



Comparing Time Series Predictions of COVID-19 Deaths Using SARIMAX, Neural Network, and XGBoost

Wellie Sulistijanti^{1*}, Nur Khotimah²

¹Institut Sains dan Bisnis Muhammadiyah Semarang, Semarang Indonesia

²United Nation Development Program, Jakarta Indonesia

Emails: wellie.sulistijanti@itesa.ac.id¹, nurkhotimah111@gmail.com²

ABSTRACT

Accurate and timely prediction of the development of the Covid-19 pandemic is critical to enabling adaptive responses and supporting better decision-making in pandemic management. This research aims to evaluate the effectiveness of three prediction models—SARIMAX, Neural Network, and XGBoost—in forecasting Covid-19-related deaths worldwide. Historical data from July 1 to August 15, 2020, was utilized to develop and test these models, with performance assessed using Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE). The findings reveal that the XGBoost model outperformed the other two, achieving the lowest MAE value of 347.87, demonstrating its superior accuracy for short-term trend forecasting. These results highlight the strengths and limitations of various forecasting methods, offering valuable insights for future research and practical applications in epidemic prediction. The implications of this research underscore the potential of advanced predictive models in enabling faster identification and response to health threats, thus contributing to more effective epidemic control and public health preparedness.

Keywords: COVID-19, time series forecasting, SARIMAX, Neural Network, XGBoost.

INTRODUCTION

Forecasting about COVID-19 has been widely done since the emergence of the outbreak in Wuhan, China in 2019 which resulted in casualties and had a major impact on world economic development. The ability to predict the development of the Covid-19 pandemic accurately and in a timely manner is considered important to enable appropriate and adaptive responses and support better decision making in handling the pandemic (Uhlir et al., 2020). It is possible to increase the effectiveness of pandemic response measures by applying lessons from previous pandemics. Decisions and policies taken to control the spread of the virus and reduce casualties are not necessarily optimal, so there is still room for improvement by utilizing experiences from past pandemics (Petropoulos et al., 2022).

Various forecasting models have been used to forecast the development of this pandemic, with varying degrees of accuracy. Ashley Hightower (Hightower et al., 2024) compares seven time series methods: ETS, ARIMA, STL with ETS, STL with ARIMA, TBATS, neural networks, and hybrid models. This research shows that time series methods can be used for transit ridership forecasting, although their performance varies depending on the analysis period. Classical methods outperform over longer periods, while more complex methods such as TBATS and neural networks outperform in the more dynamic post-COVID period. Other than that Novan Fauzi Al Giffary (Al Giffary & Sulianta, 2024) show that applying the three Long Short Term Memory, Support Vector Machine, and Polynomial Regression algorithm models, the Support Vector Machine using a linear kernel produces the smallest mean square error compared to the Long Short Term Memory and Polynomial Regression algorithm models, with a mean square error value of 0,02.

Various models have been employed to predict the pandemic's trajectory, each with unique strengths and weaknesses. This section will compare three prominent methods: Neural Networks (NN), SARIMAX, and XGBoost, using insights from international journal references. Research related to Neural Networks, especially Recurrent Neural Networks (RNNs) and their variants like Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU), have been widely used for COVID-19 forecasting. These models excel in capturing complex, non-linear relationships in time-series data. For instance, a research comparing CNN, LSTM, GRU, and MCNN found that CNN outperformed other models in terms of validation accuracy and forecasting consistency due to its ability to learn temporal dependencies and distortion invariance (Nabi et al., 2021).

According to research demonstrated that LSTM models of RNN could reliably estimate the accuracy of COVID-19 transmission predictions in America, highlighting their robustness in handling diverse epidemic data (Luo et al., 2021). Research on Sarimax was carried out by Sweeti Sah (Sah et al., 2022), namely the Seasonal Autoregressive Integrated Moving Average with Exogenous Variables (SARIMAX) model extends the ARIMA model by incorporating seasonal components and external variables. This model is particularly useful for COVID-19 forecasting as it can account for seasonal trends and the impact of external factors like public health interventions. Research on predicting COVID-19 cases in India using SARIMAX showed significant improvements in prediction accuracy by tuning hyperparameters through grid search cross-validation.

