BloEx - Journal
Britain International of
EXACT Sciences Journal
Publish: | January - May - September |
E-ISSN: 2686-1208
P-ISSN: 2686-1216
http://biarjournal.com/index.php/bioex

Performance Measures for Evaluating the Accuracy of Time Series Hybrid Model Using High Frequency Data

T.O. Olatayo[1], K. I. Ekerikevwe[2]
[1]Department of Mathematical Sciences, Olabisi Onabanjo University, Ago-Iwoye
[2]Department of Statistics, Delta State Polytechnic, Otefe-Oghara, Delta State, Nigeria
Email: kennedyekerikevwe@gmail.com

Abstract:

Given that the traditional ARIMAX model has rarely been applied to any of the climate change and environmental agents, which are the most cognate agents with associated exogenous variables; to neutralize the model for a better and enhanced prediction of the system, a distributional form of the error term that is robust and sufficient in capturing and accommodating both the external covariate(s) and high frequency data is required. This study therefore evaluates the forecasting accuracy of two forecasting models namely ARIMAX and log-ARIMAX. The monthly adjusted high frequency data recorded by four Oil and Gas companies from 2005 – 2020 were used. The forecastability of the two models was evaluated with different error matrices. The effect of Akaike Information Criterion (AIC) and the linear correlation on candidate models among the considered oil spill data tested were discussed. Results for ARIMAX and LOG-ARIMAX Models selection with respect to AIC show that log-ARIMAX is more efficient and performed better than the traditional ARIMAX model for observations characterized by kurtosis, skewness, outliers, high frequency and large fluctuation series with heavy tailed traits as seen in environmental data.

Keywords:
arimax; log-arimax; heavy tailed; high frequency; non-linear

I. Introduction

Non-linear time series models are limited because they are neither robust nor sufficient in capturing high-frequency observational time series, such as those including exogenous factor(s). To improve forecast accuracy, an autoregressive integrated moving average model, ARIMAX (p, d, q, b), for short-memory observational time series data with exogenous covariate(s) to capture times series data like interest rate, number of stocks sold, changes in monthly commodity prices, etc.) was developed. However, for long-memory frequency observations, like the climate data and ecological data; a modification will be necessary to neutralize the model for a better and improved prediction of the system.

Long-memory frequencies' observations, like the climate-measured daily/weekly/average monthly temperatures recorded, changes in daily recorded climate; currencies exchange rates; Consumer Price Index (CPI) and GDP, a modification will be necessarily needed to neutralize or log-linearize the ARIMAX (p,d,q,b) model for the betterment and improvement of the (constant rate of slope change), constant trend (constant mean), transfer function, residual process and prediction of the system. Having ascertained the power of logarithm in the process of differencing or transformation to help stabilize, eliminate (or reducing) trends, mean of the time series and seasonality signal if any when characterized with high or long memory series. So, a log function would be added to the ARIMAX(p,d,q,b)

to make it Log-ARIMAX (p,d,q,b)to neutralized the threat posed by long-memory traits that might likely affect not only the parameterization (over-parameterization or under-parameterization) and the end product of given a reliable system forecasts. Forecast generally usually emanated from the generalization of the residual processes attached to the model. However, Log-ARIMAX (p,d,q,b) would not be an exception but additional components of in-sample and out-sample forecast would be incorporated and tested via forecast indexes via AR, MA and the exogenous residual processes. The Log-ARIMAX abstraction of reality would not only make room for merging linear/non-linear regression with ARMA model for better broadening of the applicability of non-linear time series models but also going to serve as a platform of introducing Generalized Non-linear/linear time series for transfer function (otherwise called mean function and impulse weight function in Generalized Linear Model (GLM). This technique of solving residual of different residual structures of different distributional forms and different variable types will be employed to treat the white noise of the Log-ARIMAX model. Additionally, appropriate formulations for the autocorrelation structure (Partial Autocorrelation Function (PACF)) of the error term from the regression equation function (transfer function) of the long memory (highly frequency data) would be sufficiently identify and compare to ARMA and ARMAX models. Another meritorious trait of both the Log-ARIMAX and ARIMAX models would be the attachment of degree of fit (contributions of the exogenous variables) as measured by the coefficient of R-squared and its variants to fitted models; and capturing of the dynamics of seasonal variation change patterns over time. The long memory (high frequency) associated to economic, environmental, climate change, wave data (sea and ocean wavy pattern record) etc. will be a typical example of most long memory data due to their fluctuations, higher values, dependency, and switching circular traits.

II. Review of Literature

Shilpa & Sephardi (2019) in their paper modeled ARIMAX model that was incorporated with exogenous variables which is an extension of ARIMA model using STLF on a time series data of Karnataka State Demand pattern. They enhanced ARIMA model by considering hours of days of the week as the independent variables for ARIMAX model. Utility of data electrical loaded demanded from Karnataka Power Transmission Corporation Ltd. (KPTCL) website was used to develop and test the proposed forecasting model of ARIMAX. However, economic index like inflation, which has persistent and appreciable rise in the general level of prices; general price level might be response or predictive variable (Frimpong and Oteng-Abayie, 2010). Another example of long memory and conditionally covariates series exchange rate (which is the value of the domestic currency in terms of foreign currency). Exchange rate changes can affect the relative prices, thereby the competitiveness of domestic and foreign producers. Theoretically, exchange rate will have a negative or positive relationship with economic growth. This is because currency depreciation will foster a country export that will lead to an increase in Gross Domestic Product (GDP) while currency depreciation will also discourage a country import, thus leading to decrease in GDP of that country. It counts out that appreciation of exchange rate exerts positive influence on GDP and real economic growth (Aliyu, 2011). Therefore, exchange rate and GDP might be dependency or predictor for an ARIMAX or ARIMA model depending on their context. Among other congenital examples of realization that best fit the Log-ARIMAX conceptualization are environmental and climate changes of oil spillage and temperature.

Gopinath and Kavithamani (2019) constructed and fitted an ARIMAX model to production of sugarcane in India and as well to know the future values of sugarcane

production in India from 2015 to 2026. The aim of the research was centered on accurate prediction. The secondary data used was collected from Sugarcane Breeding Institute, Coimbatore in India. The ARIMAX model was introduced and ascertainment of its order, parameter and diagnostic checking was followed via Box and Jenkins method with both ARIMA and ARIMAX models given greater accuracy in comparison.

Ling et al. (2019) reported that the inappropriate spraying, untreated incidence and inexact forecast of cocoa black pod disease around the world have led to incurable losses of more than $400 million. They accounted external factors has the contributing this disease affecting cocoa. Factors like relative humidity, rainfall and temperature are the influencing external factors to this cocoa black pod disease.

