International Journal of Scientific Research and Engineering Development-– Volume 7 Issue 4, July-Aug 2024
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Estimating Revenue for TH Hotel Brand Abuja using the Bilinear Auto Regression Moving Average (BARMA) Time Series Model

Ushie, Mark Clement
Affiliate Institution: Robert Gordon University Aberdeen Scotland UK
Emails: m.ushie@rgu.ac.uk

Abstract

*The increase in revenue within a destination reflects the quality of service delivery, especially in the hotel brand sector. This research explores the application of bilinear time series models to estimate the monthly revenue data of TH hotel brand Abuja. The study utilizes quantitative secondary data, analysing revenue figures from 2017 to 2022. The aim was to identify and estimate a suitable bilinear time series model for the revenue data of the TH hotel brand up to 2030. Both descriptive and inferential statistics were employed to test the hypothesis. The Akaike Information Criterion (AIC) was used to determine the best-fitting model, revealing that the bilinear time series model (6, 0, 6, 0) is most suitable for modelling the revenue series. Additionally, the Shapiro-Wilk test for normality was conducted, showing that the null hypothesis (H0) at an alpha level of 0.05 is rejected when the p-value is less than 0.05 ($p < 0.05$). This indicates that the data tested do not come from a normally distributed population, highlighting the necessity of this research. The results demonstrate that Bilinear Fit Six (6,0,6,0) with standard error of 1.443, BIC of 1125.547, AIC of 1125.542, F- statistics of 6.769 and P-value of 8.746*10-11 provides a superior model fit and forecasting accuracy compared to other models. In contrast, the forecasting results suggest that the bilinear model predicts an increasing trend in revenue up to 2030. This indicates that the bilinear time series model will likely provide accurate forecasts, supporting rejecting the null hypothesis in favour of the alternative hypothesis.*

Keywords:

Bilinear model, Bayesian Information Criterion, Akaike Information Criteria, Revenue, Vector Autoregressive Moving Average model, Bilinear Autoregressive Average Moving Average model, Time Series, Hotel, Nigeria.

1.1 INTRODUCTION

1.2 Background of the Study

This paper aims to provide estimates based on the Bilinear Autoregressive Moving Average (BARMA) model for the performance and profitability of Transcorp Hilton Abuja. In recent years, bilinear time series models have garnered considerable attention. Introduced by Granger and Anderson (1978), the bilinear model is one of the simplest nonlinear time series models. The most general form of this model is the BARMA model. Several authors, including Subba Rao (1981), Pham (1986), Gooijger and Heauts (1986), and Mohler and Tang (1988), have developed the theory on the statistical properties of the BARMA model.

Mathews and Lee (1993) conducted further analysis concerning its stability, stationarity, and invertibility. Building on the work of these preceding authors, Subba Rao (1981) proposed estimation algorithms, which were further developed by Challapilla and Rao (1998) using the Newton-Raphson algorithm and evolutionary algorithms. Following these developments, Hili (2008) simplified bilinear structures.

The bilinear time series model is a complex yet powerful tool encompassing linear and nonlinear components of autoregressive moving average processes (Hammad, 2024). The linear part aggregates autoregressive and moving average processes, while the nonlinear part is the product of these two processes. The Elementary Bilinear (EB) time series model, a simpler version of the BARMA model, consists of only two terms: white noise and bilinear. Understanding the intricate relationship between the properties and identifiability of these models is crucial for adequately identifying more complex bilinear time series models.

The study by Hammad (2024) explores the estimation of bilinear time series models with exponential and symmetric coefficients, providing insights into the application and nuances of these models. Additionally, Cao et al. (2020) discuss model order reduction for bilinear control systems, demonstrating the equivalence of such systems to time-invariant bilinear control systems and the feasibility of applying conventional model reduction techniques. In mathematical modelling, Nascimento et al. (2023) stress the significance of defining systematic and random components in processes like the Scaled Muth–ARMA model applied to financial markets. These insights, taken together, highlight the importance of accurately delineating different components within models to enhance comprehension estimates.

Nevertheless, Maravall (1983) studied the application of bilinear models for modelling and forecasting nonlinear processes, demonstrating improvements over the ARIMA model with economic data. Terdik (1999) asserted that bilinear processes can partially capture nonlinearity in economic series data. This study introduces bilinear time series models to explain the nonlinear revenue behaviour of the TH hotel brand in Nigeria. Iwok and Etuk (2009) compared the performance of pure Vector Autoregressive Moving Average (VARMA) and Vector Bilinear Autoregressive Moving Average (VBARMA) time series models, concluding that revenue series contain both linear and nonlinear components. They also confirmed that nonlinear models such as BARMA are superior to pure linear ARMA models.

However, VARMA modeling presents some complications. First, VARMA representations are not unique, meaning many parameterizations can describe the same Data Generating Process (DGP) (Lütkepohl, 2005). An identified representation is needed for consistent estimation, requiring the researcher to choose one. This identified VARMA representation involves more integer-valued parameters than a VAR representation, which only requires a single integer parameter, the lag length. This introduces additional uncertainty during the specification stage, although procedures for VARMA models exist that can be used automatically (Hannan and Kavalieris, 1984; Poskitt, 1992).

Within this context, the primary goal is to forecast and estimate the revenue of TH hotel brands using the BARMA model. To achieve this, section two reviews the literature on BARMA, VARMA-ANN, and EC-VARMA. Section three presents the methodology based on BARMA and the associated algorithms. Section four analyzes and presents the estimation of bilinear time series parameters, fits the model, and performs forecasting. Section five concludes and makes recommendations.

1.3 statements of the problems

The hotel industry continues to experience low levels of revenue generation due to factors such as exchange rates, security concerns, management policies, and seasonality. Despite various efforts, fluctuations in revenue generation persist across both government and private sector entities. Regardless of design and operations, an organization's profitability largely depends on factors such as revenue, exchange rate, return on assets, return on equity, and the quality of services rendered.

Muio (2012) studied the impact of exchange rates and revenue over time on the financial performance of hotels in Kenya, finding a significant relationship between revenue and financial performance. However, limited research has been conducted to estimate revenue generation and its profitability using Bilinear Autoregressive Moving Average (BARMA) and Vector Autoregressive Moving Average (VARMA) models. Existing research on profitability and financial performance has a gap, as it does not consider bilinear (dependent) variables. This study aims to fill this knowledge gap, thereby justifying the need for this research.

1.4 objectives of the study

The objective of this research work is to fit a bilinear model that will estimate the revenue of TH hotel brands over time.

Research Questions

1. How do the estimates of the simulated bilinear fit (six) compare to the actual bilinear fit (six) for the revenue data of Transcorp Hilton Abuja?
2. What are the forecasted future values of the revenue observations for TH hotel brand Abuja from 2017 to 2030 using the bilinear time series model?

Hypotheses

1. **Comparison of Estimates:**

   Null Hypothesis (H0): There is no significant difference between the estimates of the simulated bilinear fit (six) and the actual bilinear fit (six).
   Alternative Hypothesis (H1): There is a significant difference between the estimates of the simulated bilinear fit (six) and the actual bilinear fit (six).

2. **Forecasting Accuracy:**

   Null Hypothesis (H0): The bilinear time series model does not provide accurate forecasts of the revenue observations for TH hotel brand Abuja from 2017 to 2030.
   Alternative Hypothesis (H1): The bilinear time series model provides accurate forecasts of the revenue observations for TH hotel brand Abuja from 2017 to 2030.

1.5 Significance of the research work

For hotel brands, this study offers several benefits. Firstly, it will assist hotels, particularly the Transcorp Hilton (TH) hotel brand, in projecting expected revenue and evaluating business performance. Secondly, the findings will be valuable for Hilton management in reporting to the

Board on the development, operation, and monitoring of services. Thirdly, employees can utilize the results to enhance strategy implementation.

For researchers and scholars, this study contributes to the existing body of knowledge on BARMA and VARMA models. It provides valuable background information for research organizations, individual researchers, and scholars interested in further exploring this area. The study will aid researchers and academicians in expanding their research into profitability ratios, especially within the hospitality industry, using models such as VAR, VMA, EC-VARMA, and BARMA. In conclusion, this study will add to the growing body of knowledge in statistical quality control and strengthen the frontiers of knowledge in this discipline.

2.0 Literature Review

2.1 Stationary Multivariate Time Series

There has been increasing attention recently on the application of bilinear time series models (Abbasimehr, Behboodi, and Bahrini, 2024; Wu and Lu, 2021). However, within the hospitality industry, there has been limited reporting on their application for modeling revenue. Ojo (2009) argues that many recently studied bilinear models fail to achieve stationarity for all nonlinear series. These challenges highlight the limitations of one-dimensional bilinear time series models, which struggle to achieve stationarity for nonlinear series.

Past studies have demonstrated the effectiveness of bilinear models in other contexts. For instance, Maravall (1983) used a bilinear model to forecast Spanish monetary data and reported nearly a 10% improvement in one-step-ahead mean square errors compared to several autoregressive moving average (ARMA) alternatives. Additionally, certain factors identified as bilinear influences suggest that most economic or financial data exhibit fluctuations when modeled with bilinear time series methods.

In Nigeria, Etuk (2011) applied a SARIMA model to the daily Nigeria-British pound exchange rate, estimating parameters for SARIMA (0,1,1)(0,1,1)7, with a period of seasonality of seven due to the weekly data frequency. Similarly, Amadi and Aboko (2013) fitted an ARIMA (2,1,2) model to Nigeria's Gross Domestic Product (GDP), using the Akaike Information Criterion (AIC) to select it as the best fit among the ARIMA models considered. In contrast, Eke et al. (2015) evaluated three-time series trend models—linear, quadratic, and exponential—on annual GDP data from 1982 to 2014. Their forecast results revealed that the exponential trend model had the lowest mean absolute percentage error.

On the other hand, Ismail and Maphol (2005) modeled and forecasted Malaysian electricity generation using a SARIMA model, while Iwueze, Nwogu, and Nlebedim (2013) adopted an ARIMA model to forecast Nigeria's external reserves from 2013 to 2024. Their initial results involved fitting an ARMA (1,0,0) model to the non-stationary data series. After differencing, they applied an ARIMA (2,1,0) model, which provided better estimates than the initial ARMA (1,0,0). Prior to this, Etuk (2012) fitted a Seasonal Autoregressive Moving Average (SARIMA) model to Nigeria's Gross Domestic Product (GDP) using data from 1980 to 2007. The fitted model was $X_t = 0.235X_{t-4} – 0.9043e_{t-1}$, effectively captured both the seasonal and non-seasonal components of the data.

In a different approach, Okereke and Bernard (2014) fitted a SARIMA (2,1,2) × (1,0,1)*(4)* model to Nigeria's quarterly GDP data. They applied a log transformation and obtained a first-order regular difference for the series. Although their model did not indicate the need for seasonal differencing, it was similar to the SARIMA (1,1,2) × (1,0,1)({12}) model fitted by Buckman and Enock (2013) to Ghana's inflation rate from 1985 to 2011. This comparison highlights the variability in model fitting across different datasets and contexts.

The assumption of linearity in time series models has often been challenged. Jan and Kuldeep (1992) noted that linear models might sometimes be inadequate, leading to the adoption of non-linear alternatives. Usoro and Omekara (2008) fitted Bilinear Autoregressive Vector (BARV) models to compare the performance of linear and bilinear time series models in fitting revenue data from a local government in Nigeria. Their analysis confirmed that the bilinear model was more suitable for fitting the revenue series which further resonates with the perspective of Akpanta and Okorie (2015) who argued that ARIMA model fit bilinear data better.

Similarly, Amadi and Aboko (2013) fitted an ARIMA (2,1,2) model to GDP data, finding that the model did not exhibit linearity, which questioned the suitability of ARIMA (2,1,2) as the best fit. Although ARIMA models are typically linear, their non-linearity in the context of GDP justified the use of SARIMA models as a multiplicative time series model in studies by Yarrington (2021). While some non-linear time series models outperform ordinary linear models, the SARIMA model was chosen based on the quarterly data frequency (with a period of (s=4)). This comparison aims to evaluate the performance of the SARIMA model against bilinear time series models in modeling Nigeria's GDP.

There have been no studies testing the performance of bilinear time series models specifically within the Nigerian hospitality industry, particularly for hotel brands. This review aims to introduce the bilinear time series model and determine its effectiveness in estimating or fitting hotel revenue data. Abu Hammad (2024) noted that in certain time series applications, univariate linear estimates can be comparatively more accurate than those obtained from bilinear models. Similarly, Aras and Kocakoç (2016) investigated the relative merits of multivariate linear processes versus univariate bilinear processes, the results indicated that linear models generally outperformed bilinear models.

Historically, studies by Volterra (1930) and Wiener (1958) examined functional series and questioned the assumption of linearity, significantly contributing to the development of non-linear models. However, Wiener's (1958) representation was deemed too general, and the statistical estimation of Wiener Kernels proved unwieldy. Subsequent authors, including Ozaki and Oda (1977), Jones (1978), Haggan and Ozaki (1980), and Tong and Lim (1980), presented various perspectives supporting the statistical estimation of the Wiener Kernels model, with many suggesting it supports non-linearity in estimation.

Recent work by Lukasz (2016) identified stable elementary bilinear time series and highlighted that a key issue in accurately estimating the coefficients of the stable elementary-bilinear model is the model's invertibility. Subba Rao (1981) had earlier addressed the elementary bilinear model in his work on the theory of time series models, deriving sufficient conditions for the variance, covariance, and invertibility of bilinear models. His application of the elementary bilinear model to sunspot data (with 512 observations) demonstrated that the series is highly non-linear, suggesting that a non-linear model may be more appropriate. For example, Ojo (2012) fitted generalized and subset integrated autoregressive moving average bilinear time series models to Wolfer sunspot numbers. The integrative approach satisfied stationary conditions and showed that the model performed better than the subset generalized integrated autoregressive bilinear model.