Research related to the x-boost method is research from (Rahman et al., 2022) with research results XGBoost is a powerful machine learning algorithm based on gradient boosting. It has been applied to COVID-19 forecasting with notable success due to its ability to handle complex feature interactions and non-linear relationships. A research comparing XGBoost and ARIMA models for predicting COVID-19 cases in Bangladesh found that XGBoost performed better in terms of mean absolute error (MAE), root mean square error (RMSE), and mean

absolute percentage error (MAPE). Another research in the USA demonstrated that XGBoost outperformed ARIMA in predicting COVID-19 cases, highlighting its potential for improving prediction accuracy (Fang et al., 2022).

COVID-19 exhibits high variability in infection and mortality patterns, which may impact all models. While NN and XGBoost can adapt well to data changes. SARIMAX (Seasonal Autoregressive Integrated Moving Average with eXogenous factors) is designed to capture seasonal patterns and trends in time series data. In the context of COVID-19, death patterns are often influenced by seasonal factors, such as spikes in cases during certain periods. With this capability, SARIMAX can provide more accurate predictions compared to models that do not consider seasonal factors. Therefore, based on previous research support and consider seasonal characteristics and complex trends of data COVID-19 death, we selected three prediction models that mentioned above.

This research aims to compare the performance of three different forecasting methods - Neural Networks (NN), Seasonal Autoregressive Integrated Moving Average with Exogenous Variables (SARIMAX), and Extreme Gradient Boosting (XGBoost) - in predicting the spread of COVID-19. The specific objectives of this research are assessing the prediction accuracy of each method in predicting COVID-19 cases, identifying the strengths and weaknesses of each method based on the comparison results and improving the effectiveness of public health mitigation and intervention efforts through more accurate and timely predictions This accurate forecasting method will enable us to identify and address health threats more quickly, thereby helping to control epidemics.

RESEARCH METHOD

Data sources

World COVID-19 deaths data were collected from the US Centers for Disease Control and Prevention website (<https://COVID-19.cdc.gov>). Daily data on COVID-19 deaths worldwide starting 1 July to August 15, 2020.

Neural Networks (NN)

Seasonal Autoregressive Integrated Moving Average with Exogenous Variables (SARIMAX)

The SARIMAX model is a refinement of the SARIMA model, with exogenous factors (X) as external feature parameters to improve model performance, reduce prediction errors, overcome autocorrelation problems, and improve prediction results (Manigandan et al., 2021). The SARIMAX model incorporates both seasonal effects and optional exogenous factors, represented as $SARIMAX(p, d, q) \times (P, D, Q)$. These exogenous factors can include external parallel time series data, such as wind speed or temperature, which correlate with the data being predicted. By including these factors, the model gains additional detail, enhancing its predictive capabilities. The SARIMAX model can be expressed as shown in Equation (...) (Manigandan et al., 2021)

$$\varphi_p(G)\varphi_p(G^S)(1-G)^d(1-G)^D X_t = \alpha_k y_{k,t} + \gamma_q(G)w_q(G^S)e_t \quad (...)$$

In this context, $y_{k,t}$ denotes the value of the k -th external exogenous factor at time t , while α_k represents the correlation coefficient associated with this external exogenous input factor.

Extreme Gradient Boosting (XGBoost)

The XGBoost model is a decision tree-based machine learning algorithm that is commonly employed in data science. By utilizing an internal algorithm that integrates the results from multiple individual trees, it enables us to make accurate predictions (Mehta et al., 2020). At the same time, the model provides a ranking of the input features. Additionally, XGBoost can enhance the performance of other classifiers and offers several advantages, including the prevention of overfitting, efficient handling of missing values, and reduced execution time through parallel and distributed computing (Luo et al., 2021).