Iflah and Parul (2020) noted that Auto-Regressive Integrated Moving Average (ARIMA) and Artificial Neural Networks (ANN) are leading linear and non-linear models in Machine learning respectively for time series forecasting. Their survey paper presents a review of recent advances in the area of Machine Learning techniques and artificial intelligence used for forecasting different events. This paper presents an extensive survey of work done in the field of Machine Learning where hybrid models are compared to the basic models for forecasting on the basis of error parameters like Mean Absolute Deviation (MAD), Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Normalized Root Mean Square Error (NRMSE). The results of their work summarize important discuss on the basis of some parameters which explain the efficiency of hybrid models or when the model is used in isolation. They concluded that hybrid model has realized accurate results as compared when the models were used in isolation yet some research papers argue that hybrids cannot always outperform individual models.

Ahmed et al. (2020) proposed a new heavy-tailed exponential distribution that accommodates bathtub, upside-down bathtub, decreasing, decreasing-constant, and increasing hazard rates for actuarial data. Actuarial measures including value at risk, tail value at risk, tail variance, and tail variance premium are derived. A computational study for these actuarial measures was conducted, proving that the proposed distribution has a heavier tail as compared with the alpha power exponential, exponentiated exponential, and exponential distributions.

Literature showed that quite a number of researchers have studied ARIMAX associated with exogenous covariate (s), using different short-memory frequency data, with little or no strength to capture long-memory (high frequency) observations with heavy tailed traits. Having in mind that conventional ARIMAX model has been rarely applied to any of the climate change and environmental agents which are the most cognate agent with associated exogenous variables and are usually characterized by kurtosis, skewness, outliers, long memory (high frequency) and large fluctuation series; a distributional form of the error term that would be robust and sufficient in capturing and accommodating both the external covariate(s) and long memory (high frequency) data would be needed to neutralize the model for a better and improved prediction of the system. Therefore, the motivation to propose and formulate a Log-ARIMAX model whose distributional form would be robust and sufficient in capturing and accommodating both the external covariate (s) and the heavy-tailed properties of long-memory (high frequency) observational time series events becomes necessary.

III. Research Methods

3.1 Autoregressive Model

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots\cdots + \phi_p y_{t-p} + \varepsilon_t \quad (1)$$
$$= \phi_0 + \sum_{i=1}^{p} \phi_i y_{t-i} + \varepsilon_t$$

3.2 Moving Average

$$\varepsilon_t = \theta_0 + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots\cdots + \theta_q \varepsilon_{t-q} \quad (2)$$
$$= \theta_0 + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}$$

3.3 Autoregressive Moving Average

Autoregressive Moving Average (ARMA) model is considered the mixture of AR and MA models, it is given by:

$$Y_t = \varphi_0 + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \quad (3)$$

$$\varphi_p(B) Y_t = \mu + \theta_q(B) \varepsilon_t \quad (4)$$

$$\varphi_p(B) \nabla^d Y_t = \varphi_0 + \theta_q(B) \varepsilon_t \quad (5)$$

3.4 Autoregressive Integrated Moving Average with Covariates "X"

According to Yang and Wang (2017) ARIMAX can be defined as:

$$Y_t = \varphi_0 + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots\cdots + \varphi_p y_{t-p} + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \cdots + \beta_p x_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \quad (6)$$

$$\varphi_p(B) Y_t = \varphi_0 + \varphi(B) x_t + \theta_q(B) \varepsilon_t \quad (7)$$

$$\varphi_p(B) \nabla^d Y_t = \varphi_0 + \varphi(B) x_t + \theta_q(B) \varepsilon_t \quad (8)$$

For ARMAX and ARIMAX respectively

3.5 Log-Autoregressive Integrated Moving Average-X (Log-ARIMAX)

For log-ARMAX

$$Y_t = \sum_{j=1}^{n} \frac{\eta_h(B) B^b}{\lambda_r(B)} X_t + \frac{\theta_q(B) \varepsilon_t}{\varphi_p(B)} \sim N\left[exp\left(\frac{\sigma^2}{2}\right), exp(2\sigma^2) - exp(\sigma^2)\right] \quad (9)$$

For log-ARIMAX

$$Y_t = \sum_{j=1}^{n} \frac{\eta_h(B) B^b}{\lambda_r(B)} X_t + \frac{\theta_q(B) \varepsilon_t}{\nabla^d \varphi_p(B)} \sim N\left[exp\left(\frac{\sigma^2}{2}\right), exp(2\sigma^2) - exp(\sigma^2)\right] \quad (10)$$

a. Criteria for Selection of Optimal Models

According to Ekerikevwe and Olatayo (2022), the following are criteria for selection of optimal models:

Final Prediction Error:

$$= \hat{\sigma}^2 \frac{n+p+q}{n-p-q} \quad (11)$$

Akaike's Information Criterion (AIC):

$$= -2\ln Likelihood\left(\hat{\varphi},\hat{\theta},\hat{\sigma}^2\right) + 2(p+q+1) \quad (12)$$

Bayesian Information Criterion (BIC):

$$= (n-p-q)\ln\frac{n\hat{\sigma}^2}{n-p-q} + n\left(1+\ln\sqrt{2\pi}\right) + (p+q)\ln\left[\frac{\sum_{t=1}^{n}X_t^2 - n\hat{\sigma}^2}{p+q}\right] \quad (13)$$

b. Forecasting Accuracy Measures

According to Olatayo and Taiwo (2016), the following are forecasting measurement. Parameters:

Mean Absolute Error

$$MAE = \frac{1}{h+1}\sum_{t=s}^{h+s}\left(\hat{X}_t - X_t\right)^2 \quad (14)$$

Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{h+1}\sum_{t=s}^{h+s}\left(\hat{X}_t - X_t\right)^2} \quad (15)$$

Mean Absolute Percentage Error

$$MAPE = \frac{100}{h+s}\sum_{t=s}^{h+s}\left|\frac{\hat{X}_t - X_t}{\hat{X}_t}\right| \quad (16)$$

Diebold and Mariano

$$S^* = \left[\frac{n+1-2h+n^{-1}h(h-1)}{n}\right]^{1/2} S \quad (17)$$

IV. Discussion

4.1 Results

Data

The data on oil spillage from four different oil companies in metric tons are presented in the table 3.1 below. These recording are from: BUNGE PETROLEUM (BG), BASE PETROLEUM (BP), CAINE ENERGY (CNE) and TULLOW OIL & GAS (TLW). This data were obtained from January 2005 to December 2020 for the purpose of validating our proposed model.

