

FORECASTING ELECTRICITY DEMAND IN GHANA WITH THE SARIMA MODEL

P. Y. Andoh^{1*}, C. K. K. Sekyere¹, L. D. Mensah^{1,2}, D. E.K. Dzebre^{1,2}

Department of Mechanical Engineering, Kwame Nkrumah University of Science and Technology (KNUST), Kumasi, Ghana¹

The Brew Hammond Energy Centre, KNUST, Kumasi, Ghana.²

[*pyandoh.coe@gmail.com](mailto:pyandoh.coe@gmail.com)

Received : 27 October 2021, Revised: 04 December 2021 , Accepted : 16 December 2021

*Corresponding Author

ABSTRACT

Demand forecasting is a challenging subject of interest to many organizations whose main focus is to improve their steady growing customer request/demand, and help in increasing their revenue generation. The story is no different in the power industry. It is quite difficult for power or electrical producers to store high quantum of the energy produced, hence this poses a challenge in estimating precisely the quantum of electrical energy in order to equate demand and supply of powers as well as reducing or eliminating the rising transmission losses. This study explores potential time series models in electricity demand prediction or forecasting for the Western Regions of Ghana. Secondary data was sourced formally from the regional headquarters of ECG to aid in research design to be able to estimate the quantum of electricity needed by consumers in the region. This was done using time series data analysis toolpak software. Results show that the models formulated are viable for future consumption forecasts and other investment in alternative power source projects in meeting these future demands. Since there are up-surgings energy demand patterns in the region, the flexibility of the formulated models can be very useful and supplementary to framing effective and efficient energy policies.

Keywords : Electricity demand, Forecasting SARIMA, Ghana

1. Introduction

Accurately modeling and forecasting electricity demand is a very important issue for decision making in deregulated electricity markets. For the efficient management of day-to-day operations of a power system, short term forecasts are very important (Jain et al., 2009), whereas for generator maintenance, acquiring energy resources, balance sheet calculations, new structural investment based on the anticipation of future demand or for long term bilateral contracts to ensure low financial risk, medium and long term forecasts are usually utilized.

Efficient and reliable electricity generation, transmission and distribution are critical to achieving accelerated growth in all sectors of Ghana's economy. As Ghana's industry base started growing rapidly in the early 2000s, there was a commensurate increase in the country's national electric energy consumption from approximately 364 kilowatt-hours per capita (kWh/capita) in 2000 to 534 kWh/capita in 2020 (Asumadu-Sarkodie and Owusu, 2016a; Energy Commission of Ghana, 2021).

Even with increases in generating capacity over the years arising from public and private sector participation, there are still severe challenges with reliable supply due to growing demand and network instabilities emanating from technical challenges with thermal generation which now constitutes over 64% of Ghana's gross generating capacity (Energy Commission of Ghana, 2021).

Electricity forecasting can be employed for demand side management. In demand side management, customers are cautioned on energy saving tips in conserving electricity during the high demand periods called the peak hours and utilize it during the low peak hours.

Statistical models such as regression models, exponential smoothing and time series models are widely used for electricity pricing and demand forecasting problems. These models generally perform better in cases of short-term forecasting and are extensively studied (Cheng et al, 2015; Trindade, 2002; Kim et al., 2002; Kyriakides and Polycarpou, 2007; Taylor and McSharry, 2007;

Weron and Misiorek 2005; Baki et al, 2005; Abdel-Aal et al, 1997). Exponential smoothing method uses a weighting factor known as the smoothing constant (parameter) that reflects the weight given to the most recent data values. Smoothing constant value lies between 0 and 1 and it determines the degree of smoothening and consequently, how responsive the model is to the fluctuations in the data (Bosq, 2015). In the electricity context, these techniques have been widely used to model and predict electricity demand and prices. In the short term forecast the techniques generally produce better results as greater weights are given to more recent observations. However, the forecast is relatively sensitive to the specification of the smoothing constant. In general, these methods are capable of accounting for both the time correlation and random nature existing in the underlying phenomenon. To account for the possible correlation among different load periods, vector autoregressive (VAR) models are also used to forecast demand series (Sulugodu & Deka 2019; Ang, 2005).

Proper demand management facilitates the planning and use of resources for positive and profitable results, and may involve marketing programs designed to increase or reduce demand in a relatively short time (Feinberg and Genethliou, 2005; Louie, 2017).