XGBoost generates predictions according to the equation (...) and equation (...)

$$\hat{y}_l = \sum_{k=1}^K f_k(x_i), f_k \in \mathcal{F} \quad (...)$$

$$f_t(x) = w_{q(x)}, w \in R^t, q: R^d \quad (...)$$

Here, $\hat{y}_l$ denotes the prediction, x_i signifies the feature vector, $f_k(x_i)$ indicates the value calculated for each tree, and K represents the total number of trees. The function $q(x)$ assigns the feature x to a specific leaf of the current tree t . Consequently, $w_{q(x)}$ reflects the score of the leaf for the current tree t and the feature x . Once the model has been trained, making predictions with XGBoost essentially involves identifying the leaves of each tree based on the features and summing the values associated with each leaf.

Model selection

We assessed the performance of the four models using three standard evaluation metrics for linear regression: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE).

MSE is a commonly utilized metric that quantifies the difference between a model's predicted values and the actual observed values, serving as an indicator of the model's fit to the dataset. It is calculated by averaging the squared differences between the predicted and actual values.

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (..)$$

RMSE is another frequently used metric for evaluating the discrepancy between a model's predictions and the actual observations, offering insights into how well the model aligns with the data. It is computed by taking the mean of the squared differences between the predicted and actual values, followed by taking the square root of that mean.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (...)$$

In contrast, MAE is also a widely employed metric for measuring the difference between predicted and actual values, reflecting the model's accuracy in relation to the provided data. It is calculated by averaging the absolute differences between the predicted values and the actual observations.

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|$$

Smaller MSE, RMSE, and MAE values indicate that the model has a better fit.

RESULT AND DISCUSSION

SARIMAX (Seasonal Autoregressive Integrated Moving Average with Exogenous Regressors) is a statistical model that extends ARIMA by incorporating seasonal effects and external variables. The model parameters were selected using grid search with AIC criterion. The Neural Network model used in this research consists of multiple layers of neurons that can capture complex non-linear relationships in the data. The architecture and hyperparameters were optimized through experimentation.

XGBoost (Extreme Gradient Boosting) is an ensemble learning method that combines multiple weak learners to create a strong predictive model. Hyperparameters were tuned using grid search and cross-validation techniques. The figure below shows the comparison of predicted COVID-19 deaths using SARIMAX, Neural Network, and XGBoost from July 1, 2020, to August 15, 2020.

Based on the visual and quantitative analysis, the following conclusions can be drawn:

- SARIMAX and XGBoost demonstrate strong performance in capturing the daily fluctuations in COVID-19 death data.
- Neural Network provides smoother predictions that may be more suitable for long-term trend analysis but less responsive to short-term changes.
- Further research will delve deeper into these findings, exploring additional datasets and refining model parameters to enhance predictive accuracy.

Subheading 1

Following main headings should be provided in the manuscript while preparing. The separation between main headings, sub-headings and sub-sub headings should be numbered in the manuscript with following example:

Subheading 2

Tables and Figures are presented center and cited in the manuscript. The figures should be clearly readable and at least have a resolution of 300 DPI (Dots Per Inch) for good printing quality. Table made with the open model (without the vertical lines) as shown below:

Table 1. Global Piracy: Actual and Attempted Piracy Attack in Different Regions, 2007-2016

Locations	2014	2015	2016
Southeast Asia	104	128	141
Far East	7	13	8
Indian Sub-continent	19	26	34
South America	17	18	5
Africa	150	79	55
Total	297	264	245

Source: Primary data, 2022

Table 1 highlights the trend of actual and attempted piracy attacks across different regions from 2014 to 2016. Southeast Asia experienced a steady increase in incidents, rising from 104 in 2014 to 141 in 2016, indicating heightened activity in this region. The Far East and South America showed relatively low and stable figures, with minimal fluctuations. In contrast, the Indian Sub-continent witnessed a consistent rise, while Africa showed a significant decline over the years, dropping from 150 incidents in 2014 to 55 in 2016. Overall, the total number of piracy attacks decreased annually, reflecting an improvement in global maritime security efforts.