Table 1. Data of Oil Spills from Four Companies

Years	BG (Metric tons)	**BP (**Metric tons)	**CNE (**Metric tons)	**TLW (**Metric tons)
2005	361.75	429.91	168.44	156.75
2006	408.25	469.21	189.96	178
2007	411.25	457.94	176.17	173.75
2008	405.25	445.83	175.56	160.25
2009	416.75	467.3	184.05	167
2010	459	492.74	204.19	186.5

2011	471.25	533.87	227.22	191.25
2012	499.5	535.99	268.27	215.5
2013	538	571.19	297.05	260
2014	496	529.21	261.6	242.5
2015	540.5	537.69	277.21	264
2016	574.5	524.97	290.84	270
2017	635	576.26	291.9	313.5
2018	667.5	541.96	292.36	298.5
2019	719.5	568.63	321.89	339
2020	737	581.96	352.34	411.5

Source: NBS Yearly Bulletin, 2020

The graph below shows the time plots of the observed data in two different time horizons.

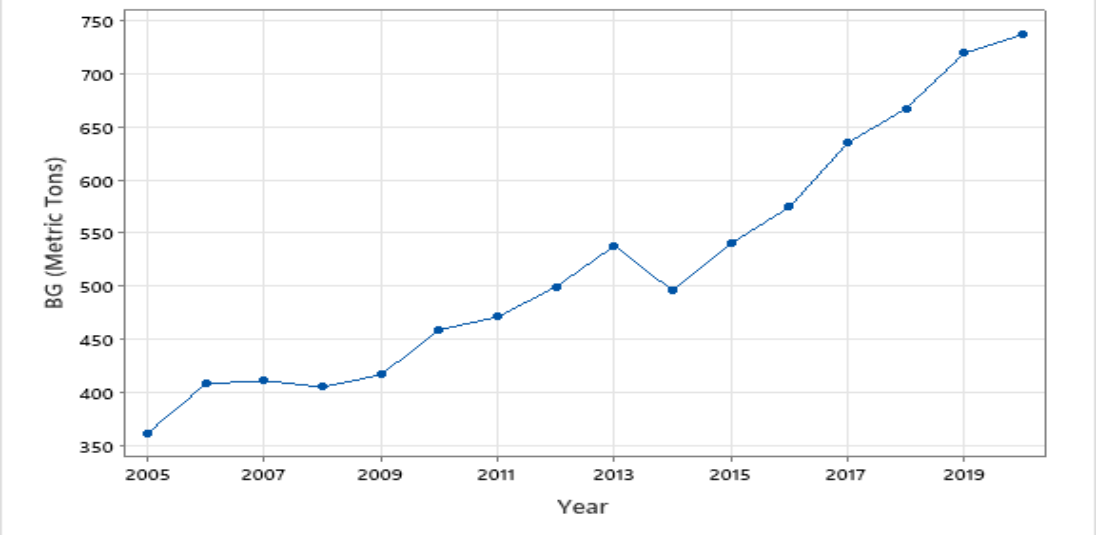

Figure 1. *A Time Plot of the Observed Data for BG (2005-2020)*

Figure 1 below shows the in-sample time plots of the selected models in the first-time regime (2005 – 2020). It is evident from the graph that all the candidate models performed well as each model followed the time plot of the observed data (BG).

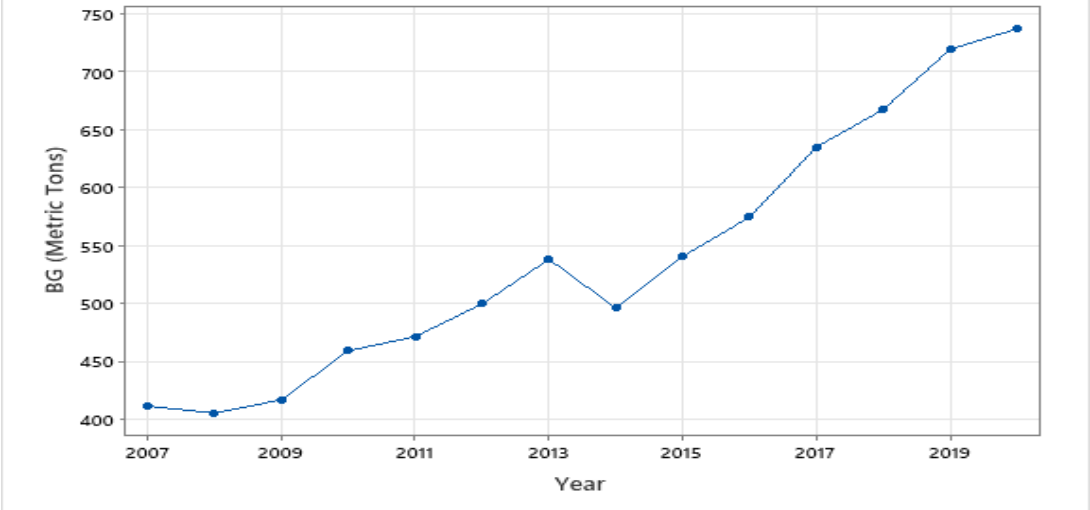

Figure 2. *A Time Plot of the Observed Data for BG (2007-2020)*

On the hand Figure 2 below shows the in-sample time plots of the selected models in the second time regime (2007 – 2020). It is evident from the graph that all the candidate models performed well as each model followed the time plot of the observed data (BG). However, in the first regime, the candidate models followed the observed data (BG) very as compared to that of the second regime (2007 – 2020).). Error metrics for the second time regime is large as compared to the first-time regime. The graph in the second time regime is not as compact as the one in the first-time regime.