The generalized subset integrated autoregressive moving average bilinear time series model, as defined in various studies—including those by Ojo and Olanrewaju (2021), Kendell and Ord (2015), and Housh (2023)—follows similar patterns as presented in Equation 2.1 below:

$$X_t = \sum_{i=1}^{p} \phi_i x_{t-1} + \sum_{j-1}^{q} \theta_j \varepsilon_{t-j} + \sum_{i=1}^{P}\sum_{j=1}^{Q} \beta_{ij} x_{t-i}\varepsilon_{t-j} + \varepsilon_t \qquad 2.1$$

where x_t is the bilinear process, x_{t-i} and ε_{t-j} are time-varying autoregressive and moving average processes, respectively, with ϕ_i and θ_j as their respective parameters; β_{ij} are the non-linear parameters of $x_{t-i}\varepsilon_{t-j}$, $1 \le i \le p$, $1 \le j \le q$ and $1 \le i \le p$, $1 \le j \le Q$; ε_t is the error term, $\varepsilon_t \sim N(0, \sigma^2_{\varepsilon_t})$. Model (1) is a difference equation for a time-varying bilinear process of order (p, q, p, Q), also known as BL (p, q, P, Q). Granger and Anderson (1978) identified a complete non-linear part of the bilinear model with p = q = 0. The bilinear approach in this stage is shown in equation 2.2 below

$$X_t = \sum_{i=1}^{P}\sum_{j=1}^{Q} \beta_{ij} x_{t-i} \varepsilon_{t-j} + \varepsilon_t \qquad 2.2$$

It is expressed in the matrix form as shown in equation 2.3 below

$$X_t = (x_{t-1}, x_{t-2}, \dots, x_{t-p})^1 (\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-Q}) \underline{\beta} + \varepsilon_t \qquad 2.3$$

where β is the matrix of parameters of order P and Q.

$$\underline{\beta} = (\beta_{ij}) \text{ , i = 1, 2, …, P; j = 1, 2, …, Q.}$$

Usoro and Omekara (2008) used the non-linear part of bilinear time series model to express pure diagonal bilinear model. From their model (2.3), if i = j, the coefficient matrix $\underline{\beta} = (\beta_{ij})$, i = 1, 2, …, P; j = 1, 2, …, Q of the nonlinear part of the bilinear model exists with elements in the principal diagonal of the matrix.

2.2 Bilinear Autoregressive Integrated Moving Average (Barima)

The general bilinear autoregressive integrated moving average time series process as suggested by Subba Rao (1981) as BARIMA (p, q, d, P, Q). This model is presented in equation (2.4) as

$$X_t = \sum_{i=1}^{p} \phi_i x_{t-1} + \sum_{j=1}^{q} \theta_j e_{t-j} + \sum_{k=1}^{P}\sum_{l=1}^{Q} \beta_{kl} x_{t-k} e_{t-l} \qquad (2.4)$$

Where x_t is the time series process, ϕ_i *and* θ_j are the parameters of the linear part of the autoregressive and moving average processes with p and q as the orders respectively, β_{kl} is the parameter of the nonlinear part of the model with P and Q as the orders respectively, d is the differencing. Granger and Anderson (1978) explained and introduced this model to examine economic and financial revenue data. While the BARIMA model is an extension of the ARIMA model with a nonlinear component, the conditions for the application of ARIMA also apply to the BARIMA model.

2.3 Linear Vector Autoregressive Moving Average (Varma) Model

Iwok and Etuk (2009), presented the general vector (VARMA) analogue to the univariate autoregressive moving average (ARMA) for the n-series as shown in equation 2.5 below:

Let x_{it}^{1} be a vector for n-dimensional time series

$$X_{it} = \sum_{r=1}^{n} \sum_{k=1}^{\max p} \gamma_{k.ir} x_{rt-k} + \sum_{s=1}^{n} \sum_{l=1}^{\max q} \lambda_{k.it} \varepsilon_{st-l} + \varepsilon_{it} \quad (2.5)$$

Where $\gamma_{k.ir}$ *and* $\lambda_{l.it}$ are the autoregressive (AR) and moving average (MA) parameters. P and q are the AR and MA orders. $\varepsilon_{it}^{1}[\varepsilon_{ii}, \varepsilon_{2t}, ..., \varepsilon_{nt}]$ Is a vector of white noise, K and l representing the lags of AR and MA models.

$\sum_{r=1}^{n} \sum_{k=1}^{\max p} \gamma_{k.ir} x_{rt-k}$ Is the vector Autoregressive (AR) part of the series

$\sum_{s=1}^{n} \sum_{l=1}^{\max q} \lambda_{k.it} \varepsilon_{st-l} + \varepsilon_{it}$ is the vector moving average and the white noise part of the series, according to Subba Rao (1981), if $\lambda_{l.it}$ = 0 for all lagged white noise, the linear vector AR(VAR) model can be isolated and written this way.

$$X_{it} = \sum_{r=1}^{n} \sum_{k=1}^{\max p} \gamma_{k.ir} x_{rt-k} + \varepsilon_{it} \quad (2.6)$$

Similarly, vector MA (VEMA) part can be obtained by setting $\gamma_{k.ir}$ = 0 and the resulting expression is:

$$X_{it} = \sum_{s=1}^{n} \sum_{l=1}^{\max q} \lambda_{k.it} \varepsilon_{st-l} + \varepsilon_{it} \qquad (2.7)$$

The expressions agree with the bilinear theory of time series Subba Rao (1981) and Akaike (1974)

2.4 Vector Non-Linear Models

Iberedem and Ette (2009) combined the vectors of the AR model and the MA model to create a combined model called the Bilinear Vector Autoregressive Moving Average Model (BIVARMA), also known as the BIVARMA model merge. The Vector Non-Linear Model (Autoregressive) process can be written as:

Given a vector elements $x_{1t}, x_{2t}, \ldots, x_{nt}$ the non- linear model for an AR process is

$$X_{it} = \sum_{r=1}^{n} \sum_{k=1}^{\max p} \sum_{l=0}^{\max q} \beta_{kl.ir} x_{it-k} \varepsilon_{it-1} + \varepsilon_{it} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.5)$$

Where $\beta_{kl.ir}$ are the bilinear parameters of the product series and $l = 0 \forall q$

Vector Non-Linear Model Moving Average (MA) process can be written as:

$$X_{it} = \sum_{s=1}^{n} \sum_{l=1}^{\max q} \sum_{k=0}^{\max p} \beta_{kl.is} x_{st-k} \varepsilon_{it-l} + \varepsilon_{it} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.6)$$

Where $\beta_{kl.is}$ are the bilinear parameters and $k = 0 \forall$ q

Etuk (2012) in his study which seek to compare the performances of the two vector models (Linear and Bilinear) comes up with a BIVARMA Model, this not only support his judgement, but further agrees with Rao (1981) and Akaike (1974) work on Bilinear Time Series Models.

The BIVARMA model equation is written as:

$$X_{it} = \sum_{r=1}^{n}\sum_{k=1}^{\max p}\gamma_{k.ir}x_{rt-k} + \sum_{s=1}^{n}\sum_{l=1}^{\max q}\lambda_{k.it}\varepsilon_{st-l} + \varepsilon_{it} + \sum_{r=1}^{n}\sum_{k=1}^{\max p}\sum_{l=0}^{\max q}\beta_{kl.ir}x_{it-k}\varepsilon_{it-1} + \varepsilon_{it} + \sum_{s=1}^{n}\sum_{l=1}^{\max q}\sum_{k=0}^{\max p}\beta_{kl.is}x_{st-k}\varepsilon_{it-l} + \varepsilon_{it} \quad (2.7)$$

2.5 Vector of Bilinear Models

It is well known that the linear autoregressive moving average models can be written in the form of a first-order vector difference equation and this vector form is known as the state space form. It is convenient to study the properties of the process when the model is in the state space form because of the Markovian nature of the model Akaike (1974). The bilinear models are represented in state space form.

$$\mathrm{X}(t) + \sum_{j=1}^{p} a_j \mathrm{X}(t-j) = e(t) + (\sum_{l=1}^{p} b_{l_1}\mathrm{X}(t-1))e(t-1) \quad (2.8)$$

Let's define the matrices

$$\mathrm{A} = \begin{bmatrix} -a_1 & -a_2 & \cdots & -a_p \\ 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \quad \mathrm{B} = \begin{bmatrix} b_{11} & b_{21} & b_{31} & \cdots & b_{p1} \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \dots\dots\dots\dots (2.9)$$

and C’ = (0, 0, 0, …, 0), H’ = (1,0,0,….,0), and let X’(t) = (X(t), X(t – 1),…, X(t – p + 1)). With this notation, we can write the model (2.8) in the form

$$\mathrm{X}(t) = \mathrm{A}x(t-1) + \mathrm{B}x(t-1)e(t-1) + \mathrm{C}e(t) \dots\dots\dots\dots\dots\dots (2.10)$$

Which is a vector of bilinear model BL (p, 0, p, 1) and denoted it by VBL (p).

2.6 Seasonal Autoregressive Integrated Moving Average (Sarima) Model

A research study conducted by Hamidreza and Leila (2012) focused on the SARFIMA model to study and predict Iran’s oil supply. Their analysis indicated that the SARFIMA (0, 1, 1)(0, -0.199, 0)(_{12}) model was the most effective for predicting the quantity of oil supply in Iran until the end of 2020.

Similarly, Fritzer, Gabriel, and Johann (2002) evaluated the performance of VAR and ARIMA models for forecasting Austrian HICP inflation. Their research investigated whether disaggregated modeling of five subcomponents of inflation provided superior results compared to modeling headline HICP inflation alone. They aimed to identify adequate VAR and ARIMA specifications that minimized the 12-month out-of-sample forecasting error. Their findings were twofold: First, VAR models outperformed ARIMA models in terms of forecasting accuracy over longer projection horizons (8 to 12 months ahead). Second, disaggregated modeling significantly improved forecasting accuracy for ARIMA models, while the VAR approach demonstrated the superiority of modeling the five subcomponents over headline HICP inflation primarily over longer periods (10 to 12 months ahead).

Previous studies by Gomez and Maravall (1998) reviewed a unified approach to automatic modeling for univariate time series. They examined ARIMA models and classical methods for fitting these models, explored objective methods for model identification, and described algorithmic procedures for automatic model identification. They also proposed an algorithm for automatic model identification in the presence of outliers. In a related study, Liela and Masoud (1998) empirically investigated the utility of SARFIMA models in energy science. Their results indicated that the SARFIMA (2,1,0)(0,0.473,0)(_{12}) model was most appropriate for predicting the consumption rates of petroleum products until the end of 2013.

2.7 Section two Summary

Bilinear time series models have gained significant attention due to their potential to handle non-linear data, although their application within the hospitality industry, particularly for modeling revenue, remains underexplored. Early studies by Subba Rao (1981) and Pham and Tran (1981) laid the foundation for bilinear models, highlighting their advantages in specific contexts. Maravall (1983) demonstrated a notable improvement in forecasting Spanish monetary data using bilinear models over traditional ARMA models. However, researchers like Ojo (2012) pointed out challenges in achieving stationarity for non-linear series, suggesting a need for further refinement of these models to enhance their applicability.

ARIMA and SARIMA models have been widely used to model economic and financial time series data, demonstrating their robustness and versatility. For example, Etuk (2012) and Amadi and Aboko (2013) applied these models to exchange rates and GDP, respectively, showing their effectiveness in capturing underlying patterns. However, studies like those of Iwueze, Etuk and Nwachukwu, J. (2013) revealed that non-linear models often outperform their linear counterparts, particularly in dealing with complex economic data. SARIMA models, in particular, have shown promise in handling seasonal data, as evidenced by the works of Okereke and Bernard (2014) and Buckman and Enock (2013).

The BARIMA model extends the traditional ARIMA framework by incorporating non-linear elements, offering a more nuanced approach to modeling complex data. Subba Rao (1981) and Granger and Anderson (1978) explored the application of BARIMA to economic data, demonstrating its superiority in certain cases over standard ARIMA models. These studies underscore the importance of accounting for non-linear dynamics to improve model accuracy and reliability, particularly in the context of economic and financial time series.

Finally, VARMA and BIVARMA models extend the ARMA approach to multivariate data, effectively capturing interactions between multiple time series. Research by Iwok and Etuk (2009) and Iberedem and Ette (2009) has been instrumental in developing these models, with BIVARMA combining linear and non-linear elements to enhance predictive performance. Additionally, SARIMA models have proven effective in various applications, such as forecasting Iran's oil supply (Hamidreza and Leila, 2012) and Austrian inflation (Fritzer et al., 2002). These studies highlight the models' ability to handle complex seasonal patterns, further validating their utility in time series analysis. However, there remains a gap in applying these advanced models to the hospitality industry, particularly in modeling hotel revenue, suggesting a promising area for future research. The following section presents the methodology underpinning the paper.

3.0 Research Methodology

3.1 Introduction

The validity and reliability of research strongly depend on the methods adopted for data collection and the administration of data collection instruments. Therefore, methodology in research is a crucial aspect that requires a standard procedure to be adopted and followed. Methodology is defined as a set of methods and principles used to perform a particular activity.

This section focuses on the statistical method used for data analysis, primarily the Bilinear Autoregressive Moving Average (BARMA) model. The BARMA model is an extension of the univariate ARIMA model, which describes relationships among several time series variables. In this model, each variable depends not only on its past values but also on the past values of other variables. The bilinear time series model, like the ordinary autoregressive moving average model, is characterized by individual autoregressive and moving average processes. The key distinction of the bilinear model is the introduction of a non-linear component.

The bilinear autoregressive and bilinear moving average models will be estimated to provide parameter estimates for the BARMA model. This model is particularly useful when dealing with data series that exhibit non-linear behavior, as demonstrated by Maravall (1983). The bilinear model consists of two parts: the linear part, which is the sum of the autoregressive and moving average processes, and the non-linear part, obtained by taking the product of the two processes. This section outlines the estimation procedures for the bilinear autoregressive moving average (BARMA) model parameters, setting the stage for a detailed analysis of the data.

3.2 GENERAL BILINEAR MODEL

The general form of the bilinear model according to Granger and Anderson (1978) is given by:

$$X_t = \sum_{i=1}^{p} a_i x_{t-i} + \sum_{j=i}^{q} c_j x_{t-j} + \sum_{k=1}^{P}\sum_{j=1}^{Q} b_{kj} x_{t-k} \varepsilon_{t-j} + \varepsilon_t \text{ ....}$$

(3.1)

Similarly, Subba Rao (1981) described the bilinear time series model BL (p, r, m ,k) as given by the difference equation:

$$X(t) + \sum_{j=1}^{p} a_j X(t-j) = \sum_{j=1}^{r} c_j e(t-j) + \sum_{i=1}^{m}\sum_{l^1}^{k} b_{ll^1} X(t-l) e(t-l') \text{ .... } (3.2)$$

Where, {e(t)} is an independent white noise process and $C_o = 1$. {X (t)} is termed the bilinear process. The autoregressive moving average model ARMA (p, r) is obtained from equation 3.2 by setting $d_{ll^1} = 0 \forall\ 1$ and l^1.