The ability to estimate demand depends on a medium of unearthing the values for demand in future crises. This explains the critical role of demand forecasting in managing every business entity. The decision of an organization to invest and/ or expand an existing or new infrastructure, increase human resource capacity and process/procedure depend on what the future depicts, (Boisseleau, 2004). The key roles or benefits of forecasting in demand management are in: stabilizing employment and production, taking management decisions, evaluating performance, planning of facilities and an optimal day-to-day operation of the plants. (Feinberg and Genethliou, 2005; Kyriakides and Polycarpou, 2007; Hahn and Nieberg, 2009; Sarkodie, 2017; Adom and Bekoe, 2012).

Hahn and Nieberg (2009), in their extensive review of methods and models for load demand forecasting found that, though may suffer from numerical instabilities, time-series methods such as the Autoregressive Integrated Moving Average Model (ARIMA) model, are commonly used in load forecasting. The ARIMA model is a special kind of a regression model which comprises an Autoregressive (AR) component and a Moving Average (MA) component that are dependent on past values and past errors, respectively. ARIMA models are comparatively good in modeling and forecasting as compared to other linear time series models. However, these models are not capable of capturing both seasonal and non-seasonal patterns in a time series. An extension of these models to capture seasonality is the Seasonal Autoregressive Integrated Moving Average (SARIMA) model. A time series is described as exhibiting seasonality if there exists a regular pattern in changes that repeats over several time intervals until the patterns repeat again. SARIMA models are well known for statistical modeling and forecasting but are, however, unable to extract nonlinear relationships within the time series data (Box and Jenkins, 1976; Meese and Geweke, 1984; Harvey, 1990; Marcellino, 2007; Atilla et al, 2007; Shitan and Peiris, 2011; Stock and Watson, 1996; Stock and Watson 2003; N.B Ahmed, 2018; Asumadu-Sarkodie and Owusu, 2016d; Box et al, 1987; Jiang et al, 2018).

A significant amount historical demand data exists for the Western regions of Ghana, however, this data which is generated within a 24-hour base has only been used for monthly billing of customers rather than gathering other insights from it to help improve on customer satisfaction and facilitate strategic decision making.

The SARIMA model is touted as having the capability to predict more accurately than traditional forecasting techniques. This study therefore seeks to evaluating the effectiveness of SARIMA in predicting electricity demand using data from the Western Region of Ghana

2. Research Methods

2.1. Case Study Area

The Western (i.e. Western and Western-North) Regions of Ghana are located at the southern part of the country, along the Atlantic Ocean. The area is regarded as one of the agro-based regions in the country due to its fertile soil and favorable rainfall pattern.

The area has a projected population of almost three million, one hundred thousand (3, 100, 000) (Ghana Statistical Service, 2021).

The area covers a total land area of approximately twenty-four thousand square kilometers (24, 000 sq. kms). The area is blessed with abundant natural resources like gold, bauxite, diamond, manganese, as well as the commercial discovery of oil and gas known as the black gold. The presence of these mineral resources makes the regions view themselves as some of the most tax resourceful regions to the government in executing its mandate to the area and the country as a whole.

Over the past decade, more than 58% of electricity generated or produced in Ghana has been utilized by the non-residential and industrial sectors (Energy Commission of Ghana, 2020), of which the quantities consumed by mining companies in the Western Region of Ghana cannot be ignored.

2.2. Data Collected

Electricity cannot be stored in large quantum, hence, one can describe its demand and utilization as just in time; that is, once produced, electricity must be transmitted and distributed promptly for its usage. The continuous usage of this electrical power allows metering and billing to be done from seconds to years, as appropriate, by the distributing company. This monitoring and metering are done through the usage of advanced smart meters, some of which are prepaid meters, whilst others are postpaid meters. Postpaid or prepaid advanced smart meters are installed at noticeable and readable points by the distribution companies in the region to capture the inflow and usage of electrical power by customers.

The meters are calibrated to record electricity usage by every customer for billing purposes. The daily usage of electricity is collated into months, and then translated into financial terms to be paid by the customers.

Domestic electrical connectivity ensures that electrical power is stepped down to a safe voltage level of about 250 volts using a transformer. Industrial or commercial users demand higher amounts of connectivity to power their high consuming electrical gadgets. It is in this vein that the distribution company decided to segregate its customers into three specialized categories to facilitate determining the connectivity quantum, decide on the appropriate connectivity apparatus, as well as, estimating the right billings for revenue generation.

The Special Load Tariff group of customers is made of high consuming industries who demand a minimum of 100 KVA of electricity, the Non-Special Load Tariff customers are electricity consumer groups consuming below the one hundred kilo voltages (<100 KVA) and due to inefficiencies and unreliability of electricity consumption data as well as the alarming rate of electricity theft among the Non-Special Tariff customers, an advanced metering device that allow customers to purchase credits before accessing electricity other than the using power before payment system (Prepaid Tariff Customers) was also introduced. The monthly consumption of electricity by the three groups of customers from January 2008 to December 2013 was used as historical data for forecasting

2.3. Data Analysis and Modelling

The Seasonal ARIMA forecasting technique is adopted for this study. The Seasonal ARIMA model is usually denoted in the form ARIMA (p, d, q) * (P, D, Q)_s where:

p is the number of non-seasonal Autoregressive component; q is the number of non-seasonal Moving Average component; d is the order of non-seasonal differencing; P is the number of seasonal AR component; Q is the number of seasonal MA component; D is order of seasonal differencing and s is time span of repeating seasonal pattern and can be represented monthly as s = 12. (Shi et al, 2011).

The seasonal ARIMA model incorporates both non-seasonal and seasonal factors in a multiplicative model. It can be written as:

ARIMA (p, d, q) X (P, D, Q) S

The general form of seasonal model SARIMA (p, d, q) (P, D, Q) s is given by:

$$\phi(B)\Phi(B)^s(1 - B)^d(1 - B)^s y_t = \theta(B)\Theta(B^s)\epsilon_t \dots\dots\dots 2.1$$

where

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \dots\dots\dots 2.2$$

$$\Phi(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_p B^{ps} \dots\dots\dots 2.3$$

$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_p B^p \dots\dots\dots 2.4$$

$$\theta(B^s) = 1 + \theta_1 B^s + \theta_2 B^{2s} + \dots + \theta_p B^{ps} \dots\dots\dots 2.5$$

and y_t is the time series observation at time period (t),

B is the backward shift operator,

ε_t is a sequence of error term with mean zero and constant variance (σ^2), $\phi(i)$ and Φ_j are the non- seasonal and seasonal AR components respectively, θ_i and Θ_j are the non- seasonal and seasonal MA components respectively.

The non-seasonal components are:

$$\text{AR: } \phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p \dots\dots\dots 2.6$$

$$\text{MA: } \theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q \dots\dots\dots 2.7$$

The seasonal components are:

$$\text{Seasonal AR: } \varphi(B^s) = 1 - \varphi_1 B^s - \dots - \varphi_p B^{ps} \dots\dots\dots 2.8$$

$$\text{Seasonal MA: } \vartheta(B^s) = 1 + \vartheta_1 B^s + \dots + \vartheta_q B^{qs} \dots\dots\dots 2.9$$

The preliminary values of autoregressive order p, the order of differencing d, the moving average order q and their corresponding seasonal parameters P, D and Q were estimated with the help of anomalies identified from a time plot of the data. When preliminary values of D and d have been fixed, the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) were checked to confirm that the values of P, Q, P and q were acceptable stating parameters for the model (Stoffer and Dhumway, 2010). The model was then tentatively generated with the estimated parameters and the corresponding standard errors. The predicted models were validated using historical data from January 2014 - December, 2020.

3. Results and Discussions

3.1 Result

The model for the three Load Tariff monthly time series was based on the ideas of the Box-Jenkins methodology. This methodology seeks to fit a Seasonal Auto Regressive Integrated Moving Average Model to the given series. The data was decomposed into its trend, seasonal and random components as presented in Figures 3.1 to 3.3

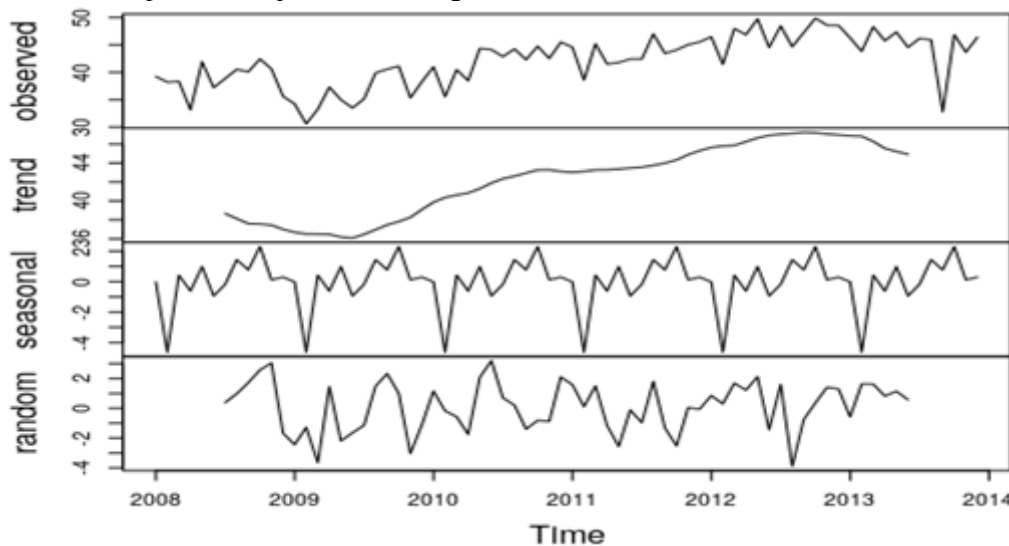


Fig. 1. Decomposition of additive time series for the Special Load Tariff Customer

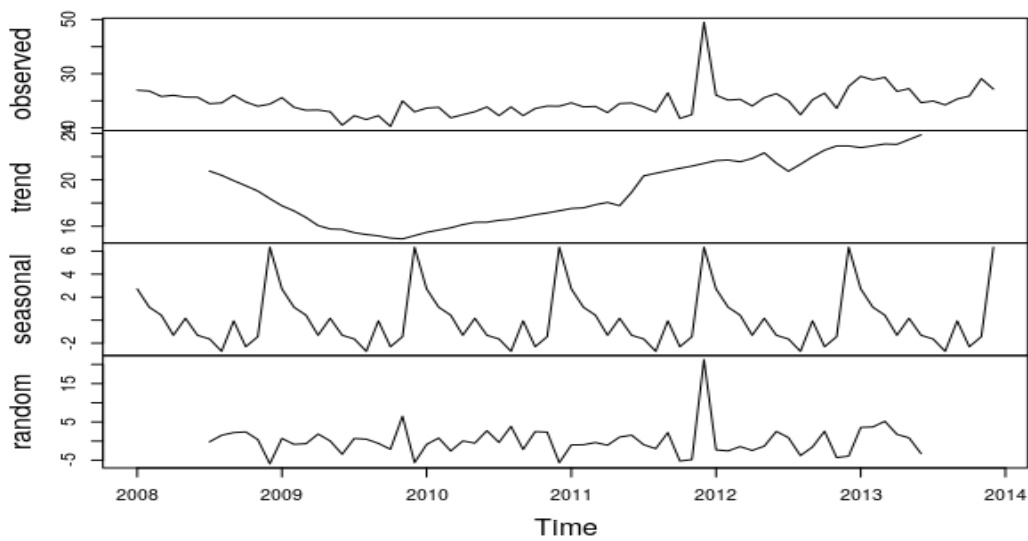
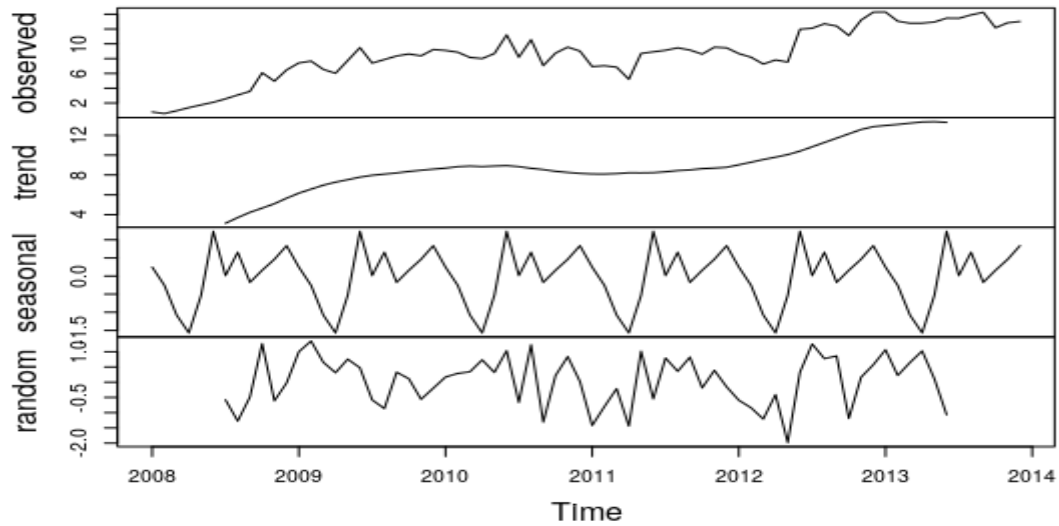


Fig. 2. Decomposition of additive time series for the Non-Special Load Tariff Customer

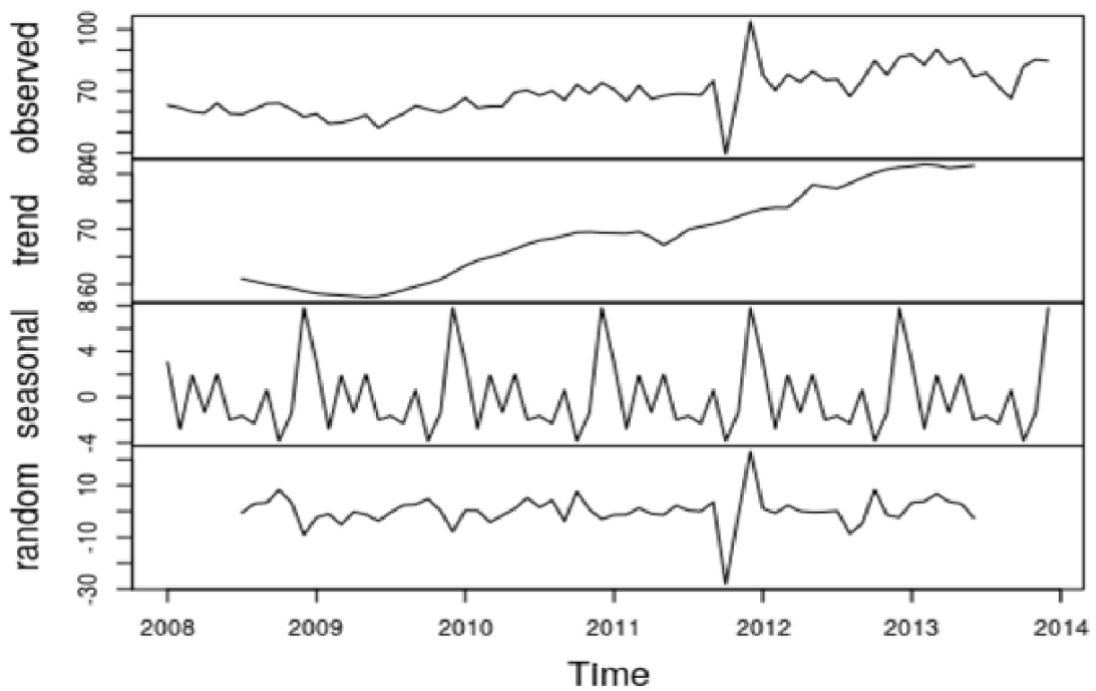


Fig. 3. Decomposition of additive time series for the Prepaid Load Tariff Customer

The results obtained by following the modeling procedure outlined in Section 2.3 are shown in Tables 1 and 2. Three distinct models were produced for each load tariff and the entire region. The adequate model for each load tariff was obtained using the F-criterion (the least or minimum sum of squared estimate of error).

Table 1 - Model for three Load Tariff Data

Type of Customer	Parameters	Models		
		SARIMA (0,1,1)×(0,0,2)	SARIMA (1,1,0)×(0,0,2)	SARIMA (1,1,1)×(0,0,2)
Special Load Tariff Data	ϕ_1		-0.4808 ± 0.20	-0.0118±0.33
	ψ_1	-0.6991 ± 0.17		-0.6930± 0.24
	Ψ_1	0.1245 ± 0.27	0.1212 ± 0.26	0.1240±0.27
	Ψ_2	0.4338± 0.38	0.4294± 0.40	0.4331±0.38
	SSE	606.64	699.96	606.74
Non-Special Load Tariff Data	ϕ_1		-0.4318 ± 0.208	-0.0692 ± 0.281
	ψ_1	-0.8044 ± 0.148		-0.0818 ± 0.160
	SSE	1671.518	2144.627	1666.005
Prepaid Load Tariff Data	δ			0.2187 ± 0.275
	ϕ_1		-0.3957 ± 0.653	-0.4158 ± 0.225
	ψ_1	-0.3589 ± 0.205	-0.0035 ± 0.208	
	Φ_1	0.3328 ± 0.232	-0.3661 ± 0.235	0.3490 ± 0.231
	SSE	93.66789	91.47522	88.54043

Table 2 - Models for the Regional Load Tariff Data

Parameters	Models		
	SARIMA (1,1,3)	SARIMA (2,1,0)	SARIMA (2,1,1)
ϕ_1	-0.1516 ±0.710304	-0.4275 ± 0.200704	0.0276 ± 0.380632
ϕ_2		-0.4874 ± 0.197568	-0.3259 ± 0.288708
ψ_1	-0.3973 ± 0.689528		-0.6291 ± 0.398076
ψ_2	-0.5061 ± 0.148372		
ψ_3	0.1959 ± 0.312228		
SSE	3607.437	3956.292	3726.168

3.2 Generation and Validated of the Predicted Model

From Table 1, the SARIMA (0,1,1) × (0,0,2) model was chosen for the special load tariff data because it has the least or minimum sum of squared estimate of error (SSE) which yields the least square of data regression line and the corresponding best model is written out as $X_t - X_{t-1} = \epsilon_t + 0.6991\epsilon_{t-1} - 0.087037\epsilon_{t-2} - 0.4338\epsilon_{t-24} - 0.3032696\epsilon_{t-25} \dots \dots \dots 3.1$

The SARIMA (1,1,1) model was chosen for the non-special load tariff data because it has the least/minimum sum of squared estimate of error (SSE) which yields the least square of data regression line with the best model. This is expressed as,

$$X_t - X_{t-1} = \epsilon_t - 0.8044\epsilon_{t-1} \dots \dots \dots 3.2$$

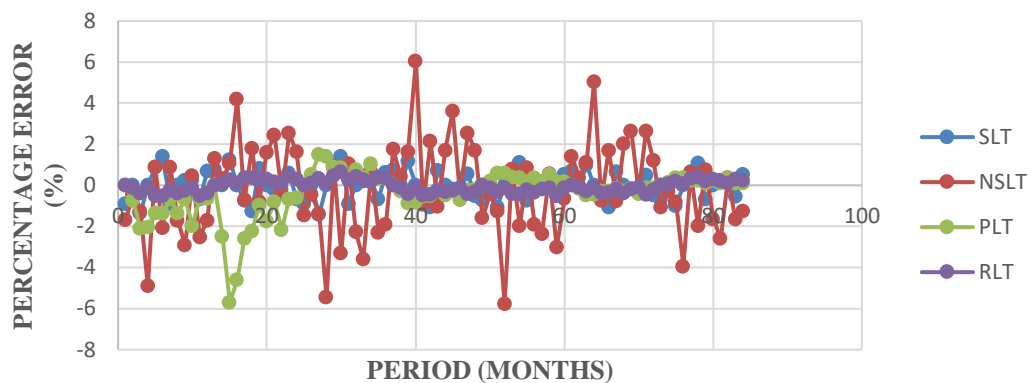
The SARIMA (1,1,0) × (1,0,0) model was chosen for the prepaid load tariff data because it has the least/minimum sum of squared estimate of error (SSE) which yields the least square of data regression line and its best model is written out as

$$X_t - 0.5842X_{t-1} - 0.4158X_{t-2} - 0.3490X_{t-12} + 0.2039X_{t-13} + 0.1451X_{t-14} = \epsilon_t + 0.2187 \dots \dots \dots 3.3$$

Finally, the SARIMA (0,1, 1) × (1,1,3) model was chosen for the regional load tariff data because it has the least/minimum sum of squared estimate of error (SSE) which yields the least square of data regression line. The best model is written out as

$$X_t - 0.8484X_{t-1} - 0.1516X_{t-2} = \epsilon_t - 0.3973\epsilon_{t-1} - 0.5061\epsilon_{t-2} + 0.1959\epsilon_{t-3} \dots \dots \dots 3.4$$

The models obtained were then validated by comparing the predicted result and the recorded results for the 84months period from January, 2014 to December, 2020 and the results for the four customers are presented in Figure 4.



NOTE: SLT for Special Load Tariff; NSLT for Non-Special Load Tariff; PLT for Prepaid load Tariff; RLT for Regional Load Tariff

Fig. 4. The percentage Error for the four Customers

For the purpose of this study, an error of ≤ 1 GW/was chosen as a criterion to determine the relative difference in results between actual electricity consumption data and predicted data. From Figure 3.4, the results indicate that the predicted values of the model are close to the true value, ranging from -6.00% to 6.00%. The predicted electricity consumption for the Special Load Tariff group of customers resulted in 53.36 ± 1.02 GW/HR @ 95% confidence level whilst the predicted electricity consumption for the Non-Special Load Tariff group of customers resulted in 25.65 ± 0.34 GW/HR @ 95% confidence level. Also, the predicted electricity consumption for the Prepaid customers had an average use of 14.95 ± 0.67 GW/HR @ 95% confidence level as a monthly amount of 101.09 ± 0.23 GW/HR @ 95% confidence level was used for the whole region. The models obtained are, therefore, adequate to be used to forecast monthly electricity consumption for all the three Load Tariff Customers. In conclusion, it can be summarized that, with a confidence level of 95%, models developed produced a reliable result that can help guide in the future estimation of electricity to be used by the various categories of customers as well as a means of revenue projection and expansion of the company's customer network base.

5. Conclusion

In conclusion, it can be summarized that, with a confidence level of 95%, models developed produced reliable results that can help guide in the future estimation of electricity to be used by the various categories of customers as well as a means of revenue projection and expansion of the company's customer network base. These time series models have been proven to expose and expatiate electricity demand for the region in future years or months. The models have been refined and validated to ensure consistency in outcome results.

References

- Abdel-Aal, R. E., & Al-Garni, A. Z. (1997). Forecasting monthly electric energy consumption in eastern Saudi Arabia using univariate time-series analysis. *Energy*, 22(11), 1059-1069.
- Adom, P. K., & Bekoe, W. (2012). Conditional dynamic forecast of electrical energy consumption requirements in Ghana by 2020: a comparison of ARDL and PAM. *Energy*, 44(1), 367-380.
- Ahmed, N. B. (2018). *A Comparative Analysis of Forecast Performance between Sarima and Setar Models Using Macroeconomic Variables in Ghana* (Doctoral dissertation, University of Ghana).

- Ang, B. W. (2005). The LMDI approach to decomposition analysis: a practical guide. *Energy policy*, 33(7), 867-871.
- Asumadu-Sarkodie, S., & Owusu, P. A. (2016). The potential and economic viability of wind farms in Ghana. *Energy sources, Part A: Recovery, utilization, and environmental effects*, 38(5), 695-701.
- Asumadu-Sarkodie, S., & Owusu, P. A. (2016). Forecasting Nigeria's energy use by 2030, an econometric approach. *Energy Sources, Part B: Economics, Planning, and Policy*, 11(10), 990-997.
- Aslanargun, A., Mammadov, M., Yazici, B., & Yolacan, S. (2007). Comparison of ARIMA, neural networks and hybrid models in time series: tourist arrival forecasting. *Journal of Statistical Computation and Simulation*, 77(1), 29-53.
- Billah, B., Hyndman, R. J., & Koehler, A. B. (2005). Empirical information criteria for time series forecasting model selection. *Journal of Statistical Computation and Simulation*, 75(10), 831-840.
- Boisseleau, F. (2004). The role of power exchanges for the creation of a single European electricity market: market design and market regulation.
- Bosq, D. (2015). Models Associated with Extended Exponential Smoothing. *Communications in Statistics-Theory and Methods*, 44(3), 468-475.
- Box G.E.P. and Jenkins G.M. (1976). *Time Series Analysis: Forecasting and Control*. Holden-Day, revised edition.
- Box, G. E., Pierce, D. A., & Newbold, P. (1987). Estimating trend and growth rates in seasonal time series. *Journal of the American Statistical Association*, 82(397), 276-282.
- Cheng, C., Sa-Ngasoongsong, A., Beyca, O., Le, T., Yang, H., Kong, Z., & Bukkapatnam, S. T. (2015). Time series forecasting for nonlinear and non-stationary processes: a review and comparative study. *Iie Transactions*, 47(10), 1053-1071.
- Energy Commission of Ghana (2020), 2020 Energy Outlook for Ghana, pp 33.
- Energy Commission of Ghana (2021), National Energy Statistics – 2021, pp 47.
- Feinberg, E. A., & Genethliou, D. (2005). Load forecasting. In *Applied mathematics for restructured electric power systems* (pp. 269-285). Springer, Boston, MA.
- Ghana Statistical Service, Population by Regions [accessed 2021 20 Oct]; Available from: <https://www.statsghana.gov.gh/regionalpopulation.php?population=MTQ1MTUyODEyMC43MDc1&&Western®id=7>
- Hahn, H., Meyer-Nieberg, S., & Pickl, S. (2009). Electric load forecasting methods: Tools for decision making. *European journal of operational research*, 199(3), 902-907.
- Harvey, A. C. (1990). *Forecasting, structural time series models and the Kalman filter*.
- Jain, A., Srinivas, E., & Rauta, R. (2009, December). Short term load forecasting using fuzzy adaptive inference and similarity. In *2009 World Congress on Nature & Biologically Inspired Computing (NaBIC)* (pp. 1743-1748). IEEE.
- Jiang, S., Yang, C., Guo, J., & Ding, Z. (2018). ARIMA forecasting of China's coal consumption, price and investment by 2030. *Energy Sources, Part B: Economics, Planning, and Policy*, 13(3), 190-195.
- Kim, J., Hwang, M., Jeong, D. H., & Jung, H. (2012). Technology trends analysis and forecasting application based on decision tree and statistical feature analysis. *Expert Systems with Applications*, 39(16), 12618-12625.
- Kyriakides, E., & Polycarpou, M. (2007). Short term electric load forecasting: A tutorial. *Trends in Neural Computation*, 391-418.
- Louie, H. M. (2017). Time-series modeling of aggregated electric vehicle charging station load. *Electric Power Components and Systems*, 45(14), 1498-1511.
- Marcellino, M. (2007). *A comparison of time series models for forecasting GDP growth and inflation*. Bocconi University, Italia.
- Meese, R., & Geweke, J. (1984). A comparison of autoregressive univariate forecasting procedures for macroeconomic time series. *Journal of Business & Economic Statistics*, 2(3), 191-200.
- Sarkodie, S. A. (2017). Estimating Ghana's electricity consumption by 2030: An ARIMA forecast. *Energy Sources, Part B: Economics, Planning, and Policy*, 12(10), 936-944.

- Shi, J., Qu, X., & Zeng, S. (2011). Short-term wind power generation forecasting: Direct versus indirect ARIMA-based approaches. *International Journal of Green Energy*, 8(1), 100-112.
- Shitan, M., & Peiris, S. (2011). Time series Properties of the class of generalized first-order autoregressive processes with moving average errors. *Communications in Statistics-Theory and Methods*, 40(13), 2259-2275.
- Stock, J. H., & Watson, M. W. (1996). Evidence on structural instability in macroeconomic time series relations. *Journal of Business & Economic Statistics*, 14(1), 11-30.
- Stock, J. H. and Watson, M. W. (2003). Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature*, 41(3):788–829.
- Sulugodu, B., & Deka, P. C. (2019). Evaluating the performance of CHIRPS satellite rainfall data for streamflow forecasting. *Water Resources Management*, 33(11), 3913-3927.
- Taylor, J. W., & McSharry, P. E. (2007). Short-term load forecasting methods: An evaluation based on european data. *IEEE Transactions on Power Systems*, 22(4), 2213-2219.
- Trindade, A. A. (2002). Time-Series Forecasting.(Book Reviews). *Journal of the American Statistical Association*, 97(459), 920-921.
- Weron, R., & Misiorek, A. (2005, May). Forecasting spot electricity prices with time series models. In Proceedings of the European electricity market EEM-05 conference (pp. 133-141).