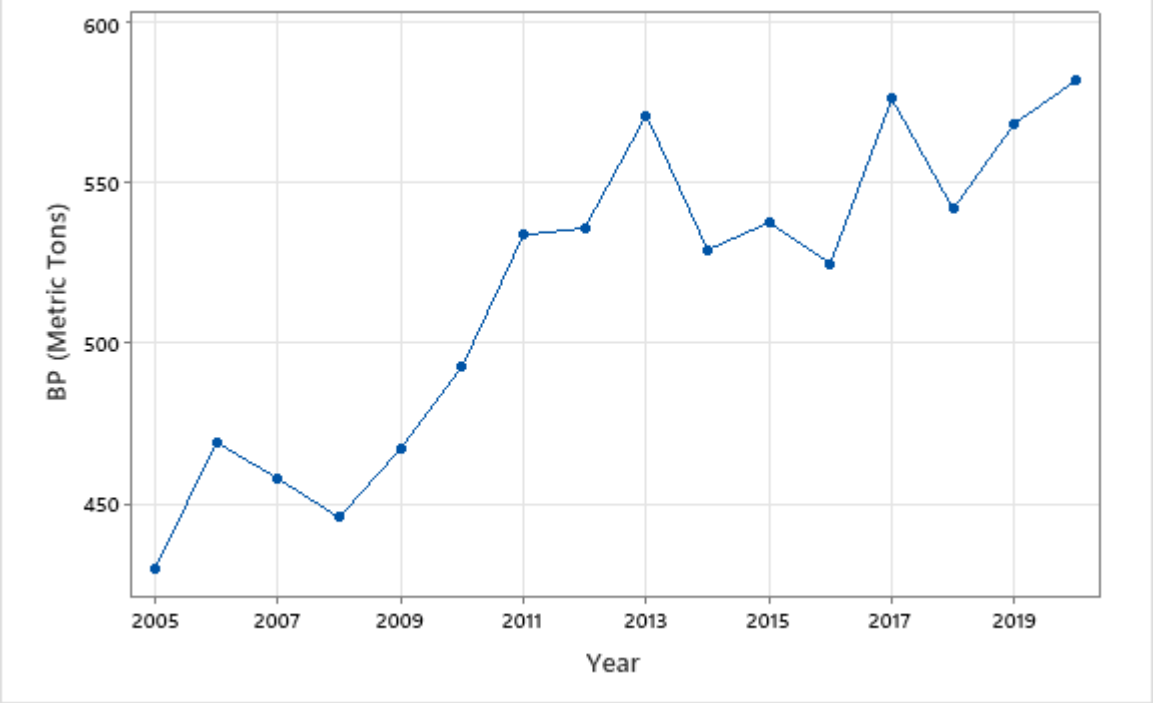

Figure 3. *A Time Plot of the Observed Data for BP (2005-2020)*

Figure 3 below shows the in-sample time plots of the selected models in the first-time regime (2005 – 2020). It is evident from the graph that all the candidate models performed well as each model followed the time plot of the observed data (BP).

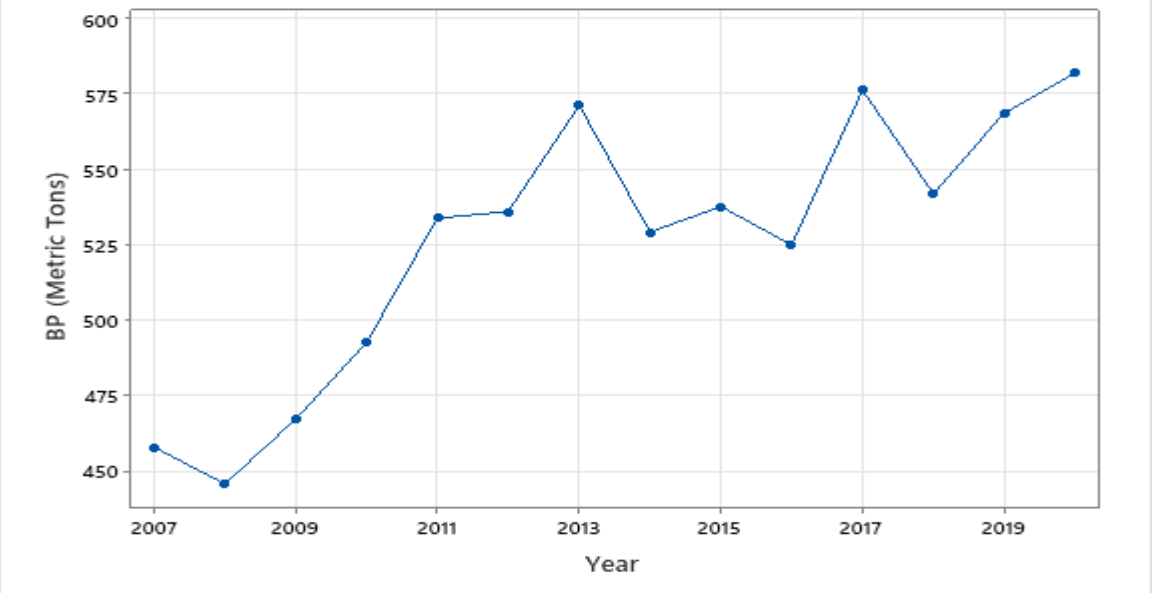

Figure 4. *A Time Plot of the Observed Data for BP (2007-2020)*

On the hand Figure 4 below shows the in-sample time plots of the selected models in the second time regime (2007 – 2020). It is evident from the graph that all the candidate models performed well as each model followed the time plot of the observed data (BP). However, in the first regime, the candidate models followed the observed data (BP) very as compared to that of the second regime (2007 – 2020).

Fundamental measures that will be very instrumental in the forecastability of the candidate methods are the Akaike Information Criterion (AIC) and the linear correlation between the considered Oil spills given the respective history of data. These Oil spills are: BUNGE LIMITED (BG GRP), BP, CAINE ENERGY (CNE) and TULLOW OIL (TLW). The linear correlation between the considered Oil spills given the six-year history of monthly adjusted close price data is given below in Table 2.

Table 2. Correlation Matrix –First Time Horizon (2005 - 2020)

	BG	BP	CNE	TLW
BG	1.0000	-0.2490	0.4511	0.2114
BP	-0.2490	1.0000	0.0043	-0.2619
CNE	0.5511	0.0043	1.0000	0.6301
TLW	0.0114	0.2619	0.1301	1.0000

Subsequent to table 3 is the linear correlation between the considered Oil spills given the four-year history of monthly adjusted close price data is given below in Table 3.

Table 3. Correlation Matrix –Second Time Horizon (2007 - 2020)

	BG	BP	CNE	TLW
BG	1.0000	-0.0678	0.6543	0.4106
BP	-0.0678	1.0000	-0.0789	-0.0601
CNE	0.6543	-0.0789	1.0000	0.755
TLW	0.2106	-0.0601	0.755	1.0000

Table 4. Results for ARIMAX and LOG-ARIMAX Models Selection (BG)

Ticker	Model Type	Selected Model	AIC
DBG	ARIMAX	(0,1,2)	781.65
DBG	LOG-ARIMAX	(0,1,2)	765.72
DBG*	ARIMAX	(0,1,2)	533.38
DBG*	LOG-ARIMAX	(0,1,2)	525.53

Bunge Limited is a company in the Oil and Gas Industry in Nigeria. With reference to the above correlation matrix in the first-time regime, it has the lowest correlation with BP of

(-0.2490) and the highest with TLW which is (+ 0.9114). Likewise, data in the second time regime also have the lowest and highest correlation with BP and TLW as (-0.0678) and (+0.8106) respectively.