Parameter estimation of bilinear processes has been studied for particular cases by Bouzaachane et al. (2006). Meanwhile, Boonchai and Eivind (2005) provided the general form of multivariate bilinear time series models as follows:

$$X(t) = \sum a_i . X(t-i) + \sum m_j . e(t-j) + \sum\sum\sum_{dij} . X(t-i) . e_d(t-j) + e(t) \text{ ..... } (3.3)$$

Here, the state X (t) and noise e (t) are n-vectors, and the coefficients a_i, m_j *and* b_{diij} are n by n matrices. If all $b_{diij} = 0$, we have a class of well-known vector ARMA – models. By this Zielinski (2000) and Bielinski (2008) established the matrix form of vector autoregressive time series as:

$$X_t = \sum_{k=1}^{\max pi} \gamma_k x_{t-k} + u_t \qquad (3.4)$$

3.3 Stationery and Convergence of General Bilinear Model (p, d, q, r, s)

For general S, it is not easy to provide an infinite series representation for each x_t. For this general case, we exhibit the process (x_t, t∈ Z) as an almost sure limit of a sequence

{ { $S_{n,t,}$ $t \in$ Z } , n ≥ 1 } of stationary process.

Theorem

Let { e_t, t ∈ Z} be a sequence of independent identically distributed random variables defined on a probability space (Ω, l, R, P) such that E e_t = 0 E $e_t^2 = \sigma^2 < \infty$.Ψ, $B_1, B_2, \ldots, B_q$ be q+1 matrices each of order p x p .

$$\Gamma_1 = (\Psi \otimes \Psi + \sigma^2((\Theta \otimes \Theta + B \otimes B - 2\Theta \otimes B))$$
$$\Gamma_i = \sigma^2[Bi \otimes (\Psi^{j-i}B_1 + \Psi^{j-2}B_2 + \ldots + \Psi B_{j-1})$$
$$+ (\Psi^{i-1}B_1 + \Psi^{i-2}B_2 + \ldots + \Psi B_{i-1}) \otimes B_j$$
$$+ B_i \otimes (\Theta^{j-1}B_1 + \Theta^{j-2}B_2 + \ldots + \Theta B_{j-1})$$
$$+ (\Theta^{i-1}B_1 + \Theta^{i-2}\Theta_2 + \ldots + \Theta B_{i-1}) \otimes B_j$$
$$+ (B_j \otimes B_j)],\ j = 2,3,\ldots s.$$

Suppose all the eigenvalues of the matrix

$$\underset{p^2q\times p^2q}{L} = \begin{bmatrix} \Gamma_1 & \Gamma_2 & \cdots & \Gamma_{q-1} & \Gamma_q \\ I_{p^2} & 0 & \cdots & 0 & 0 \\ 0 & I_{p^2} & \cdots & 0 & 0 \\ 0 & 0 & I_{p^2} & \cdots & 0 \end{bmatrix}$$

Have moduli less than unity, i.e, $\rho(L) = \lambda \le 1$. Let C be given column vector

Then there exists a vector valued strictly stationary process (X_t, t∈ Z) confirming to the bilinear model

$$X_t = \Psi X_{t-1} - \Theta e_{t-1} + \sum_{l-1}^{s} B_1 X_{t-1} e_{t-1} + ce_t \text{ for every Z. } \ldots\ldots \quad (3.5)$$

3.4 Bilinear Autoregressive (Bar) Model

Given the difference equation

$$X_t = \sum_{i=1}^{p} \phi_i x_{t-i} + \sum_{j=i}^{q} \theta_j x_{t-j} + \sum_{i=1}^{P} \sum_{j=1}^{Q} X_t \beta_{ij} x_{t-i} \varepsilon_{t-j} + \varepsilon_t \quad (3.6)$$

X_{t-1} is the distributed lag component of the autoregressive process with coefficient ϕ_i, ε_{t-j} is the distributed lag component of the moving average process with coefficient θ_j. The sum of the two processes makes up the linear part of the model (3.6). The nonlinear component is the product of the two processes. The model is thus linear in X's and also in the ε's separately, but not in both as product of the two series. In the general ARMA model.

If q = Q = 0, $\Rightarrow$ j = 0, the model becomes

$$X_t = \sum_{i=1}^{p} \phi_i x_{t-i} + \sum_{i=1}^{P} \beta_{i0} x_{t-i} \varepsilon_{t-j} + \varepsilon_t \ldots\ldots (3.7)$$

Where β_{i0} are the coefficient of the non- linear autoregressive part of the bilinear model.

$\phi_i, X_{t-i}, \varepsilon_t$.

Where, the first part on the R.H.S is the autoregressive process, while the second is the moving average process, where ε_t is strict white noise,

CASE 1: If j = 0, the model becomes

$$X_t = \sum_{i=1}^{p} \phi_i x_{t-i} + \varepsilon_t \quad \text{...(3.8)}$$

This implies $\beta_0 = 0$

Model (3.8) is pure autoregressive model.

CASE 2: If I = 0, the the model becomes

$$X_t = \sum_{i=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t \quad \text{..(3.9)}$$

This implies $\theta_j = 0$

Model (3.9) is a pure moving average model.

From model (3.6),

$$X_t = \sum_{i=1}^{p} \phi_i x_{t-i} + \sum_{j=i}^{q} \theta_j x_{t-j} + \sum_{i=1}^{P} \sum_{j=1}^{Q} \beta_{ij} x_{t-i} \varepsilon_{t-j} + \varepsilon_t$$

If j = 0, the above model becomes

$$X_t = \sum_{i=1}^{p} \phi_i x_{t-i} + \sum_{i=1}^{P} \sum_{j=1}^{Q} x_{t-i} \varepsilon_t + \varepsilon_t \quad \text{.. (3.10)}$$

This implies $\beta_0 = 0$

The model is the Bilinear Autoregressive (BAR) model.

This is so because the component of the moving average process is eliminated from the general model. The ε_t in the models is the usual white noise, assumed to be independent and identically distributed with mean zero and variance δ_e^2, (Usoro and Omekara, 20 08). Normally, the value of ε_t , being the error values are obtained as the deviation of estimates from the original series

X_t is a variable that represents the original values of the revenue series. While $X_{t-1}, X_{t-2}, X_{t-3}\ldots$ are the time lag of variable X_t. However, in the research work, we will apply the Bilinear Autoregressive (BAR) Model (3.10) to fit the revenue data of TH Abuja.

3.5 Application Of Newton–Raphson In Estimating Bilinear Models

The joint density of the of $(e_m, e_{m+1}\ldots\ldots, e_n)$ where m = max (r, s) is given by

$$\frac{1}{2\pi\sigma_e^2{}^{\left(\frac{n-m+1}{2}\right)}}\exp\left(\frac{1}{-2\sigma_e^2}\right)\sum_m^n e_t^2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.11)$$

Since the Jacobian of the transformation from $(e_m, e_{m+1}\ldots\ldots, e_n)$ to $(x_m, x_{m+1},\ldots., x_n)$ is unity, the likelihood function of $(x_m, x_{m+1},\ldots., x_n)$ is the same as the joint density function of

$(e_m, e_{m+1}\ldots\ldots, e_n)$. Maximising the likelihood function is the same as minimising the

Q(G) ,

Where $Q(G) = \sum_m^n e_t^2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.11)$

Concerning the parameter $G' = (\Psi_1,\ldots., \Psi_p; \theta_1, \theta_2,\ldots\ldots, \theta_q; \mathrm{B}_{11},\ldots\ldots., \mathrm{B}_{rs})$

Then the partial derivatives of Q(G) are given by

$$\frac{dQ(G)}{dG_i} = 2\sum_{t=m}^n e_t \frac{de_t}{dG_t} \quad (I = 1, 2, \ldots.., R) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.12)$$

$$\frac{d^2Q(G)}{dG_I\, dG_j} = 2\left(\sum_{t=m}^n e_t \frac{de_t}{dG_t}\frac{de_t}{dG_j} + \sum_{t=m}^n e_t \frac{d^2 e_t}{dG_i\, dG_j}\right)$$

Where these partial derivative of e(t) satisfy the recursive equations

$$\frac{de_t}{d\Psi_i} + \sum_{j=1}^s w_t(t)\frac{de_{t-j}}{d\Psi_i} = X_{t-i} \text{ if } I = 1, 2, \ldots\ldots., p$$

$$\frac{de_t}{d\theta_i} + \sum_{j+1}^{s} w_t(t)\frac{de_{t-j}}{d\theta_i} = e_{t-i}$$

$$\frac{de_t}{dB_{kmi}} + \sum_{j=1}^{s} w_j(t)\frac{de_{t-j}}{dB_{kmi}} = -X_{t-K}\, e_{t-m} \quad (\text{ k = 1, 2, ……r; } m_i = 1, 2, \text{……., s)……….}$$

(3.13)

$$\frac{d^2 e_t}{d\Psi_i d\Psi_i^{'}} + \sum_{j=1}^{s} w_j(t)\frac{d^2 e_{t-j}}{d\Psi_i d\Psi_i^{'}} = 0 \text{ (i = 0, 1, 2, …….., p)}$$

$$\frac{d^2 e_t}{d\theta_i d\theta_i^{'}} + \sum_{j=1}^{s} w_j(t)\frac{d^2 e_{t-j}}{d\theta_i d\theta_i^{'}} = 0 \text{ (i, = 0, 1, 2, …….q) ………..(3.14)}$$

$$\frac{d^2 e_t}{d\Psi_i dB_{kmi}^{'}} + \sum_{j=1}^{s} w_j(t)\frac{d^2 e_{t-j}}{dB_{kmi} d\varphi_i^{'}} + x_{t-k}\frac{d^2 e_{t-mi}}{d\Psi_i} = 0 \text{ (i, = 0,1,2..,p; k = 1, 2,.., r;}$$

m = 1,2,..,s)

$$\frac{d^2 e_t}{d\theta_i dB_{kmi}^{'}} + \sum_{j=1}^{s} w_j(t)\frac{d^2 e_{t-j}}{dB_{kmi} d\theta_i^{'}} + x_{t-k}\frac{d^2 e_{t-mi}}{d\theta_i} = 0 \text{ (i, = 1,2,…,q; k = 1,2,….,r;}$$

m=1,2,…,s)

$$\frac{d^2 e_t}{d\Psi_i d\theta_i^{'}} + \sum_{j=1}^{s} w_j(t)\frac{d^2 e_{t-j}}{d\Psi_i d\theta_i^{'}} = 0$$

$$\frac{d^2 e_t}{dB_{kmi} dB_{kmi}^{'}} + \sum_{j=1}^{s} w_j(t)\frac{d^2 e_{t-j}}{dB_{kmi} dB_{kmi}^{'}} + X_{t-k}^{'}\frac{d^2 e_{t-mi}}{dB_{kmi}} = -x_{t-k}\frac{de_{t-m}}{d\,B_{kmi}^{'}} \text{ where}$$

(k, k' = 1,2,……,r; m_i m_i' = 1,2,…,s)…………………………………………………….. (3.15)

$w_j(t) = \sum_{j=1}^{s} B_{ij} X_{t-j}$ we assume e_t = 0 (t = 1,2,……, m – 1) and also

$\frac{de_t}{dG_i}=0$, $\frac{d^2 e_t}{dG_i dG_J}=0$ (i , j = 1,2,…..,R; t = 1,2,…,m – 1) from e_t = 0 (t = 1,2,……, m – 1)

$\frac{de_t}{dG_i}=0$ $\frac{d^2 e_t}{dG_i dG_J}=0$, and $\frac{de_t}{dB_{kmi}} + \sum_{j=1}^{s} w_j(t)\frac{de_{t-j}}{dB_{kmi}}$ = -$X_{t-K}\, e_{t-m}$ …….. (3.16)

It follows that the second order derivatives with respect to Ψ_i = (i = 0,1,2,….,p) and θ_i (i = 0,1,2,….q) are zero.

To evaluate the first and second order derivatives using the recursive equation above

V ' (G) = $\frac{dQ(G)}{dG_1}, \frac{dQ(G)}{dG_2}, \ldots \ldots \frac{dQ(G)}{dG_K}$ and let H(G) = [$\frac{d^2 Q(G)}{dG_i dG_j}$] be a matrix of second partial derivatives as in (Krzanowski 1998). Expanding V(G), near G = G in Taylor series, we obtain.

$V(G)_{G=G} = 0 =$ V(G) + H(G)(G' – G) and thus obtain the Newton – Raphson iterative equation $\hat{G} - G = -H^{-1}(G)V(G)$, and thus obtain an iterative equation given by

$G^{k+1} = G^{(k)} - H^{-1}(G^{(k)})$ where $G^{(k)}$ is the set of estimates obtained at the K^{th} stage of iteration. The estimates obtained by the above equations usually converge. For starting the iteration, we need to have good sets of initial values of the parameters.

If L denotes the likelihood function of (X(m), X(m + 1),…X(n)), then we have, approximately,

$$\frac{1}{n}\frac{d^2 \log L}{dGdG'} = \frac{1}{2\sigma_e^2}\frac{1}{n}\frac{d^2 Q(G)}{dGdG'}$$

Let

$$-\frac{1}{n}\frac{d^2 \log L}{dGdG'}$$

Converge stochastically to I(G), where I(G) is Fisher's information matrix. Then as $n \to \infty$,

$$\frac{1}{n}\frac{dQ(G)}{dGdG'} \quad \ldots\ldots (3.17)$$

Converges stochastically to $2\sigma_e^2$ I(G). Further as $n \to \infty$, it can be shown that (G'' – G) has approximately a multivariate normal distribution with mean 0 and variance matrix equal to $I^{-1}(\theta)$.

3.6 INITIAL ESTIMATES OF BILINEAR MODEL.

To obtain a good set of estimates we must have a good set of initial values to start the iteration. We have tried two methods of estimating the initial values. The first method consists of fitting an AR(p) model (with constant a_0 present) and taking these coefficients as initial values for the autoregressive part of the bilinear model BL(p,0,p,q), and setting $b_{ij} = 0$ (I = 1, 2, …, p; j = 1, 2,…, q).

The second method is as follows (with constant a_0 included): when a bilinear model of order BL (p, 0, p, q – 1) is fitted, we take for the initial values the estimates obtained from BL (p, 0, p, q – 1) or BL (p – 1,0, p – 1, q), whichever has the smallest residual sum of squares. If for the initial values, the coefficients from the model BL (p, 0, p, q – 1) are chosen, the rest of the p coefficients are equal to zero. Similarly, if the model BL (p -1, 0, p – 1, q) is chosen, the rest of the q + 1 coefficients are set to zero. However, another alternative to the Newton-Raphson technique is to use the Marquardt algorithm where only the first-order derivatives are necessary. For fitting the bilinear models, both these procedures are used.

