcohol or drugs, and drivers' age, fog level, rainfall, etc. Furthermore, comparing multivariate dynamic models with machine learning methods could result in better forecasting improvements. Such extensions would further increase model robustness and provide deeper insight into the underlying causes of road traffic fatalities.

## References

1) Abdulqader, Q., Tareq, H. M. & Hasan, A. K. (2020). *Building a Mathematical SARIMA Model for Forecasting the Number of Monthly Injured People by Traffic Accidents in Erbil City*. 62. https://www.researchgate.net/publication/341099823

DOI: https://doi.org/10.36325/ghjec.v22i1.20796

2) Ahmed, M. M. & Mahmood, S. H. (2023). Building SARIMA Models to Analysis and Forecast Time Series Data of Road Traffic Accidents in the Kurdistan Region of Iraq. *Zanco Journal of Humanity Sciences*, *27*(1). https://doi.org/10.21271/zjhs.27.1.23

3) Ahmed, N. M. , & Abdulqader, Q. M. (2023). Enhancing the ARIMAX model by using the bivariate wavelet denoising: Application on road traffic accidents. *Iraqi Journal Of Statistical Sciences*, *20*(2), 15-29. https://doi.org/10.33899/iqjoss.2023.0181146

4) Ali, taha, Hayawi, H. & Hamza, H. (2025). Bayesian Time Series Modelling with Wavelet Analysis for Forecasting Monthly Inflation. *Iraqi Journal Of Statistical Sciences*, *22*(1), 181-194. https://doi.org/10.33899/iqjoss.2025.187792

5) Andayani, N. , Sumertajaya, I. M., Ruchjana, B. N. & Aidi, M. N. (2015). *Rice Price Modeling In Six Province Of Java Island Using VARMAX Model*. 17-19. nurita.andayani@gmail.com, Yogyakarta State University, 17-19 May 2015

6) Boddu, Y., Manimaran A, Arunkumar B & Ramkumar D. (2017). *Design of an Iterative Dual Metaheuristic VARMAx Model Enhancing Efficiency of Time Series Predictions. VOLUME XX, 2017.* https://doi.org/DOI10.1109/ACCESS.2024.3454540

7) Box, G. E. P. & Jenkins, G. M. J. (1976). *Time Series Analysis: Forecasting and Control*. Holden-day, Oakland California. Library of Congress Catalog card number 768713. Printed in the United State of Amrica.

8) Box, G. E. P., JENKINS, G. M., REINSEL, G. C. & LJUNG, G. M. (2016). *Time Series Analysis: Forecasting and Control* (Fifth Edition). Published by John Wiley & Sons, Inc., Hoboken, New Jersey. Published simultaneously in Canada. Printed in the United States of America.

9) Box, George., Reinsel, Gregory. & Jenkins, Gwilym. (2008). *Time Series Analysis: Forecasting and Control* (Fourth Edition). John Wiley & Sons, INC., publication, Printed in the United States of America. Published by John Wiley & Sons, Inc., Hoboken, New Jersey. Published simultaneously in Canada.

10) Chatfield, Chris. (1995). *Chatfield_The_Analysis_of_Time_Series__An_Introduction_5ed*. Email: cc@aths.bath.ac.uk

11) Chatfield, Christopher. (2001). *Time-series forecasting*. Chapman & Hall/CRC.

12) Chukwutoo, C. I. & Onwurah, U. O. ,. (2018). Road traffic accidents prediction modelling: An analysis of Anambra State, Nigeria. *Accident Analysis and Prevention*, *112*, 21-29. https://doi.org/10.1016/j.aap.2017.12.016

13) Cryer, J. D. & Chan, K.-S. (2008). *Springer Texts in Statistics: Time Series Analysis With Applications in R* (Second Edition). (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA. Printed on acid-free paper.

14) Feng Feng & Liping Fu, Ph. D. , P. E. (2015). Winter Road Surface Condition Forecasting. *Journal of Infrastructure Systems*, *21*(3). https://doi.org/10.1061/(asce)is.1943-555x.0000241

15) Hamilton, J. Douglas. (1994). *Time Series Analysis James D. Hamilton*. Princeton University Press.

16) Haydier, E. , Albarwari, N. H. & Ali, T. H. (2023). The Comparison Between VAR and ARIMAX Time Series Models in Forecasting. *Iraqi Journal Of Statistical Sciences*, *20*(2), 249-262. https://doi.org/10.33899/iqjoss.2023.181260

17) Helmut Lütkepohl. (2005). *New Introduction to Multiple Time Series Analysis*. Library of Congress Control Number: 2005927322.Springer Berlin Heidelberg New York, Department of Economics European University Institute

DOI: https://doi.org/10.36325/ghjec.v22i1.20796

Villa San Paolo Via della Piazzola 43 .50133 Firenze Italy. E-mail: helmut.luetkepohl@iue.it

18) Mahmood, N. H., Kadir, D. H. & Alzawbae, O. M. M. e. (2024). Building a Statistical Model to Forecast Traffic Accidents for Death and Injuries by Using Bivariate Time Series Analysis. *Zanco Journal of Humanity Sciences*, *28*(1). https://doi.org/10.21271/zjhs.28.1.18

19) Ostermark, R., Saxen, H. & Henrik. (1996). VARMAX-modelling of blast furnace process variables. European Journal of Operational Research. *European Journal of Operational Research*, *90*, 85-101.

20) Parvareh, M., Karimi, A., Rezaei, S., Woldemichael, A., Nili, S., Nouri, B. & Nasab, N. E. (2018). Assessment and prediction of road accident injuries trend using time-series models in Kurdistan. *Burns & Trauma*, *6*. https://doi.org/10.1186/s41038-018-0111-6

21) Peng, B., Meng, Y., Shi, D., Dai, M., Zhou, H. & Chen, T. (2021). Forecast of the trend of heavy fog based on the VARIMAX model. *E3S Web of Conferences*, *257*. https://doi.org/10.1051/e3sconf/202125703013

22) Rojas, I., Pomares, H. & Valenzuela, O. (2017). *Contributions to Statistics Time Series Analysis and Forecasting Selected Contributions from ITISE 2017*. http://www.springer.com/series/2912

23) Shumway, Robert, H. & Stoffer, D. S. (2005). *Time Series Analysis and Its Applications with R Examples, Springer Texts in Statistics* (George Casella, Stephen Fienberg, & Ingram Olkin, Éd.; Second Edition). Springer Science & Business Media, LLC, 233 Spring Street, New York, NY 10013, USA, Printed in the United States of America. (MVY).

24) Shumway, Robert H. & Stoffer, D. S. S. (2011). *Time Series Analysis and Its Applications With R Examples. Third edition*. www.springer.com/series/417

25) Tibshirani, R. (2023). *Lecture 6: Autoregressive Integrated Moving Average Models Introduction to Time Series, Fall 2023 Ryan Tibshirani*.

26) Tsay, R. S. (2005). *Analysis of Financial Time Series Second Edition* (Second Edition). Published by John Wiley & Sons, Inc., Hoboken, New Jersey. Published simultaneously in Canada. gsbwww.uchicago.edu/fac/ruey.tsay/teaching/fts2

27) Tsay, R. S. (2014). *Multivariate Time Series Analysis*. Published by John Wiley & Sons, Inc., Hoboken, New Jersey Published simultaneously in Canada, Printed in the United States of America. www.wiley.com.

28) Tsirigotis, L., Vlahogianni, E. I. & Karlaftis, M. G. (2011). Does Information on Weather Affect the Performance of Short-Term Traffic Forecasting Models? *International Journal of Intelligent Transportation Systems Research*, *10*(1), 1-10. https://doi.org/10.1007/s13177-011-0037-x

29) Warsono, Russel, E., Wamiliana, Widiarti & Usman, M. (2019). Vector autoregressive with exogenous variable model and its application in modeling and forecasting energy data: Case study of PTBA and HRUM energy. *International Journal of Energy Economics and Policy*, *9*(2), 390-398. https://doi.org/10.32479/ijeep.7223

30) Wei, W. W. S. (2006). *Time series analysis: univariate and multivariate methods. (2nd ed.)*. Greg Tobin, Pearson Addison Wesley, Library of Congress Cataloging-in-Publication Data.

31) Wei, W. W. S. (2019). *Multivariate Time Series Analysis and Applications* (David J. Balding, Noel A. C. Cressie, Garrett M. Fitzmaurice, Geof H. Givens, Harvey Goldstein, Geert Molenberghs, David W. Scott, Adrian F. M. Smith, & Ruey S. Tsay, Éd.). John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK, Set in 10/12pt Times by SPi Global, Pondicherry, India. http://www.wiley.com/go/wsps

32) Zivot, E. & Wang, J. (2006). *Modelling Financial Time Series with S-PLUS, Second Edition*. http://faculty.washington.edu/ezivot/ModelingFinancialTimeSeries.htm

**DOI:** https://doi.org/10.36325/ghjec.v22i1.20796