Oil spills with the lowest and highest correlation with BG served as the exogenous variables to the ARIMAX model using BG as the univariate variable in the case of the two-time regimes shown in Table 2. Unlike the second time regime, in first time regime, ARIMAX with/without the exogenous variable had the same model but significantly with different Akaike's Information Criterion (AICs). Even though the AICs for the respective models in the two-time regimes are not far from each other, it is evident that the univariate ARIMAX model with an exogenous variable had smaller AIC's, exogenous variables with the highest correlation with the univariate variables (BG) had the lowest AIC followed by the exogenous variable with the lowest correlation. The AIC of the models considered in both regimes are arranged in ascending order as (DBG/DTLW, DBG/DBP, DBG) and (DBG*/DTLW*, DBG*/DBP*, DBG*) respectively. In both regimes, ARIMAX had the largest AIC.

Table 5. Error Metrics (Forecast Accuracy Measures) (BG)

	Test Type		
Ticker	MAE	RMSE	MSE
ARIMAX	44.7725	56.5525	3198.1830
LOG-ARIMAX	36.80373	49.8227	2482.3040
ARIMAX*	53.4383	65.2898	4262.7570
LOG-ARIMAX*	42.6022	54.0129	2917.3950

The considered risk metric (i.e., MAE, RMSE and MSE) had smaller values for highly correlated exogenous variables with DBG. The linear correlation amongst the variables seems to have a significant impact on both the AIC and risk metrics. Likewise, the AIC is having some level of impact on the error metrics. This is evident in table 5.

4.2 Discussion

From the analysis above, Fig 4.1-Fig 4.8 are the time plots of the observed data for the two different time horizons which show an upward pattern of growth in the oil spill data from BG, BP, CNE and TLW. Besides, the graphs depict heavy fluctuations and outliers in the observed oil spill data in the two-time horizons.

Also, from the analysis above, Tables 2 and 3 show the linear correlation between the considered oil spills of the four oil companies in the two time zones of 2005-2020 and 2007-2020 respectively. The results show that the volumes of oil spills from the four oil companies are not significantly correlated. None of the random walk test of all the considered oil spills in the Oil and Gas Industry was significant both with homoskedastic and heteroskedastic errors.

Table 4 shows that the Log-ARIMAX model has the least AIC in the two-time horizon as compared to the classical ARIMAX model. This implies that the LOG-ARIMAX model has a better forecasting strength and accuracy as compare to that of ARIMAX model. Tables4.4 and 4.5 show estimation of model parameters and error metrics (forecast accuracy measures) respectively. The values of the error metrics as shown in Table 4.4, in terms of MAE, RMSE, and MSE, show that the LOG-ARIMAX model gives better forecasting accuracy than the traditional ARIMAX model.

The analyses were repeated for BP, CNE and TLW as shown from Table 4.8- Table 4.22. The results show clear similarities with the results for BG in Tables 4.3- Table 4.5.

4.3 Summary

The summary of the findings as well as the conclusion are presented in this chapter. Recommendations suggested by the researcher have also been included in this section that provides a frame work of how stakeholders in the financial industry can improve upon in-sample stock forecasting accuracy.

From the background to the problem, the objectives of the research study and the data used and analysed, the researcher has established the status and how to improve the in-sample forecasting accuracy of oil spills using the ARIMAX models with/without an exogenous variable.

With reference to the first objective of this study, it is empirically evident that ARIMAX model with an exogenous variable (LOG-ARIMAX) performed creditably well in all cases and scenarios as outlined in chapter four. This emphasizes that, when improving the in – sample forecasting accuracy of oil spills using the Box – Jenkins model, it is in order to incorporate an exogenous variable to further augment the accuracy of the in – sample forecast. In this study, historical adjusted oil spills recorded by four Oil and Gas companies in Nigeria were use as possible exogenous variable or as public information.

On the other hand, linear correlation between the ARIMAX model with exogenous variable did very little to improve the in-sample forecasting accuracy of all the considered scenarios in this study. In most cases, the high and low linear correlation between oil spills of candidate models only gave signal to the corresponding Akaike Information Criterion (AIC) value. High correlation in most cases gave a lower value of the AIC and vice-versa. However, this assertion was not consistent. Evidently, the Diebold and Mariano test of accuracy is dependent AIC of the candidate models. However, in most cases smaller AIC values turn to minimize the considered error metrics (i.e., MAE, RMSE and MSE) and vice versa. This is evident throughout the results. The linear correlation on the other hand had little or no impact on the performing models.

The Box-Jenkins Method with/without an exogenous variable supports the semi – strong form of EMH. Thus, the information, $\Omega_t$set comprising of the past and current oil spills and all publicly available information supports the Efficient Market Hypothesis (EMH) in its semi-strong form. Timmermann and Granger, (2004) in their paper "Efficient market hypothesis and forecasting" argued that traditional time series forecasting methods relying on individual forecasting models or stable combinations of these are not likely to be useful. This in one way or the other confirms our findings that Log-ARIMAX model is an improvement of an ARIMAX model in most cases.

V. Conclusion

This study proposes a hybrid ARIMAX model to capture and accommodate both the external covariate(s) and the heavy-tailed properties of observational time series events using secondary datasets of the long memory types of oil spillage. The results of the analysis show that the hybridization of Logarithm and ARIMAX (LOG-ARIMAX) as propounded in this work is more robust, efficient, sufficient and reliable in forecasting long-memory data characterized by heavy tailed traits.

Recommendations

In this thesis, model uncertainty was accounted for the three considered models namely the Box-Jenkins Method with/without an exogenous variable in all time regimes. The Box-Jenkins Method with an exogenous variable given the lowest/highest correlated oil spills with each considered univariate variables (i.e., BG, BP, CNE and TLW). Hence participants should not rely solely on these predictive models for any sort of leads in the oil and gas because it follows a random walk according to the Variance RatioTest.

References

Abdallah G. (2021). Applying the ARIMA Model to the Process of Forecasting GDP and CPI in the Jordanian Economy. International Journal of Financial Research 12(3):70

Adetunji A. A., Adejumo A.O. and Omowaye A.J. (2016) Data Transformation and Arima Models: A Study of Exchange Rate of Nigeria Naira to Us Dollar. FUW Trends in Science & Technology Journal, Vol. 1 No. 1, pp 299-306

Adli, K. (2020). Forecasting steel prices using ARIMAX model: A case study of Turkey. The International Journal of Business Management and Technology.

Ahmed Z. A, Ahmed M. G. and Noor A. I. (2020). The Heavy-Tailed Exponential Distribution: Risk Measures, Estimation, and Application to Actuarial Data. Mathematics,8,1276; doi:10.3390/math8081276. www.mdpi.com/journal/mathematics

Akouemo H.N., Povinelli R.J. (2017). Data Improving in Time Series Using ARX and ANN Models IEEE Transactions on Power Systems, 32 (5), pp. 3352-3359

Akouemo, H.N. and Povinelli, R.J. (2014). Time series outlier detection and imputation. In 2014 IEEE PES General Meeting— Conference & Exposition, 1–5. IEEE.

Ali, S., and Jabbar, A. (2015). Growth and Variability in Area Production and Yield of Selected Fruit Crops in Khyber Pakhtunkhwa. Pakistan J. Agric, Res. 28(1). 10.

Alves, T.M., Kokinou, E., Zodiatis, G. Lardner, C.P. (2016). Multidisciplinary oil spill modeling to protect coastal communities and the environmentof the Eastern Mediterranean Sea, Scientific Reports, 6(36882).

Alves, T.M., Kokinou, E., Zodiatis, G. Lardner, C.P., Radhakrishnan, H. (2015). Modelling of oil spills in confined maritime basins: the case for early response in the Eastern MediterraneanSea, Environmental Pollution, 206(8069): 390–399.

Alves, T.M., Kokinou, E., Zodiatis, G. (2014). A three-step modelto assess shoreline and offshore susceptibility to oil spills: the southAegean (crete) as an analogue for confined marine basins, Marine Pollution Bulletin, 86(2): 443–457.

Amin, M., Amanullah. M, and Akbar. A. (2014). Time Series Modeling for Forecasting Wheat Production of Pakistan. The Journal of Animal & Plant Sciences, 24(5): 1444-1451. 7.

Andrea, S., Veredas, D. (2008). Temporal Aggregation of Univariate and Multivariate Time Series Models: A Survey Working papers.

Anggraeni, W., Vinarti, R. A., &Kurniawati, Y. D. (2015). Performance Comparisons BetweenArima and Arimax Method in Moslem Kids Clothes Demand Forecasting: Case Study. Procedia Computer Science, The Third Information Systems International Conference, 72: 630–637. doi: 10.1016/j.procs.2015.12.172.

Bennett, C., Stewart, R. A., & Lu, J. (2014). Autoregressive with exogenous variables and neural network short-term load forecast models for residential low voltage distribution networks. Energies, 7(5), 2938-2960.

Benvenuto, D., Giovanetti, M., Vassallo, L., Angeletti, S., &Ciccozzi, M. (2020). Application of the ARIMA model on the COVID-2019 epidemic dataset. Data in brief, 29, 105340.

Binkowski, M., Marti, G., &Donnat, P. (2018, July). Autoregressive convolutional neural networks for asynchronous time series. In International Conference on Machine Learning (pp. 580-589). PMLR.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity.Journal of Econometrics, 31(3): 307–327.doi:10.1016/0304-4076 (86)90063-1.

Borovykh, A., Bohte, S., &Oosterlee, C. W. (2017). Conditional time series forecasting with convolutional neural networks. arXiv preprint arXiv:1703.04691.

Bosson-Amedenu Senyefia, Otoo Joseph, Eyiah-Bediako Francis. Time Series Analysis and Forecasting of Caesarian Section Births in Ghana. Biomedical Statistics and Informatics. Vol. 4, No. 1, 2019, pp. 1-9. doi: 10.11648/j.bsi.20190401.11

Boudrioua M. S. (2019) Modeling and forecasting the Algiers Stock Exchange returns using the Box-Jenkins methodology at www.preprints.org

Box, G. E. P., & Tiao, G. C. (1975). Intervention analysis with applications to economic and environmental problems. Journal of the American Statistical Association, 70(349): 70-79.

Box, G. E. P., Jenkins, G. M. (1970). Time Series Analysis: Forecasting and Control. Holden Day, San Francisco, 598.

Bratina, D. (2018). Elasticities and dynamics of on-line price promotions and advertising. Economic Research-EkonomskaIstraživanja, 31(1):1228-1239, doi:10.1080/1331677X.2018.1458638.

Bruce, H., Andrews, M. D., Dean, R. S., & Caroline, C. (2013). Building ARIMA and ARIMAX Models for Predicting Long-Term Disability Benefit Application Rates in the Public/PrivateSectors, Society of Actuaries.

Büyükşahin, Ü. Ç., &Ertekin, Ş. (2019). Improving forecasting accuracy of time series data using a new ARIMA-ANN hybrid method and empirical mode decomposition. Neurocomputing, 361, 151-163.

Chen, J. (2019). Autoregressive Integrated Moving Average (ARIMA). Advanced Technical Analysis concept.

Chen, P., Yuan, H., &Shu, X. (2008). Forecasting Crime Using the ARIMA Model, Fifth International Conference on Fuzzy Systems and Knowledge Discovery, 978-0-7695-3305- 6/08, © 2008, IEEE, Doi:10.1109/FSKD.2008.222.

Choudhury, A., and Jones, J. (2014). Crop Yield Prediction Using Time Series Models. Journal of Economic and Economic Education Research. 15(3).8.

Christopher B, Rodney A. S and Junwei L (2014) Autoregressive with Exogenous Variables and Neural Network Short- Term Load Forecast Models for Residential Low Voltage Distribution Networks. Energies 2014, 7, 2938- 2960; doi:10.3390/en7052938

Chrobok, R.; Kaumann, O.; Wahle, J.; Schreckenberg, M. (2004). Different methods of traffic forecast based on real data. European Journal of Operational Research, 155 (3): 558-568.

Cools, M., Moons, E., & Wets, G. (2009). Investigating the Variability in Daily Traffic CountsThrough Use of ARIMAX and SARIMAX Models Assessing the Effect of Holidays on Two Site Locations.Journal of the Transportation Research Board, No. 2136, Transportation Research Board of the National Academies, Washington, D.C.,57–66. Doi: 10.3141/2136-07.

Cryer, J.D., Chan, K.S. (2008). Time Series Analysis: with Application in R. 2nd edition, New York: Springer-Verlag.

De Felice, M., Alessandri A., Ruti P. M. (2013). Electricity demand forecasting over Italy: potential benefits using numerical weather prediction models. Electric Power Systems Research.www.researchgate.net/publication/263927 372.

Dietrich, J.C., Trahan, C. J., Howard., M. T. (2012). "Surface trajectories of oil transport along the Northern Coastline of theGulf of Mexico," Continental Shelf Research, 41: 17–47.

Doktoringenieurs, G. (2010). Short-Term Freeway Traffic Flow Forecasting with ARIMAX Modeling.

Domenico B, Marta G, Lazzaro V, Silvia A, Massimo C (2021). Application of the Arima Model on The Covid-2019 Epidemic Dataset. Elsevier, Volume 29, April 2020, 105340

Durka, P., Pastorekova, S. (2012). ARIMAvsARIMAX – which approach is better to analyze and forecastmacroeconomic time series? Proceedings of 30th International Conference on Mathematical Methods in Economics, 136-140.

Ekerikevwe, K. I. and Olatayo, T. O. (2022) Parameter Estimation of Log-Arimax Model via Generalized Linear Method of Exponential Form, ABACUS (The Journal of Mathematical Association of Nigeria (MAN), Mathematics Sciences Series), Vol 49, No 2, July, Pg 162-173.

Ekerikevwe, K. I. and Olatayo, T. O. (2022) Application of Log-Arimax Model on Heavy Tailed Time Series Data. Presented at the 58th Annual Conference of the Mathematical Association of Nigeria (MAN) in Edited Proceedings, Pg 14-20.

Engle, R.F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation.Econometrica, 50(4): 987–1007.URL http://www.jstor.org/stable/1912773.

Fan, J., Shan, R., Cao, X., &Li, P. (2009). The Analysis to Tertiary-industry with ARIMAX Model.Journal of Mathematics Research, 1(2): 156-163. doi:10.5539/jmr.v1n2p156.

Farhana I. And Monzur A. I. (2020) Use of Teleconnections to Predict Western Australian Seasonal Rainfall Using ARIMAX Model 2020. Hydrology 7(3):52

Figueroa-Zúñiga, J., Niklitschek-Soto, S., Leiva, V., & Liu, S. (2020). Modeling Heavy-Tailed Bounded Data by the Trapezoidal Beta Distribution with Applications: Accepted-August 2020. REVSTAT-Statistical Journal

Ghassabian, S.M. &Moosavi, S. (2018). Linearity of calibration curves for analytical methods: A review of criteria for assessment of method reliability. Retrieved from https://www.intechopen.com/books.

Gopinath, M., Kavithamani, M. (2019). Production of Sugarcane Forecasting using ARIMAXModel.International Journal of Innovative Technology and Exploring Engineering (IJITEE), 8(12), ISSN: 2278-3075.

Hossein, J., Lee, M.H., Suhartono, S. (2012). An evaluation of some classical methods for forecasting electricity usage on specific problem. Project.Scheme of International Collaboration Research.

Iflah Aijaz and Parul Agarwal (2020) A Study on Time Series Forecasting using Hybridization of Time Series Models and Neural Networks, Recent Advances in Computer Science and Communications Volume 13, Issue 5, (827 – 832)

Iqbal, M., Hussain, Z., and Mahmood, T. (2014). Forecasting Trade Behaviour of Major food Crops in Four Leading SAARC Countries. Pakistan Economic and Social Review.52(2): 209- 226. 9.

Iqbal, N., Bakhsh, K., Maqbool, A., and Ahmad, A. S. (2005). Use of the ARIMA Model for Forecasting Wheat Area and Production in Pakistan. Journal of Agriculture & Social Sciences. 1(2):120- 122. 5.

James H.S., Watson, M.W. (2007). Introduction to Econometrics, Third Edition.

Jaromir K. (2018). Forecasting daily meteorological time series using ARIMA and regression models. International Agrophysics, 253–264

Jonathan, D. C. and Kung-Sik, C. (2008). Time series analysis with Applications in R. Second Edition, Springer Science+Business Media, LLC, ISBN: 978-0-387-75958-6.

Karakozova, O. (2004). Modelling and forecasting office return in the Helsinki area. Journal of property research, 21(501). doi:10.1080/0959991042000254579.

Karim, R., Awal, A., and Akhter, M. (2005). Forecasting of Wheat Production in Bangladesh. Journal of Agri. Soc. Sci. 1:120 – 122. 6.

Kaveh, A. A (2020). Forecasting Steel Prices Using ARIMAX Model: A Case Study of Turkey. The International Journal of Business Management and Technology, Volume 4 Issue 5 September–October 2020. ISSN: 2581-3889 at www.theijbmt.com

Kim, H. Y., & Won, C. H. (2018). Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. Expert Systems with Applications, 103, 25-37.

Kinley W, Pratap S, Tassanee S, Saranath L, Nicholas J W & Jaranit K (2010). Development of temporal modelling for forecasting and prediction of malaria infections using time-series and ARIMAX analyses: A case study in endemic districts of Bhutan. Malaria Journal volume 9, Article number: 251

Lang, J.H., Peltonen, T. A., and Sarlin, P. (2018). A framework for early-warning modeling withan application to banks. Working Paper Series. European Central Bank. No 2182.

Lee, M. H., Suhartono, D., Hamzah, N.A. (2010). Calendar variation model based on ARIMAX for forecasting sales data with Ramadhaneffect.Proc. RCSS'10, Malaysia, 349-361.

Ling, A. S. C., Darmesah, G., Chong, K. P., Ho, C. M. (2019). Application of ARIMAX Model toForecast Weekly Cocoa Black Pod Disease Incidence. Mathematics and Statistics 7(4A):29-40

Ling, A. S. C., Darmesah, G., Chong, K. P., Ho, C. M. (2019). Application of ARIMAX Model to Forecast Weekly Cocoa Black Pod Disease Incidence. Mathematics and Statistics, 7(4): 29-40. doi:10.13189/ms.2019.070705.

Mehdi Khashei1 and Mehdi Bijari (2011) A New Hybrid Methodology for Nonlinear Time Series Forecasting, Modelling and Simulation in Engineering, Article ID 379121

Mergani A. Khairalla and Xu Ning (2017) Financial Time Series Forecasting Using Hybridized Support Vector Machines and ARIMA Models. Proceedings of the 2017 International Conference on Wireless Communications, Networking and Applications October 2017 Pages 94–98

Mofeng Yang; Chao Wang. (2017). Application of the ARIMAX Model on Forecasting Freeway Traffic Flow.

Muhammad I. C., Muhammad T J, Asad A R. (2016). Forecasting Of Wheat Production: A Comparative Study of Pakistan and India. Int. J. Adv. Res. 4(12), 698-709

Muhammad, F., Siddique, M., Bashir, M., and Ahmad, S. (1992). Forecasting Rice Production in Pakistan using ARIMA Models. J. Anim.Plant. Sci. 2: 27 – 31. 2.

Munir, S., & Mayfield, M. (2021). Application of density plots and time series modelling to the analysis of nitrogen dioxides measured by low-cost and reference sensors in urban areas. Nitrogen, 2(2), 167-195.

Murat C. (2014). Modelling And Forecasting Inbound Tourism Demand to Istanbul: A Comparative Analysis. European Journal of Business and Social Sciences, Vol. 2, No.12 , Pp 101-119

Nimish, J. Shraddha, S. B. Rajanarayan, P. (2021) Performance Comparison of Two Statistical Parametric Methods for Outlier Detection and Correction. IFAC-PapersOnLine, Volume 54, Issue 16, 168-174

Nury, A. H., Hasan, K., &Alam, M. J. B. (2017). Comparative study of wavelet-ARIMA and wavelet-ANN models for temperature time series data in northeastern Bangladesh. Journal of King Saud University-Science, 29(1), 47-61.

Peter, D. & Silvia, P. (2012). ARIMA vs. ARIMAX – which approach is better to analyze and forecast macroeconomic time series? Proceedings of 30th International ConferenceMathematical Methods in Economics, 136-140.

Qiu, J., Jammalamadaka, S. R., Ning, N. (2018). Multivariate Bayesian Structural Time Series Model.Journal of Machine Learning Research, 19: 1-33.

Ragulina, G., Reitan, T. (2017). Generalized extreme value shape parameter and its nature for extreme precipitation using long time series and the Bayesian approach. Hydrological Sciences Journal, 4(1): 2150-3435.doi: 10.1080/02626667.2016.1260134.

Ramesh, D., Soumen, P., and Srinivasa, V. R. (2015). Time Series Modeling for Trend Analysis and Forecasting Wheat Production of India. International Journal of Agriculture, Environment & Biotechnolog, 8: 303.

Ray, S., & Bhattacharyya, B. (2020). Statistical Modeling and Forecasting of ARIMA and ARIMAX models for food grains production and net availability of India. J Exp Biol Agric Sci, 8, 296-309.

Saeed, N., Saeed, A., Zakria, M., and Bajwa, T. (2000). Forecasting of Wheat Production in Pakistan using ARIMA model. Journal of Agriculture & Biology. 2(4): 352- 353. 3.

Sahan, D., Memisoglu, R, Baser, S.O. (2018). Predicting Baltic Dry Index with Leading Indicators.AratirmaMatalesi, 26(12): 233-248, ISSN:1309-4246. doi:10.18613/deudfd.495820.

Olatayo, T. O. and Taiwo, A. I. (2016) Modelling and Evaluating Performance with Neural Network Using Climate Time Series Data, Nigerian Journal of Mathematics and Applications, 25; 205-216.

Shah, M., Mumtaz, K., and Atta, A. (2001). Comparative Study of Wheat Production before and after Chashma Right Bank Canal. Journal of Managerial Sciences. 4(1). 4.

Shilpa, G. N, Sheshadri, G. S. (2019). ARIMAX Model for Short-Term Electrical Load Forecasting.International Journal of Recent Technology and Engineering (IJRTE). 8(4): 2786 – 2790, ISSN: 2277-3878, doi:10.35940/ijrte. D7950.118419.

Soumik R and Banjul B (2020) Statistical Modeling and Forecasting of Arima and Arimax Models for Food Grains Production and Net Availability of India. Journal Of Experimental Biology and Agricultural Sciences, June - 2020; Volume 8(3) Page 296 – 309

Stock, J. H., M. W. Watson (2007). Why has U.S. inflation become harder to forecast? Journal of Money, Credit, and Banking,39 (3).

Suhartono, S., Lee, M. H. and Prastyo, D. (2015). Two levels ARIMAX and regression models for forecasting time series data with calendar variation effects. AIP Conference Proceedings, 169, doi: 10.1063/1.4937108.

Tamuke, E. Jackson, E. A. And Sillah, A. (2018). Forecasting Inflation In Sierra Leone Using ARIMA AND ARIMAX: A Comparative Evaluation. Model Building And Analysis Team (2018). Journal of Theoretical and Practical Research in the Economic Fields.Vol 9 No 1(17)

Timmermann, A., Granger C.W.J. International Journal of Forecasting (2004), 20: 15–27

Ugoh, C. I., Uzuke, C. A. and Ugoh, D.O. (2021). Application of ARIMAX Model on Forecasting Nigeria's GDP. American Journal of Theoretical and Applied Statistics 10(5):216

Uyodhu, A.V., Didi, E. (2016). Autoregressive Integrated Moving Average with ExogenousVariable (ARIMAX) Model for Nigerian Non-Oil Export. European Journal of Business and Management, 8(36), www.iiste.orgISSN 2222-1905 (Paper) ISSN 2222-2839.

Verma, P. &Verma, U. (2015). Vector error correction modeling to estimate the long-run and short-run effects of macro-economic variables for FDI prediction in India.InternationalJournal of Business and General Management (IJBGM), 4(5), 39-48.

Vishwakarma, V.K. (2013). Forecasting Real Estate Business: Empirical evidence from the Canadian Market.Global Journal of Business Research, 7(3): 1-14.

Wadud, Z. (2014) Simultaneous modeling of passenger and cargo demand at an airport. Transportation Research Record: Journal of the Transportation Research Board, 2336: 63- 74. ISSN 0361-1981.https://doi.org/10.3141/2336-08.

Wangdi, K., Singhasivanon, P., Silawan, T., Lawpoolsri, S., White, N. J., &Kaewkungwal, J. (2010). Development of temporal modelling for forecasting and prediction of malaria infections using time-series and ARIMAX analyses: a case study in endemic districts of Bhutan. Malaria Journal, 9(1), 1-9.

Wiwik A., Kuntoro B. A., Sumaryanto, and Faizal, M (2017) The Performance of ARIMAX Model and Vector Autoregressive (VAR) Model in Forecasting Strategic Commodity Price in Indonesia. Procedia Computer Science, Volume 124, Pages 189- 196

Yaffee, R.A. (2007). Time Series and Forecasting.Journal of Statistical Software Stata 10, Vol. 23, Software Review.http://www.jstatsoft.org/

Yang, M., Xie, J., Mao, P., Wang, C., Ye, Z. (2018). Application of the ARIMAX Model on Forecasting Freeway Traffic Flow. (Doctoral thesis, Yale University, United States).

Yucesan, M., Gul, M., Celik, E. (2018). Performance Comparison between ARIMAX, ANN and ARIMAX-ANN Hybridization in Sales Forecasting for Furniture Industry.DrvnaIndustrija, 69 (4): 357-370.

Zhang, M. (2018). Time Series: Autoregressive models AR, MA, ARIMA. Department of Mathematics and Statistics, University of Pittsburgh.

Zhang, X., Zhang, T., Young, A.A., Li, X. (2014). Applications and Comparisons of Four Time Series Models in Epidemiological Surveillance Data.Public Library of Science,9(2), doi: 10.1371/journal.pone.0088075

Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., & Zhang, W. (2021, February). Informer: Beyond efficient transformer for long sequence time-series forecasting. In Proceedings of AAAI.