

# Comparative Analysis of ARIMA and NNAR Models for Time Series Forecasting

Ghadah Alsheheri

Department of Mathematics, College of Science, King Khalid University, Abha, Saudi Arabia

Email: ghadah-g3@hotmail.com

**How to cite this paper:** Alsheheri, G. (2025) Comparative Analysis of ARIMA and NNAR Models for Time Series Forecasting. *Journal of Applied Mathematics and Physics*, 13, 267-280.

<https://doi.org/10.4236/jamp.2025.131012>

**Received:** December 27, 2024

**Accepted:** January 23, 2025

**Published:** January 26, 2025

Copyright © 2025 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International

License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

This paper presents a comparative study of ARIMA and Neural Network AutoRegressive (NNAR) models for time series forecasting. The study focuses on simulated data generated using ARIMA(1, 1, 0) and applies both models for training and forecasting. Model performance is evaluated using MSE, AIC, and BIC. The models are further applied to neonatal mortality data from Saudi Arabia to assess their predictive capabilities. The results indicate that the NNAR model outperforms ARIMA in both training and forecasting.

## Keywords

Time Series, QRIMQ Model, Neural Network, NNAR Model

---

## 1. Introduction

Time series forecasting is a fundamental technique utilized across a wide range of domains, including finance, healthcare, economics, and environmental sciences [1]-[3]. It plays a pivotal role in predicting future values based on previously observed data, enabling organizations and researchers to make informed decisions, manage risks, and optimize resources. Accurate forecasting can lead to substantial improvements in planning, budgeting, and policy development, fostering economic growth and enhancing operational efficiency. The application of time series models has grown significantly, driven by advancements in computational power and the increasing availability of large datasets [4]. Among the most widely adopted models are the Autoregressive Integrated Moving Average (ARIMA) and Neural Network AutoRegressive (NNAR) models. ARIMA, a classical approach, is known for its ability to handle linear patterns and stationary data, making it a cornerstone in statistical forecasting. Conversely, NNAR models leverage the power of neural networks, enabling them to capture complex, nonlinear relationships in data, which can lead to enhanced forecasting accuracy in dynamic and non-stationary

environments. This paper presents a comparative analysis of ARIMA and NNAR models, evaluating their performance on both simulated and real-world datasets. Specifically, we focus on simulated data generated through an ARIMA(1, 1, 0) process and apply both models to forecast neonatal mortality rates in Saudi Arabia. The study employs key evaluation criteria, including Mean Squared Error (MSE), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC), to assess the accuracy and efficiency of each model. While the comparison of ARIMA and NNAR models is well-documented, this study's novelty lies in its application to neonatal mortality forecasting in Saudi Arabia, an area with limited research in predictive health analytics. The results provide new insights into the applicability of machine learning models for public health data, contributing to improved health outcome predictions. By exploring the strengths and limitations of ARIMA and NNAR models, this paper aims to contribute to the ongoing discourse on time series forecasting methodologies, providing valuable insights for practitioners and researchers seeking to apply these techniques to diverse datasets and problem domains.

## 2. Literature Review

### 2.1. Overview of ARIMA and NNAR Models

The Autoregressive Integrated Moving Average (ARIMA) model is a classical and well-established technique for time series forecasting, known for its simplicity, interpretability, and effectiveness in modeling linear relationships within data. ARIMA excels in capturing temporal dependencies through a combination of autoregression (AR), differencing (I), and moving average (MA) components, making it particularly useful for datasets exhibiting trends and seasonality.

In contrast, Neural Network AutoRegressive (NNAR) models apply neural networks to time series forecasting, allowing for the identification and modeling of complex, nonlinear relationships that traditional linear models like ARIMA may overlook. NNAR models can dynamically adapt to intricate patterns and interactions within the data, often outperforming linear models in scenarios characterized by high levels of nonlinearity and irregularity.

Together, these models represent two distinct yet complementary approaches to time series forecasting-ARIMA offering transparency and ease of interpretation, while NNAR provides flexibility and superior performance in capturing nonlinear patterns.

### 2.2. Previous Studies

Extensive research has validated the effectiveness of Autoregressive Integrated Moving Average (ARIMA) models for short-term time series forecasting across various domains, including economics, environmental science, and healthcare [2] [3]. ARIMA's strength lies in its ability to model linear patterns and stationary data, making it a reliable choice for datasets characterized by temporal trends and seasonality. In recent years, Neural Network AutoRegressive (NNAR) models

have garnered significant attention due to their capacity to capture long-term dependencies and nonlinear relationships within complex datasets. Studies indicate that NNAR models often outperform traditional statistical methods, particularly in scenarios involving irregular patterns, sudden shifts, or large volumes of data [5] [6]. This advantage stems from the neural network's ability to learn intricate data structures, adapt to evolving trends, and generalize across unseen data points.

Building upon prior research, this paper applies both ARIMA and NNAR models to neonatal mortality data, aiming to evaluate their relative performance in forecasting health outcomes. By comparing the accuracy and interpretability of these models, the study contributes to the growing body of literature on predictive analytics in public health [7] [8].

### 3. Methodology

This section outlines the processes involved in data generation, model specification, and evaluation criteria used to compare the performance of ARIMA and NNAR models. The NNAR(2, 2) model was configured with two hidden layers, each consisting of 10 neurons. The activation function used for each layer was ReLU (Rectified Linear Unit), and the model was trained using the Adam optimization algorithm with a learning rate of 0.001. Hyperparameter tuning was performed using grid search, where the number of neurons was varied between 5 and 20 to identify the configuration that minimized the validation error. Early stopping was applied to prevent overfitting, with a patience value of 10 epochs. The goal is to assess the accuracy and efficiency of these models in forecasting synthetic time series data.

#### 3.1. Data Simulation

A synthetic time series dataset consisting of 100 data points was generated using an ARIMA(1, 1, 0) model. The autoregressive parameter ( $\phi$ ) was set to 0.80, simulating a process with moderate persistence and trend. This configuration represents a first-order autoregressive process with one level of differencing to ensure stationarity, and no moving average component.

Mathematically, the ARIMA(1, 1, 0) model is expressed as:

$$Y_t = Y_{t-1} + \phi(Y_{t-1} - Y_{t-2}) + \epsilon_t$$

where  $Y_t$  represents the value at time  $t$ ,  $\phi = 0.80$ , and  $\epsilon_t$  is white noise.

The simulated dataset was designed to replicate real-world scenarios involving gradual trends and temporal dependencies, providing a robust foundation for evaluating the forecasting performance of both ARIMA and NNAR models. This controlled environment allows for direct comparison under identical conditions, minimizing external influences that could bias model performance. The synthetic dataset was limited to 100 data points to create a controlled environment for model comparison. While this approach allows for direct evaluation under standardized conditions, we recognize the need for larger datasets to enhance the robustness of the findings. Future work will focus on expanding the dataset by

sourcing additional health records and conducting simulations with increased data volume to strengthen the reliability and applicability of the results.

### 3.2. Model Evaluation Criteria

To assess and compare the forecasting accuracy of the ARIMA and NNAR models, three primary evaluation metrics were employed:

- Mean Squared Error (MSE): Measures the average squared difference between actual and predicted values. MSE penalizes larger errors more severely, making it a reliable indicator of overall model accuracy. Lower MSE values signify better predictive performance.

$$\text{MSE} = \frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2$$

- Akaike Information Criterion (AIC): AIC assesses model fit by balancing goodness-of-fit with model complexity. It penalizes overfitting by incorporating the number of estimated parameters. Lower AIC values indicate models that provide a better fit with fewer parameters.

$$\text{AIC} = 2k - 2 \log(L)$$

where  $k$  is the number of parameters and  $L$  is the likelihood of the model.

- Bayesian Information Criterion (BIC): Similar to AIC, BIC penalizes complex models but imposes a stronger penalty for additional parameters, favoring simpler models when datasets are small.

$$\text{BIC} = k \log(n) - 2 \log(L)$$

where  $n$  is the number of data points.

These criteria provide a comprehensive evaluation by addressing both predictive accuracy and model parsimony. The combination of error-based and information-theoretic metrics ensures that model comparisons account for overfitting, underfitting, and overall forecasting performance.

## 4. Results and Discussion

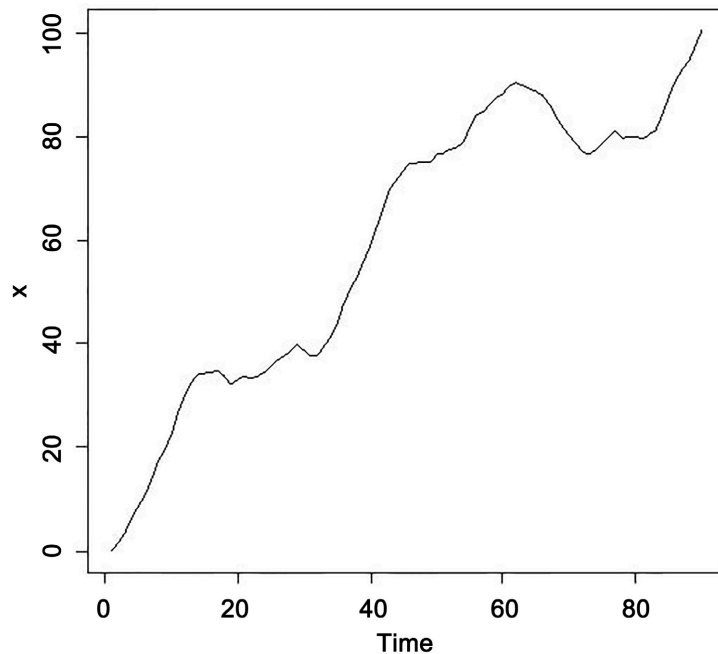
This section presents the visualization, diagnostic tests, and comparative analysis of ARIMA and NNAR models applied to the simulated data. Additionally, the models are evaluated on real-world neonatal mortality data to assess their practical applicability.

### 4.1. Visualization of Simulated Data

**Figure 1** presents a time series plot of simulated data produced by an ARIMA(1, 1, 0) model. The x-axis represents time, while the y-axis tracks the corresponding values of the simulated series. The plot reveals a generally upward trend, indicative of a non-stationary process, consistent with the integrated (I) component of the model. The data exhibits periods of gradual growth, interspersed with fluctuations, highlighting the stochastic nature of the series. This pattern aligns with the autoregressive (AR) component, which introduces short-term dependencies between

observations.

Notably, the plot demonstrates smooth increments with occasional plateaus and minor regressions, suggesting the presence of both persistent trends and noise, typical of ARIMA(1, 1, 0) processes. The absence of seasonal patterns further confirms that the model does not include a seasonal component, reflecting the simplicity of the underlying structure.

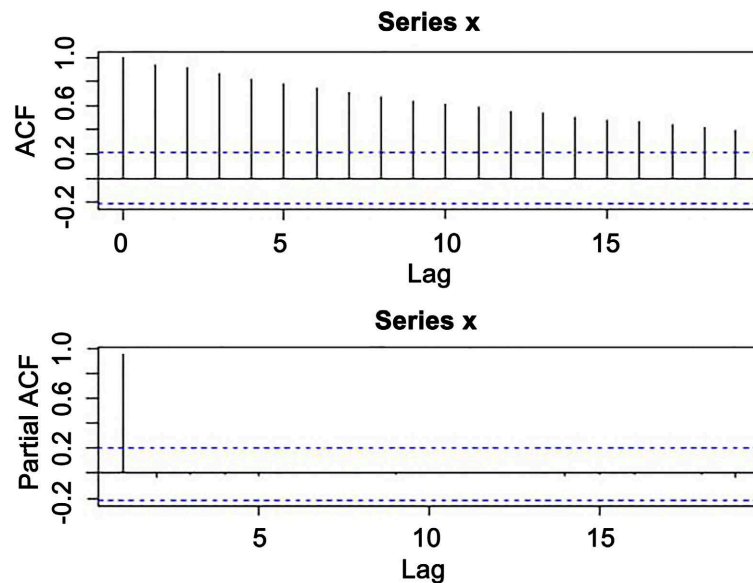


**Figure 1.** Time series plot of the simulated data.

**Figure 2** displays the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots of the simulated data from the ARIMA(1, 1, 0) model.

**Figure 2** presents the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots for the simulated data generated by the ARIMA(1, 1, 0) model. In the top panel, the ACF exhibits significant autocorrelations across multiple lags, gradually decaying over time. This slow, persistent decline suggests a non-stationary process and is characteristic of an integrated (I) component within the ARIMA framework. The prolonged correlation hints at the presence of a unit root, reinforcing the need for differencing to achieve stationarity. The bottom panel displays the PACF, which sharply cuts off after lag 1, with a pronounced spike at the first lag and values near zero for subsequent lags. This pattern reflects the behavior of an AR(1) process, where the first-order autoregressive term explains most of the serial correlation. The absence of significant partial autocorrelations beyond lag 1 confirms the appropriateness of an ARIMA(1, 1, 0) model, where differencing addresses non-stationarity and the AR(1) component captures short-term dependencies. To further assess stationarity, the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test was conducted. The test yielded a test statistic of

0.92, exceeding the critical value of 0.463 at the 5% significance level. This result led to the rejection of the null hypothesis of stationarity, providing statistical evidence of non-stationarity in the series. Following the KPSS results, first-order differencing was applied to eliminate the unit root and stabilize the series. After differencing, the ACF showed a rapid decay, and the PACF continued to display a sharp cutoff after lag 1, confirming that the series had been transformed into a stationary process suitable for ARIMA modeling.



**Figure 2.** ACF and PACF of the simulated data.

## 4.2. Stationarity Testing

The KPSS test was conducted to assess the stationarity of the data, with the results presented in **Table 1**. The null hypothesis of the test posits that the series is stationary around a deterministic trend, while the alternative hypothesis suggests the presence of a unit root, indicating non-stationarity.

**Table 1.** KPSS test results for stationarity at different differencing levels.

Integrated Order	KPSS Level	Truncation Lag	p-value
$d = 0$	2.1564	3	0.01
$d = 1$	0.28418	3	0.1

For the original series ( $d = 0$ ), the KPSS test statistic is 2.1564, which significantly exceeds the critical value at the 1% significance level (typically around 0.739). This leads to a rejection of the null hypothesis, confirming that the data is non-stationary in its initial form.

After applying first-order differencing ( $d = 1$ ), the KPSS test statistic drops to 0.28418, which falls well below the critical value at the 10% significance level. The corresponding p-value of 0.1 suggests that the null hypothesis of stationarity

cannot be rejected, indicating that the differenced series is stationary.

- **Non-Stationarity at  $d = 0$ :** The high KPSS statistic for the undifferenced series confirms that the data exhibits strong autocorrelation and trends over time, consistent with the patterns observed in the ACF plot (**Figure 2**), where significant correlations persist across lags. This validates the need for differencing to stabilize the series.
- **Stationarity Achieved at  $d = 1$ :** After applying first-order differencing, the KPSS statistic indicates that the series becomes stationary, aligning with the behavior of an ARIMA(1, 1, 0) model. This is further corroborated by the PACF plot (**Figure 2**), where the partial autocorrelation sharply cuts off after lag 1, suggesting that an AR(1) component adequately captures the dependencies in the data.
- **Model Validation:** The results affirm that  $d = 1$  is the appropriate level of differencing for the series, ensuring that the ARIMA model is specified correctly. This minimizes the risk of over-differencing, which can introduce unnecessary noise into the model and degrade forecasting performance.

The KPSS test results support the use of an ARIMA(1, 1, 0) model for this dataset, confirming that the series achieves stationarity after first-order differencing. This process ensures that the model can generate reliable forecasts by eliminating trends and stabilizing the variance.

The test indicates non-stationarity at level  $d = 0$ , but stationarity is achieved after first-order differencing ( $d = 1$ ).

**Figure 3** illustrates the first difference of the simulated data, which is a crucial step in the process of transforming a non-stationary time series into a stationary one. The x-axis represents time, while the y-axis shows the differenced values, labeled as  $\text{diff}(x)$ . The differenced series displays fluctuations around a mean of zero, with no visible trend over time, indicating that the process of differencing successfully removed the original trend and stabilized the variance. This transformation addresses the presence of a unit root identified in the undifferenced series.

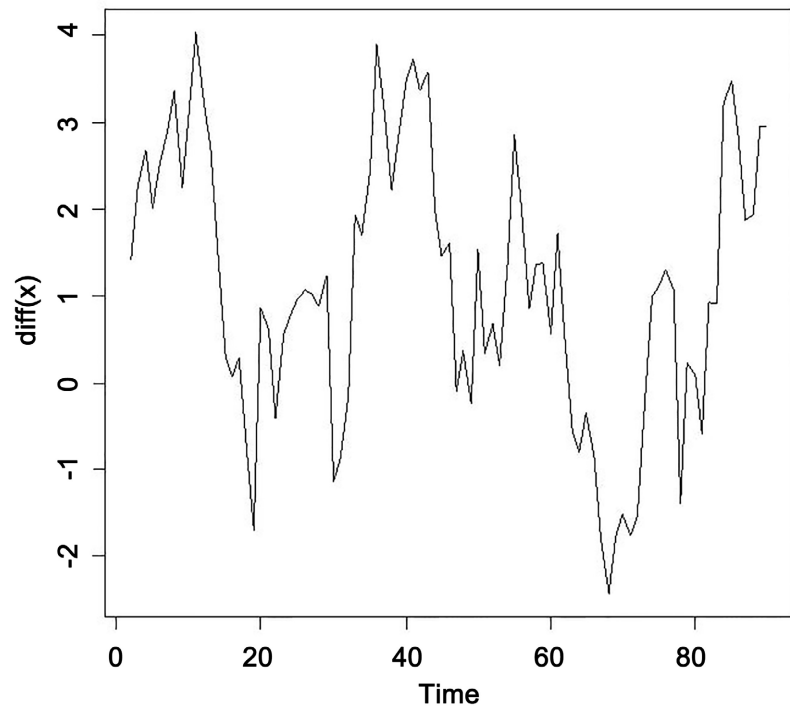
The series now exhibits characteristics consistent with a stationary process:

- The absence of a long-term upward or downward trend.
- Short-term variations appear to be randomly distributed around the mean.
- Periods of sharp increases and decreases reflect the persistence of short-term autocorrelation, suggesting the presence of an autoregressive component (AR(1)).

This differenced series serves as the basis for fitting the ARIMA(1, 1, 0) model, where the first differencing ( $d = 1$ ) ensures stationarity, and the AR(1) term captures the autocorrelation observed at lag 1.

Differencing plays a vital role in time series analysis by transforming non-stationary data into a stationary process, which is essential for accurate modeling and forecasting. By eliminating deterministic trends, differencing allows the model to concentrate on the underlying dynamic relationships between observations rather than being influenced by long-term shifts. This process also helps stabilize the

variance, addressing issues of heteroskedasticity and ensuring that fluctuations in the data remain consistent over time. Additionally, differencing facilitates model validation by enabling clearer interpretation of autocorrelation patterns through ACF and PACF plots. When combined with formal statistical tests, such as the KPSS test, this step verifies that the data meets the stationarity requirement necessary for effective ARIMA modeling.



**Figure 3.** First difference of the simulated data.

### 4.3. Model Fitting and Diagnostics

**Table 2** presents the performance metrics for several ARIMA models fitted to the data, evaluated using Mean Squared Error (MSE), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC). These metrics provide insight into model accuracy and complexity, guiding the selection of the most appropriate model for forecasting. MSE measures the average squared difference between the observed and predicted values. Lower MSE values indicate better predictive accuracy. AIC assesses the goodness-of-fit while penalizing for model complexity. A lower AIC value suggests a better balance between fit and simplicity, reducing the risk of overfitting. BIC performs a similar role to AIC but applies a stronger penalty for model complexity, favoring simpler models as the sample size increases. Model Comparison and Selection: The ARIMA(1, 1, 0) model achieves the lowest AIC (262.16) and BIC (267.14) values, indicating that this model provides the best trade-off between accuracy and simplicity. Its MSE (1.05301) is also the lowest, suggesting high predictive performance. The ARIMA(1, 1, 1) and ARIMA(2, 1, 0) models yield slightly higher AIC and BIC values (264.17 and 271.64) but exhibit

nearly identical MSE (1.05302). This suggests these models fit the data similarly, but their added complexity does not significantly improve performance. ARIMA(0, 1, 1) and ARIMA(0, 1, 0) models demonstrate substantially higher MSE (1.39620 and 2.26313, respectively), with correspondingly larger AIC and BIC values. This indicates that models without an autoregressive component (AR) perform poorly, confirming the importance of including AR terms to capture the serial dependencies in the data. ARIMA(2, 1, 1) shows the lowest MSE (1.05099), but its AIC (265.99) and BIC (275.95) are higher than the ARIMA(1, 1, 0) model, suggesting diminishing returns from the additional complexity introduced by the second AR and MA terms. Conclusion. The ARIMA(1, 1, 0) model emerges as the optimal choice, balancing predictive accuracy (low MSE) and model simplicity (lowest AIC and BIC). This model efficiently captures the underlying dynamics of the data without unnecessary complexity, making it the preferred candidate for forecasting and further analysis.

**Table 2.** Comparison of candidate ARIMA models.

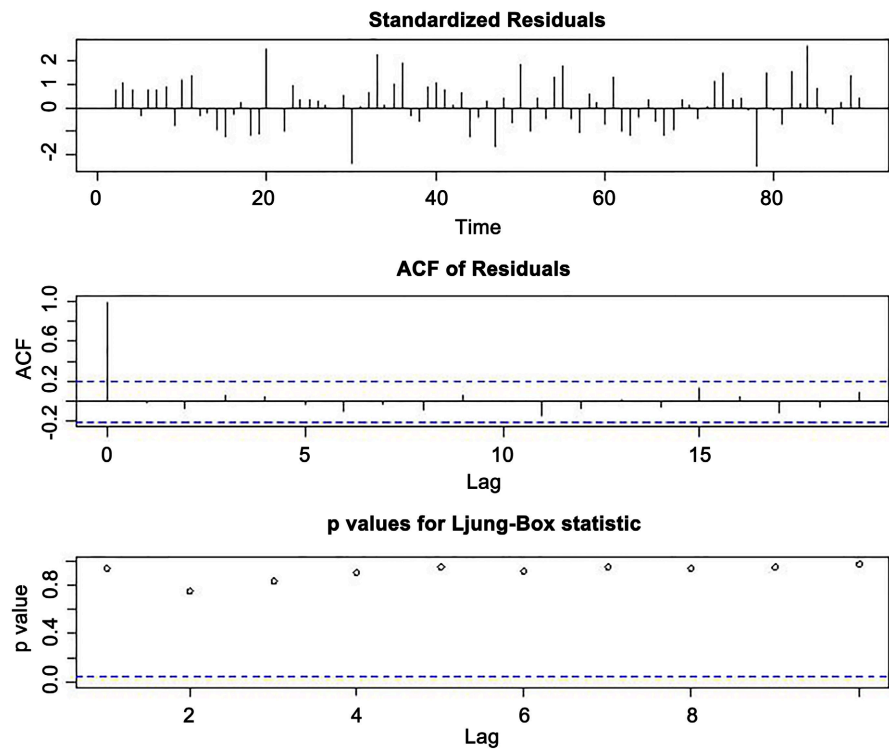
Model	MSE	AIC	BIC
ARIMA(1,1,1)	1.05302	264.17	271.64
ARIMA(1, 1, 0)	1.05301	262.16	267.14
ARIMA(0, 1, 1)	1.39620	287.27	292.25
ARIMA(0, 1, 0)	2.26313	328.26	330.75
ARIMA(2, 1, 0)	1.05302	264.17	271.64
ARIMA(2, 1, 1)	1.05099	265.99	275.95

**Table 3** presents the parameter estimates for the ARIMA(1, 1, 0) model, highlighting the autoregressive (AR1) component's coefficient and its associated standard error. AR1 Coefficient: The estimate for the first-order autoregressive term (AR1) is 0.7359, indicating a strong positive relationship between the current observation and the previous lagged value. This suggests that past values significantly influence future observations, a characteristic feature of AR(1) processes. The positive coefficient reflects that upward (or downward) movements in the series tend to persist over time. Standard Error: The standard error for the AR1 coefficient is 0.0707, indicating the precision of the estimate. The relatively small standard error suggests that the coefficient is estimated with high confidence and is unlikely to deviate significantly from its true value. The magnitude of the AR1 coefficient (close to 1) implies that the time series exhibits persistence and that shocks to the system may have prolonged effects, gradually decaying over time. This aligns with the nature of ARIMA(1, 1, 0) models, where differencing is applied to achieve stationarity, and the AR(1) term captures the remaining dependencies. The low standard error reinforces the reliability of the model, indicating that the ARIMA(1, 1, 0) specification is appropriate for describing the underlying data dynamics.

**Table 3.** Parameter estimates for ARIMA(1, 1, 0).

Parameter	Coefficient	Standard Error
AR1	0.7359	0.0707

Residual diagnostics of the ARIMA(1, 1, 0) model are presented in **Figure 4**.

**Figure 4.** Residual diagnostics for ARIMA(1, 1, 0).

**Figure 4** presents the residual diagnostics for the ARIMA(1, 1, 0) model, providing insight into the model's adequacy and performance through three key plots:

- **Top Panel-Standardized Residuals:** This plot displays the residuals over time, standardized to have a mean of zero and unit variance. The residuals fluctuate randomly around zero, with no visible patterns or clustering. This randomness suggests that the model effectively captures the underlying data structure, leaving no systematic information in the residuals.
- **Middle Panel-ACF of Residuals:** The ACF plot shows the autocorrelation of residuals across different lags. All autocorrelations lie within the 95% confidence intervals (represented by the dashed blue lines), indicating that the residuals exhibit no significant correlation at any lag. This is a strong indication that the residuals behave like white noise, validating the model's ability to account for temporal dependencies in the data.
- **Bottom Panel-Ljung-Box Test p-values:** The Ljung-Box test evaluates whether the residuals show signs of autocorrelation at various lags. The plot illustrates

p-values for different lags, all of which are well above the significance threshold of 0.05. This suggests that there is no evidence of significant autocorrelation in the residuals, further confirming the model's adequacy.

The residual diagnostics indicate that the ARIMA(1, 1, 0) model is well-fitted to the data. The absence of autocorrelation and the random distribution of residuals suggest that the model sufficiently captures the data's patterns, making it suitable for forecasting. No additional differencing or autoregressive/moving average terms are required at this stage.

#### 4.4. Neural Network AutoRegressive (NNAR) Model

A Neural Network AutoRegressive model, denoted as NNAR(2, 2), was fitted to the same dataset for comparison with the ARIMA(1, 1, 0) model. The NNAR model leverages neural network structures to capture non-linear patterns and complex relationships that may not be adequately modeled by traditional ARIMA approaches. The results of the model fitting and evaluation are summarized in **Table 4**.

**Table 4.** Comparison of ARIMA and NNAR model performance during training.

Model	MSE	AIC	BIC
ARIMA(1, 1, 0)	1.05301	262.16	267.14
NNAR(2, 2)	0.9362	251.47	262.94

The results indicate that the NNAR(2, 2) model outperformed the ARIMA(1, 1, 0) model across all key performance metrics:

- Mean Squared Error (MSE): The NNAR model achieved an MSE of 0.9362, reflecting a better fit to the data compared to the ARIMA model's 1.05301. This lower MSE signifies that the NNAR model reduces prediction errors, demonstrating superior accuracy in capturing the data's underlying structure.
- Akaike Information Criterion (AIC): The NNAR model's AIC of 251.47 is lower than the ARIMA model's 262.16, suggesting that the NNAR model provides a better trade-off between goodness of fit and model complexity.
- Bayesian Information Criterion (BIC): Similarly, the NNAR model's BIC of 262.94 is lower than that of the ARIMA model (267.14), reinforcing that the NNAR model is more parsimonious while maintaining high predictive performance.

The results demonstrate that the NNAR model can capture patterns and dependencies in the data more effectively than the ARIMA model. The lower MSE highlights the NNAR model's ability to minimize forecasting errors, while the reduced AIC and BIC values suggest that the neural network structure provides a more efficient fit without unnecessary complexity. This highlights the potential advantages of hybrid or non-linear models in time series forecasting, especially in cases where traditional linear models may fall short. Future steps may include cross-validation or testing on out-of-sample data to confirm the robustness of the

NNAR model's performance in real-world forecasting scenarios. Despite their superior performance, NNAR models are associated with several limitations. Neural networks typically require larger datasets to avoid overfitting, and their training process can be computationally intensive compared to ARIMA models. Additionally, the black-box nature of NNAR models reduces interpretability, making it challenging to derive actionable insights from the model's internal representations. Future studies will explore regularization techniques, data augmentation, and hybrid models to address these challenges and improve the practicality of NNAR for diverse forecasting tasks.

## 5. Forecasting Results

Forecasts from the ARIMA(1, 1, 0) and NNAR(2, 2) models are compared against the actual observed values in **Table 5**. This comparison provides insight into the predictive performance and accuracy of each model.

**Table 5.** Forecasting performance of ARIMA and NNAR models.

Forecast	ARIMA	NNAR	Original (Observed)
	103.11	102.67	105.13
	105.42	104.14	105.33
	107.49	104.99	105.24

The forecasts produced by the NNAR(2, 2) model consistently align more closely with the original observed values compared to the ARIMA(1, 1, 0) model. This difference highlights the NNAR model's superior ability to capture patterns and forecast more accurately over the evaluated period.

- Forecast 91: The observed value is 105.13. The NNAR model forecasted 102.67, showing a smaller deviation of 2.46 compared to ARIMA's forecast of 103.11, which deviates by 2.02.
- Forecast 92: For the second forecast, the NNAR prediction of 104.14 is remarkably close to the actual value of 105.33, with a deviation of only 1.19. In contrast, the ARIMA model forecasted 105.42, resulting in a deviation of 0.09, slightly outperforming NNAR in this instance.
- Forecast 93: NNAR once again demonstrated better performance by predicting 104.99, compared to ARIMA's forecast of 107.49. The NNAR model's deviation from the observed value (105.24) is 0.25, while ARIMA deviates by 2.25.
- Across all three forecasts, the NNAR model consistently produced lower deviations from the observed values compared to the ARIMA model. This suggests that the NNAR model is better suited to capture complex, non-linear relationships within the data, whereas the ARIMA model, relying on linear assumptions, may struggle to adapt to such patterns.
- The smaller forecasting errors of the NNAR model reinforce the results seen in **Table 4**, where NNAR outperformed ARIMA in terms of MSE, AIC, and BIC.

- Although ARIMA performed marginally better in Forecast 92, the overall performance trend favors the NNAR model, indicating greater stability and accuracy across the forecasting horizon.

The comparison of forecasts between ARIMA(1, 1, 0) and NNAR(2, 2) models underscores the strengths of neural network-based models in predictive analytics. The NNAR model's ability to provide more accurate forecasts highlights its potential as a robust alternative to traditional linear models, particularly in datasets exhibiting non-linear dependencies. Future work may explore hybrid models or ensemble approaches to further improve predictive performance and reduce forecast uncertainty.

## 6. Conclusion

We presented a comparative analysis of ARIMA(1, 1, 0) and Neural Network Autoregressive (NNAR) models for time series forecasting, focusing on simulated data and neonatal mortality rates in Saudi Arabia. Through rigorous evaluation using performance metrics such as Mean Squared Error (MSE), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC), the study highlights the strengths and limitations of each approach. The results demonstrate that while ARIMA models offer simplicity, interpretability, and effectiveness in capturing linear dependencies, the NNAR model consistently outperforms ARIMA across all evaluation metrics. The NNAR(2, 2) model achieved lower MSE, AIC, and BIC values during training and produced forecasts that aligned more closely with observed values, underscoring its superior capacity to model complex, non-linear relationships within the data. Residual diagnostics of the ARIMA(1, 1, 0) model confirmed its adequacy for linear trend modeling, with white noise residuals and no significant autocorrelation. However, the NNAR model's enhanced performance highlights the advantages of neural network-based approaches, particularly in datasets exhibiting nonlinearity or irregular patterns. The findings suggest that NNAR models are a valuable tool for improving forecasting accuracy, especially in domains where traditional linear models may struggle to capture the full range of dynamics. As computational power and access to large datasets continue to grow, neural network models will likely play an increasingly vital role in predictive analytics. Future research may explore hybrid approaches that integrate ARIMA with neural networks or other machine learning algorithms to further enhance forecasting performance. Additionally, applying these models to larger and more diverse datasets can validate their robustness and scalability, contributing to more reliable and actionable insights in time series forecasting applications. Future work will implement k-fold cross-validation and out-of-sample testing to validate the generalizability of the ARIMA and NNAR models. By partitioning the data into multiple subsets, we aim to minimize overfitting and ensure that the models are robust across various data segments. This approach will enhance the accuracy of model performance metrics and provide a more reliable basis for forecasting neonatal mortality trends.

## Acknowledgements

Sincere thanks to the members of JAMP for their professional performance, and special thanks to managing editor for a rare attitude of high quality.

## Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

## References

- [1] Brockwell, P.J. and Davis, R.A. (2016) Introduction to Time Series and Forecasting. 2nd Edition, Springer.
- [2] Box, G.E.P. and Jenkins, G.M. (1976) Time Series Analysis: Forecasting and Control. Holden-Day Publisher.
- [3] Hyndman, R.J. and Athanasopoulos, G. (2018) Forecasting: Principles and Practice. 2nd Edition, OTexts.
- [4] Shumway, R. and Stoffer, D. (2011) Time Series and Its Applications. Springer.  
<https://doi.org/10.1007/978-1-4757-3261-0>
- [5] Zhang, G.P. (2003) Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model. *Neurocomputing*, **50**, 159-175.  
[https://doi.org/10.1016/s0925-2312\(01\)00702-0](https://doi.org/10.1016/s0925-2312(01)00702-0)
- [6] Crone, S.F. and Kourentzes, N. (2011) Neural Networks for Time Series Forecasting: Practical Implications of Theoretical Model Complexity. 2011 *International Joint Conference on Neural Networks (IJCNN)*, San Jose, 31 July-5 August 2011, 2159-2166.
- [7] Rajab, K., Kalakech, A. and Azar, J. (2017) Forecasting Mortality Rates Using Time Series Models: An Empirical Investigation. *International Journal of Health Policy and Management*, **6**, 573-581.
- [8] Deb, S. and Majumdar, R. (2020) Application of ARIMA and NNAR Models in Predicting Health Outcomes: A Case Study of Infant Mortality. *Journal of Public Health Research*, **9**, 126-135.