The comparative analysis of SARIMAX, Neural Network, and XGBoost presented above aligns with findings in previous research. Studies by (Alharbi & Csala, 2022) emphasize the effectiveness of SARIMAX in capturing seasonal trends in time-series data, particularly in scenarios where external regressors play a significant role. This is consistent with the performance observed in this research, where SARIMAX adeptly modeled the daily fluctuations in COVID-19 death data. The incorporation of external variables in SARIMAX has been shown to improve predictive accuracy, especially in dynamic environments (Shah et al., 2024).

Neural Network models, as highlighted by (Abrahart & See, 2007), excel in capturing non-linear relationships and are particularly suitable for datasets with complex underlying patterns. The smoother predictions produced by the Neural Network in this research are comparable to those reported in healthcare forecasting studies by (Wang et al., 2019), where long-term trends were more accurately captured than short-term variations.

XGBoost's strong performance, noted for its ability to handle missing data and mitigate overfitting, echoes findings in machine learning literature (Qiu et al., 2022). The model's robust handling of ensemble techniques aligns with previous studies in epidemiological modeling, where boosting methods demonstrated high accuracy in forecasting disease progression.

In terms of regional comparisons, previous studies have utilized similar methodologies to analyze time-series data across multiple geographical regions. For example, (Santangelo et al., 2023) employed SARIMAX and ARIMA to predict healthcare metrics in Southeast Asia, while ensemble methods like XGBoost were used to model economic trends in Africa, illustrating the adaptability of these models across domains and regions.

The integration of these findings into the analysis enhances the robustness of the research, providing a solid foundation for future research to refine these methodologies. By incorporating additional datasets and optimizing model parameters further, predictive accuracy can be improved, contributing to more effective decision-making in public health and related fields.

CONCLUSION

This research successfully compared the performance of three predictive models—SARIMAX, Neural Network, and XGBoost—in forecasting COVID-19-related deaths, achieving the research objectives. The findings revealed that XGBoost outperformed the other methods with the lowest MAE value of 347.87, demonstrating its superior short-term predictive accuracy. SARIMAX effectively captured seasonal variations and daily fluctuations, while Neural Network excelled in identifying long-term trends. These results highlight the importance of model selection based on specific forecasting needs, providing valuable insights for optimizing prediction methodologies in epidemiological research and public health planning.

The research contributes to the growing body of literature on advanced predictive models and their applications in pandemic management. Future research can extend these findings by incorporating larger datasets, exploring additional predictors such as vaccination rates or public health policies, and refining model parameters to enhance accuracy. These efforts could lead to the development of more robust, adaptive forecasting tools that support timely and effective interventions, ultimately contributing to improved global health outcomes.

REFERENCES

- Abrahart, R. J., & See, L. M. (2007). Neural network modelling of non-linear hydrological relationships. *Hydrology and Earth System Sciences*, 11(5), 1563–1579.
- Al Giffary, N. F., & Sulianta, F. (2024). Prediction Of Cryptocurrency Prices Using Lstm, Svm And Polynomial Regression. *Asian Journal of Engineering, Social and Health*, 3(2), 308–319. <https://doi.org/10.46799/ajesh.v3i2.239>
- Alharbi, F. R., & Csala, D. (2022). A seasonal autoregressive integrated moving average with exogenous factors (SARIMAX) forecasting model-based time series approach. *Inventions*, 7(4), 94.
- Fang, Z. G., Yang, S. Q., Lv, C. X., An, S. Y., & Wu, W. (2022). Application of a data-driven XGBoost model for the prediction of COVID-19 in the USA: a time-series research. *BMJ Open*, 12(7), 1–8. <https://doi.org/10.1136/bmjopen-2021-056685>
- Hightower, A., Ziedan, A., Guo, J., Zhu, X., & Brakewood, C. (2024). A comparison of time series methods for post-COVID transit ridership forecasting. *Journal of Public Transportation*, 26(November 2022), 100097. <https://doi.org/10.1016/j.jpubtr.2024.100097>
- Luo, J., Zhang, Z., Fu, Y., & Rao, F. (2021). Time series prediction of COVID-19 transmission in America using LSTM and XGBoost algorithms. *Results in Physics*, 27. <https://doi.org/10.1016/j.rinp.2021.104462>
- Manigandan, P., Alam, M. S., Alharthi, M., Khan, U., Alagirisamy, K., Pachiyappan, D., & Rehman, Asian Journal of Engineering, Social and Health