3.7 DETECTING STATIONARITY

As stated earlier in this chapter, stationary can be detected from a time plot by mere inspection of the plot. A more comprehensive way of detecting stationarity is the unit root test. Examples of the unit root test in the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) unit root test show that the null hypothesis is:

1. H_0 :The series has no unit root (i.e. the process is stationary)

   Against

2. H_1 : The series has a unit root (the process is not stationary)

Decision rule: reject H_0 if the p-value of the test statistics is less than α

Where α = level of significance

The Phillips-Peron test: In this test, the null hypothesis is:

1. H_0 : The series has a unit root (the process is not stationary)

   Against

2. H_1 : The series has no unit root (the process is stationary)

Decision rule: reject H_0 if the p-value is less than α

The Dickey-Fuller test: just like the Phillips Peron test, the null hypothesis is:

1. H_0 : The series has a unit root (the process is not stationary)

   Against

2 H_1 : The series has no unit root (the process is stationary)

Decision rule: reject H_0 if the p-value is less than α.

3.8 Parameter Estimation

This is the next step after model selection. At this stage, the parameters of the model are estimated using maximum likelihood estimation. After estimating the parameters, some penalty function statistics are adopted to check the adequacy of the selected model. These functions include:

The Akaike information criteria (AIC)

The Schwarz Bayesian information criteria (BIC)

The corrected Akaike information criteria (AIC).

These 3 functions are based on the principle of parsimony. On this basis, the model with the smallest AIC and BIC is deemed to have residuals which resembles a white noise process.

$$AIC = -2\log L + 2k$$

$$BIC = -2In(L) + In(n)$$

Where L is the likelihood, n is the number of residuals, k is the number of free parameters.

Models with smaller AIC and BIC are regarded as adequate.

Each parameter estimate reports a standard error for that parameter. Using the parameter estimate and its standard error, a test for statistical significance (t-value) is then conducted. For statistically significant parameters, the absolute values of the *t*-ratios are expected to be greater than 1.96 or 2 for the parameters to be maintained in the model whereas parameters that are not significant are trimmed or removed from the model. Another criterion for detecting a very good model is that the AR and MA parameters must be l i.e. between -1 and 1. If this condition is not satisfied, then a new model is considered for estimation.

3.9 Diagnostic Checking

This is a test performed to check whether the fitted model follows a white noise process, if it does, then the model does adequately fit the data. Some simple tests for diagnostic checking include:

I. The sample autocorrelation function: The sample autocorrelation of a white noise process follows a normal distribution with mean zero (0) and variance $1/n$. To have a white noise process, about 95% of the sample autocorrelations must fall between the bounds $\pm 1.96/\sqrt{n}$. if this condition does not hold, then the model does not follow a white noise process. Hence the model does not fit the data adequately well.

II. Portmanteau test: instead of checking to see whether each sample autocorrelation function r_k falls inside the bounds defined in (i), we can consider a single statistic

$$Q = n\sum_{k=1}^{K} r_k^2 ,$$

Where $K > p+q$ but smaller than n and r_k is the sample autocorrelation of the residual series. If the fitted model is correct, then

$$Q \sim \chi^2_{(K-p-q)}$$

So we can base a test on this; we can reject H_0 at level α if $Q > \chi^2_{(K-p-q)}(1-\alpha)$.

An improved statistic is the modified Box-Pierce statistic which is given by

$$\tilde{Q} = n(n+2)\sum_{k=1}^{K}\frac{r_k{}^2}{n-k}$$

The distribution of $\tilde{Q}$ is closer to $\chi^2_{(K-p-q)}$ than Q.

Where n the length of the series, K is the maximum lag being considered, p and q are the order of the AR and MA processes. The hypothesis to be tested is formulated in the form.

H_0: The set of autocorrelations for residual is white noise (model fit data quite well)

H_1: The set of autocorrelations for residual is different from white noise.

The decision is to reject the null hypothesis (H_0) if $\tilde{Q} > \chi^2_{\alpha(h-p-q)}$, but if otherwise, H_0 cannot be rejected

4.0 Data Analysis and Results Presentation

4.1 Introduction

Data analysis is a crucial component of any research project, as it transforms raw data into meaningful insights that address the research aims and objectives outlined in Chapter One. This chapter focuses on applying the statistical methodology described in Chapter Three, specifically the Bilinear Autoregressive Moving Average (BARMA) model, to fit the algorithm to the time series data. The primary goal is to estimate, analyse, and interpret the revenue data of Hotel TH Abuja, Nigeria, over a period of 26 years. Through this analysis, the study aims to uncover patterns and trends that can inform future decision-making and strategic planning for the hotel industry.

4.2 Model Identification

The analysis of the revenue data was conducted using R Version 3.4.1 and IBM SPSS Statistics. The aim of employing these packages is to estimate and fit a simple Bilinear Model to the revenue data of TH Abuja. However, to make statistical judgments and inferences about the features of the secondary data set, we must make some statistical assumptions about the feature and structure of the revenue data since it follows a time series. These assumptions include stationarity and normality. Stationarity occurs in a time series when the mean value of the series remains constant over time. Normality occurs when the p-value is less than the chosen alpha (α), which is 0.05 at a 95% confidence interval. To demonstrate the normality test, summary statistics were performed, and the features of the data were displayed.

4.3 Descriptive Statistics

Descriptive statistics are brief descriptive coefficients that summarize a given raw data set, which can be either a representation of the entire population or a sample of it. Descriptive statistics are broken down into measures of central tendency and measures of variability (spread). Measures of central tendency include the mean, median, and mode, while measures of variability include the standard deviation, variance, minimum and maximum values, and the kurtosis and skewness of the distribution.

Table 4.1 Descriptive Statistics (Tr)

	statistics	Standard error
Mean	4.623	0.147
95% confidence interval for mean lower bound	4.334	
95% confidence interval for mean upper bound	4.911	
5% trimmed mean	4.365	
Median	4.100	
Variance	6.971	
Standard deviation	2.697	
Minimum	0.300	

Maximum	15.700	
Range	15.400	
Interquartile range	1.950	
Skewness	1.724	0.135
Kurtosis	3.054	0.270
Number of observation	324	

From Table 4.1 above, the number of valid observations for revenue data is 324. The range, which is 15.40, provides basic details about the spread of the data by giving the difference between the lowest and the highest values. The minimum value of the data set is 0.30, while the maximum value is 15.7. The mean, which measures the central value of the discrete set of numbers, is 4.62, and the mean standard error, which measures how far the sample mean of the data is likely to be from the true population mean, is 0.15. The variance is 6.97, indicating how spread out the data set is within the population.

The skewness of the distribution is 1.72, indicating the degree and direction of asymmetry. The kurtosis of the distribution is 3.05, reflecting the presence of outliers in the data set (Westfall, 2014). Finally, the standard deviation is 2.64, indicating the spread of the data set; the larger the standard deviation, the more spread out the observations are. The 95% confidence interval for the mean has a lower bound of 4.33 and an upper bound of 4.91. The median and interquartile range are 4.10 and 1.95, respectively.

4.4 Test of Normality

In statistics, normality tests are used to determine if a data set is well-modeled by a normal distribution and to compute how likely it is for a random variable underlying the data set to be normally distributed (Ojo and Olanruwaju 2021). More precisely, the tests are a form of model selection and can be interpreted in several ways, depending on one's interpretation of probability. In hypothesis testing, data are tested against the null hypothesis that it is normally distributed. A graphical tool for assessing normality is the normal probability plot, a quantile-quantile plot (Q-Q plot) of the standardized data against the standard normal distribution. Another way of testing normality is to compare a histogram of the data to a normal probability curve. These plots are easy to interpret and also have the benefit of easily identifying outliers.

4.5 Shapiro-Wilk Test of Normality (Tr)

The Shapiro-Wilk test is a parametric test statistic that checks whether a series is normally distributed within a given population. The null hypothesis of the test is that the series or population is normally distributed. If the p-value is less than the chosen alpha level, the null hypothesis is rejected, indicating that there is evidence that the population or the data are not from a normally distributed population.

Table 4.2 Test Of Normality

Kolmogorov – Smirnov			Shapiro – Wilk		
Statistics	Df	Significance	Statistics	Df	significance
0.230	324	0.00	0.822	324	0.00

From Table 4.2, the results from two well-known tests of normality, namely the Kolmogorov-Smirnov Test and the Shapiro-Wilk Test, are presented. The Shapiro-Wilk Test is more appropriate for small sample sizes (<50 samples) but can also handle sample sizes as large as 2000. For this reason, we will use the Shapiro-Wilk test as our numerical means of assessing normality. We can see from the table that the total revenue (TR) was not normally distributed.

This is because the significance value of the Shapiro-Wilk Test is lower than 0.05, indicating that the data deviate from a normal distribution.

The underpinning hypotheses are:

- **H0 (Null hypothesis):** The data is normally distributed.
- **H1 (Alternative hypothesis):** The data is not normally distributed.

If the chosen alpha level is 0.05 and the p-value is less than 0.05, then the null hypothesis that the data are normally distributed is rejected. If the p-value is greater than 0.05, then the null hypothesis is not rejected.

Figure 4.1 Bar Chart of Revenue Data

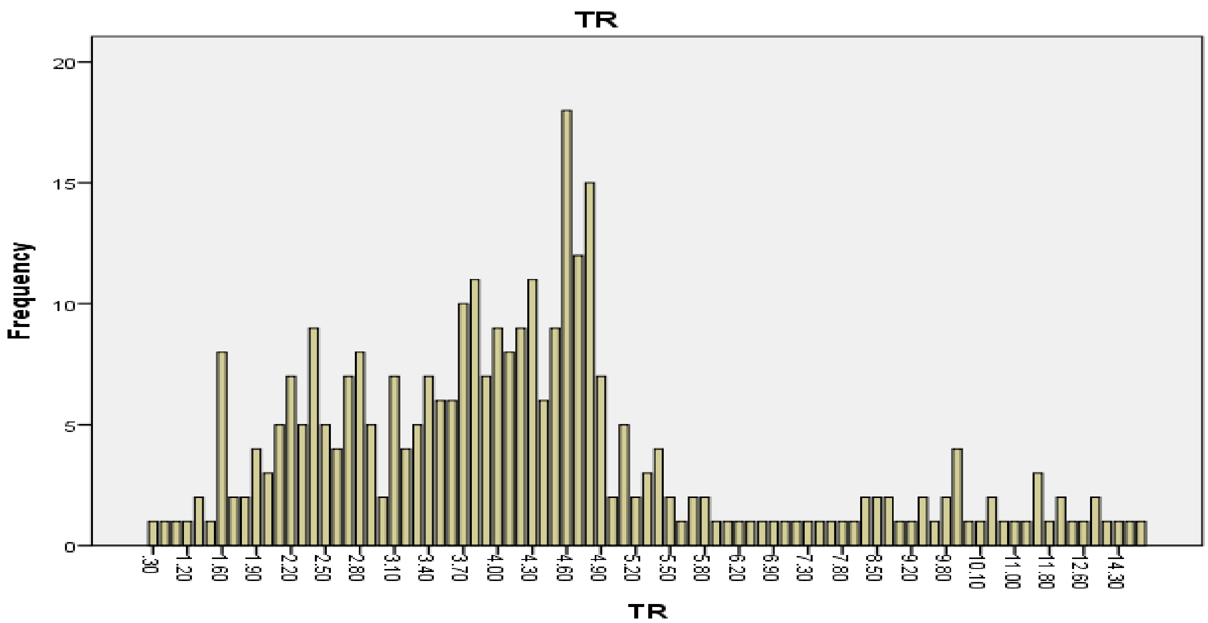
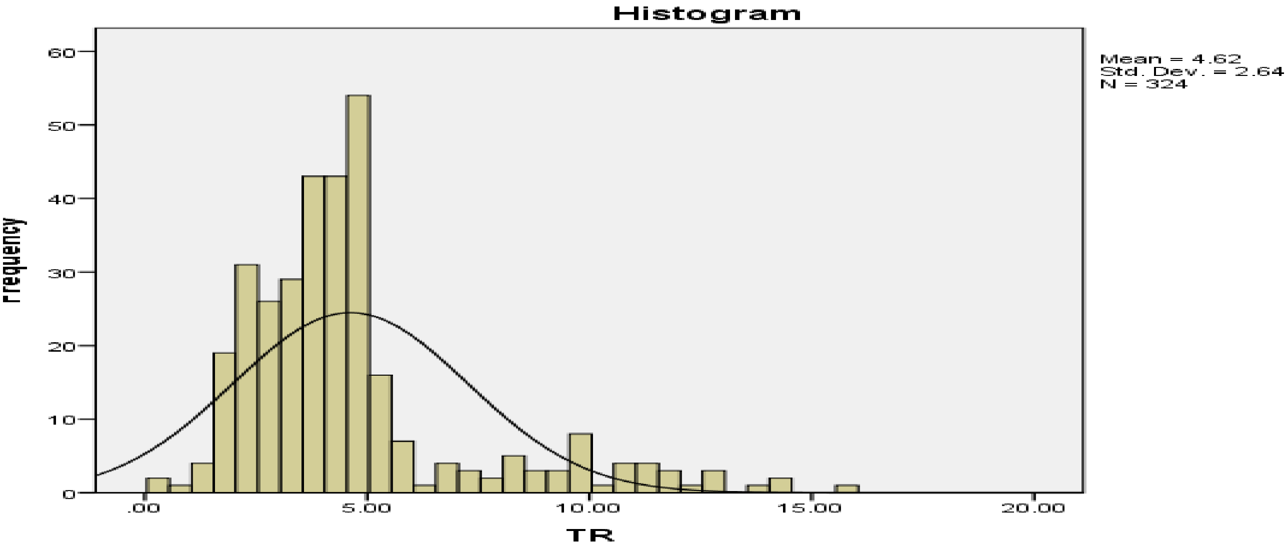


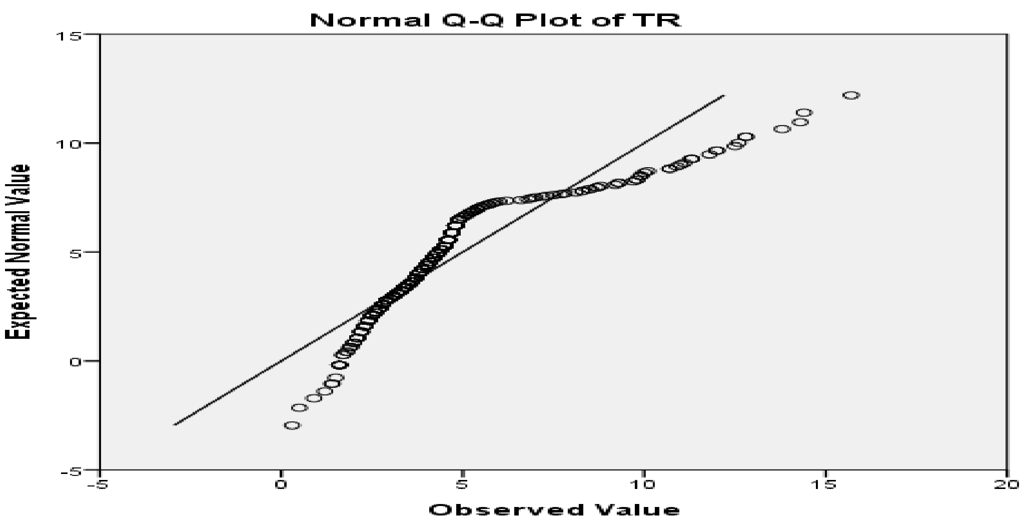
Figure 4.1 above, the bar shows plateau features which might be called a bimodal distribution, where several processes with normal distributions are combined. Because there are many peaks close together, the top of the distribution resembles a plateau, further supporting the case for a bilinear model.

Figure 4.2: Skewed Histogram of Revenue Data



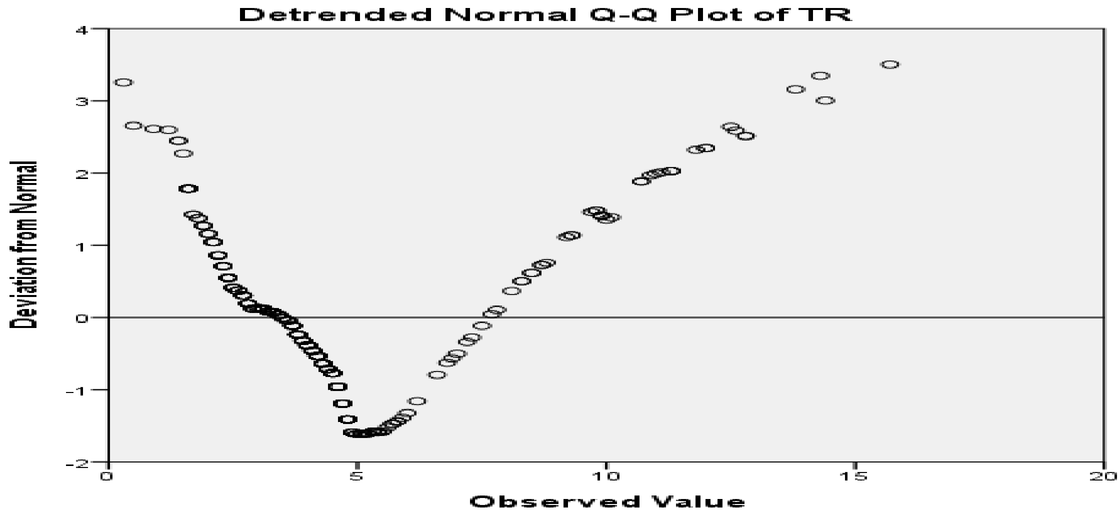
From Figure 4.2 above, the distribution peak is off-centre toward the limit, and a tail stretches away from it. The distribution of the revenue data is skewed to the right, meaning the mean is less than the median, indicating positive skewness. However, it is obvious that the distribution is not normally distributed at ground level because skewness measures the degree and direction of asymmetry. A symmetric distribution, such as a normal distribution, has a skewness of 0.

Figure 4.3 Normal Plot for Tr



From Figure 4.3, the normal plot shows that the black line indicates the values your sample should adhere to if the distribution were normal. The dots represent the actual data. If the dots fall exactly on the black line, then your data are normal. Since the data deviates from the trend, the data is non-normal and not normally distributed.

Figure 4.4 Detrended Normal Plot for Tr



The detrended Q-Q plots in Figure 4.4 above show systematic deviations from normality. Notice that the overall shape of the detrended plot is parabolic (U-shaped). The deviations from normality are relatively large. The Y-axis of the detrended normal Q-Q plot indicates that the deviations range in magnitude from -1.8 to 3.8.

Figure 4.5 Sequence Plot for the Original Data

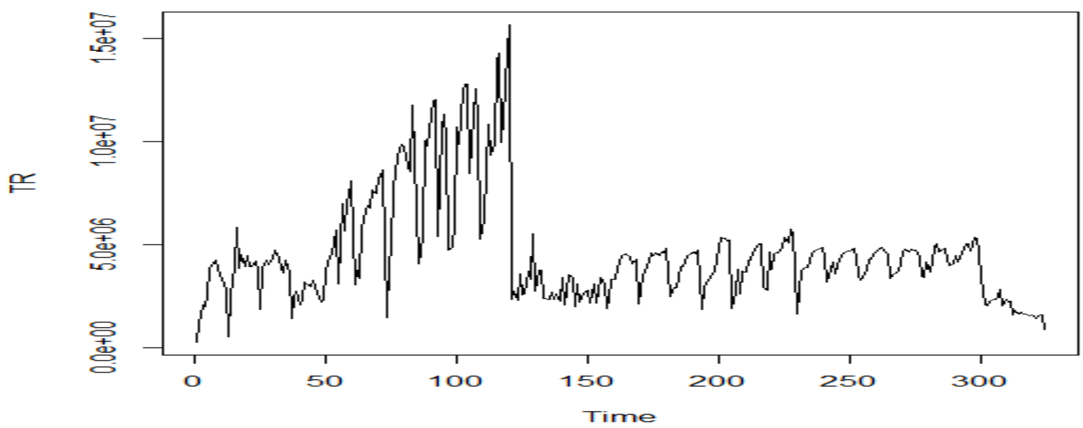
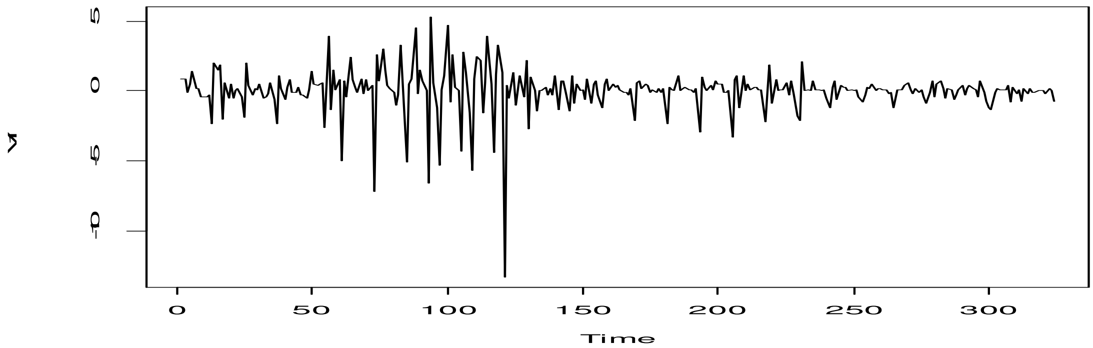


Figure 4.5 is a time series plot of the monthly revenue of the TH hotel brand over a period of 26 consecutive years. In this time series plot, the revenue is plotted against time. The series does not exhibit a consistent trend (upward or downward) over the time span. It appears to slowly wander, reaching a peak and then declining. There is no evident seasonality, despite

the data being monthly. The series also shows some outliers, making it difficult to determine whether the variance is constant.

Figure 4.6 Revenue Plot at Difference 1



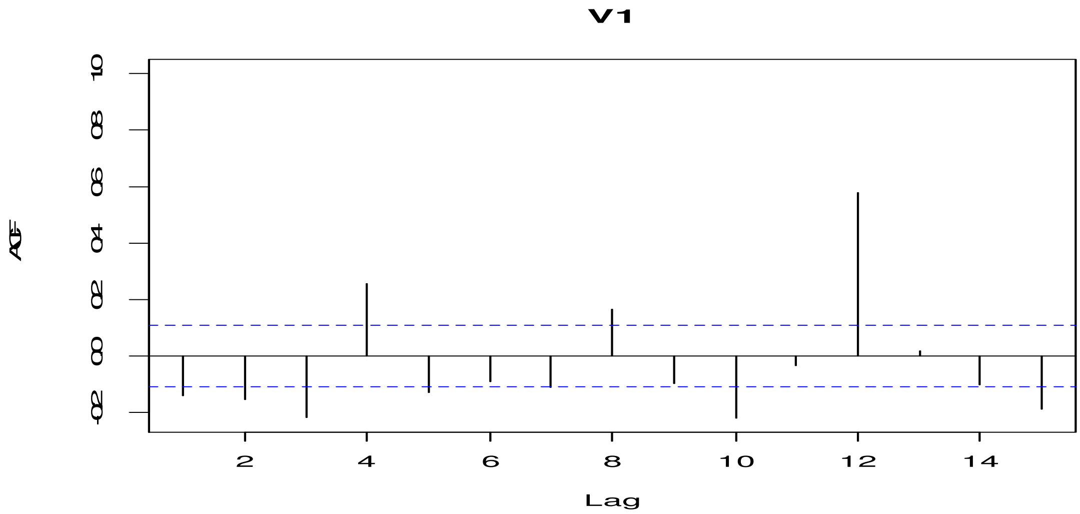
From Figure 4.6, after the first difference, the series achieves stationarity. This indicates that the mean, variance, and autocorrelation are constant over time, which facilitates easier and more accurate predictions.

AUGUMENTED DICKEY- FULLER TEST (tr) at 0 level.

$h_0 : \sigma_1 = \sigma_0$ vs $h_1 : h_0$ not true

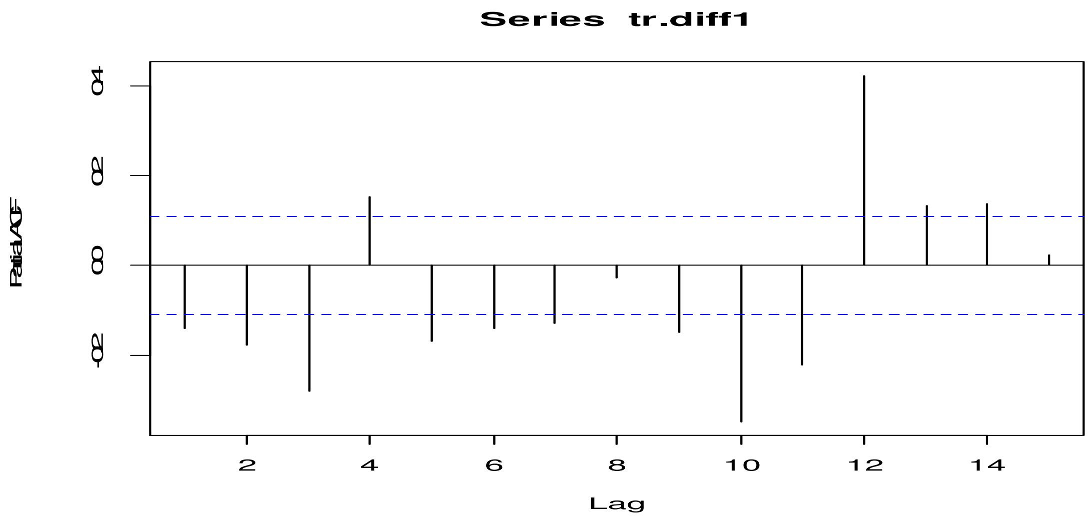
The null hypothesis is that the variable contains a unit root, while the alternative hypothesis is that the variable was generated by a stationary process. In Figure 4.1, the ADF test statistic is -3.0936, with a lag order of 6 and a P-value of 0.1154. Since the P-value is greater than 0.05, we do not reject the null hypothesis, indicating that the variable has a unit root. Therefore, we difference the data, as shown in Figure 4.2. At the first difference, the ADF test statistic is -10.263, with a lag order of 6 and a P-value of 0.01, indicating that the variable is now stationary.

Figure 4.7 Autocorrelation Function for X_t



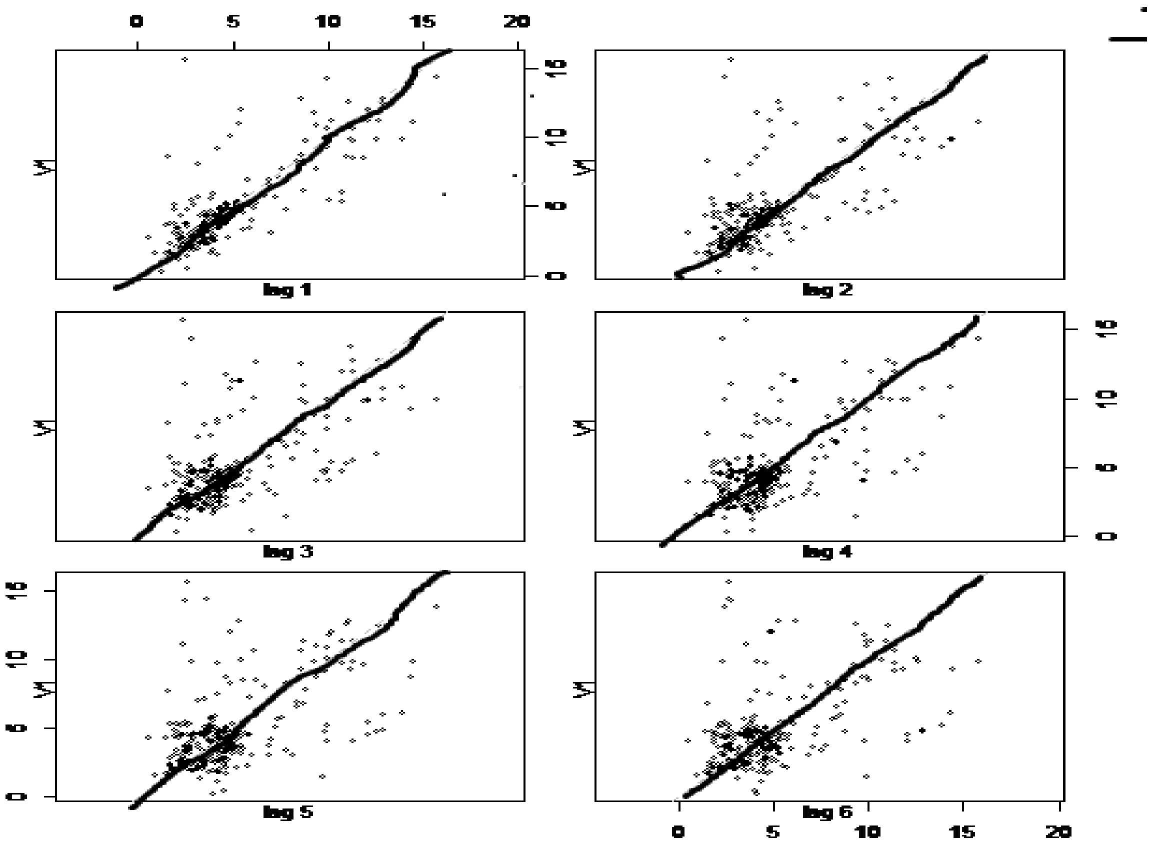
In Figure 4.3, the autocorrelation function shows spikes from the first lag to the fourth lag, followed by a gradual decay with both positive and negative lags. This pattern suggests an AR bilinear model.

figure 4.8 Partial autocorrelation function for tr. diff.1



In Figure 4.8 above, the partial autocorrelation function of the first differenced series (tr. Diff.1) shows a spike at the first negative lag, increases negatively to the third lag, and then decays positively at lag four. This positive decay continues until lag seven before it decays in a sine wave pattern. This pattern suggests a possible MA bilinear model for this data.

Figure 4.9 Residual Plot for the Six Lags



From Figure 4.9 above, the residual plot shows how randomly dispersed a variable is along the fit, as noted by Milinski (2012). If the points in a residual plot are concentrated around the horizontal axis, a linear model is appropriate for the data. Otherwise, a non-linear model is more appropriate. However, the residual plot above shows six positive typical concentrations along the trend before dispersing positively around the points, which suggests a better fit for a non-linear model.

4.6 Model Estimation

We estimate a tentative bilinear model for a particular time series data by evaluating the parameters of all the bilinear BARMA (p, P, Q, q) models, following criteria such as P-Value, Stationary R^2, Normalized BIC, and AIC, as well as Q-statistics for each identified model. The most important summary statistics to measure the goodness of fit are the Stationary R^2, normalized maximum likelihood function, standard error of the estimated value, and Q-statistics. Among the evaluated models, BARMA (6,0,6,0) emerges as the most appropriate model because it has the lowest BIC and AIC, the lowest P-value, and the highest adjusted R^2. Therefore, BARMA (6,0,6,0) is deemed suitable and the best for modelling the revenue of TH htel brand.

4.7 Estimate of Bilinear (Barma)

To present the application of the models proposed, the model contains two parts: linear and non-linear. The regression estimates obtained produce the following models for six bilinear coefficients:

$$X_{1t} = 0.001787\, x_{1t-1} - 0.142774\, x_{1t-2} - 0.020304\, x_{1t-3} \qquad 1$$

$$X_{2t} = -0.0006899\, x_{2t-1} - 0.1701088\, x_{2t-2} - 0.1817295\, x_{2t-3} - 0.0255165\, x_{2t-4}$$

$$- 0.026257\, x_{2t-5} \qquad 2$$

$$X_{3t} = -0.0001474\, x_{3t-1} - 0.2280780\, x_{3t-2} - 0.2328829\, x_{3t-3} - 0.294578\, x_{3t-4}$$

$$-0.0694322\, x_{3t-5} - 0.0361457\, x_{3t-6} - 0.0421758\, x_{3t-7} \qquad 3$$

$$X_{4t} = -0.003550\, x_{4t-1} - 0.179650\, x_{4t-2} - 0.196120\, x_{4t-3} - 0.256440\, x_{4t-4}$$

$+ 0.14663\, x_{4t-5} - 0.01768\, x_{4t-6} - 0.0115\, x_{4t-7} - 0.03442\, x_{4t-8}$

$+ 0.01662\, x_{4t-9}$ 4

$X_{5t} = -0.00758\, x_{5t-1} - 0.16518\, x_{5t-2} - 0.27001\, x_{5t-3} - 0.31964\, x_{5t-4}$

$+0.07649\, x_{5t-5} - 0.22047\, x_{5t-6} - 0.06693\, x_{5t-7} - 0.09161\, x_{5t-8}$

$-0.07271\, x_{5t-9} + 0.00148\, x_{5t-10} - 0.04133\, x_{5t-11}$ 5

$X_{6t} = -0.00895\, x_{6t-1} - 0.18887\, x_{6t-2} - 0.24305\, x_{6t-3} - 0.38078\, x_{6t-4}$

$+ 0.03756\, x_{6t-5} -0.26602\, x_{6t-6} -0.23191\, x_{6t-7} + 0.03234\, x_{6t-8}$ 6

$-0.12921\, x_{6t-9} - 0.14800\, x_{6t-10} -0.033980\, x_{6t-11} - 0.042550\, x_{6t-12} -0.04558\, X_{6t-13}$

As seen above, these models are linear in states but non-linear jointly, as the name 'bilinear' implies. The estimates provided by the models (1-6) are displayed in Tables 4.3 to 4.9.1 and are found to be accurate, as evidenced by the closeness between the actual and estimated values.

Table 4.3 Bilinear Estimates of the Parameters (Fit 1)

Type	Estimate	Std error	T-value	Pr. (> /t/)
x_{1t-1}	0.001787	0.088640	0.020	0.9839
x_{1t-2}	-0.142774	0.056321	-2.535	0.0117 *
x_{1t-3}	-0.020304	0.062773	-0.323	0.7466

Significance codes: $0^{***}, 0.001^{**}, 0.01^{*}$

The model of the bilinear fit one (1,0,1,0):

$$X_{1t} = 0.001787\, x_{1t-1} - 0.142774\, x_{1t-2} - 0.020304\, x_{1t-3}$$

From Table 4.3 above, the results display the bilinear estimates of the parameters obtained from the regression analysis conducted in R. The first and second lags are not significant, indicating that these components do not affect the model. Therefore, the final estimates are shown in Table 4.3.1 below:

Table 4.3.1 Bilinear Final Estimate of the Parameters (Fit 1)

Type	Estimate	Std error	T-value	Pr. (> /t/)
x_{1t-2}	-0.142774	0.056321	-2.535	0.0117 *

Significance codes: 0***,0.001**, 0.01*

The model from the final bilinear estimate Table 4.3.1: $X_{1t} = -0.142774\, x_{1t-2}$

Table 4.5 bilinear estimates of the parameters (fit 2)

Type	Estimate	Std error	T-value	Pr.(>/t/)
x_{2t-1}	-0.0006899	0.0876822	-0.008	0.99373
x_{2t-2}	-0.1701088	0.0562369	-3.025	0.00270 **
x_{2t-3}	-0.1817295	0.0562081	-3.233	0.00136 **
x_{2t-4}	-0.0255165	0.0620304	-0.411	0.68110
x_{2t-5}	-0.0262571	0.0620483	-0.423	0.67246

Significance codes: 0***,0.001**, 0.01*

The equation of the bilinear fit 2:

$$X_{2t} = -0.0006899\, x_{2t-1} - 0.1701088\, x_{2t-2} - 0.1817295\, x_{2t-3} - 0.0255165\, x_{2t-4} - 0.0262571\, x_{2t-5}$$

From Table 4.5 above, the results show the estimates of the parameters obtained from the regression analysis using R software. The first, fourth, and fifth lags are not significant. Therefore, these components are removed, as their non-linearity has no significant effect on the model. The final estimates of the parameters are presented in Table 4.5.1 below.

Table 4.5.1 Final Estimates of the Parameters (Fit 2)

Type	Estimate	Std error	T-value	Pr.(>/t/)

x_{2t-2}	-0.1701088	0.0562369	-3.025	0.00270 **
x_{2t-3}	-0.1817295	0.0562081	-3.233	0.00136 **

Significance codes: 0***,0.001**, 0.01*

The model from the final estimates is: $X_{2t} = -0.1701088x_{2t-2} - 0.1817295x_{2t-3}$

Table 4.6 Bilinear Estimates of the Parameters (Bilinear Fit 3)

Type	Estimate	Std error	T-value	Pr.(>/t/)
x_{3t-1}	-0.0001474	0.0844549	-0.002	0.999
x_{3t-2}	-0.2280780	0.0551959	-4.132	0.0000464 ***
x_{3t-3}	-0.2328829	0.0549340	-4.239	0..0000297 ***
x_{3t-4}	-0.2945783	0.0554842	-5.309	0.0000212 ***
x_{3t-5}	-0.0694322	0.0602523	-1.152	0.250
x_{3t-6}	-0.0361457	0.0597161	-0.605	0.545
x_{3t-7}	-0.0421758	0.0597502	-0.706	0.481

Significance codes: 0***,0.001**, 0.01*

The equation of the bilinear fit:

$$X_{3t} = -0.0001474x_{3t-1} - 0.2280780x_{3t-2} - 0.2328829x_{3t-3} - 0.2945783x_{3t-4}$$

$$-0.0694322x_{3t-5} - 0.0361457x_{3t-6} - 0.0421758x_{3t-7}$$

From Table 4.6 above, the results show the estimates of the parameters obtained from the regression analysis using R software. The first, fifth, sixth, and seventh lags are not significant. Therefore, these components are removed, as their non-linearity has no significant effect on the model. The final estimates of the parameters are presented in Table 4.6.1 below.

Table 4.6.1 Final Estimates of the Parameters (Fit 3)

Type	**Estimate**	**Std error**	**T-value**	**Pr.(>/t/)**

x_{3t-2}	-0.2280780	0.0551959	-4.132	0.0000464 ***
x_{3t-3}	-0.2328829	0.0549340	-4.239	0..0000297 ***
x_{3t-4}	-0.2945783	0.0554842	-5.309	0.0000212 ***

Significance codes: 0***,0.001**, 0.01*

The final model equation is:

$$X_{3t} = -0.2280780\ x_{3t-2} - 0.2328829\ x_{3t-3} - 0.2945783\ x_{3t-4}$$

Table 4.7 bilinear estimates of the parameters (fit 4)

Type	Estimate	Std error	T-value	Pr.(>/t/)
x_{4t-1}	-0.00355	0.08416	-0.042	0.966385
x_{4t-2}	-0.17965	0.05858	-3.067	0.002356 **
x_{4t-3}	-0.19612	0.05696	-3.443	0.000656 ***
x_{4t-4}	-0.25644	0.05766	-4.447	0.0000122 ***
x_{4t-5}	0.14663	0.06133	2.391	0.017416 *
x_{4t-6}	-0.01768	0.06374	-0.277	0.781689
x_{4t-7}	-01150	0.06031	-0.191	0.848848
x_{4t-8}	-0.03442	0.05957	-0.578	0.563831
x_{4t-9}	0.01662	0.05960	0.279	0.780615

Significance codes: 0***,0.001**, 0.01*

The equation of the bilinear fit 4:

$$X_{4t} = -0.00355\, x_{4t-1} - 0.17965\, x_{4t-2} - 0.19612\, x_{4t-3} - 0.25644\, x_{4t-4} + 0.14663\, x_{4t-5}$$

$$-0.01768\, x_{4t-6} - 01150\, x_{4t-7} - 0.03442\, x_{4t-8} + 0.01662\, x_{4t-9}$$

From Table 4.7 above, the results are the estimates of the parameters obtained from the regression analysis using R software. The first, sixth, seventh, eighth, and ninth lags are not significant. Therefore, these components are removed, as their non-linearity has no significant effect on the model. The final estimates of the parameters are in Table 4.7.1 below.

Table 4.7.1 bilinear final estimates of the parameters (fit 4)

Type	Estimate	Std error	T-value	Pr.(>/t/)
x_{4t-2}	-0.17965	0.05858	-3.067	0.002356 **
x_{4t-3}	-0.19612	0.05696	-3.443	0.000656 ***
x_{4t-4}	-0.25644	0.05766	-4.447	0.0000122 ***
x_{4t-5}	0.14663	0.06133	2.391	0.017416 *

Significance codes: 0***,0.001**, 0.01*

The new final equation of the estimate is as below:

$$X_{4t} = -0.17965\, x_{4t-2} - 0.19612\, x_{4t-3} - 0.25644\, x_{4t-4} + 0.14663\, x_{4t-5}$$

Table 4.8 Bilinear Estimates of the Parameters (Fit 5)

Type	Estimates	Std error	T-value	Pr. (>/t/)
x_{5t-1}	-0.007580	0.082856	-0.091	0.927166
x_{5t-2}	-0.165183	0.057839	-2.856	0.004590**
x_{5t-3}	-0.270008	0.059658	-4.526	$8.67x10^{-6}$ ***
x_{5t-4}	-0.319638	0.059264	-5.393	$1.46x10^{-7}$***
x_{5t-5}	0.076488	0.063295	1.208	0.227834
x_{5t-6}	-0.220465	0.062248	-3.542	0.000461***
x_{5t-7}	-0.066928	0.064522	-1.037	0.300429
x_{5t-8}	-0.0910605	0.063434	-1.444	0.149752
x_{5t-9}	-0.072712	0.059600	-1.220	0.223416
x_{5t-10}	0.001475	0.058746	0.025	0.979983
x_{5t-11}	-0.41342	0.059245	-0.698	0.485833

Significance codes: 0***,0.001**, 0.01*

The equation of the bilinear fit 5 is:

$$X_{5t} = -0.007580\, x_{5t-1} - 0.165183\, x_{5t-2} - 0.270008\, x_{5t-3} - 0.319638\, x_{5t-4} + 0.076488\, x_{5t-5}$$

$$-0.220465\, x_{5t-6} - 0.066928\, x_{5t-7} - 0.0910605\, x_{5t-8} - 0.072712\, x_{5t-9} + 0.00148\, x_{5t-10}$$

$$-0.41342\, x_{5t-11}$$

From table 4.8 above, the result is the estimates of the parameters obtained from the regression analysis using R software. The first, seven, eight, nine, ten and eleven lags are not significant. Therefore, the components are removed, since its non-linearity has no significant effect in the model. The final estimates of the parameters are in Table 4.8.1 below.

Table 4.8.1 Bilinear Final Estimates of the Parameters (Fit 5)

Type	Estimates	Std error	T-value	Pr. (>/t/)
x_{5t-2}	-0.165183	0.057839	-2.856	0.004590**
x_{5t-3}	-0.270008	0.059658	-4.526	8.67×10^{-6} ***
x_{5t-4}	-0.319638	0.059264	-5.393	1.46×10^{-7}***
x_{5t-6}	-0.220465	0.062248	-3.542	0.000461***

Significance codes: 0***,0.001**, 0.01*

The new estimate equation of bilinear fit 5 is as below:

$$X_{5t} = -0.165183\, x_{5t-2} - 0.270008\, x_{5t-3} - 0.319638\, x_{5t-4} + 0.076488\, x_{5t-5} - 0.220465\, x_{5t-6}$$

Table 4.9 Bilinear Estimates of the Parameters (Fit 6)

Type	Estimates	Std error	T-value	Pr. (>/t/)
x_{6t-1}	-0.008946	0.081846	-0.109	0.913036
x_{6t-2}	-0.188866	0.057505	-3.284	0.001144 ***
x_{6t-3}	-0.243049	0.059563	-4.081	0.000577 ***
x_{6t-4}	-0.380775	0.061198	-6.222	0.000002 ***
x_{6t-5}	0.037589	0.063575	0.591	0.54448

x_{6t-6}	-0.2666020	0.062918	-4.228	0.000031 ***
x_{6t-7}	-0.231911	0.069058	-3.358	0.000886 ***
x_{6t-8}	0.032347	0.070792	0.547	0.648053
x_{6t-9}	-0.129214	0.063979	-2.020	0.044315 *
x_{6t-10}	-0.147999	0.062886	-2.353	0.019250 *
x_{6t-11}	-0.033976	0.058924	-0.577	0.564636
x_{6t-12}	-0.042549	0.058489	-0.727	0.467509
x_{6t-13}	-0.045581	0.058530	-0.779	0.436733

Significance codes: 0***,0.001**, 0.01*

The model of the bilinear fit six (6,0,6,0):

$$X_{6t} = -0.008946\, x_{6t-1} - 0.188866\, x_{6t-2} - 0.243049\, x_{6t-3} - 0.380775\, x_{6t-4}$$

$$+ 0.0375589\, x_{6t-5} - 0.266020\, x_{6t-6} - 0.231911\, x_{6t-7} + 0.03234\, x_{6t-8}$$

$$-0.129214\, x_{6t-9} - 0.147999\, x_{6t-10} - 0.033976\, x_{6t-11} - 0.042549\, x_{6t-12}$$

$$-0.04558\, x_{6t-13}$$

4.7

From Table 4.9 above, the results are the estimates of the parameters obtained from the regression analysis using R software. The first, fifth, eighth, eleventh, twelfth, and thirteenth lags are not significant. Therefore, these components are removed, as their non-linearity has no significant effect on the model. The final estimates of the parameters are shown in Table 4.9.1 below.

Table 4.9.1 Bilinear Final Estimates of the Parameters (Fit$_6$)

Type	Estimates	Se. Coefficient	t- value	Pr(>/t/)
X_{6t-2}	-0.188866	0.057505	-3.284	0.001144 ***

X_{6t-3}	-0.243049	0.059563	-4.081	0.000577 ***
X_{6t-4}	-0.380775	0.061198	-6.222	0.000002 ***
X_{6t-6}	-0.2666020	0.062918	-4.228	0.000031 ***
X_{6t-7}	-0.231911	0.069058	-3.358	0.000886 ***
X_{6t-9}	-0.129214	0.063979	-2.020	0.044315 *
X_{6t-10}	-0.147999	0.062886	-2.353	0.019250 *

Significance codes: 0***,0.001**, 0.01*

The model from the final estimates is:

$$\hat{X}_{6t} = -0.188866X_{6t-2} - 0.243049X_{6t-3} - 0.380775X_{6t-4}$$
$$-0.266020X_{6t-6} - 0.231911X_{6t-7} - 0.129214X_{6t-9}$$

Table 4.10 Akaike Information Criterion for the Six Models

Models	Specification	AIC	BIC	Residual Error	F-statistics
Blrft1	(1,0,1,0)	1189.124	1204.147	1.576	3.213
Blrft2	(2,0,2,0)	1179.503	1202.019	1.556	4.289
Blrft3	(3,0,3,0)	1153.193	1183.188	1.497	7.791
Blrft4	(4,0,4,0)	1148.332	1185.794	1.489	6.654
Blrft5	(5,0,5,0)	1135.861	1180.777	1.464	6.785
Blrft6	(6,0,6,0)	1125.542	1125.542	1.443	6.769

From Table 4.10 above, the Akaike information criterion has shown a smaller value of AIC in Bilinear fit six (6) and BIC in Bilinear fit six (6) with both having smaller values of residual error as compared to the rest fit. Hence, the bilinear model fit 6 is chosen as against the rest models.

Table 4.11 Goodness of Fit for Bilinear Models at Sample Size 0f >300

Model	BLRFT1	BLRFT2	BLRFT3	BLRFT4	BLRFT5	BLRFT6
Res.err.	1.576	1.556	1.497	1.489	1.464	1.443
R^2	0.020	0.052	0.132	0.149	0.184	0.2142
Adj. R^2	0.014	0.040	0.115	0.127	0.157	0.183
F(stacs)	3.213	4.287	7.791	6.654	6.785	6.769
P- value	0.042	0.002	$8.12*10^{-8}$	$5.04*10^{-8}$	$1.52*10^{-9}$	$8.75*10^{-11}$

From Table 4.11, we can evaluate the performance of the six models based on the residual variance associated with each model. Notably, the residual variance for Bilinear fit six (6, 0, 6, 0) is smaller compared to the other bilinear fit models. This indicates that Bilinear fit six provides a better fit for the data, as it accounts for less unexplained variability. This finding is consistent with the results reported by Kaur, Parmar, and Singh (2023), where bilinear models demonstrated lower residual variances compared to alternative models. Their study highlights the effectiveness of bilinear models in capturing complex relationships within the data, which is further corroborated by our results.

In this study, Bilinear fit six is identified as the most effective model for estimating the revenue data of the TH hotel brand. Its superior performance, as indicated by the smaller residual variance, suggests that it provides a more accurate representation of the underlying revenue trends. This model’s performance could be attributed to its ability to better capture the non-linear interactions in the data, which other models may not fully account for. Further analysis could involve comparing Bilinear fit six with additional models not considered in this study or exploring different criteria for model evaluation to ensure robustness. Nonetheless, the evidence strongly supports the use of Bilinear fit six for the estimation of the TH brand’s revenue data.

4.8 Diagnostic Checks and Residual Variance

Diagnostic checks are essential tools for model identification. One key measure used is the coefficient of determination, which indicates the proportion of variability in the dependent variable explained by the model. The F-statistic, used for hypothesis testing in regression analysis, follows an F-distribution with p and T−p−1 degrees of freedom under the null hypothesis. The null hypothesis is rejected if the F-statistic exceeds the critical value (William, Sweeney, and Anderson, 2006).

Furthermore, the F-statistic is commonly used to compare statistical models fitted to a dataset, helping to identify which model best represents the population from which the data were sampled. For the bilinear model (6, 0, 6, 0), the F-statistic is 18.9 with 12 and 289 degrees of freedom, and the p-value is < 2.2e-16. This F-statistic is significant given the critical values of 1.75 at a 0.05 significance level and 1.95 at a 0.01 significance level, indicating that the model's performance is statistically significant.

Table 4.12 Test for Residual Variance In TR

Model	f- statistics	p-value	df
Bil. Fit six (6,0,6,0)	18.9	$< 2.2\ e^{-16}$	289

Additionally, as shown in Table 4.12, after fitting the six models, the residual variances were calculated for each model. Among these, the bilinear model (6, 0, 6, 0) exhibited the smallest residual variance. This smaller variance suggests that Bilinear Fit Six provides a better fit to the data compared to the other five models.

Table 4.13 Aic/Bic For The Simulated Bilinear Models

Models	Specific.	AIC	BIC	RES Er.	F-STAT.

Blrft1	(1,0,1,0)	1644.489	1659.393	3.50	44.07
Blrft2	(2,0,2,0)	1600.418	1622.759	3.27	36.72
Blrft3	(3,0,3,0)	1587.577	1617.34	3.22	27.36
Blrft4	(4,0,4,0)	1571.846	1609.016	3.15	23.22
Blrft5	(5,0,5,0)	1559.156	1603.731	3.11	19.97
Blrft6	(6,0,6,0)	1540.628	1592.574	3.03	18.90

From Table 4.13, it is observed that the information criteria for the simulated random variable of the revenue data indicate that Bilinear Fit Six (6) achieves the smallest values for both the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). Additionally, Bilinear Fit Six (6) has a a lower residual error and a more favourable F-statistic compared to the other models. Given these results, Bilinear Fit Six (6) is selected over the other models. The smaller residual variance of Bilinear Fit Six (6) (6, 0, 6, 0) further supports this choice, aligning with findings from Usoro (2017), who also reported that bilinear models can outperform other models in terms of residual variance.

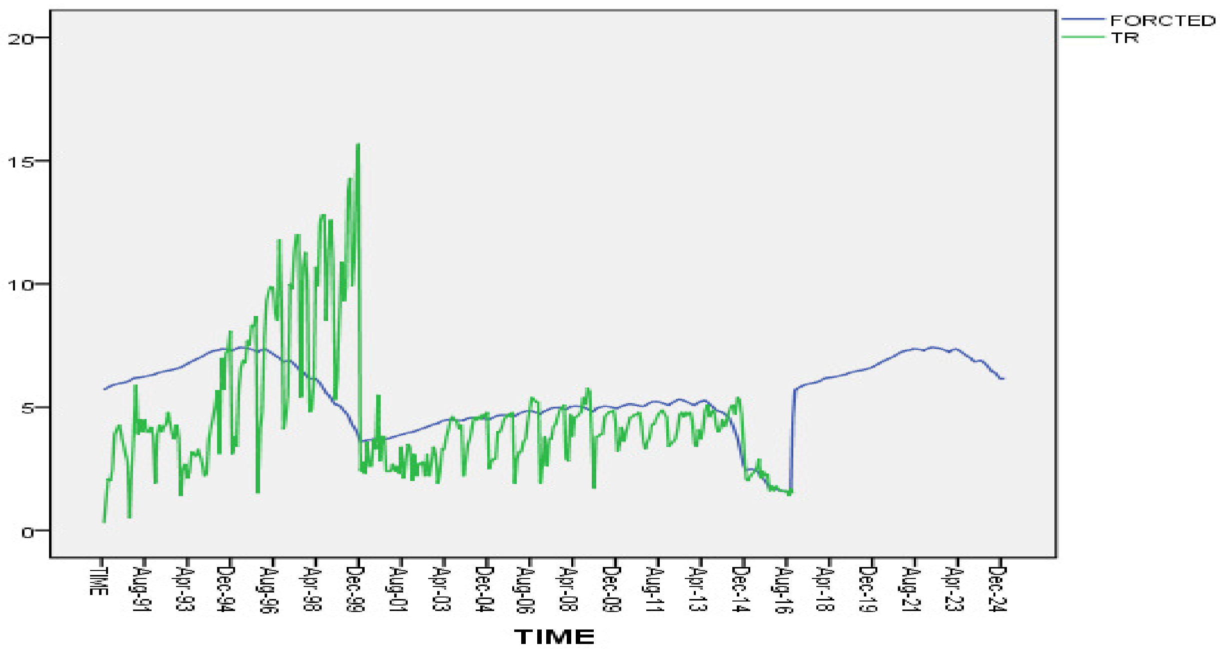
Table 4.14 Comparison Between Actual and Simulated Bilinear Models

	ACTUAL AIC/BIC BILINEAR MODEL FOR (TR)					**SIMULATED AIC/BIC BILINEAR MODELS (TR)**				
	SPECIF.	AIC	BIC	RE	F.STAT.	SPECIF.	AIC	BIC	RE	F.STAT.
FIT1	1,0,1,0	1189.12	1204.15	1.58	3.213	1,0,1,0	1644.49	1659.39	3.50	44.07
FIT2	2,0,2,0	1179.50	1202.02	1.56	4.287	2,0,2,0	1600.42	1622.76	3.27	36.72
FIT3	3,0,3,0	1153.19	1183.19	1.50	7.790	3,0,3,0	1587.58	1617.34	3.22	27.36
FIT4	4,0,4,0	1148.33	1185.79	1.49	6.654	4,0,4,0	1571.85	1609.02	3.15	23.22

FIT5	5,0,5,0	1135.86	1180.78	1.46	6.785	5,0,5,0	1559.16	1603.72	3.10	19.97
FIT6	6,0,6,0	1125.5	1177.7	1.44	6.770	6,0,6,0	1540.63	1592.57	3.03	18.90

From Table 4.14 above, the model selection criteria, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), are techniques based on simple fit to estimate the likelihood of a model to predict or estimate future values (Akaike, 1974). A good model is one that has the minimum AIC/BIC among all the other models. A lower AIC or BIC value indicates a better fit. Considering the simulated and actual values of all the AIC and BIC, the simulated value is higher in both criteria, which further shows that the actual data is a better fit than the simulated data. Specifically, the bilinear fit six (6,0,6,0) has the smaller BIC and AIC in the simulated data, which again shows that the actual data fit agrees. Conclusively, bilinear fit six in both simulated data and actual data are both smaller in AIC and BIC, which agrees with Akaike (1974) and Subba Rao (1981).

Figure 4.10 Forecasted Chart for Revenue



From the forecasted chart in Figure 4.10 above, the revenue from 1990 increases up to 1992 and then reduces drastically in 1993. However, from 1994, there is a sharp increase in revenue generation, which could be attributed to changes in service delivery to their esteemed guests. In 1999, with the arrival of democracy, the revenue reached an all-time high. This could be attributed to the influence of foreign investment in the country and Abuja. However, this did not last as the revenue dropped sharply in 2000, continuing until 2003. Thereafter, the revenue showed upward and downward oscillations until 2014, attributed to the increase in exchange rates and market volatility due to insecurity, which significantly affected the business.

In 2015, the economic meltdown affected all hospitality outfits, with TH hotel brand Abuja suffering its fair share, resulting in a drastic drop in revenue to an all-time low. In 2016, things did not change much because the increase in exchange rates made all the gains in revenue irrelevant. The forecasted trend shows that the revenue will keep increasing. Despite COVID-19 disrupting business worldwide, the hotel brand maintained its revenue, which has increased. Estimates show that this will remain so until 2025, as shown above. This trend is expected to continue increasing until 2030, suggesting that management is investing more in training, renovation, and adapting to economic changes.

4.9 Addressing the Research Questions and Hypotheses

Comparison of Estimates:

Research Question:

How do the estimates of the simulated bilinear fit (six) compare to the actual bilinear fit (six) for the revenue data of TH hotel brand Abuja?

Results:

From Table 4.13, we observe that the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) for the simulated bilinear fit (six) are higher than those for the actual bilinear fit (six). Specifically, Bilinear Fit Six (6,0,6,0) exhibits the lowest

AIC and BIC values among the models evaluated. According to Akaike (1974) and Rao (1980), a lower AIC or BIC value indicates a better model fit, suggesting that the actual bilinear fit provides a more accurate representation of the revenue data compared to the simulated fit. This finding aligns with the results of similar studies, such as those by Maghsoodi (2023), Ojo (2010), and Usoro (2017), which also demonstrated the superiority of bilinear models in fitting complex time series data.

Forecasting Future Values:

Research Question:

What are the forecasted future values of the revenue observations for TH hotel brand Abuja from 2017 to 2025 using the bilinear time series model?

Results:

The forecasting analysis, as illustrated in Figure 4.1, shows that revenue for TH hotel brand Abuja is projected to increase steadily until 2025. This trend is attributed to improvements in service delivery, economic factors, and management strategies. The historical data trends, including the impact of political and economic changes (e.g., the democratic transition in 1999 and economic downturns in 2015), support the forecasted increase in revenue. The bilinear time series model effectively captures these trends, providing forecasts that are consistent with historical observations and the management's strategic actions.

Hypotheses - Comparison of Estimates:

Null Hypothesis (H0): There is no significant difference between the estimates of the simulated bilinear fit (six) and the actual bilinear fit (six).

Alternative Hypothesis (H1): There is a significant difference between the estimates of the simulated bilinear fit (six) and the actual bilinear fit (six).

