Statistical Analysis of the Impact of Maritime Transport Gross Domestic Product on Nigeria's Economy

K. P. Oyeduntan, K. Oshinubi

***Abstract*—**Nigeria is referred as the 'Giant of Africa' due to high population, land mass and large economy. However, it still trails far behind many smaller economies in the continent in terms of maritime operations. As we have seen that the maritime industry is the sparkplug for national growth, because it houses the most crucial infrastructure that generates wealth for a nation, it is worrisome that a nation with six seaports lag in maritime activities. In this research, we have studied how the Gross Domestic Product (GDP) of the maritime transport influences the Nigerian economy. To do this, we applied Simple Linear Regression (SLR), Support Vector Machine (SVM), Polynomial Regression Model (PRM), Generalized Additive Model (GAM) and Generalized Linear Mixed Model (GLMM) to model the relationship between the nation's Total GDP (TGDP) and the Maritime Transport GDP (MGDP) using a time series data of 20 years. The result showed that the MGDP is statistically significant to the Nigerian economy. Amongst the statistical tool applied, the PRM of order 4 describes the relationship better when compared to other methods. The recommendations presented in this study will guide policy makers and help improve the economy of Nigeria.

***Keywords*—**Economy, GDP, maritime transport, port, regression.

I. Introduction

ANALYZING the impact of maritime transport gross domestic product on Nigeria's economy provides valuable insights into the country's maritime sector's contribution to overall economic growth and development. Numerous studies have identified the maritime industry as a key driver of national progress and advancement due to its hosting of vital infrastructure that generate value for the country [1]. Even though Nigeria is the largest economy in Africa with a GDP of $432 Billion US dollars in 2020 [2], it still dawdles behind many smaller economies in Africa in terms of ports and maritime operations. For example, in 2020, Nigeria's container port throughput was reported at 1,528,520 TEU compared to Morocco at 6,980,958 TEU, South Africa at 4,029,000 TEU, and Egypt at 5,928,454 TEU [3]. It has been proved that over 90% of the volume of global trade is conveyed by sea and the economic growth of any country depends greatly on its performance in the international market [4]. However, it would not be possible to convey these marketable goods and perform effectively in the global market without efficient maritime transport [5]. This shows that the maritime sector is an important component in facilitating international trade.

Fig. 1 shows Map of Nigeria, which is the focus of this study, it is a globally identified maritime country that is endowed with 13,000 square kilometers of navigable water area. It is in West Africa and borders Niger in the north, Chad within side in the northeast, Cameroon in the east, the Republic of Benin within side from the west, and borders the Gulf of Guinea to the south, masking a place of 923,769 km^2 with a populace of over 200 million [6].

Nigeria has six seaports that it uses for the purpose of maritime transport. For many years, these ports have been bedeviled by infrastructure and logistics deficiency, traffic congestion, security issues, policy and summersaulting regulations, intersecting functions, and replicating roles among the Ministries, Departments and Agencies [7]. These had jagged the growth, development and efficiency of this vital cash-cow sector that strongly holds the beacon of hope and survival of a growing economy like Nigeria. No doubt, these lingering challenges has created avenues for gross misallocation of resources and imposes tremendous costs to the economy, distorting development policies and undermining confidence of foreign and local maritime stakeholders. Considering the current state of ports and maritime activities in Nigeria, it is important to investigate if the maritime transport contributes significantly to the nation's GDP.

Only a few studies have investigated the relationship between the maritime industry and Nigeria's GDP. For example, [4] modeled the econometric impact of shipping on the Nigerian economy and tested the significance. The study adopted analytical method of regression analysis with a time series data of 10 years; from 2003-2012. The study concludes that shipping has made an enormous impact on the Nigerian economy over the period covered in the study.

A study on the impact of maritime logistics factors on the Nigerian GDP was also conducted by [8]. The study examined maritime logistics factors on Nigeria's GDP with the view of evaluating the relationship as well as influence of the logistics factors (Vessel Movement (VM), and Cargo Throughput) in term of costs on Nigeria's GDP. The result showed that, though the influence of the logistics factors except cargo throughput on Nigeria GDP is negatively insignificant, they contributed 29.1% to the improvement of Nigeria GDP.

K. P. Oyeduntan is with Medlog Transports and Logistics & International Logistics and Port Management Department, EM Normandie Business School, Le Havre, 76600 France (e-mail: kehindepeter.oyedu08@em-normandie.fr).

K. Oshinubi is with Laboratory AGEIS EA 7407, Team Tools for e-Gnosis Medical, Faculty of Medicine, University Grenoble Alpes (UGA), 38700 La Tronche, France (corresponding author, phone: +33 7 66 76 05 29; e-mail: kayode.oshinubi@univ-grenoble-alpes.fr).

Fig. 1 Map of Nigeria [6]

Reference [9] investigated the relationship between the cargo throughput and the maritime GDP, the data were modeled in a SLR, and it showed that the economy of the Nigerian maritime transport depends on the cargo throughput by 68%. Further results indicated that, the concession only had impact on the Nigerian port till 2011, afterwards, the growth rate for both the cargo volume and maritime GDP has been diminishing.

In this paper, we will use different statistical methods to investigate the relationship between the maritime GDP and the nation's overall GDP. Analytical tools such as, SLR, Polynomial regression, SVMs, Generalized Additive Method and GLMM will be used to analyze the correlation between them and predict the TGDP. The main objective of this study is thus to evaluate the connection between the national GDP and the MGDP and to identify which statistical technique best describe the relationship. It is expected that, apart from contributing to literature on the scope of maritime transport, this paper will also assist Nigerian maritime policy makers to make effective administration and management decisions.

This research work is thus divided into four sections: following the introduction, Section II describes the methods used to process the data, whilst Section III discusses the results and demonstrate the efficiency of the models in comparison to each other. Finally, Section IV concludes this paper by highlighting major findings from our research and proposed some perspectives to improve the maritime transport industry.

II. Materials and Methods

The data used in this study were obtained from the Central Bank of Nigeria statistical bulletin [10]. Based on the availability of the information, and to position our study on contemporary trends, we processed data from 1999-2018. The statistical methods employed are SLR, Polynomial Regression, SVMs, General Additive Method and GLMM. The simulation, data visualization and computation were done in R environment. The data and the code used for the analysis link are provided in the data availability section.

