Science in Information Technology Letters
Vol. 2., No. 1, May 2020, pp. 9-15
ISSN 2722-4139
http://pubs2.ascee.org/index.php/sitech

Forecasting electrical power consumption using ARIMA method based on kWh of sold energy

Gianika Roman Sosa [a,1], Moh. Zainul Falah [a,2], Dika Fikri L [a,3], Aji Prasetya Wibawa [a,4,*], Anik Nur Handayani [a,5], Jehad A. H. Hammad [b,6]

[a] Master of Electrical Engineering, Faculty of Engineering, Universitas Negeri Malang, Indonesia
[b] Faculty of Technology and Applied Sciences, Department of Computer Information Systems, Al-Quds Open University, Palestine
[4] aji.prasetya.ft@um.ac.id*; [5] aniknur.ft@um.ac.id; [6] jhammad@qou.edu
* Corresponding Author

ARTICLE INFO

Article history
Received March 9, 2021
Revised April 18, 2021
Accepted May 10, 2021

Keywords
Forecasting
ARIMA
Electrical energy consumption
MAPE

ABSTRACT

Customer demand for electrical energy continues to increase, so electrical energy infrastructure must be developed to fulfill it. In order to generate and distribute electrical energy cost-effectively, it is crucial to estimate electrical energy consumption reasonably in advance. In addition, it is necessary to ensure that customer demands can be met and that there is no shortage of electricity supply. This study aims to determine the estimated long-term electricity use with a historical Energy Sold (T1) database in kW accumulated over several periods from 2008 to 2017. The ARIMA method with the Seasonal-ARIMA (SARIMA) pattern is used in forecasting analysis. The ARIMA method was chosen because it is considered appropriate for forecasting linear and univariate time-series data. The results of this study indicate that the MAPE (%) error rate is relatively low, with a result of 7,966, but the R-Square reaches a value of -0.024 due to the lack of observational data.

1. Introduction

Electrical energy is vital in everyday life [1]. The existence of electrical energy can help humanity in meeting their needs [2]. Electricity can be used as one of the primary sources of energy in every activity, whether household, business, industrial, technology, or educational. Customer needs for electrical energy continue to increase [3]. Electrical energy infrastructure must be developed to meet customers' increasing demand for electrical energy [4]. In other words, the development of the electricity industry must be able to keep up with the increasing demand for electrical energy every year [5], [6]. Therefore, to create and distribute electrical energy cost-effectively, it is necessary to estimate the use of electricity well before it is distributed.

Electricity consumption has become a critical topic for many countries [7]. The increasing demand in the market is driven mainly by the increasing population and industrialization of the country. Forecasting long-term electricity consumption is significant in the planning, analyzing, and managing of electric power systems to ensure a sustainable, safe and economical electricity supply [8]. Studies focusing on forecasting electricity consumption using different techniques are substantial for developing countries such as Turkey, India, Indonesia, etc.

There are several essential points to establishing an accurate consumption forecasting model. Initially, it is crucial to accurately identify indicators that strongly impact a country's consumption and add these indicators to the forecast model before making a forecast. Each country may have different model inputs according to their respective circumstances, and the impact and amount of these inputs on consumption can significantly affect forecast performance. Second, the critical point is to choose the appropriate modeling methodology. The non-linear relationship between most of the input and output variables and the difficulty in expressing this mathematical relationship is one of the challenges in this field. Criteria for improving forecast performance are improving over time, not theoretical criteria in model selection. Third, the critical point is that the methodology used must be able to respond to future events. In other words, it must be able to produce forecasts.

Consumers in the Indonesian electricity sector are divided into four major segments: residential or household, industrial, commercial, and public [9]. As reported by PLN in its 2010 annual report (PLN, 2011), the commercial sector ranks first with an average growth of 10.45% in electricity sales 2006 – 2010 [10], followed by housing and industry, with 9.14% and 3.86% respectively. In 2010, the largest source of electricity sales revenue still came from the industrial and residential tariff groups. In 2010 total revenue from electricity sales increased by 14.20% to Rp102,974 billion from Rp90,712 billion in 2009. This increase was due to the increase in electricity tariffs which took effect on July 1, 2010. Based on this fact, the management of the electricity sector and its implication is believed to have a strong relationship between PLN (operator of the electricity sector and the government (regulator). Concerning the impact of economic growth on the development of the electricity sector [11], [12], the growth of electricity consumption in the housing sector will appear to be very influential.

The latest research published in 2018 regarding predicting energy needs using time series data related to coal or electricity needs has been widely used in the ARIMA method. The ARIMA algorithm is the component in forming a hybrid in predictions using time series data that is most widely used in this study [13]–[15]. The ARIMA algorithm is a stochastic model (a mathematical model in which symptoms can be measured with a varying degree of certainty) [16]. ARIMA is one of the most popular models for modeling using time series data [17] over the last 3 decades. The ARIMA model assumes that future values of the time series have a linear relationship with current and past values . The ARIMA model approach may not be adequate for complex non-linear real-world problems with white noise.

Regressive analysis, Auto-Regressive Iterative Moving Average (ARIMA), and Seasonal Autoregressive Iterative Moving Average (SARIMA) have been presented as econometric solutions. Bianco et al. [8], [18] predict Italy's electricity consumption with a linear regression model and examine the effect of economic and demographic variables. The variables are historical consumption, population, GDP (Gross Domestic Product) per capita, and GDP on annual electricity consumption in Italy to increase the long-term consumption forecasting model. Erdoğdu [19] uses the ARIMA model to analyze electricity demand, and Tunç et al. [20] predict the demand for electricity consumption by regression analysis. Pappas et al. [21] proposed an ARIMA model for Greek electricity consumption, and the model results were compared with several analytical time series models. Ediger et al. [22] use ARIMA and SARIMA methods for forecasting electricity demand. Also, several studies project energy consumption by traditional methods. Egelioglu et al. [23] studied the effect of economic variables on annual electricity consumption.

The ARIMA method was chosen because it is considered appropriate for forecasting linear and univariate time-series data. This study aims to determine the Forecasting of Electrical Power Consumption based on kWh of Energy Sold using the AutoRegressive Integrated Moving Average (ARIMA) method.

2. Method

2.1. Time Series Analysis and Forecasting

Time series data is a kind of data collected in a specific order over a certain period [24]. The current observation (Z_t), influenced by previous words (Z_{t-k}), provides the basis for the time series. Time series analysis is used to understand better and explain specific mechanisms, anticipate future values, and optimize control systems. Forecasting is the process of predicting what will happen in the future over a relatively long period. On the other hand, divination is a state expected to develop. Predicting this requires precise historical data that can be used to assess future situations [25].

2.2. ARIMA Method

ARIMA (Autoregressive Integrated Moving Average) method, often known as Box-Jenkins, is a time series analysis method. This method is based on a mixture of Autoregressive (AR) and Moving Average (MA) models from George Box and Gwilym Jenkins [26]. The Box-Jenkins procedure uses four steps to construct the ARIMA model: model identification, parameter estimation, diagnostic examination, and forecasting. Generally, the ARIMA model (p, d, q) can be written in (1).

$$\phi_P(B)(1-B)^d Z_t = \theta_q(B)a_t \tag{1}$$

If there are non-seasonal and seasonal effects, the model formed is multiplicative ARIMA (p, d, q) (P, D, Q)^s, or what is known as Seasonal ARIMA (SARIMA).

Where,
p, d, q: The non-seasonal part of the model
(P, D, Q) : Seasonal part of model s: Number of periods per season
The seasonal effect/SARIMA is mathematically written in (2).

$$\Phi_P(B^s)\phi_p(B)(1-B)^d(1-B^s)^D Z_t = \theta_q(B)\Theta_Q(B^S)a_t \tag{2}$$

Where,

$_pB$: AR Non-Seasonal
Φ_pB^S : AR Seasonal
$(1-B)^d$: differencing non-seasonal
$(1-B^s)^D$: differencing seasonal
$\theta_q(B)$: MA Non-Seasonal
$\Theta_q(B^S)$: MA Seasonal

2.3. Model Evaluation

The evaluation phase of the forecasting model is carried out using an out-sample approach. MAPE and R-Square (R2) are used to measure the accuracy of the forecasting model [27] used with the (3) and (4).

$$MAPE = \frac{1}{n}\left|\frac{Y_t - \acute{Y}_t}{Y_t}\right| x100\% \quad (3)$$

$$R^2 = 1 - \frac{\sum(Y_t - \acute{Y}_t)^2}{\sum(Y_t - \bar{Y})^2} x100\% \quad (4)$$

MAPE is an error measurement that calculates the percentage deviation between actual and forecast information. Mean Absolute Percentage Error (MAPE) is calculated using the absolute error in each period divided by the actual observed value for that period. After that, average the absolute percentage error. The coefficient of determination, or R^2 (or R-two), is another metric we can use to evaluate a model. However, it has the advantage of being scale-free - it does not matter if the output value is very large or very small; R^2 will always be between -∞ and 1.

2.4. Data Sources and Research Variables

The data used in this study came from PT. PLN (Persero) Network Service Unit (UPJ) Srengat, East Java, for 10 years since 2008. Attributes or research variables include January 2008 to December 2017 (YEAR) and the amount of electrical energy sold each month (T1) in kWh units. The use of the attribute number T1 is defined as a univariate type.

3. Results and Discussion

The data is processed using Python 3 with the Google Collaboratory editor. The time series plot of the electrical energy data sold from PLN for 10 years is shown in Fig. 1. The data shows conditions that do not meet the mean and variance stationarity. By using various forms of non-seasonal (d=1) and seasonal (D=1, S=12) differencing transformations.

At the parameter estimation stage, the prediction model has significant univariate parameters. Furthermore, residual diagnostic checks were carried out using the Ljung-Box test and the Jarque-Bera normality test to determine the feasibility of the ARIMA model. From the best conjecture model, ARIMA (1,1,0) $(0,1,1)^{12}$ met the residual assumptions shown in Fig. 2 but did not meet the normality assumption. Outliers cause the residual abnormality, so an outlier detection procedure is needed in this case.

By performing outlier detection procedures, it is necessary to re-estimate the parameters by modeling seasonal outliers (SARIMAX). After analysis, significant parameters were obtained for the model, and all residual assumptions were met.

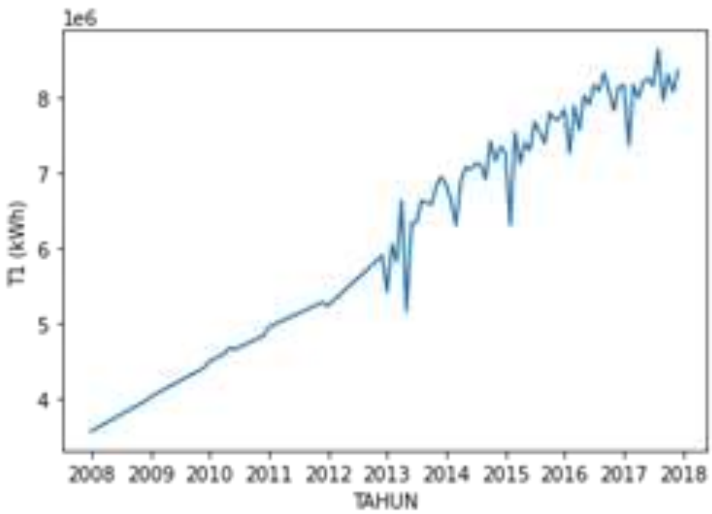

Fig. 1. Time series plot of the amount of energy sold during 2008-2017

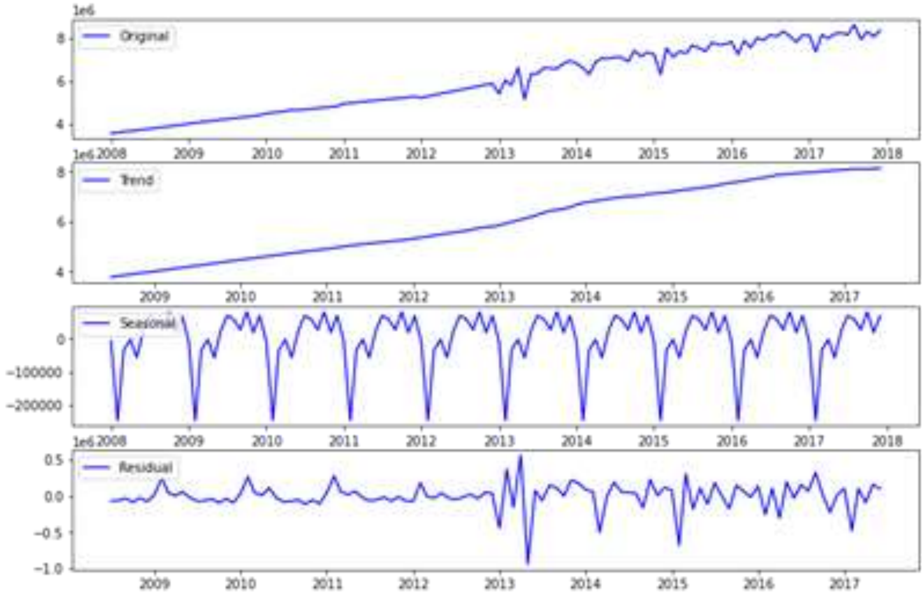

Fig. 2. Time series Decomposition

Decomposition in forecasting is a method that uses four main components in predicting future values, and these components include observed data (original), trend, seasonal (seasonal), and residual or irregular components. Decomposition isolates these components and then rearranges them into seasonal effects, cyclical effects, trend effects, and errors.

From data processing with training and testing data scenarios, which is 70:30, the results of testing data with the best-guessed model are obtained as shown in Fig. 3, where T1 is the amount of electrical energy sold. The results of forecasting trials in the 2018-2021 range using the best guess model of ARIMA are shown in Fig. 4.

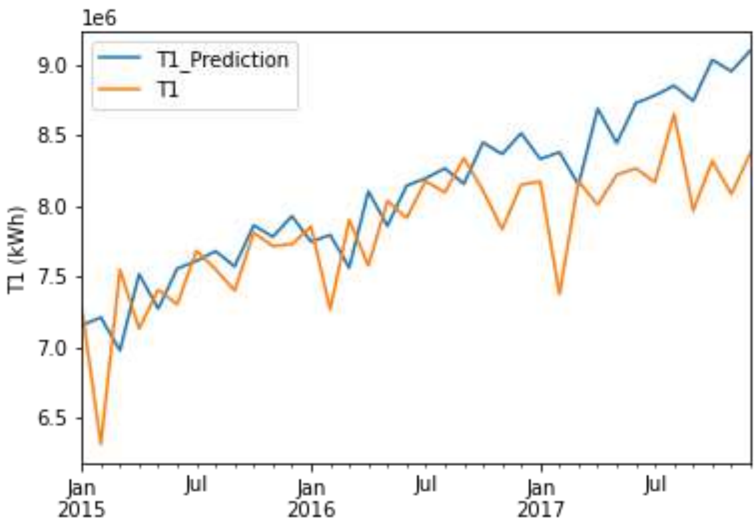

Fig. 3. Prediction vs. actual value of electrical energy sold (T1)

From the ARIMA method results, the MAPE (%) and R-Square (R2) values show the results of 7.966 and -0.024. The MAPE value shows a small error, so the model is considered accurate. However, the R2 value reaches minus, which means the ratio between the amount of data/observations and the number of variables is too small.

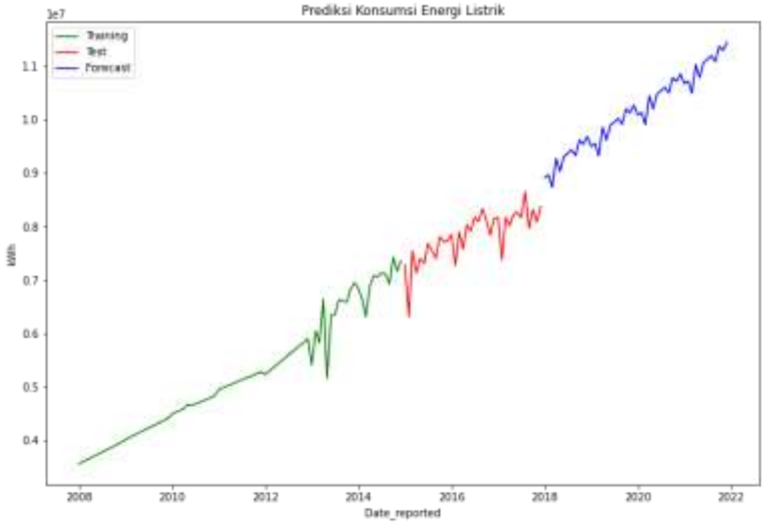

Fig. 4. The results of the electrical energy forecasting experiment sold (T1)

4. Conclusion

Based on the research results on electricity consumption in the Network Service Unit (UPJ) Srengat, East Java, it can be done using the ARIMA method with a monthly seasonal pattern (S=12). The most suitable conjecture model was found, ARIMA (1,1,0) (0,1,1)12, used as a training: testing data set with a 70:30 scenario. The MAPE (%) error rate is relatively low at 7,966, but the R-Square value reaches -0.024 due to the lack of observational data. So the ARIMA method is considered suitable for predicting electrical energy consumption.

References

[1] K. M. Putra and S. Agustina, “Penentuan Kapasitas Motor Listrik sebagai Penggerak Generator pada Pembangkit Listrik Energi Listrik,” Sriwijaya University, 2018. Available: Google Scholar.

[2] I. Al Rasyid, “Analisis Perilaku Konsumsi Energi Listrik Ditinjau Dalam Perspektif Ekonomi Islam (Studi Pada Masyarakat Kelurahan Tanjung Baru Kecamatan Kedamaian Kota Bandar Lampung),” UIN Raden Intan Lampung, 2020. Available: Google Scholar.

[3] S. I. Kristianto and B. Mudakir, “Analisis Konsumsi Listrik Rumah Tangga di Kecamatan Tembalang,” Fakultas Ekonomika dan Bisnis, 2015. Available: Google Scholar.

[4] E. H. S. Harahap, “Analisis Prakiraan Permintaan dan Penyediaan Energi Listrik Tahun 2019-2023 di Kabupaten Padang Lawas,” Unuversitas Islam Negeri Sultan Syarif Kasim Riau, 2019. Available: Google Scholar.

[5] J. D. Jenkins, M. Luke, and S. Thernstrom, “Getting to Zero Carbon Emissions in the Electric Power Sector,” *Joule*, vol. 2, no. 12, pp. 2498–2510, Dec. 2018, doi: 10.1016/j.joule.2018.11.013.

[6] J. Koskela, A. Rautiainen, and P. Järventausta, “Using electrical energy storage in residential buildings – Sizing of battery and photovoltaic panels based on electricity cost optimization,” *Appl. Energy*, vol. 239, pp. 1175–1189, Apr. 2019, doi: 10.1016/j.apenergy.2019.02.021.

[7] R. R. Kumar and K. Alok, “Adoption of electric vehicle: A literature review and prospects for sustainability,” *J. Clean. Prod.*, vol. 253, p. 119911, Apr. 2020, doi: 10.1016/j.jclepro.2019.119911.

[8] V. Bianco, O. Manca, S. Nardini, and A. A. Minea, “Analysis and Forecasting of Nonresidential Electricity Consumption in Romania,” *Appl. Energy*, vol. 87, no. 11, pp. 3584–3590, Nov. 2010, doi: 10.1016/j.apenergy.2010.05.018.

[9] I. A. Rahardjo, M. Subekti, P. Parjiman, and D. Rosyanti, “Analysis of electric load forecasting using combined method (Study case 2019 – 2029 in PT. PLN (Persero) UP3 Sukabumi),” *{IOP} Conf. Ser. Mater. Sci. Eng.*, vol. 1098, no. 4, p. 42026, Mar. 2021, doi: 10.1088/1757-899x/1098/4/042026.

[10] Y. Tanoto and M. Praptiningsih, “Factors Decomposition of Indonesia’s Household Electricity Consumption,” *Eng. J.*, vol. 17, no. 2, pp. 19–28, Apr. 2013, doi: 10.4186/ej.2013.17.2.19.

[11] M. Widiyanti, I. Sadalia, Z. Zunaidah, N. Irawati, and E. Hendrawaty, “Determining Firm’s Performance: Moderating Role of CSR in Renewable Energy Sector of Indonesia,” *Polish J. Manag. Stud.*, vol. 19, no. 2, pp. 432–441, Jun. 2019, doi: 10.17512/pjms.2019.19.2.37.

[12] R. Kurniawan and S. Managi, “Economic Growth and Sustainable Development in Indonesia: An Assessment,” *Bull. Indones. Econ. Stud.*, vol. 54, no. 3, pp. 339–361, Sep. 2018, doi: 10.1080/00074918.2018.1450962.

[13] H. do N. Camelo, P. S. Lucio, J. B. V. Leal Junior, P. C. M. de Carvalho, and D. von G. dos Santos, “Innovative hybrid models for forecasting time series applied in wind generation based on the combination of time series models with artificial neural networks,” *Energy*, vol. 151, pp. 347–357, May 2018, doi: 10.1016/j.energy.2018.03.077.

[14] Y. Shao and Y.-S. Tsai, “Electricity Sales Forecasting Using Hybrid Autoregressive Integrated Moving Average and Soft Computing Approaches in the Absence of Explanatory Variables,” *Energies*, vol. 11, no. 7, p. 1848, Jul. 2018, doi: 10.3390/en11071848.

[15] C. Wang, B.-B. Li, Q.-M. Liang, and J.-C. Wang, “Has China’s coal consumption already peaked? A demand-side analysis based on hybrid prediction models,” *Energy*, vol. 162, pp. 272–281, Nov. 2018, doi: 10.1016/j.energy.2018.08.031.

[16] R. Adhikari and R. K. Agrawal, “An Introductory Study on Time Series Modeling and Forecasting,” Feb. 2013, doi: 10.48550/arXiv.1302.6613.

[17] Y. Mao, A. Pranolo, A. P. Wibawa, A. B. Putra Utama, F. A. Dwiyanto, and S. Saifullah, “Selection of Precise Long Short Term Memory (LSTM) Hyperparameters based on Particle Swarm Optimization,” *2022 Int. Conf. Appl. Artif. Intell. Comput.*, pp. 1114–1121, May 2022, doi: 10.1109/ICAAIC53929.2022.9792708.

[18] V. Bianco, O. Manca, and S. Nardini, “Electricity consumption forecasting in Italy using linear regression models,” *Energy*, vol. 34, no. 9, pp. 1413–1421, Sep. 2009, doi: 10.1016/j.energy.2009.06.034.

[19] E. Erdogdu, “Electricity demand analysis using cointegration and ARIMA modelling: A case study of Turkey,” *Energy Policy*, vol. 35, no. 2, pp. 1129–1146, Feb. 2007, doi: 10.1016/j.enpol.2006.02.013.

[20] M. Tunç, Ü. Çamdali, and C. Parmaksizoğlu, “Comparison of Turkey’s electrical energy consumption and production with some European countries and optimization of future electrical power supply investments in Turkey,” *Energy Policy*, vol. 34, no. 1, pp. 50–59, Jan. 2006, doi: 10.1016/j.enpol.2004.04.027.

[21] S. S. Pappas, L. Ekonomou, D. C. Karamousantas, G. E. Chatzarakis, S. K. Katsikas, and P. Liatsis, “Electricity demand loads modeling using AutoRegressive Moving Average (ARMA) models,” *Energy*, vol. 33, no. 9, pp. 1353–1360, Sep. 2008, doi: 10.1016/j.energy.2008.05.008.

[22] V. S. Ediger and S. Akar, “ARIMA forecasting of primary energy demand by fuel in Turkey,” *Energy Policy*, vol. 35, no. 3, pp. 1701–1708, Mar. 2007, [Online]. Available: https://ideas.repec.org/a/eee/enepol/v35y2007i3p1701-1708.html.

[23] F. Egelioglu, A. A. Mohamad, and H. Guven, “Economic variables and electricity consumption in Northern Cyprus,” *Energy*, vol. 26, no. 4, pp. 355–362, Apr. 2001, doi: 10.1016/S0360-5442(01)00008-1.

[24] C. Zhang *et al.*, “A Deep Neural Network for Unsupervised Anomaly Detection and Diagnosis in Multivariate Time Series Data,” *Proc. AAAI Conf. Artif. Intell.*, vol. 33, pp. 1409–1416, Jul. 2019, doi: 10.1609/aaai.v33i01.33011409.

[25] J. Fattah, L. Ezzine, Z. Aman, H. El Moussami, and A. Lachhab, “Forecasting of demand using ARIMA model,” *Int. J. Eng. Bus. Manag.*, vol. 10, pp. 1–9, Jan. 2018, doi: 10.1177/1847979018808673.

[26] A. L. Schaffer, T. A. Dobbins, and S.-A. Pearson, “Interrupted time series analysis using autoregressive integrated moving average (ARIMA) models: a guide for evaluating large-scale health interventions,” *BMC Med. Res. Methodol.*, vol. 21, no. 1, p. 58, Dec. 2021, doi: 10.1186/s12874-021-01235-8.

[27] N. Khan, A. Arshad, M. Azam, A. H. Al-marshadi, and M. Aslam, “Modeling and forecasting the total number of cases and deaths due to pandemic,” *J. Med. Virol.*, vol. 94, no. 4, pp. 1592–1605, Apr. 2022, doi: 10.1002/jmv.27506.