Results:

The results suggest a significant difference between the simulated and actual bilinear fits, as indicated by the smaller AIC and BIC values for the actual data compared to the simulated data. This supports the rejection of the null hypothesis and acceptance of the alternative hypothesis.

Forecasting Accuracy:

Null Hypothesis (H0): The bilinear time series model does not provide accurate forecasts of the revenue observations for TH hotel brand Abuja from 2017 to 2030.

Alternative Hypothesis (H1): The bilinear time series model provides accurate forecasts of the revenue observations for TH hotel brand Abuja from 2017 to 2030.

Results:

The forecasting results suggest that the bilinear model predicts an increasing trend in revenue up to 2030. This indicates that the bilinear time series model is likely to provide accurate forecasts, supporting the rejection of the null hypothesis in favour of the alternative hypothesis.

Two key variables informed the preference for the BARMA (6,0,6,0) model. The preference for models with lower AIC and BIC values is supported by the work of Akaike (1974) and Schwarz (1978), who established these criteria for model selection based on trade-offs between model fit and complexity. Our results confirm that Bilinear Fit Six aligns with these criteria, emphasizing its robustness. In terms of its forecasting accuracy, previous research, including that by Usoro (2017), highlights the effectiveness of bilinear models in capturing complex patterns in time series data. This study corroborates those findings by showing that the bilinear model provides accurate forecasts of future revenue, aligning with established literature on forecasting techniques.

4.10 Limitations and Assumptions

One limitation of the bilinear model is its assumption of linearity in the parameters, which may not fully capture all non-linearities in the data. While the bilinear model performs well, it may not account for all potential variables affecting revenue. In this context, the forecasting model assumes that past trends and management strategies will continue unchanged (Maghsoodi 2023). Any significant deviation in economic conditions, political stability, or company policies could impact the accuracy of the forecasts. However, regarding data accuracy, the accuracy of both the model fit and forecasts is dependent on the quality and completeness of the input data. Missing or inaccurate historical data can lead to less reliable estimates and predictions.

5.0 Conclusion and Discussion

The results suggest that the Bilinear Fit Six (6,0,6,0) model is highly efficient in predicting revenue. The model demonstrates a good fit to the data, as indicated by the low standard error (1.443), BIC (1125.547), and AIC (1125.542) values, which minimise the unexplained variance. The F-statistic of 6.769 and the P-value of 8.746×10^{-11} provide strong evidence for the model's robustness and predictive power.

The bilinear model's forecast of a rising revenue trend until 2030 aligns with the conclusions of other studies, such as Abu Hammad (2024). These studies consistently emphasise the efficacy of bilinear models in capturing intricate relationships within time series data. Yu's (2024) research demonstrates that bilinear models tend to outperform linear models in forecasting accuracy, especially in situations involving non-linear interactions. While the results support rejecting the null hypothesis in favour of the alternative hypothesis, it also aligns with established research. It enhances the credibility of the bilinear model's forecasts.

The congruity of results with established literature, exemplified by the works of Yu (2024) and Mill (2019), underscores the dependability of bilinear models for revenue prediction. The studies have shown that bilinear models can accurately represent the changing patterns in revenue data, making them well-suited for predicting future trends over extended periods of time.

It is essential to highlight the significance of consistently monitoring and updating the model parameters to ensure its relevance and accuracy in response to changing conditions. This methodology guarantees the model's adaptability and ability to consistently produce accurate predictions, instilling confidence in its long-term dependability. Future research could explore alternative models or additional variables to enhance forecast accuracy and model robustness.

References

Abbasimehr, H., Behboodi, A. and Bahrini, A., 2024. A novel hybrid model to forecast seasonal and chaotic time series. *Expert Systems with Applications*, *239*, p.122461.

Akaike, H. (1974). A new look at the statistical model identification, IEEE Transactions on

Amadi, E.C., & Aboko, S.A. (2013). *ARIMA Model Fitting for Nigeria's Gross Domestic Product*. Nigerian Journal of Economics and Statistical Studies, 11(3), 112-130.

Amadi, E.C., & Aboko, S.A. (2013). *ARIMA Model Fitting for Nigeria's GDP*. Nigerian Journal of Economics and Statistical Studies, 11(3), 112-130.

Amadi, I. U. & Aboko, S. I.(2013) The fitting of ARIMA model in forecasting Nigeria Gross Domestic Product, journal of physical science and innovation. 5(2).

Automatic Control AC-19: 716–723.

Bielinska, E. (2008). *Hybrid Linear–Bilinear Prediction Model in Economic Forecasting*. Economic Modelling, 25(3), 450-460.

Boonchai, S., & Eivind, K. (2005). *General Multivariate Bilinear Time Series Models*. Journal of Time Series Analysis, 26(4), 123-138.

Bouzaachane, S., Zouhri, A., & Benaicha, M. (2006). *Estimation of Bilinear Time Series Models: A Case Study*. Journal of Statistical Research, 10(2), 45-62.

Buckman, J., & Enock, D. (2013). *SARIMA Modeling of Ghana's Inflation Rate*. International Journal of Forecasting, 29(2), 248-261.

Cao, X., Benner, P., Duff, I., & Schilders, W. (2020). Model order reduction for bilinear control systems with inhomogeneous initial conditions. International Journal of Control, 94(10), 2886-2895.

Challapilla, R., & Rao, T.S. (1998). *Newton-Raphson Algorithm for Bilinear Time Series Models*. Journal of Computational and Graphical Statistics, 7(1), 85-101.

Eke, D.O., Okoro, E.N., & Nwachukwu, J.O. (2015). *Comparative Analysis of Time Series Models on Nigeria's Gross Domestic Product*. Journal of Economic Modeling, 17(4), 234-245

Etuk, E.E. (2011). *Seasonal ARIMA Modeling of Nigeria-British Pound Exchange Rate*. African Journal of Mathematics and Statistics, 8(2), 78-91.

Etuk, E.E. (2012). *Seasonal ARIMA Modeling of Nigeria's Gross Domestic Product*. African Journal of Mathematics and Statistics, 9(2), 101-118.

Fritzer, G., Gabriel, K., & Johann, H. (2002). *Forecasting Austrian HICP Inflation: A Comparison of VAR and ARIMA Models*. European Journal of Economic and Social Systems, 15(3), 211-224

Gomez, V., & Maravall, A. (1998). *Automatic Modeling for Univariate Time Series: A Unified Approach*. Journal of Time Series Analysis, 19(4), 357-377.

Gooijger, J.G., & Heauts, P.A. (1986). *Bilinear Time Series Models: A General Survey*. Statistics & Probability Letters, 4(1), 23-29.

Granger, C.W.J., & Anderson, A.P. (1978). *An Introduction to Bilinear Time Series Models*. Vandenhoeck & Ruprecht.

Haggan, M., & Ozaki, M. (1980). *On the Estimation of Nonlinear Time Series Models*. Journal of Statistical Planning and Inference, 4(2), 113-127.

Hamidreza, M., & Leila, S. (2012). *Predicting Iran's Oil Supply Using SARFIMA Models*. Energy Economics, 34(6), 1295-1305.

Hammad, M. (2024). Novel insights into estimation of bilinear time series models with exponential and symmetric coefficients. Symmetry, 16(4), 405.

Hannan, E.J., & Kavalieris, L. (1984). *Automatic Identification Procedures for Vector Autoregressive Moving Average Models*. Journal of Time Series Analysis, 5(4), 287-297.

Hili, W. (2008). *Simplifying Bilinear Structures in Time Series Analysis*. Journal of Time Series Econometrics, 3(2), 159-176.

Ismail, M., & Maphol, J. (2005). *Modeling and Forecasting Malaysian Electricity Generation Using SARIMA Models*. Energy Economics, 27(6), 947-963.

Housh, M., 2023. Optimizing bilinear multi-source water supply systems using mixed-integer linear programming approximations: An analysis of the Israeli seawater desalination array. *Advances in Water Resources*, *178*, p.104498.

Akpanta, A.C. and Okorie, I.E., 2015. ARI (p, d) Modelling and forecasting of Nigeria's External Reserves. *World Journal of Probability and Statistics Research*, *1*(1), pp.1-9.

Iwueze, I.S., Nwogu, E.C. and Nlebedim, V.U., 2013. Time series modeling of Nigeria external reserves. *CBN Journal of Applied Statistics*, *4*(2), pp.111-128.

Yarrington, C.S., 2021. *Review of forecasting univariate time-series data with application to water-energy nexus studies & proposal of parallel hybrid SARIMA-ANN model*. West Virginia University.

Aras, S. and Kocakoç, İ.D., 2016. A new model selection strategy in time series forecasting with artificial neural networks: IHTS. *Neurocomputing*, *174*, pp.974-987.

Abu Hammad, M.M., Alkhateeb, R., Laiche, N., Ouannas, A. and Alshorm, S., 2024. Comparative Analysis of Bilinear Time Series Models with Time-Varying and Symmetric GARCH Coefficients: Estimation and Simulation. *Symmetry*, *16*(5), p.581.

Iwok, I.A., & Etuk, E.H. (2009). *Comparative Performance of Pure Vector Autoregressive Moving Average and Vector Bilinear Autoregressive Moving Average Time Series Models*. Journal of Statistical Computation and Simulation, 79(6), 689-700.

Iwueze, E., Etuk, E., & Nwachukwu, J. (2013). *Forecasting Nigeria's External Reserves Using ARIMA Models*. Nigerian Journal of Economics and Statistics, 12(1), 78-95.

Jan, P., & Kuldeep, S. (1992). *Non-linear Time Series Models: Challenges and Alternatives*. Journal of Time Series Analysis, 13(4), 235-247.

Jones, J. (1978). *Nonlinear Time Series Models: A Review and Extension*. Journal of Econometrics, 8(3), 187-205.

Kaur, J., Parmar, K.S. and Singh, S., 2023. Autoregressive models in environmental forecasting time series: a theoretical and application review. *Environmental Science and Pollution Research*, *30*(8), pp.19617-19641.

Kendell, D., & Ord, J.K. (2015). *Time Series Analysis: Theory and Practice*. Cambridge University Press.

Krzanowski, W. J. (1998). *Principles of Multivariate Analysis: A User's Perspective*. Oxford University Press

Liela, A., & Masoud, H. (1998). *Empirical Study of SARFIMA Models in Energy Science*. Energy Studies Review, 9(2), 245-260.

Lukasz, R. (2016). *Stable Elementary Bilinear Time Series: Estimation and Applications*. Journal of Time Series Analysis, 37(2), 184-202.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer.

Maghsoodi, A.I., 2023. Cryptocurrency portfolio allocation using a novel hybrid and predictive big data decision support system. *Omega*, *115*, p.102787.

Maravall, A. (1983). *Forecasting Spanish Monetary Data Using Bilinear Models*. International Journal of Forecasting, 1(4), 37-45.

Mathews, J., & Lee, C.E. (1993). *Stability and Stationarity of Bilinear Time Series Models*. Annals of the Institute of Statistical Mathematics, 45(3), 449-471.

Mills, T.C., 2019. *Applied time series analysis: A practical guide to modeling and forecasting*. Academic press.

Mohler, G.O., & Tang, B. (1988). *Theoretical Foundations of Bilinear Time Series Models*. Communications in Statistics - Theory and Methods, 17(11), 3793-3813.

Muio, J. (2012). *Impact of exchange rates on the financial performance of hotels in Kenya*. Journal of Hotel and Business Management, 1(2), 45-53.

Nascimento, A., Lima, M., Bakouch, H., & Qarmalah, N. (2023). Scaled muth–arma process applied to finance market. Mathematics, 11(8), 1908.

Ojo, J.A. (2009). *Challenges in Bilinear Time Series Models: Achieving Stationarity for Nonlinear Series*. Journal of Statistical Research, 13(1), 88-102.

Ojo, J.A. (2012). *Generalized and Subset Integrated Autoregressive Moving Average Bilinear Time Series Models Applied to Sunspot Numbers*. Journal of Statistical Research, 14(2), 89-105.

Ojo, J.F. and Olanrewaju, R.O., 2021. Review of Family of Autoregressive Integrated Moving Average Models in the Comportment of Autocorrelation Function for Non-Seasonal Time Series Data. *International Journal of Mathematical Sciences & Applications*, *19*(1), pp.79-89.

Okereke, C., & Bernard, S. (2014). *Fitting SARIMA Models to Nigeria's Quarterly GDP Data*. Journal of Economic Modelling, 17(3), 245-258.

Ozaki, M., & Oda, N. (1977). *Statistical Estimation of Wiener Kernels: A Review*. Journal of Time Series Analysis, 8(2), 114-127.

Pham, D.T. (1986). *Estimation of Parameters of Bilinear Time Series Models*. Journal of Time Series Analysis, 7(1), 61-71.

Poskitt, D.S. (1992). *Automatic Order Determination in Multivariate Autoregressive Moving Average Models*. Journal of the Royal Statistical Society: Series B (Methodological), 54(3), 519-531.

Schwarz, G. (1978). Estimating the dimension of a model, Annals of Statistics 6: 461– 464.

Subba Rao, T. (1980). *The Theory of Bilinear Time Series Models*. Journal of the Royal Statistical Society: Series B (Methodological), 42(3), 388-400.

Subba Rao, T. (1981). *On the Theory of Bilinear Time Series Models*. Journal of the Royal Statistical Society: Series B (Methodological), 43(2), 244-255.

Tong, H., & Lim, S. (1980). *Nonlinear Time Series Analysis: Theory and Methods*. Journal of Time Series Analysis, 2(4), 245-259.

Usoro, A. E. (2015). Comparative Analysis of Linear and Bilinear Time Series Model, American

Usoro, E. A. & Omekara O. C. (2008) Pure diagonal bilinear moving average vector models, Global Journal of Mathematical Sciences. 137 – 145.

Usoro, E., & Omekara, J. (2008). *Comparative Performance of Linear and Bilinear Time Series Models in Revenue Data*. Statistical Modelling, 8(1), 91-105.

Volterra, V. (1930). *Theory of Functionals and of Integral and Integro-Differential Equations*. Dover Publications.

Wiener, N. (1958). *Nonlinear Problems in Random Theory*. Wiley.

Wu, X. and Lu, Q., 2021. Financial asset yield series forecasting based on risk-neutral fuzzy bilinear regression and probabilistic neural network. *Journal of Intelligent & Fuzzy Systems*, *40*(6), pp.11829-1184.

Yu, G., 2024. Enhancing Accounting Informatization Through Cloud Data Integrity Verification: A Bilinear Pairing Approach. *Journal of the Knowledge Economy*, pp.1-18.

Zielinski, R. (2000). *Nonlinear Minimum-Variance Control with Applications to Time Series Models*. Control and Decision, 16(4), 559-574.