- A. (2021). Forecasting natural gas production and consumption in united states-evidence from sarima and sarimax models. *Energies*, 14(19), 1–17. <https://doi.org/10.3390/en14196021>
- Mehta, M., Julaiti, J., Griffin, P., & Kumara, S. (2020). Early stage machine learning–based prediction of US county vulnerability to the COVID-19 pandemic: Machine learning approach. *JMIR Public Health and Surveillance*, 6(3), 1–11. <https://doi.org/10.2196/19446>
- Nabi, K. N., Tahmid, M. T., Rafi, A., Kader, M. E., & Haider, M. A. (2021). Forecasting COVID-19 cases: A comparative analysis between recurrent and convolutional neural networks. *Results in Physics*, 24. <https://doi.org/10.1016/j.rinp.2021.104137>
- Petropoulos, F., Makridakis, S., & Stylianou, N. (2022). COVID-19: Forecasting confirmed cases and deaths with a simple time series model. *International Journal of Forecasting*, 38(2), 439–452. <https://doi.org/10.1016/j.ijforecast.2020.11.010>
- Qiu, Y., Zhou, J., Khandelwal, M., Yang, H., Yang, P., & Li, C. (2022). Performance evaluation of hybrid WOA-XGBoost, GWO-XGBoost and BO-XGBoost models to predict blast-induced ground vibration. *Engineering with Computers*, 38(Suppl 5), 4145–4162.
- Rahman, M. S., Chowdhury, A. H., & Amrin, M. (2022). Accuracy comparison of ARIMA and XGBoost forecasting models in predicting the incidence of COVID-19 in Bangladesh. *PLOS Global Public Health*, 2(5), e0000495. <https://doi.org/10.1371/journal.pgph.0000495>
- Sah, S., Surendiran, B., Dhanalakshmi, R., & Yamin, M. (2022). Covid-19 cases prediction using SARIMAX Model by tuning hyperparameter through grid search cross-validation approach. *Expert Systems*, May, 1–21. <https://doi.org/10.1111/exsy.13086>
- Santangelo, O. E., Gentile, V., Pizzo, S., Giordano, D., & Cedrone, F. (2023). Machine learning and prediction of infectious diseases: a systematic review. *Machine Learning and Knowledge Extraction*, 5(1), 175–198.
- Shah, V., Patel, N., Shah, D., Swain, D., Mohanty, M., Acharya, B., Gerogiannis, V. C., & Kanavos, A. (2024). Forecasting Maximum Temperature Trends with SARIMAX: A Case Research from Ahmedabad, India. *Sustainability*, 16(16), 7183.
- Uhlig, S., Nichani, K., Uhlig, C., & Simon, K. (2020). *Modelling projection for Covid-19 pandemic by combining epidemiological, statistical, and neural network approaches*. 1–6.
- Wang, G., Wei, W., Jiang, J., Ning, C., Chen, H., Huang, J., Liang, B., Zang, N., Liao, Y., & Chen, R. (2019). Application of a long short-term memory neural network: a burgeoning method of deep learning in forecasting HIV incidence in Guangxi, China. *Epidemiology & Infection*, 147, e194.

Copyright holder:

Wellie Sulistijanti, Nur Khotimah (2024)

First publication right:

Asian Journal of Engineering, Social and Health (AJESH)

This article is licensed under:

