INTERNATIONAL JOURNAL OF PROFESSIONAL BUSINESS REVIEW
ISSN: 2525-3654
OPEN ACCESS

VECTOR AUTOREGRESSIVE INTEGRATING MOVING AVERAGE (VARIMA) MODEL OF COVID-19 PANDEMIC AND OIL PRICE

Aras Jalal Mhamad Karim[A], **Nawzad Muhammed Ahmed**[B]

ARTICLE INFO

Article history:

Received 21 November 2022

Accepted 30 January 2023

Keywords:

Long-Term;
VARMA Model;
Multivariate Time Series Model.

ABSTRACT

Purpose: A coronavirus associated with severe respiratory syndrome has created Coronavirus Disease 2019 (COVID-19), a highly contagious illness that affects the entire world population. On the other hand, COVID-19 is having a direct impact on human life because of its proliferation. So, the study's goal is to forecast and analyze the impact of the COVID-19 pandemic and the oil price utilizing multiple time series analysis methods (VARIMA model).

Theoretical framework: Recent literature has reported that the multivariate time series is robust model for forecasting and analyzing dynamic relationship between series, while the univariate ARIMA model has been generalized to include vector variables, that is an extension of its capabilities. The VAR (p) model analyzes the interdependence between two or more series but does not take into account the impact of shocks at various time variable delays.

Design/methodology/approach: This study uses VARMA (p, q) model which links a set of variables to their prior iterations as well as those of other variables and shocks to those same variables. Sample data concerning the COVID-19 pandemic and oil price was globally provided. It contains daily observations of them variables for the years 2020-2022.

Findings: The best model is VARIMA (2,1,2), and the results shown that the oil price is not only influenced by itself but also influenced by the Covid-19 pandemic. Moreover, the standard error grows over time of the forecast.

Research, Practical & Social implications: The best model is sound for short-term forecasting but unstable for long-term forecasting. Future researchers can integrate factors across areas. Include tourism demand and industry variables in modeling.

Originality/value: Collecting COVID-19 pandemic data and oil price series in a modern model that is a multivariate time series model with a high predicted level of model accuracy between these variables in order to predict and analyze the effects between them series and estimate the interaction between these two series with the most recent data is the value of this study, and then offers merchants the chance to comprehend the forecasting of oil price throughout the covid-19 effects as well as the associated risks.

Doi: https://doi.org/10.26668/businessreview/2023.v8i1.988

[A] *Assistant Professor, Department of Statistics and Informatics, College of Administration and Economics, Suleimani University, Suleimani, Kurdistan, Iraq. E-mail: aras.mhamad@univsul.edu.iq*
Orcid: https://orcid.org/0000-0003-0924-0230
[B] *Professor, Department of Statistics and Informatics, College of Administration and Economics, Suleimani University, Suleimani, Kurdistan, Iraq. E-mail: nawzad.mahmud@univsul.edu.iq*
Orcid: https://orcid.org/0000-0003-4878-0168

MODELO AUTOREGRESSIVO VETORIAL INTEGRANDO A MÉDIA MÓVEL (VARIMA) DA PANDEMIA COVID-19 E O PREÇO DO PETRÓLEO

RESUMO

Objetivo: Um coronavírus associado à síndrome respiratória grave criou o Coronavirus Disease 2019 (COVID-19), uma doença altamente contagiosa que afeta toda a população mundial. Por outro lado, a COVID-19 está tendo um impacto direto na vida humana por causa de sua proliferação. Portanto, o objetivo do estudo é prever e analisar o impacto da pandemia da COVID-19 e o preço do petróleo utilizando múltiplos métodos de análise de séries temporais (modelo VARIMA).

Estrutura teórica: A literatura recente relatou que a série temporal multivariada é um modelo robusto para prever e analisar a relação dinâmica entre as séries, enquanto o modelo ARIMA univariado foi generalizado para incluir variáveis vetoriais, ou seja, uma extensão de suas capacidades. O modelo VAR (p) analisa a interdependência entre duas ou mais séries, mas não leva em conta o impacto de choques em vários atrasos de variáveis temporais.

Projeto/método/abordagem: Este estudo utiliza o modelo VARMA (p, q) que liga um conjunto de variáveis a suas iterações anteriores, bem como as de outras variáveis e choques a essas mesmas variáveis. Foram fornecidos dados de amostra relativos à pandemia COVID-19 e ao preço do petróleo a nível mundial. Ele contém observações diárias dessas variáveis para os anos de 2020-2022.

Conclusões: O melhor modelo é o VARIMA (2,1,2), e os resultados mostraram que o preço do petróleo não é influenciado apenas por si mesmo, mas também influenciado pela pandemia da COVID-19. Além disso, o erro padrão cresce ao longo do tempo da previsão.

Pesquisa, implicações práticas e sociais: O melhor modelo é sólido para previsões de curto prazo, mas instável para previsões de longo prazo. Os futuros pesquisadores podem integrar fatores entre as áreas. Incluir a demanda turística e as variáveis da indústria na modelagem.

Originalidade/valor: A coleta de dados pandêmicos COVID-19 e séries de preços de petróleo em um modelo moderno que é um modelo de séries temporais multivariadas com um alto nível de precisão prevista entre essas variáveis, a fim de prever e analisar os efeitos entre essas séries e estimar a interação entre essas duas séries com os dados mais recentes é o valor deste estudo, e então oferece aos comerciantes a chance de compreender a previsão do preço do petróleo ao longo dos efeitos covid-19, bem como os riscos associados.

Palavras-chave: Longo Prazo, Modelo VARMA, Modelo de Série Temporal Multivariada.

MODELO VECTORIAL AUTORREGRESIVO INTEGRADOR DE MEDIAS MÓVILES (VARIMA) DE LA PANDEMIA DE COVID-19 Y EL PRECIO DEL PETRÓLEO

RESUMEN

Propósito: Un coronavirus asociado al síndrome respiratorio grave ha creado la Enfermedad por Coronavirus 2019 (COVID-19), una enfermedad altamente contagiosa que afecta a toda la población mundial. Por otra parte, el COVID-19 está teniendo un impacto directo en la vida humana debido a su proliferación. Así, el objetivo del estudio es pronosticar y analizar el impacto de la pandemia COVID-19 y el precio del petróleo utilizando métodos de análisis de series temporales múltiples (modelo VARIMA).

Marco teórico: La literatura reciente ha reportado que las series de tiempo multivariadas es un modelo robusto para pronosticar y analizar la relación dinámica entre series, mientras que el modelo ARIMA univariado ha sido generalizado para incluir variables vectoriales, que es una extensión de sus capacidades. El modelo VAR (p) analiza la interdependencia entre dos o más series, pero no tiene en cuenta el impacto de las perturbaciones en los distintos retardos de las variables temporales.

Diseño/metodología/enfoque: Este estudio utiliza el modelo VARMA (p, q) que relaciona un conjunto de variables con sus iteraciones anteriores, así como las de otras variables y las perturbaciones de esas mismas variables. Se proporcionaron globalmente datos muestrales relativos a la pandemia COVID-19 y al precio del petróleo. Contiene observaciones diarias de dichas variables para los años 2020-2022.

Resultados: El mejor modelo es VARIMA (2,1,2), y los resultados muestran que el precio del petróleo no sólo está influido por sí mismo, sino también por la pandemia Covid-19. Además, el error estándar crece con el tiempo. Además, el error estándar crece con el tiempo de la previsión.

Investigación, implicaciones prácticas y sociales: El mejor modelo es sólido para las previsiones a corto plazo, pero inestable para las previsiones a largo plazo. Los futuros investigadores pueden integrar factores de distintos ámbitos. Incluir variables de la demanda turística y de la industria en la modelización.

Originalidad/valor: Reunir los datos de la pandemia COVID-19 y las series del precio del petróleo en un modelo moderno que es un modelo multivariante de series temporales con un alto nivel de precisión del modelo previsto entre estas variables para predecir y analizar los efectos entre ambas series y estimar la interacción entre estas dos series con los datos más recientes es el valor de este estudio, y ofrece a los comerciantes la posibilidad de comprender la previsión del precio del petróleo a lo largo de los efectos de la covid-19, así como los riesgos asociados.

Palabras clave: Largo Plazo, Modelo VARMA, Modelo Multivariante de Series Temporales.

INTRODUCTION

Beginning with the emergence of COVID-19, which poses a direct threat to humanity worldwide, this predicament is also attributable to the lockdown measures enforced by the impacted nations. The entire or partial shutdown enacted by nations across the globe halted the majority of commercial operations until the outbreak was contained (Menhat et al., 2021). Although these regulations contributed to preventing the rapid spread of the COVID-19 pandemic, they wreaked havoc on the world economy and precipitated the worst recession since the Great Depression (Menhat et al., 2021). At the same time, the COVID-19 epidemic caused significant panic and pressure on the financial markets (Akhtaruzzaman et al., 2021; Benkraiem et al., 2018; Chang & Zhang, 2016). As the stock market is a barometer of economic progress (Benkraiem et al., 2018), investors' worry intensified throughout the epidemic (Akhtaruzzaman et al., 2021) as the stock market is a barometer of economic development (Benkraiem et al., 2018). So, whether the COVID-19 epidemic has no effect on global oil markets. therefore, the purpose of the study is to forecast and assess the influence of the COVID-19 pandemic and oil price utilizing multiple time series analysis, namely the vector autoregressive integrating moving average (VARIMA) model. This study employs the theoretical framework of the VARIMA, therefore we pose three questions: (1) How can the relationship between the COVID-19 epidemic and oil price series be predicted? (2) Do COVID-19 epidemic effect oil price? (3) How consistent and robust are the various classical model approaches for evaluating the effects from questions (1) and (2)? To identifying the causes of the COVID-19 pandemic's effects on oil prices, the researchers used a VARIMA model to determine the relationship between the two variables.

Multiple time analysis was created to examine two or more simultaneous series of analyses; multiple analyses enable: (i) a shared understanding of a dynamic relationship between individuals. (ii) Establishing the causality relationship within sets. (iii) Determining the relationships between the sets. An important step in processing of VARMA (p, q) model is estimating the parameters. This form of estimation is usually done using least squares or quasi-maximum likelihood approaches. The complexity of model selection, which entails choosing an appropriate model from a group of potential models to characterize the available data, complements the challenge of parameter estimation. The quantity of parameters makes the selection of p and q crucial, $(p + q + 3) d^2$, where d is the number of the series, grew dramatically with p and q, causing statistical difficulties. The estimation of the parameters will be inconsistent if orders below the true orders of the VARMA (p, q) models are chosen, and it is likely to be inaccurate if orders above the true orders are chosen (Mainassara, 2010; Nasser et al., 2009; Rahim et al., 2020).

LITERATURE REVIEW

The VARMA models (vector autoregressive moving average) are the consequence of the wold decomposition theorem for multivariate stationary series, as demonstrated by (Athanasopoulos & Vahid, 2007). (Athanasopoulos & Vahid, 2008; Kascha & Trenkler, 2011) investigated the accuracy of forecasts generated using VARMA models, obtaining a positive result. This method, which is an extension of the Box and Jenkins (1970) method for constructing univariate forecasting models, and the model will be constructed using the construction methodology of the Tiao and Box (1981) multivariate time series model, are used when long lags of each variable are required. Feunou & Fontaine, 2009) represented the yield curve with a VARMA model, removing the limits on co-integration. Dufour & Pelletier, 2022) have presented a modified information criterion for determining the VARMA orders, which is merely a modification of the criterion developed by Hannan and Rissanen (1982). As demonstrated by (Mainassara, 2010), selecting a VARMA (p, q) order that is too small results in inconsistent estimators, whereas a VARMA (p, q) order that is too large reduces the accuracy of forecasts. (Kascha & Trenkler, 2011) expand the representations of Dufour and Pelletier (2008) that were valid for non-stationary series, beginning with the most recent notable results linked to VARMA models and presenting an approach for specification and estimate of co-integrated series (Menhat et al., 2021).

Recent years have seen a gradual shift from theoretical modeling to empirical research on the COVID-19 pandemic effects on oil prices. In order to achieve the predicted level of model accuracy, the use of correlated high-dimensional time series data presents numerous complexities and challenges. In order to anticipate and assess the relationship between the COVID-19 pandemic and oil price, the researchers attempt to develop a theoretical model with the best expected level of model accuracy, namely the VARIMA model. Consequently, our contribution consists of collecting COVID-19 pandemic and oil price in a modern model that is a multivariate time series model with a high predicted level of model accuracy between these variables in order to predict and analyze the effects between them series and estimate the interaction between these two series with the most recent data.

MATERIAL AND METHODOLOGY

VARMA Model

To analyze the interrelationship between two or more series, can be use The VAR (p) model, however it is neither especially sparse nor does it take into account the impact of shocks at various lags time. The VARMA (p, q) model is relating a set of series Z_t to earlier iterations of both its own and other series as well as earlier iterations of shocks to both of these series. The following below illustrates the VARMA model of autoregressive and moving average:

$$Z_t = C_t + \Phi_1 Z_{t-1} + \cdots + \Phi_p Z_{t-p} + \mathrm{a}_t + \Theta_1 a_{t-1} + \cdots + \Theta_q a_{t-q} \qquad \ldots\ldots\ldots\ldots\ldots (2.1)$$

For all t greater than the initial time origin, the model (2.1) is valid. Z_t is a vector of (k x 1), (Φ_1 ... Φ_p) are matrices of (k x k) autoregressive coefficient, (Θ_1 ... Θ_q) are matrices of (k x k) coefficient of moving average, C_t is a (k x 1) constants vector and a_t is a white noise process vector with zero mean and covariance matrix $E(\mathrm{a}_t \acute{\mathrm{a}}_t) = \Sigma_a$. In backshift/lag order, the model (2.1) can be stated as:

$$\left(I - \Phi_1 L - \cdots - \Phi_p L^p\right)\mathrm{Z}_t = \mathrm{C}_t + \left(I + \Theta_1 L + \cdots + \Theta_q L^q\right)\mathrm{a}_t \qquad \ldots\ldots\ldots\ldots\ldots (2.2)$$

$$\text{Or } \ \Phi(L)\mathrm{Z}_t = \mathrm{C}_t + \Theta(L)\mathrm{a}_t$$

The diagonal components of these matrices, $\Phi(L)$ and $\Theta(L)$, are known as the autoregressive and moving average structures in each series, whilst the off-diagonal portions of these matrices describe the causal relationships between the various pairings of series and shocks. Each component of Z_t describes how the current value of a particular series compares to both its own prior values and the prior values of the other series. If each series is independent of the others, each series can be represented by a different ARMA model since $\Phi(L)$, $\Theta(L)$, and $\Sigma_a$will all be diagonal matrices (Montgomery & Moe, 2002).

Assume there is a bivariate series with k = 2. The prior values of the second series Z_2 have an impact on the first series Z_1 but have no impact on the prior values of Z_1 and Z_2 if both matrices $\Phi(L)$ and $\Theta(L)$are upper triangular. A unidirectional relation was observed between the series if both matrices $\Phi(L)$ and $\Theta(L)$ are lower triangular. A dynamic relation exists between the series if the matrices $\Phi(L)$, $\Theta(L)$, and Σ_a are all fully filled. Both invertible and stationary versions of the VARMA (p, q) process are possible. It is invertible if all of the roots of the determinantal polynomial $|\Theta(L)|$ are outside the unit circle, but constant if none of the roots are outside the unit circle. It is possible to express the VARMA (p, q) process as:

$$\Pi(L)Z_t = \left(I - \Theta_1 L + \cdots + \Theta_q L^q\right)^{-1} C_t + a_t \qquad \text{............... (2.3)}$$

where Π_i are (k x k) coefficients matrices which are obtained from the autoregressive relation:

$$I - \sum_{i=1}^{\infty} \Pi_i L^i = [\Theta(L)]^{-1}\Phi(L) \qquad \text{............... (2.4)}$$

If equation (2.4) multiplies by $\Theta(L)$ from the left by, then:

$$\left(I + \Theta_1 L + \cdots + \Theta_q L^q\right)\left(I - \sum_{i=1}^{\infty} \Pi_i L^i\right) = \Phi(L) = I - \Phi_1 L - \cdots - \Phi_p L^p \text{ with } \Theta_i = 0 \text{ for } i > q$$

Thus, when comparing lag coefficients, the Π_i values are obtained. In general, the Π_i value is:

$$\Pi_i = \Phi_i + \Theta_i - \sum_{j=1}^{i} \Theta_{i-j} \Pi_j, \quad i = 1, 2, \ldots \qquad \text{............... (2.5)}$$

Wei (2006) noted that since $\frac{1}{|\Theta(L)|}\Theta + (L)\Phi(L)Z_t = a_t + (\Theta + (L))$ is the adjoint matrix), a finite VAR ($\ddot{p}$) model can be used to represent the process if the determinantal polynomial $\Theta(L)$ is independent of L, with $\ddot{p} \leq (k-1)q$ (Wei, 2006).

When the process of VARMA (p, q) is stationary, then it could be written as:

$$\begin{aligned} Z_t &= \Phi(1)^{-1} C_t + \Lambda(\mathrm{L}) a_t \\ &= \left(I - \Phi_1 - \cdots - \Phi_p\right)^{-1} C_t + \Lambda(\mathrm{L}) a_t \qquad \text{............... (2.6)} \\ &= \mu + \sum_{i=0}^{\infty} \Lambda_i a_{t-i} \end{aligned}$$

The matrices of Λ_i , $i = 1, 2, \ldots$ are obtained from the moving average relation

$$[\Phi(L)]^{-1}\Theta(L) = \sum_{i=1}^{\infty} \Lambda_i L^i \qquad \ldots\ldots\ldots\ldots\ldots (2.7)$$

When (2.7) multiplies by $\Phi(L)$ from the left, then:

$$\Theta(L) = \Phi(L)(\sum_{i=1}^{\infty} \Lambda_i L^i)$$

By comparing lag coefficients, the Λ_i values are obtained. In general, the Λ_i value is

$$\Lambda_i = \Theta_i + \sum_{j=1}^{i} \Phi_j \Lambda_{i-j}, \quad i = 1, 2, \ldots \text{ with } \Phi_j = 0 \text{ for } j > p \text{ and } \Theta_i = 0 \text{ for } i > q \quad \ldots.(2.8)$$

When the determinantal polynomial $|\Phi(L)|$ is independent of L, Hence the process can be described by a model of VMA ($\ddot{q}$) of order at most q̈ is less than or equal to $(k-1)p$ (Wei, 2006). This is not always the case in multivariate time series models because the inverse $[\Phi(L)]^{-1}$ can be expressed as a determinant and an adjoint matrix as:

$$[\Phi(L)]^{-1} = \frac{1}{|\Phi(L)|}\Phi + (L) \qquad \ldots\ldots\ldots\ldots\ldots (2.9)$$

Given that $|\ \Phi(L)|$ has a constant value and that $\Phi + (L)$ order is a finite order of autoregressive matrix polynomial, the inverse matrix $[\Phi(L)]^{-1}$ will also have a finite order.

SPECIFICATION OF VARMA MODELS

VARMA specification is of utmost importance to appropriately describe the values of p and q, as the number of parameters drastically increases as p and q grow, making statistical analysis challenging. Unlike univariate processes, vector ARMA (p, q) models cannot be specified using a specific approach. It is challenging to determine the ordering of the VARMA (p, q) model from of the autocorrelation function, the partial autocorrelation function, and the cross-correlation function whenever there are more than two time series because there will be a lot of parameters involved (Lütkepohl & Poskitt, 1996), it was highlighted that fitting a number of alternative models may be necessary until a good model is identified because determining the cut offs of these functions can be subjective. When attempting to assess if a VARMA (p, q) model is parsimonious, It is imperative to emphasize that selecting the smallest orders for the autoregressive and moving average parameters is not always appropriate

(Lütkepohl & Poskitt, 1996). There are different methods for specifying VARMA (p, q) models. Methods for expressing the two primary forms of VARMA modeling, the echelon form, the final equations form, and a different method called the Scalar Component Method for describing the VARMA (p, q) model (Ahmed, 2007; Ahmed, 2018). The primary objective of the definition of the final form of autoregressive equations is to determine the orders of p and q in the equation.

$$\varphi(L)\mathrm{Z}_t = \Theta(L)\mathrm{a}_t \qquad \text{......................... (2.10)}$$

It is assumed that the mean μ has been eliminated before the specification of this stage. $\mathrm{Z}_t = (\mathrm{Z}_{1t}, \dots, \mathrm{Z}_{kt})$ is a K dimensional system, $\varphi(L) = I - \varphi_1 L - \cdots - \varphi_p L^p$ is a one-dimensional scalar operator and

$$\Theta(L) = I + \Theta_1 L + \cdots + \Theta_q L^q$$

This formula (2.10) suggests that each component is represented by a univariate ARIMA model $\varphi(L)\mathrm{Z}_{kt} = \overline{\theta_k(L)}\mathrm{v}_{kt} \;\; k = 1, \dots, k$

Where v_{kt} is single white noise and $\overline{\theta_k(L)}$ is an operation with a greatest level of q. This means that the same autoregressive operative will apply to each component series Z_{it} , $i = 1, \dots, k$, meanwhile the degree of the moving average operation will not exceed q. Therefore, in order to provide the final equations representation, it is necessary to first specify each of the univariate models $\varphi_k(L)\mathrm{Z}_{kt} = \theta_k(L)\mathrm{v}_{kt}$. $\varphi_k(L)$ and $\theta_k(L)$ operators are defined as follows:

$$\varphi_k(L) = I - \varphi_{k1} L - \cdots - \varphi_{kp_k} L^{pk} \qquad \text{......................... (2.11)}$$

$$\theta_k(L) = I + \theta_{k1} L + \cdots + \theta_{kq_k} L^{qk} \qquad \text{......................... (2.12)}$$

After this, a common $\varphi(L)$ operator must be identified. This is accomplished by multiplying the individual AR polynomials $\varphi(L) = \varphi_1(L), \dots, \varphi_k(L)$ (Ahmed, 2020; Ahmed, 2007). The degree of this operator $\varphi(L)$ is $p = \sum_{i=1}^{k} p_i$. Next, we provide the moving average operator $\overline{\theta_k(L)} = \theta_k(L) \prod_{i=1}^{k} \varphi_i(L)$, $i \neq k$, which has the degree $q_k + \sum_{i=1}^{k} p_i$. The degree of the combined moving average operator is determined by selecting the maximum degree of all individual operators. If $\varphi_k(L)$ contains common factors, then the $\varphi(L)$ degree will be smaller than $\sum_{i=1}^{k} p_i$, whilst the joint moving average operator degree will be smaller than the maximum of $q_k + \sum_{i=1}^{k} p_i$.

The final version of autoregressive equations leads to VARMA models with several parameters. This may lead to imprecise parameter estimates (Ahmed, 2007). It is possible to set restrictions on the autoregressive and moving average operators $\varphi(L)$ and $\Theta(L)$, however doing so will make the modeling process cumbersome. The current study used the Extended Cross Correlation Matrix (ECCM) to determine the model order. The test hypothesis is as follows (Asraa et al., 2018; Rahim et al., 2020):

$$H_0: \rho_{j+1}\left(w_{it}^{(j)}\right) = \cdots = \rho_1\left(w_{it}^{(j)}\right) = 0 \text{ (Order model is nonsignificant)}$$

$$H_1: \rho_v\left(w_{it}^{(j)}\right) \neq 0 \; for \; v > j \text{ (Order model is significant)}$$

Maximum Likelihood Estimation Method of Vector Autoregressive Moving Average Model

The stationary and invertible of VARMA (1,1) expression with a zero mean is the most basic example of the VARMA (p, q) expression. which will be examined first. The estimating process is based on Lütkepohl's (2005) method (Lütkepohl, 2005; Mhamad, 2019; Nasser et al., 2009).

$$Z_t = \Phi_1 Z_{t-1} + a_t + \Theta_1 a_t \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.13)$$

Assuming a sample of size T, $Z_1, \ldots, Z_T$ can be shown as follows:

$$\eta_1 \begin{bmatrix} z_1 \\ \cdot \\ \cdot \\ \cdot \\ z_T \end{bmatrix} + \begin{bmatrix} -\Phi_1 z_0 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix} = \breve{\Theta}_1 \begin{bmatrix} a_0 \\ \cdot \\ \cdot \\ \cdot \\ a_T \end{bmatrix} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.14)$$

$$\text{Where } \eta_1 = \begin{bmatrix} I_k & & \cdots & \cdots & 0 & 0 \\ -\Phi_1 & I_k & \cdots & & 0 & 0 \\ 0 & -\Phi_1 & & & 0 & 0 \\ \vdots & \vdots & & & \vdots & \vdots \\ 0 & 0 & & & 0 & 0 \\ \vdots & 0 & \ddots & \vdots & 0 & 0 \\ \vdots & \vdots & & & \vdots & \vdots \\ 0 & 0 & & & I_k & \\ 0 & 0 & \cdots & 0 & -\Phi_1 & I_k \end{bmatrix}, \text{ and } \breve{\Theta}_1 = \begin{pmatrix} \Theta_1 & I_k & 0 & \cdots & 0 \\ & \Theta_1 & & & \\ & & \ddots & & \vdots \\ 0 & 0 & \cdots & \Theta_1 & I_k \end{pmatrix}$$

Now since $\begin{bmatrix} a_0 \\ \cdot \\ \cdot \\ \cdot \\ a_T \end{bmatrix}$ is a white noise process.

Lütkepohl (2005) noted that the likelihood function is (Lütkepohl, 2005):

$$L(\Phi_1, \Theta_1, \Sigma_a)\alpha |\Sigma_a|^{\frac{-T}{2}} exp\left\{\frac{-1}{2}(\eta_1 \overline{\overline{\Theta_1}}^{-1} Z\acute{)} (I_T \otimes \Sigma_a{}^{-1}) \overline{\overline{\Theta_1}}^{-1} \eta_1 Z\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.15)$$

The following is a test hypothesis for a diagnostic model of white noise evaluation:

H0: Γi = 0 (data Errors together have white noise)
H1: Γi ≠ 0 (data Errors together not white noise)

FORECASTING THE VARMA MODEL

The forecasting concept for the model of VARMA (p, q) is comparable to that of the model of VAR(p). Consider ARMA (p, q) vector model (Lütkepohl, 2005; Mhamad, 2019; Nasser et al., 2009).

$$Z_t = c_t + \sum_{i=1}^{p} \mathbf{\Phi}_i Z_{t-i} + a_t + \sum_{i=1}^{q} \mathbf{\Theta}_i a_{t-i} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.16)$$

where the intercept $c_t = (I - \mathbf{\Phi}_1 - \mathbf{\Phi}_2 - \cdots - \mathbf{\Phi}_p)\mu$

It is necessary to assume that each component of the vector at is independent of the others (Lütkepohl, 2005; Mhamad, 2019). Given that the potential white noise processing a $a_{t+h}h > 0$ is unrelated of the historical and current values $Z_t, Z_{t-1}, \ldots$, the value of $E(a_{t+h}|Z_t, Z_{t-1}, \ldots)$ is 0 if conditional expectations are provided to the both sides of the equation (2.16).

The step-ahead forecast Z_{t+h} predictor with the minimum mean square error is:

$$\hat{Z}_t(h) = E(Z_{t+h}|Z_t, Z_{t-1}, \ldots)$$
$$= c_t + \Phi_1 \hat{Z}_t(h-1) + \cdots + \Phi_p \hat{Z}_t(h-p) + \Theta_1 E(a_{t+h-1}) + \cdots + \Theta_q E(a_{t+h-q}) \quad \ldots(2.17)$$
$$h = 1, \ldots, q, \hat{Z}_t(h) = Z_{t+h} \; for \; h \le 0 \; and \; E(a_{t+j}) = a_{t+j} \; for \; h \le 0 \text{ (Wei, 2006)}$$

if $h > q$, then the Equation (2.17) becomes

$$\hat{Z}_t(h) = c_t + \Phi_1 \hat{Z}_t(h-1) + \cdots + \Phi_p \hat{Z}_t(h-p) \qquad \ldots\ldots\ldots\ldots\ldots (2.18)$$

Lütkepohl (2005) highlighted that the forecast for period h, $\hat{Z}_t(h)$, is likewise calculable using the infinite VAR representation: $\hat{Z}_t(h) = \sum_{i=1}^{\infty} \Pi_i \hat{Z}_t(h-i)$. The infinite VMA representation can also yield $\hat{Z}_t(h)$. The expression Z_{t+h} represents the future value $Z_{t+h} = \sum_{i=0}^{\infty} \Psi_i a_{t+h-i}$. Since $E(a_{t+h}|Z_t, Z_{t-1}, \ldots) = 0$ $h > 0$, the predictor with the smallest mean square error is $\hat{Z}_t(h) = \sum_{i=1}^{\infty} \Psi_i a_{t+h-i}$ (Dufour & Pelletier, 2022).

$Z_{t+h} - \hat{Z}_t(h)$ is distributed as a normal distribution with a mean zero and covariance matrix for forecasting errors $\sum_{i=1}^{\infty} \Psi_i \Sigma_a \acute{\Psi}_i$ (Feng et al., 2021; Lütkepohl & Poskitt, 1996; Lütkepohl, 2005; Mhamad, 2019). The white noise sequence $a_{t+h-i}, i = 1, 2, \ldots, q$ of the model in Equation (2.18), must be created recursively utilizing the previous values $Z_t, Z_{t-1}, \ldots$ from Equation

$$a_s = Z_t - \sum_{i=1}^{p} \Phi_i Z_{s-i} - \sum_{i=1}^{q} \Theta_i a_{s-i} \qquad \ldots\ldots\ldots\ldots\ldots (2.19)$$

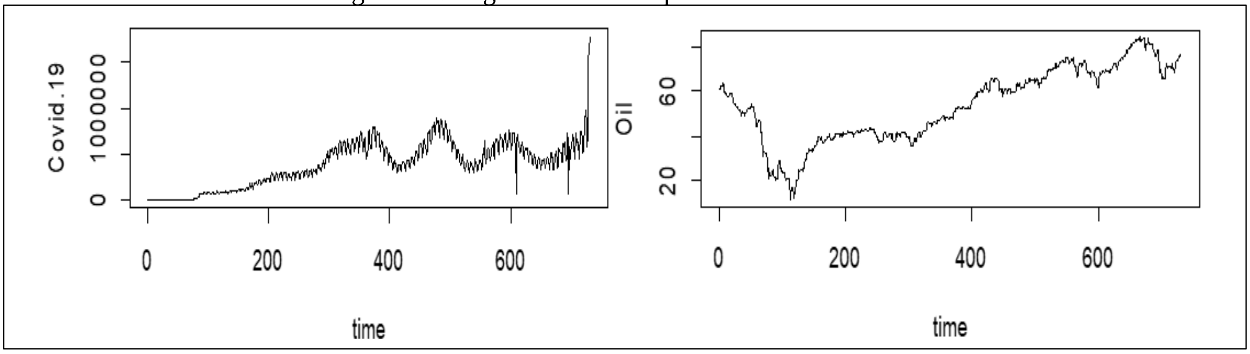
One can get this sequence by choosing proper starting values, such as $a_0, \ldots, a_{1-q}$ and $Z_0, \ldots, Z_{1-p}$ (Lütkepohl, 2005).

RESULTS AND DISCUSSION

Population and Sample

In this study, we use daily observations of the COVID-19 pandemic, for which records only containing cases are used in the statistical modeling process and are seasonally adjusted. These observations are obtained from the WHO website (covid19.who.int), along with another variable, oil price, which is not seasonally adjusted and obtained from (www.macrotrends.net). The sample period is twenty-four months, from Jan 2020 to Dec 2021. The natural time series figure of the two series during the sample period are shown in Fig 3.1.

Figure 3.1: original time series plots of the variables



Source: Prepared by the authors (2022).

we used the Augmented Dickey-Fuller test on the series. The outcomes of applying the ADF test to levels and determining the stationary by taking the logarithm with the initial difference between the two series are shown in Table 3.1.

Since every series utilized for analysis has to be stationary in order to build an effective model, therefore, the unit-root structure of the data should be examined. Even though figure (3.1) offers us a general sense of the stationarity structure of the series, to test for unit-roots for the series we have applied the Augmented Dickey-Fuller test. Table 3.1 displays the outcomes of applying the ADF test to levels and by taking the logarithm with the first difference in the two series, the stationary is calculated.

Table 3.1: Unit root results of lag variables

Variables	Lags	Test Value	p- value
COVID-19	Level	-1.933	0.317
	First	-26.909	0.000
Oil price	Level	-1.309	0.627
	First	-8.742	0.000

Source: Prepared by the authors (2022).

Through not rejecting the unit-root hypothesis at all levels, the ADF test results in table (3.1) imply that all variables are non-stationary; however, they are all stationary after taking the logarithm and first difference. In our study, we employ the transformation and series differenced methods for variables. The subsequent step of the analysis is model order selection, as indicated in the table (3.2).

Table 3.2: VARMA order selection results

	0	1	2	3	4	5	6
0	0.0001	0.0002	0.0003	0.0001	0.0001	0.0000	0.0739
1	0.0002	0.1251	0.0556	0.9308	0.0135	1.0000	0.6924
2	0.0053	0.0378	0.0462	0.9655	1.0000	0.9999	0.9876
3	0.0227	0.9996	0.9935	0.7931	1.0000	1.0000	0.9987
4	0.1112	0.7572	1.0000	1.0000	0.9798	0.9998	0.9852
5	0.2441	1.0000	1.0000	1.0000	1.0000	0.7832	0.9501
6	0.9832	0.9999	1.0000	1.0000	0.9999	0.9928	0.9801

Source: Prepared by the authors (2022).

Table (3.2) displays the vector model identification done by VARIMA. Significant orders are those with a P-value less than α (0.05). At the modeling step, four models (VARIMA (1,1,1), VARIMA (2,1,1), VARIMA (2,1,2), and VARIMA (3,1,1)) can be examined. The VARIMA models (2,1,2) with the minimum AIC and BIC values are those with an AIC value of -15.61059 and a BIC value of -15.535. Then it may be argued that VARIMA (2, 1, 2) is the best model. Table 3.3 displays the final model and parameter estimates for the VARIMA (2, 1, 2) model.

Table 3.3: Model with VARIMA (2,1,2)

Estimate	Std. Error	t value	Pr(>|t|)
-0.00176	0.001255	-1.399	0.16170
0.002662	0.001369	1.944	0.05187
0.554828	0.196273	2.827	0.00470
-0.28279	0.187321	-1.51	0.13114
1.159863	0.448614	2.585	0.00973
-0.4749	0.190784	-2.489	0.01280
-0.62726	0.134618	-4.66	0.00000
-0.66769	0.221853	-3.01	0.00262
0.26627	0.228505	1.165	0.24391
-1.1545	0.446331	-2.587	0.00969
0.426764	0.192049	2.222	0.02627
0.613024	0.134344	4.563	0.00001

Source: Prepared by the authors (2022).

From table (3.3) note that all parameters are significant, according to the VARIMA (2,1,2) model, with the exception of three parameters. The VARIMA (2,1,2) model is:

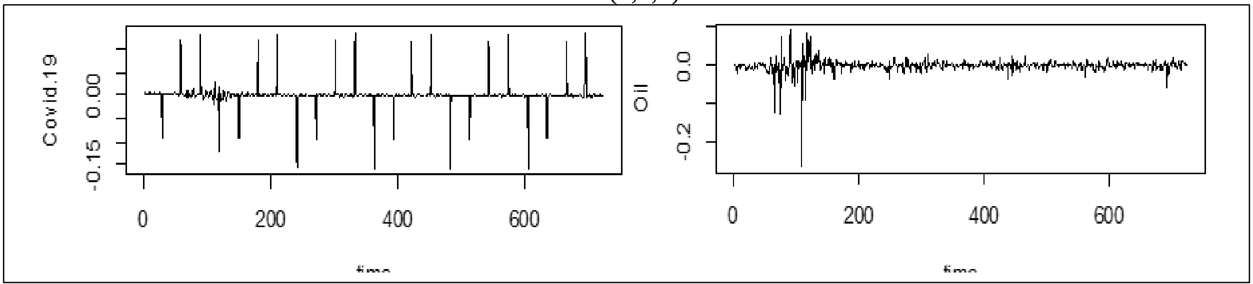
$$z_t = \begin{bmatrix} -0.00176 \\ 0.00266 \end{bmatrix} + \begin{bmatrix} 0.56 & 0.00 \\ 0.00 & 0.00 \end{bmatrix} z_{t-1} + \begin{bmatrix} -0.28 & 1.16 \\ -0.48 & -0.63 \end{bmatrix} z_{t-2} + a_t - \begin{bmatrix} 0.67 & 0.00 \\ 0.00 & 0.00 \end{bmatrix} a_{t-1} - \begin{bmatrix} -0.27 & 1.15 \\ -0.43 & -0.61 \end{bmatrix} a_{t-2}$$

where the standard errors of the estimates and the covariance matrix of residual is

$$\widehat{\Sigma_a} = \begin{bmatrix} 0.000320 & -0.000005 \\ -0.000005 & 0.00502 \end{bmatrix}$$

As Ljung–Box statistics have demonstrated in the table (3.4), Based on the residual covariance matrix, can be accept the hypothesis that there aren't any cross-correlations in the residuals. Figure 3.2 depicts the residual time plots of the fitted VARIMA (2,1,2) model. As previously stated, the data contain some outlier findings.

FIGURE 3.2: the VARIMA (2,1,2) model residual series



Source: Prepared by the authors (2022).

The residuals of the VARIMA (2,1,2) model have been subjected to the Portmanteau test, table (3.4) presents the degrees of freedom DF and the test statistic values in which it is insignificant for all lagged except for four lags where the p value and their values are more than the significance level of 0.05.

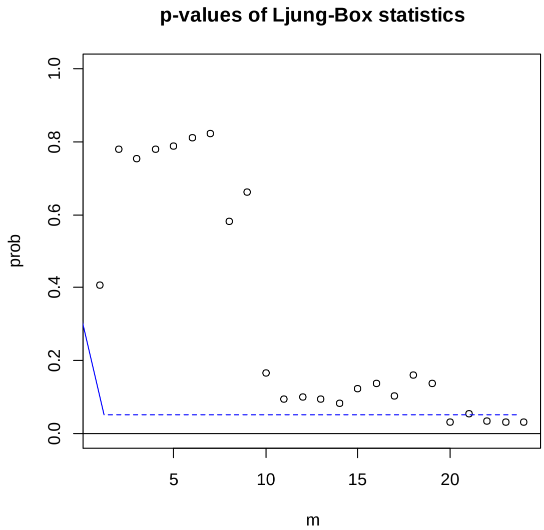
Table (3.4): Value statistics of Ljung–Box test for the model

mx	Q(m)	df	p-value	m	Q(m)	df	p-value
1	4	4	0.41	13	65.91	52	0.09
2	4.8	8	0.78	14	71.15	56	0.08
3	8.39	12	0.75	15	72.85	60	0.12
4	11.48	16	0.78	16	76.51	64	0.14
5	14.79	20	0.79	17	83.16	68	0.1
6	17.85	24	0.81	18	83.95	72	0.16
7	21.03	28	0.82	19	89.59	76	0.14
8	29.74	32	0.58	20	105.45	80	0.03
9	31.95	36	0.66	21	106.03	84	0.05
10	48.62	40	0.16	22	113.92	88	0.03
11	56.78	44	0.09	23	118.99	92	0.03
12	61	48	0.1	24	123.88	96	0.03

Source: Prepared by the authors (2022).

The null hypothesis was accepted which is stated that there is no autocorrelation in the residuals, as indicated by the P-values in Table (3.4) that are more than the level of significance (α = 0.05) as depicted in Figure (3.3). In other words, residuals are random. The conclusion is that there is no autocorrelation in the data, and the model has qualified white noise. Consequently, VARIMA (2, 1, 2) models can be utilized to forecast covid-19 and oil price.

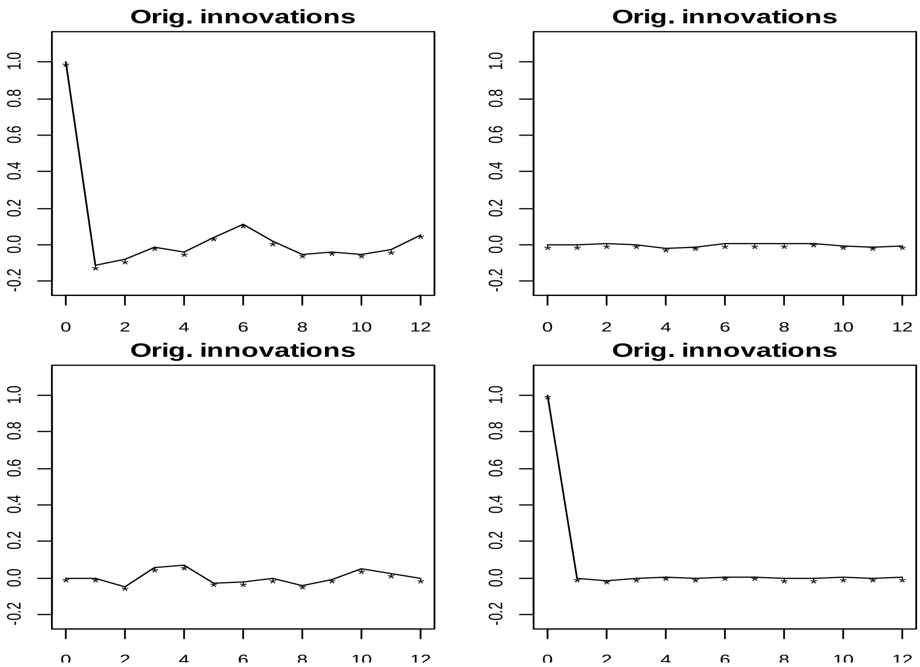
Figure 3.3 multivariate Ljung–Box *p*-values for VARIMA (2,1,2) model residuals



Source: Prepared by the authors (2022).

Figure 3.3 depicts the residuals' multivariate ljung–box p-values. The model has successfully captured the dynamic dependence of the data, as seen by the plot. The hypothesis cannot be rejected which is stated that the residuals show no evidence of cross-correlation

FIGURE 3.4 VARIMA (2,1,2) model impulse response functions



Source: Prepared by the authors (2022).

Figure 3.4 demonstrates the impulse response functions of the VARIMA (2,1,2) model, which incorporates innovative improvements. The two series have an overall favorable effect on one another. For instance, the VARIMA (2,1,2) model predicts that the Oil price's impulse responses to its own shocks will exhibit two successive increases following the initial abrupt decrease.

Granger causality is employed to examine two hypotheses. The first test examines the hypothesis that the Oil price iis influenced only by itself and not by the Covid-19 pandemic. The second experiment tests the null hypothesis that the Covid-19 pandemic is influenced only by itself and not by the Oil.

Table (3.5): Wald Test for Granger Causality

Test	Groups	Chi-sq	Prob.
1st test	First group: COVID	5.583812	0.0181
	Second group: OIL		
2nd test	Group 1 variable: OIL	1.090488	0.2964
	Group 2 variable: COVID		

Source: Prepared by the authors (2022).

Based on the findings of the test of granger causality, table (3.5) reveals that Test 1, Chi-square is equal to 5.58 with P = 0.018; Consequently, the zero hypothesis can be rejected, and can be conclude that the Oil is influenced not only by itself but also by the Covid-19 pandemic cases. Although. For Test 2, Chi-square is 1.09 and P = 0.296; therefore, the null hypothesis cannot be rejected, and we infer that the Covid-19 pandemic is exclusively influenced by itself and not by Oil.

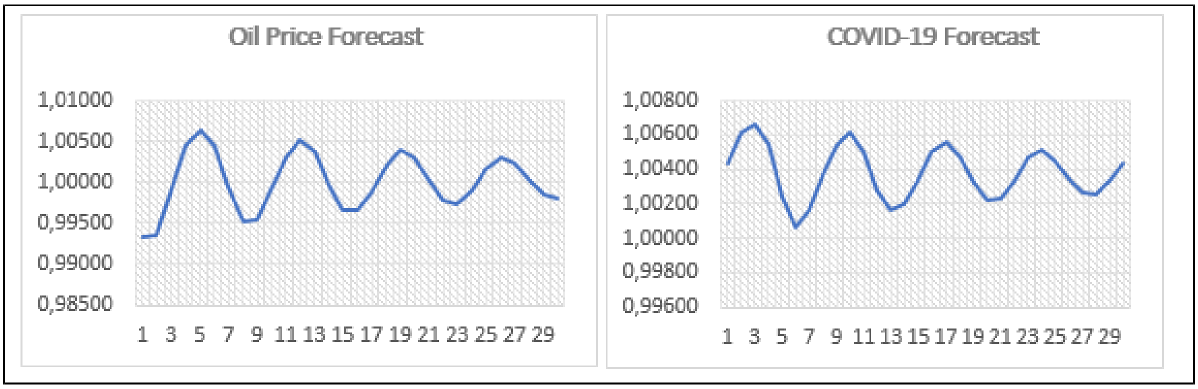
Table (3.6). Forecasts for VARIMA (2,1,2)

	Forecast		Standard errors			Forecast		Standard errors	
	Covid.19	Oil	Covid.19	Oil		Covid.19	Oil	Covid.19	Oil
1	0.99339	1.00434	0.01789	0.02241	16	0.99652	1.00499	0.01835	0.02254
2	0.99352	1.00610	0.01800	0.02241	17	0.99886	1.00562	0.01836	0.02254
3	0.99924	1.00659	0.01806	0.02243	18	1.00208	1.00468	0.01838	0.02254
4	1.00443	1.00542	0.01806	0.02245	19	1.00394	1.00316	0.01838	0.02254
5	1.00626	1.00238	0.01808	0.02249	20	1.00297	1.00222	0.01839	0.02254
6	1.00445	1.00064	0.01810	0.02249	21	1.00015	1.00228	0.01841	0.02254
7	0.99941	1.00168	0.01821	0.02250	22	0.99778	1.00334	0.01841	0.02255
8	0.99512	1.00363	0.01821	0.02250	23	0.99733	1.00464	0.01841	0.02255
9	0.99537	1.00537	0.01824	0.02251	24	0.99897	1.00511	0.01842	0.02255
10	0.99896	1.00620	0.01825	0.02251	25	1.00151	1.00451	0.01843	0.02255
11	1.00291	1.00499	0.01827	0.02252	26	1.00300	1.00343	0.01843	0.02255
12	1.00505	1.00275	0.01828	0.02253	27	1.00241	1.00260	0.01844	0.02255
13	1.00370	1.00163	0.01831	0.02253	28	1.00042	1.00256	0.01845	0.02255
14	0.99977	1.00202	0.01834	0.02253	29	0.99852	1.00336	0.01845	0.02256
15	0.99668	1.00336	0.01834	0.02253	30	0.99799	1.00434	0.01845	0.02256

Source: Prepared by the authors (2022).

VARIMA (2,1,2) model was used to forecast, and table (3.6) displays the forecasting values for oil prices from 1-Jan 2022 to 31-Jan 2022. From the first to the third period, the forecasting values increase, then drop from the fourth to the eighth period, and then jump from the ninth to the tenth period. and continue as the same fluctuation of forecasting to the last period, indicating that oil prices are stabilizing after the outbreak of the covid-19 pandemic, standard error values for oil price and the covid-19 pandemic case begin to increase from the first to the last period, in ascending sequence, this demonstrates that the volatility of the two-forecasting series rises as the length of the forecasting period increases. This indicates that the variance will increase if the model has been used to forecast for a long time. The forecasting series are depicted in figure (3.5).

Figure (3.5). Forecast of oil price and covid-19 pandemic case



Source: Prepared by the authors (2022).

CONCLUSIONS

Coronavirus disease 2019 is a contagious disease that has a huge effect on the global economy and finances. However, as a result of this virus's proliferation, the impact of the virus on human existence is expanding. Consequently, the aim of the study is to forecast and examine the covid-19 pandemic effects on oil prices using model of (VARIMA). The VARMA (p, q) model was used in the current investigation to relate a set of series to their prior values and those of other series as well as to the prior magnitudes of shocks to those series and other series. In this study, COVID-19 pandemic cases and oil prices were used as a case study; they contain daily international observations for the years 2020 to 2022. This study focuses on how to determine the optimal model and employ it for forecasting the data. The best model is VARIMA (2,1,2), that contains nine significant parameters and three non-significant parameters, despite the fact that no cross correlations in the residual based on the results of the residual covariance matrix. In other words, residuals are random, there is no autocorrelation in the data, and the model has been qualified white noise. Consequently, VARIMA (2, 1, 2) models can be utilized to forecast covid-19 and oil price. However, the two series have a positive effect on one another. For illustrate, the impulse responses of the oil price to its own shocks show two subsequent increments for the VARIMA (2,1,2) model following the initial sharp decrement. The results of Granger causality indicate that the Oil is not only influenced by itself but also by the covid 19 pandemic cases, but the covid 19 pandemic is only influenced by itself and not by the Oil. The results of the forecast demonstrated that the standard error increases over time; the standard error in the first period is extremely modest compared to the prediction of the means, but increases with time to the 30-point forecast. This shows that the model is reliable for short-term forecasting, however long-term forecasting results are unstable (due to a larger standard error). Future researchers can use different variables in different fields and discover integrating between them. One avenue that might also be considered is including the tourism demand and industry covariates together in modelling.

REFERENCES

Ahmed, B. K., Rahim, S. A., Maaroof, B. B., & Taher, H. A. (2020). Comparison Between ARIMA And Fourier ARIMA Model To Forecast The Demand Of Electricity In Sulaimani Governorate. QALAAI ZANIST JOURNAL, 5(3), 908-940. https://doi.org/10.25212/lfu.qzj.5.3.36

Ahmed, N. M. (2007). MODELING CROSS SECTION DATA CONTAINS EQUALITY CONSTRAINTS WITH APPLICATION. Journal of Administration and Economics, (64).

Ahmed, N. M. (2007). Forecasting Water Inflow and Electric Power Generation in Darbandi-khan Station By Using FFNN, and Adaptive Method In Time Series (Doctoral dissertation, Phd. Thesis of Philosophy in Statistics, University of Suleimani, Kurdistan Region–Iraq).

Ahmed, N. M., & Hamdeen, A. O. (2018). Predicting Electric Power Energy, Using Recurrent Neural Network Forecasting Model. Journal of University of Human Development, 4(2), 53-60. https://doi.org/10.21928/juhd.v4n2y2018.pp53-60

Akhtaruzzaman, M., Boubaker, S., & Sensoy, A. (2021). Financial contagion during COVID–19 crisis. Finance Research Letters, 38, 101604. ArticleDownload PDFView Record in ScopusGoogle Scholar. https://doi.org/10.1016/j.frl.2020.101604

Asraa, A., Rodeen, W., & Tahir, H. (2018). Forecasting the Impact of Waste on Environmental Pollution. International Journal of Sustainable Development and Science, 1(1), 1-12. DOI: 10.21608/IJSRSD.2018.5142

Athanasopoulos, G., & Vahid, F. (2007). A complete VARMA modelling methodology based on scalar components. *Journal of Time Series Analysis*, *29*(3), 533-554. https://doi.org/10.1111/j.1467-9892.2007.00568.x

Athanasopoulos, G., & Vahid, F. (2008). VARMA versus VAR for macroeconomic forecasting. *Journal of Business & Economic Statistics*, *26*(2), 237-252. https://doi.org/10.1198/073500107000000313

Benkraiem, R., Lahiani, A., Miloudi, A., & Shahbaz, M. (2018). New insights into the US stock market reactions to energy price shocks. Journal of International Financial Markets, Institutions and Money, 56, 169-187. ArticleDownload PDFView Record in ScopusGoogle Scholar. https://doi.org/10.1016/j.intfin.2018.02.004

Chang, C. P., & Zhang, L. W. (2016). Do natural disasters increase financial risks? An empirical analysis. Bulletin of Monetary Economics and Banking, 23, 61-86. DOI: 10.21098/bemp.v23i0

Dufour, J. M., & Pelletier, D. (2022). Practical methods for modeling weak VARMA processes: identification, estimation and specification with a macroeconomic application. *Journal of Business & Economic Statistics*, *40*(3), 1140-1152. https://doi.org/10.1080/07350015.2021.1904960

Feng, G. F., Yang, H. C., Gong, Q., & Chang, C. P. (2021). What is the exchange rate volatility response to COVID-19 and government interventions? Economic Analysis and Policy, 69, 705-719. ArticleDownload PDFView Record in ScopusGoogle Scholar. https://doi.org/10.1016/j.eap.2021.01.018

Feunou, B., & Fontaine, J. (2009). *A no-arbitrage VARMA term structure model with macroeconomic variables*. Working Paper, Duke University.

Kascha, C., & Trenkler, C. (2011). Cointegrated VARMA models and forecasting US interest rates. *Available at SSRN 1957103*.

Lütkepohl, H & Poskitt, D. S. (1996). Specification of Echelon-Form VARMA Models , Journal of Business and Economic Statistics 14(1) : 69 – 79.

Lütkepohl, H. (2005). New introduction to multiple time series analysis. Springer Science & Business Media. Berlin.

Mainassara, Y. B. (2010). Selection of weak VARMA models by Akaıke's information criteria.

Menhat, M., Zaideen, I. M. M., Yusuf, Y., Salleh, N. H. M., Zamri, M. A., & Jeevan, J. (2021). The impact of Covid-19 pandemic: A review on maritime sectors in Malaysia. Ocean & Coastal Management, 209, 105638. https://doi.org/10.1016/j.ocecoaman.2021.105638.

Mhamad, A. J. (2019). Using regression Kriging to analyze groundwater according to depth and capacity of wells. UHD Journal of Science and Technology, 3(1), 39-47. https://doi.org/10.21928/uhdjst.v3n1y2019.pp39-47

Montgomery, A.L & Moe, W. W. (2002). Should Music Labels Pay for Radio Airplay? Investigating the Relationship Between Album Sales and Radio Airplay, August 2002.

Nasser, A. A., Rahi, A. K., & Ahmad, N. M. (2009). Using Feed-Forward Neural Network (FFNN) In Time Series Forecasting. Journal of Administration and Economics, (74).

Rahim, S. A., Salih, S. O., Hamdin, A. O., & Taher, H. A. (2020). Predictions the GDP of Iraq by using Grey–Linear Regression Combined Model. The Scientific Journal of Cihan University–Sulaimaniya, 4(2), 130-139. https://doi.org/10.25098/4.2.25

Wei, W. W. S. (2006). Time Series Analysis, Univariate and Multivariate Methods: Second Edition, Addison-Wesley, Redwood City, CA.