A. Simple Linear Regression

In order to examine the relationship between a quantitative result and a single quantitative illustrative variable, SLR is the foremost considered investigation strategy [11]. The two factors that are involved in SLR analysis are depute X-dependent variable and Y-independent variable. The equation which shows the relationship between these two variables can be expressed using the structural model:

$$E(Y) = \beta_0 + \beta_1 X + \varepsilon \quad (1)$$

where β_0 is the intersection of the regression line; β_1 is the slope; $E(Y)$ is the anticipated value of Y for a given value of X; ε – Error is used to explain variation in Y that cannot be explained by a linear relationship between X and Y. If ε is not present, this means that knowing X information would determine the value of Y.

B. Support Vector Machines

Support vector regression is a supervised machine learning model which draws a hyperplane relating the data points and creates a boundary of possible data points (high and low) in future [12]. The goal of SVM is the same as the goal of the classification problem, which is to find the maximum amplitude, that is, to minimize errors. In the case of regression, the tolerance margin (epsilon) is fixed as an approximate value of the SVM. The main goal is to minimize the error by personalizing the hyperplane to maximize the amplitude, keeping in mind that some error is tolerated.

C. Polynomial Regression

When there exists a non-linear relationship between dependent and independent variables, the dependent variable Y is modelled as a polynomial of degree n in X [12], [13]. The equation of polynomial takes this form:

$$E(Y) = \beta_0 + \beta_1 X + \beta_2 X^2 + \ldots\ldots\ldots + \beta_n X^n + \varepsilon \quad (2)$$

where n is referred as the degree of the polynomial. If we increase the value for n, the model can fit non-linear relationships better, but in practice n is usually not greater than 3 or 4. Beyond this point, the model becomes too flexible and overfits the data. For this research, we shall use the polynomial of order 4.

D. Generalized Linear Mixed Model

An improvement of (3) which takes into consideration random effect is given as:

$$Y_i = X_i\beta + Z_i b + \epsilon_i \quad (4)$$

where X_i are the covariates, β is a constant, $\epsilon_i \sim N(0, P\sigma 2)$ is a residual error vector, $b \sim N(0, \varphi\emptyset)$ when $\varphi\emptyset$ is positive definite covariance matrix for random effects b and it also depends on some parameters $\emptyset$, Z_i which will be the main target of statistical inference about the random effects is a matrix of fixed coefficients describing how the response variable y_i depends on the random effects and P is a positive definite matrix used to model residual correlation, but it is often simply an identity matrix [14].

E. Generalized Additive Model

The main advantage of GAMs over traditional regression methods is their capability to model non-linear relationships which is a common feature of many data sets [15]. It does this between a response variable and multiple covariates using non-parametric smoothers. The general formula of a GAM is:

$$E(Y) = \beta_0 + f(X_t|\beta_j) + \varepsilon \quad (5)$$

where the change of Y over time is represented by the smooth function $f(X_t|\beta_j)$ with inputs as the covariates X_t and parameters $\beta_{(<}$ and ε represents the residual error.

III. Results

This section explains the out-turn from the statistical techniques used to analyze the data. The data were modelled using five statistical tools – SLR, SVMs, Polynomial Regression, GLMM and General Additive Method. The results from these models were subjected to further testing and finally compared. Amongst the different methods, the Polynomial Regression Method described appropriately the relationship between the GDP of the Maritime Transport industry and the national economy with minimum error.

A. Simple Linear Regression Model

TABLE I
RESULT ON ANALYSIS BY SIMPLE REGRESSION MODEL

Residuals:				
Min	1Q	Median	3Q	Max
-11092129	-5666586	-591674	5970119	12933338
Coefficients:				
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-22080916	3560910	-6.2	7.5e-06 ***
MGDP	15088	653	23.1	7.8e-15 ***
Residual standard error: 7240000 on 18 degrees of freedom				
Multiple R-squared: 0.967		Adjusted R-squared: 0.966		
F-statistic: 535 on 1 and 18 DF		p-value: 7.77e-15		

From the analysis of Table I, using simple regression analysis, the quantitative relationship existing between the GDP of the Maritime Transport and the GDP of the Nigerian economy over the period under study (1999-2018) can be expressed by:

$$TGDP = 15088MGDP - 22080916 + \varepsilon \quad (6)$$

Equation (7) represents the contribution trend of maritime transport to the development of the Nigerian economy over the 20 years period covered in this research. The P-value, 7.77e-15 shows that the MGDP is statistically significant in the contribution to the overall GDP. When building regression models, we hope that this p-value is less than 0.005 because it indicates that the predictor variable is useful for predicting the value of the response variable.

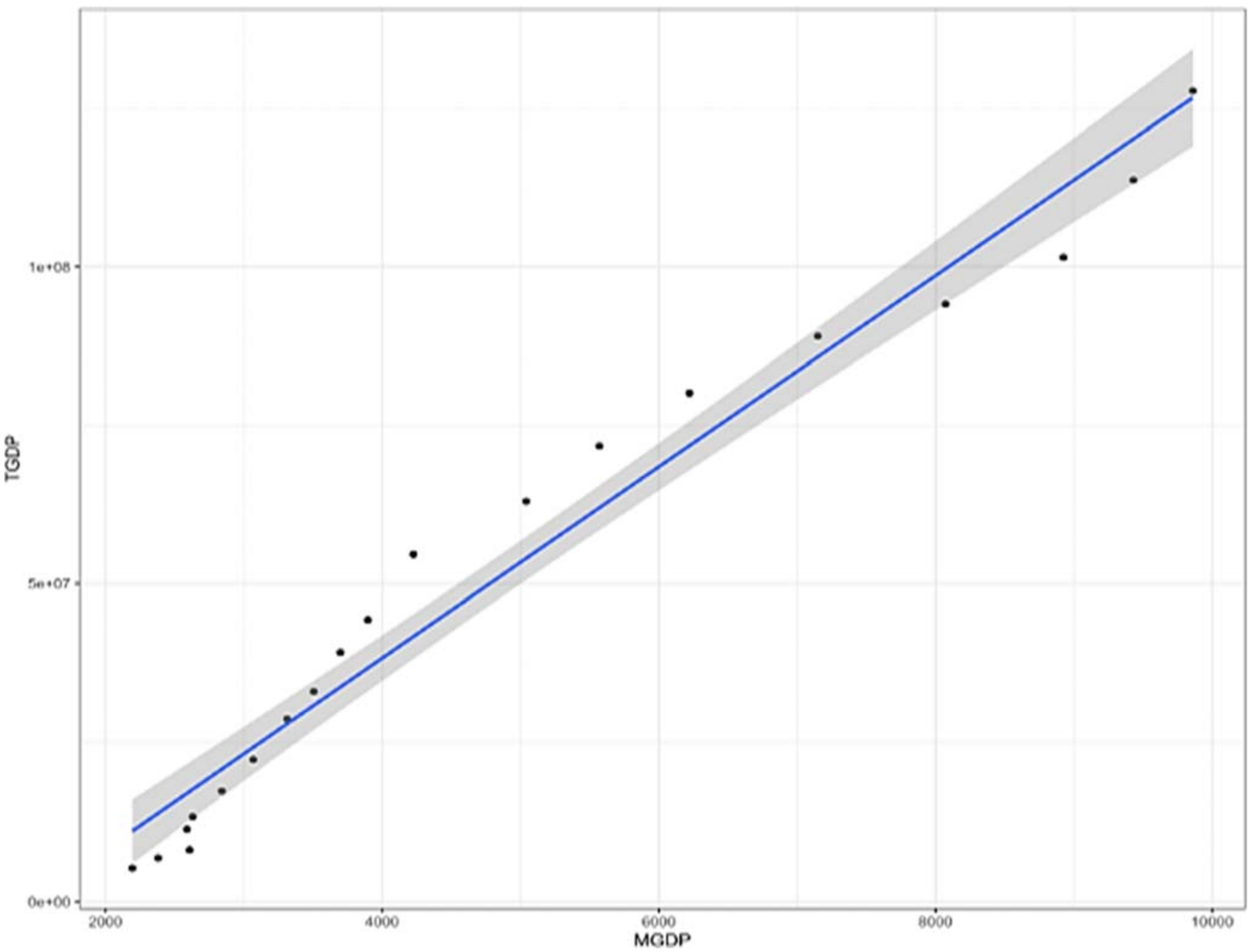


Fig. 2 Linear relationship between MGDP and TGDP

From Fig. 2, we observe that for the period under review, an increase in the GDP of the maritime transport leads to the increase in the overall GDP of the country. Hence, we can deduce that the Nigerian economy depends on the growth of the maritime industry. With this analysis, the RMSE = 7.335e+13. We further performed Jarque - Bera Normality Test to check the autocorrelation between both data, we found the asymptotic p-value = 0.5199.

The linear residual vs fitted graph indicated in Fig. 3 shows that, linearity does not seem to hold reasonably well for these data. The graph looks like a quadratic curve. For the normal Q-Q test, we can see that the points do not lie approximately on the line, this means that the data are not normally distributed. Hence, we can see that the SLR does not appropriately describe and predict the relationship between MGDP and TGDP.

B. Support Vector Machine Model

We created the model of our data in SVM. We performed an epsilon regression and did not set any value for epsilon and cost; however, it took the default value of 0.1 and 1 respectively. But, to improve the performance of the support vector regression, we carried out the hyperparameter optimization by doing a grid search.

The code returned the best parameters and performance as; Epsilon: 0, Cost: 5 and best performance: 2.442332e+13. From Fig. 4, we can see that the darker region is the best fit for our model because the RMSE is closer to zero in the darker regions. We obtained a RMSE of 3.4e+6 and R-squared of 0.987223.

In Fig. 5, it is shown that our first model in red and the tuned model in blue. We can see that the SVR model takes a better fit than the SLR model. We can justify this by comparing the RMSE. We further performed the Jarque-Bera Normality Test and obtained asymptotic p-value of 0.007441.

C. Polynomial Regression Model

The polynomial regression analysis is presented in Table II. We decided to use polynomial of order 4 and tested 40% of the data while 60% were trained. The prediction result is in Fig. 7 (b).

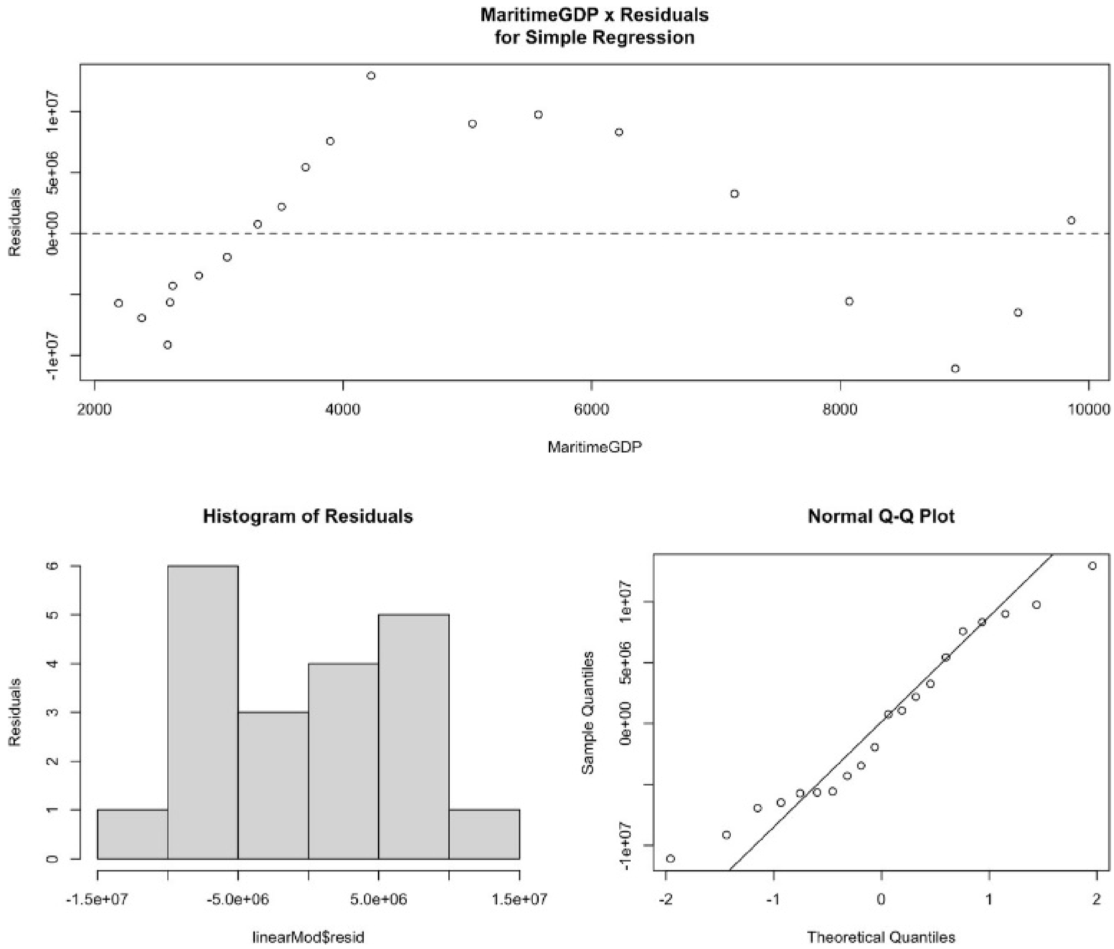


Fig. 3 Linear regression Plots

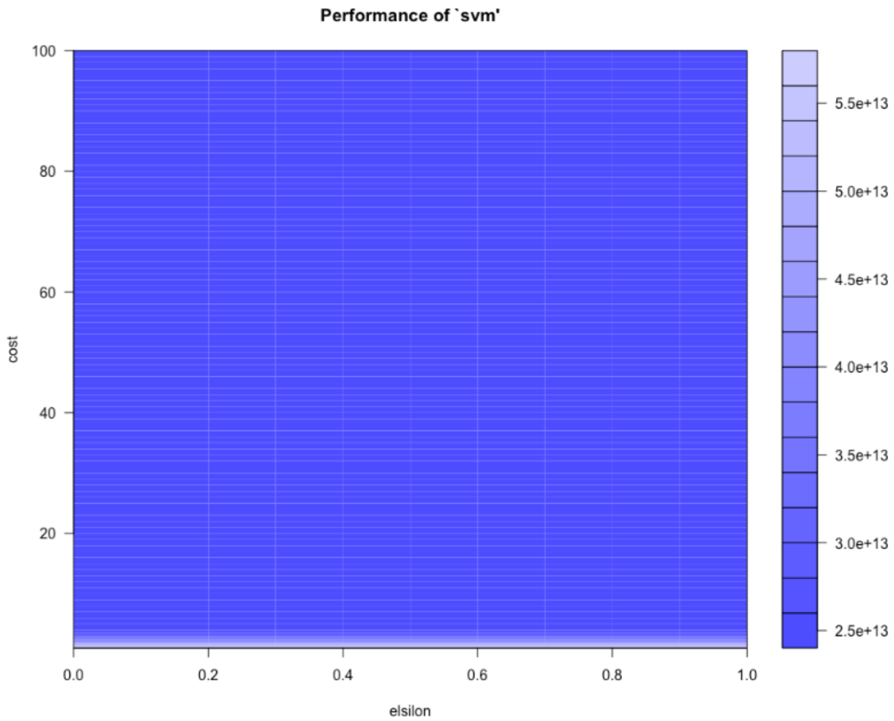


Fig. 4 SVM grid search

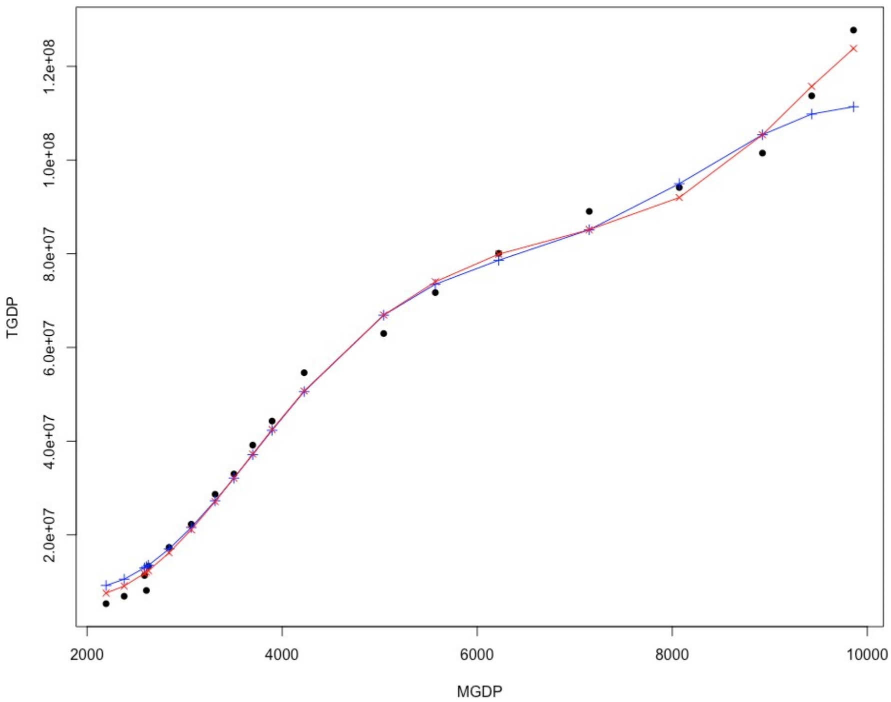


Fig. 5 The SVR model

TABLE II
RESULT ON ANALYSIS BY PRM

Residuals:				
Min	1Q	Median	3Q	Max
-3904810	-674279	-155051	947361	4387719
Coefficients:				
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-3.930e+05	2.067e+07	-0.019	0.985083
poly(LINEAR$MGDP, 4, raw = TRUE)1	-2.819e+04	1.728e+04	-1.631	0.123614
poly(LINEAR$MGDP, 4, raw = TRUE)2	1.919e+01	5.007e+00	3.833	0.001630 **
poly(LINEAR$MGDP, 4, raw = TRUE)3	-2.868e-03	5.996e-04	-4.784	0.000241 ***
poly(LINEAR$MGDP, 4, raw = TRUE)4	1.364e-07	2.529e-08	5.395	7.44e-05 ***
Residual standard error: 2062000 on 15 degrees of freedom				
Multiple R-squared: 0.9978			Adjusted R-squared: 0.9972	
F-statistic: 1698 on 4 and 15 DF			p-value: < 2.2e-16	

There exists a quantitative relationship between the two data and the polynomial equation can be written as:

$$TGDP = -3.93e^{+05} - 2.819e^{+04}MGDP + 19.2MGDP^2 - 0.00287MGDP^3 + 1.363e^{-07}MGDP^4 + \varepsilon \quad (8)$$

Equation (9) represents the polynomial trend of contribution of the MGDP to the national GDP in the period under study. The R-squared = 0.9978 and P-value identified was < 2.2e-16. This shows that MGDP is statistically relevant in the contribution to the TGDP. We further performed the Pearson correlation accuracy test to quantify the relationship between the two variable and we obtained 0.9883674. This indicates a nearly perfect positive correlation.

From Fig. 6, we observed that the polynomial of order 4, in blue line, best fit the data compared to other degrees because it almost captures all the data on the plot.

The residual versus fitted plot in Fig. 7 (a) shows a random and unpredictable pattern. This means that the PRM is valid. The prediction of the test data shows a good fit in Fig. 7 (b) which validates that the polynomial regression of order 4 perfectly predicts the data.

D. Generalized Linear Mixed Model

The purpose of this analysis is to see how the random term significantly affects the response variable which is the TGDP and from the result presented in Fig. 8, there is no significant effect.

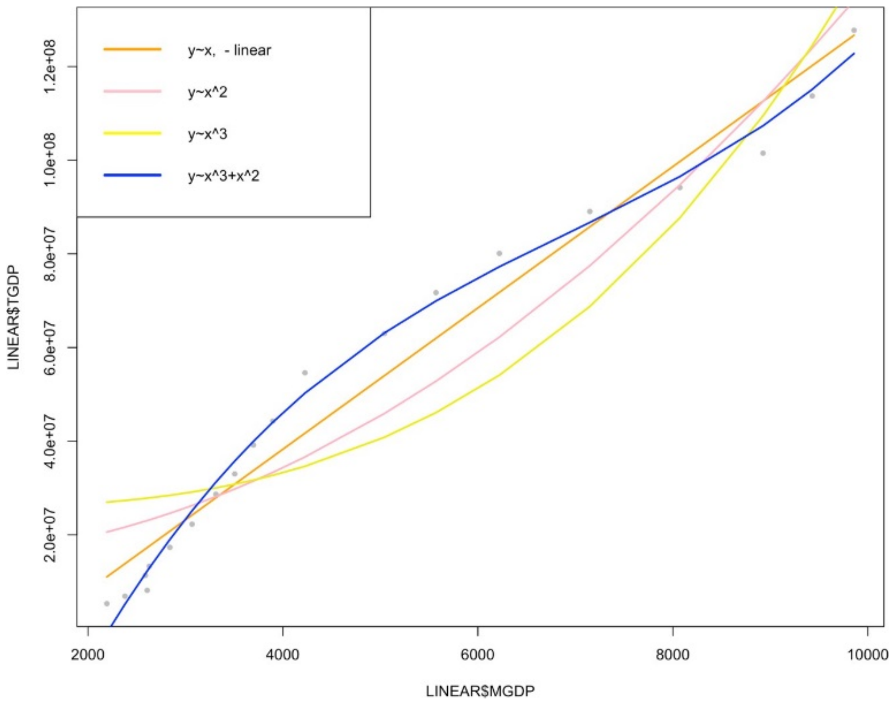


Fig. 6 The polynomial model

E. Generalized Additive Model

TABLE III
GENERALIZED ADDITIVE GAUSSIAN MODEL RESULT

Parametric coefficients:				
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	51259553	298045	172	<2e-16 ***
Approximate significance of smooth terms:				
	Edf	Ref.df	F	p-value
s(LINEAR$MGDP)	8.04	8.952	1818	<2e-16 ***
R-sq.(adj) = 0.999		Deviance explained = 99.9%		
-ML = 322.25		Scale est. = 1.7766e+12		n = 20

TABLE IV
GAM GAMMA MODEL RESULT

Parametric coefficients:				
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	17.36942	0.02173	799.3	< 2e-16 ***
Approximate significance of smooth terms:				
	Edf	Ref.df	F	p-value
s(LINEAR$MGDP)	6.452	7.538	264.2	< 2e-16 ***
R-sq.(adj) = 0.999		Deviance explained = 99.2%		
-ML = 337.01		Scale est. = 0.0094454		n = 20

We fit the model in GAM, and it assumes family gaussian and gamma. Our family gaussian result shows that the model intercept is statistically significant with p-value less than 2e-16. The EDF 8.04 indicates that the relationship between the data is not a straight line, but a wiggly curve. We can see that from the partial effect plot in Fig. 9.

We performed the model GAM check and derived the QQ plot of residuals and histogram of residuals. We observed from the QQ plot of residuals that it is close to a straight line; this shows that the model is fit. We can also see that the data are evenly distributed around zero. A perfect model would form a straight line and this we can see in Fig. 10 (b). The RMSE is 9.88e+5.

With the Gamma family, our result shows that the model intercept is statistically significant with p-value less than 2e-16. The EDF 6.452 indicates that the relationship between the data is not a straight line, but a wiggly curve. We can see that from the partial effect plot in Fig. 11.

We performed the model GAM check and derived the QQ plot of residuals and histogram of residuals. We observed from the QQ plot of residuals that it is close to a straight line; this shows that the model is fit. We can also see that the data are evenly distributed around zero. A perfect model would form a straight line on the plot and this we can see in Fig. 10 (a).

IV. DISCUSSION, CONCLUSION AND PERSPECTIVES

A. Discussion and Perspectives

Having assessed the significance of the maritime transport industry to the Nigerian economy, it is not an overstatement to say that the maritime industry has contributed to the economic growth of Nigeria. However, this sector could have contributed more but some factors have impeded the growth of this vital cash cow. We hereby recommend these perspectives and action plans in order to improve and cause a boom in this industry.

- The security of the port should be prioritized. The Nigerian government needs to work judiciously with the Nigerian Maritime Administration and Safety Agency (NIMASA) integrating all other national security bodies to manage the insecurity issues that had led to loss of lives, damage and loss of maritime investment. On this note, the implementation of the International Ship and Port Facility

Security (ISPS) should be compulsory for all ships, seaports and terminals.

- The government should invest in modern infrastructural facilities. One of the factors contributing to cargo delay and high vessel turnaround time is lack of modern cargo handling equipment. This delay has contributed to cargo diversion and made port users prefer to use other neighboring seaports.

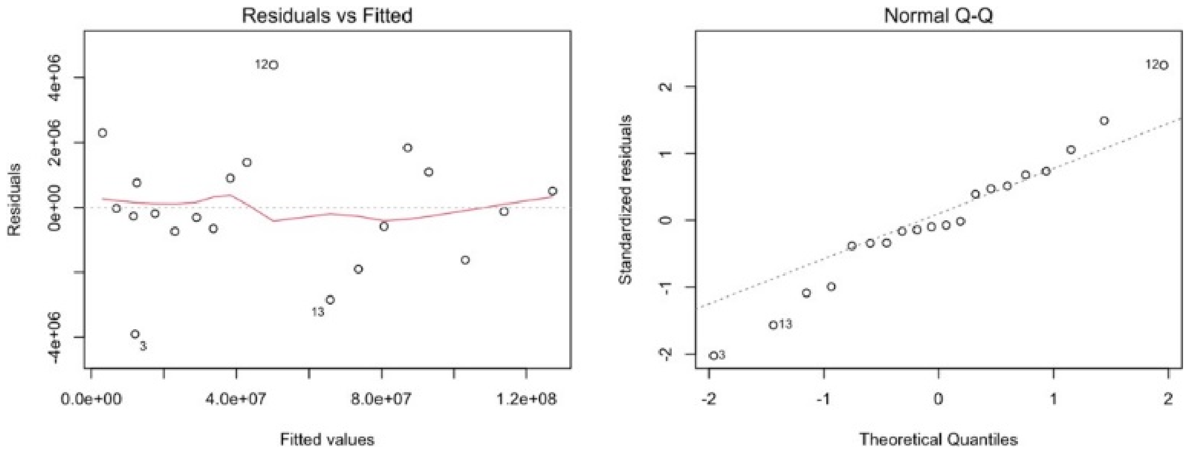


Fig. 7 (a) Polynomial regression plots

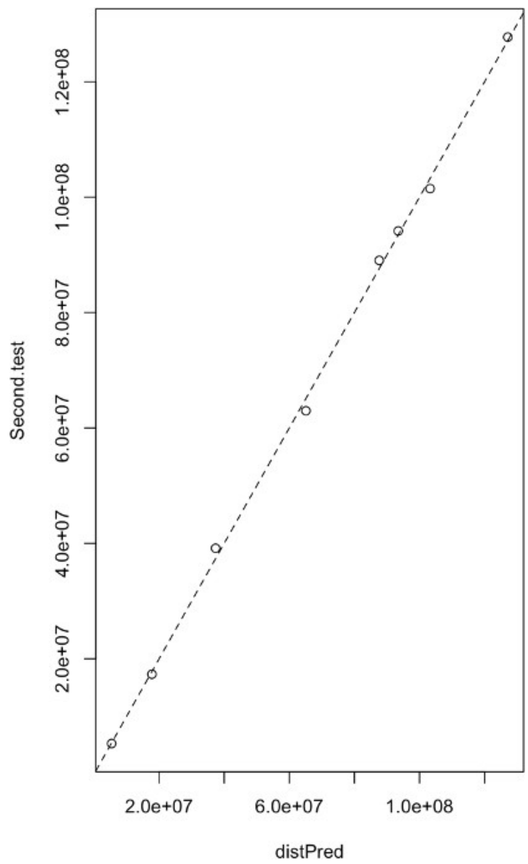


Fig. 8 (b) Test data prediction

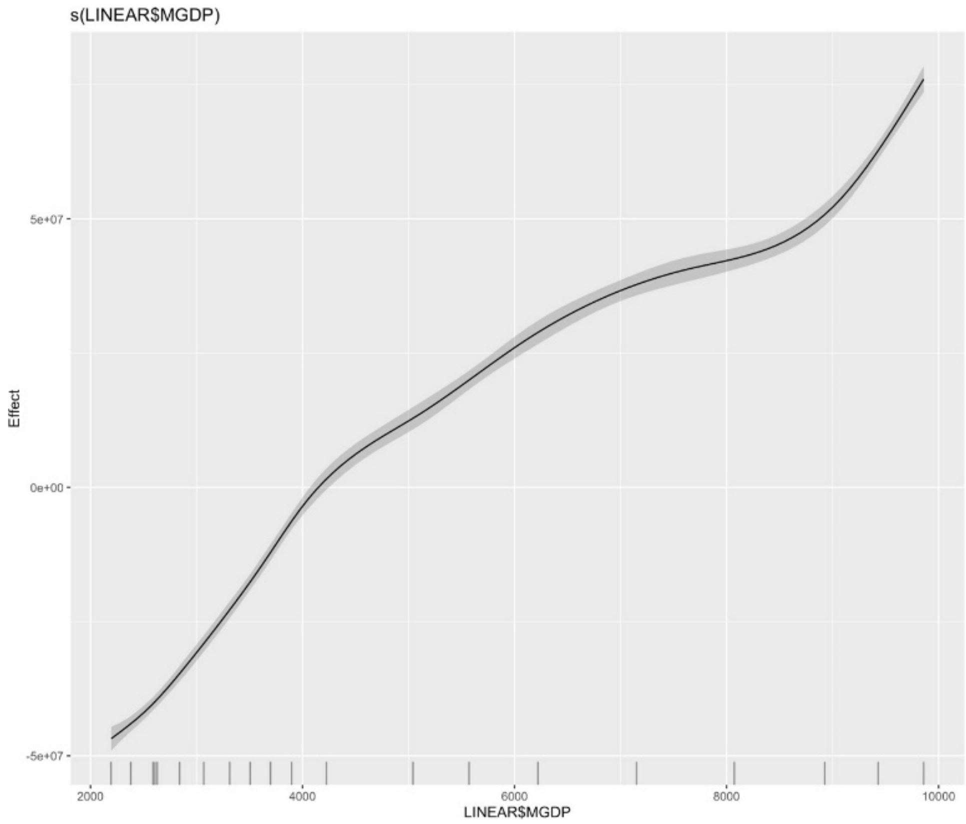


Fig. 9 GLMM partial effect plot

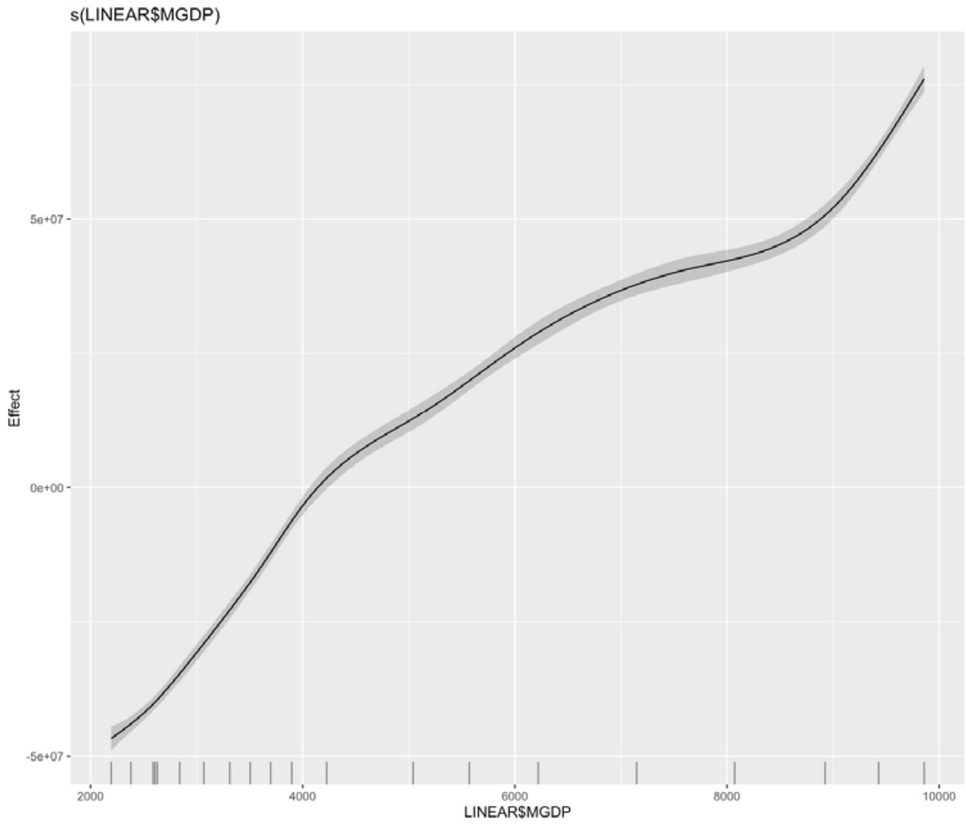
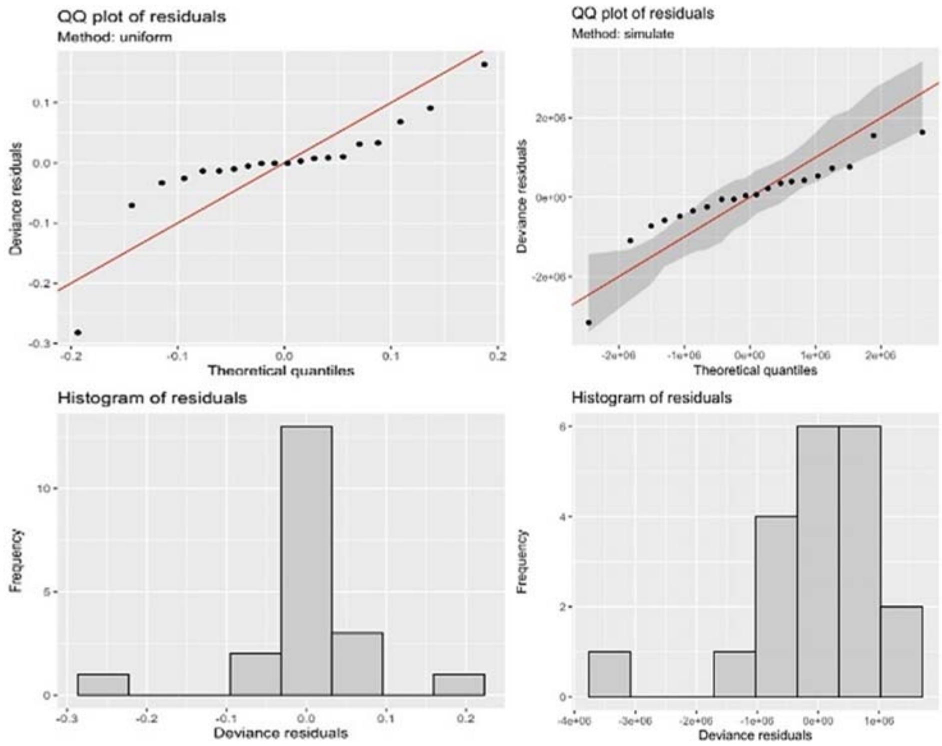


Fig. 10 Family Gaussian partial effect plot



(a) (b)

Fig. 11 (a) General Additive Model Plot (Family Gamma); (b) General additive model plot (Family Gaussian)

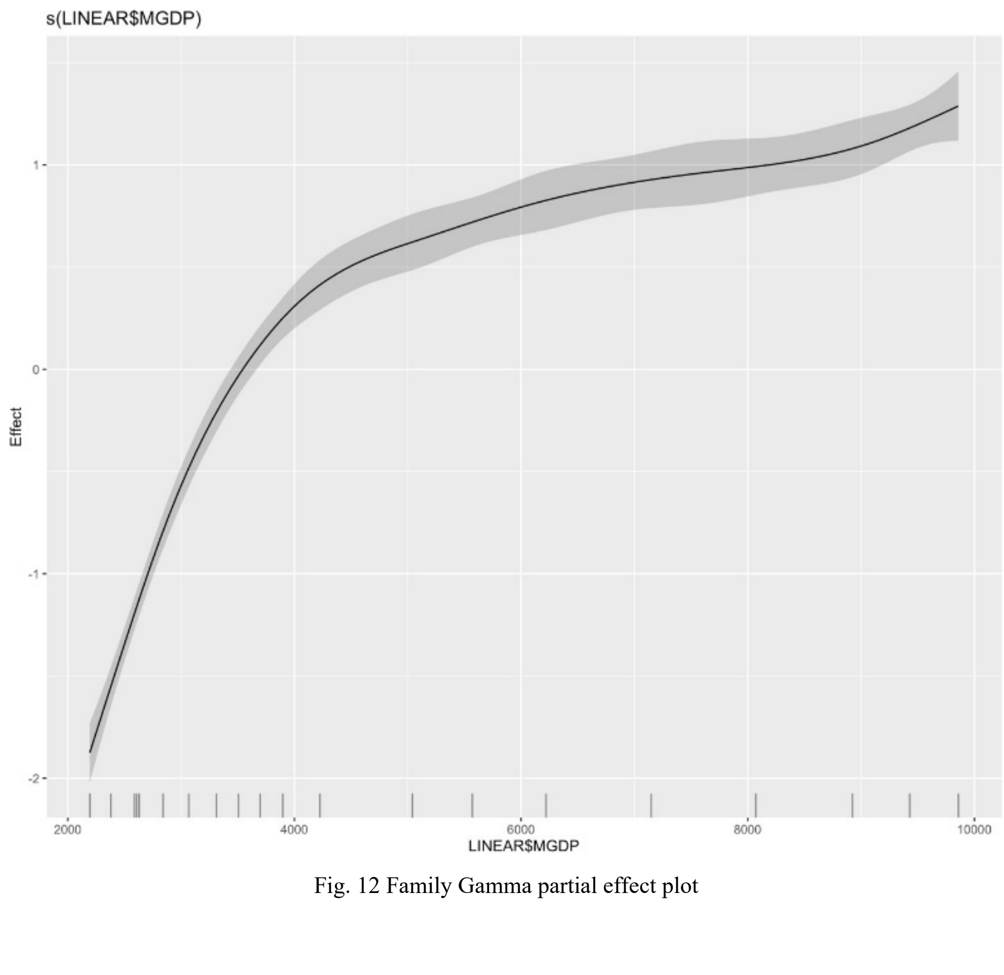


Fig. 12 Family Gamma partial effect plot

- The intermodal transport system and inland container depot should be effectively implemented. The Nigerian government should invest in other means of transportation like the rail and road to promote effective operations of the Nigerian ports. The creating of good road and rail networks, and ICDs will decongest the ports and enable fast and efficient operations of the port.
- The eastern ports should be patronized. Amongst the six seaports in Nigeria, only two is functioning near their optimal ability, the other four located in the southern part of Nigeria have been underutilized. If these ports are patronized, it will decongest the Lagos ports that are always congested.
- Right professionals should be employed to lead the affairs of the maritime ministry to bring about the change in the industry. The Nigerian maritime industry has been bedeviled with corruption and favoritism.

B. Findings and Conclusions

We have analyzed the impact of the MGDP to the Nigerian Economy. From our statistical applications, we deduced that the maritime industry contributes significantly to the nation's overall GDP. We observed a positive correlation of almost 0.99 from both data, this indicates that an increase in the GDP of the Maritime Transport industry leads to the increase in the overall GDP. Five different statistical analysis tools were used to model these data, and they all postulated a positive relationship between the MGDP and TGDP.

Aside validating the posit of the contribution of the maritime transport to the TGDP of Nigeria, this paper proposed the polynomial regression of order 4 model as the preferred prediction model and that it better represent the relationship between the MGDP and the Nigerian economy and thus, can form the basis upon which future contribution of the maritime transport industry to the Nigerian economy can be projected.

Data Availability

The data and code used for the analysis can be found at: https://github.com/Honkay/Nigeria-Maritime-and-GDP-Analysis.git

Conflict of Interest

All authors declare that there are no known conflicts of interest.

References

[1] O. H.K, O. P, E. U.M and E. G.C, "The Impact of Port Congestion on the Nigerian Economy," International Journal of Scientific Research and Management, vol. VII, pp. 3431-3437, 2015.
[2] World Bank, "GDP (Current US$) - Nigeria," 2020. (Online). Available: https://data.worldbank.org/indicator/NY.GDP.MKTP.CD?end=2020&locations=NG&start=2020 &view=bar. (Accessed 3 January 2022).
[3] United Nations, "Review of Maritime Transport," United Nations Publication, Geneva, 2020.
[4] N. T.C, O. L and I. C.C, "Econometric Modeling of the Impact of Shipping on Nigerian Economy," International Journal of Research in Commerce and Management, vol. I, no. 1, pp. 58-61, 2016.
[5] O. R.C and E. University, "Port Revenue Performance and Economic Growth: The Nigerian Ports Authority Experience, 2010 - 2019," LOGI - Scientific Journal of Transport and Logistics, vol. II, pp. 1-11, 2020.
[6] World Atlas, "Map of Nigeria," World Atlas, 2022. (Online). Available: https://www.worldatlas.com/maps/nigeria. (Accessed 25 February 2022).
[7] LCCI, "Nigeria: Reforming the Maritime Ports," The Lagos Chamber of Commerce & Industry, Lagos, 2016.
[8] R. B.A, S. M.O and L.-F. S, "Maritime Logistics and Gross Domestic Product: A Study of Nigerian Seaport," Journal of Academic Research in Economics, vol. III, pp. 382-408, 2021.
[9] K. P. Oyeduntan, "Cargo Throughput and Nigeria Maritime GDP in the Pre-Concession and Post-Concession Era," International Journal of Research in Commerce and Management Studies, vol. IV, no. 2, pp. 62-76, 2022.
[10] CBN, "Central Bank of Nigeria Statistical Bulletin," Central Bank of Nigeria, Abuja, 2018.
[11] H. Seltman, Experimental Design and Analysis, Pittsburgh: Carnegie Mellon University, 2018.
[12] K. Oshinubi, M. Rachdi and D. J, "Analysis of Daily Reproduction Rates of COVID-19 Using Current Health Expenditure as Gross Domestic Product Percentage (CHE/GDP) Across Countries," Healthcare, vol. IX, p. 1247, 2021.
[13] K. Oshinubi, M. Rachdi and D. J, "Modelling of COVID-19 pandemic vis-à-vis some socio-economic factors," Frontiers in Applied Mathematics and Statistics, vol. VII, p. 786983, 2022.
[14] T. W. Brady, B. W. Kathleen and T. G. Andrzej, Linear Mixed Models: A Practical Guide Using Statistical Software, New York: Chapman and Hall/CRC, 2014.
[15] J. Kong, E. Tekwa and S. Gignoux-Wolfsohn, "Social, Economic and Environmental Factors Influencing the Basic Reproduction Number of COVID-19 Across Countries," PLOS ONE, vol. XXVI, no. 6, 2021.