International Journal of Computational and Experimental Science and ENgineering (IJCESEN)

Vol. 11-No.3 (2025) pp. 5508-5521
http://www.ijcesen.com

ISSN: 2149-9144

Research Article

Forecasting Non-Gaussian Time Series with TB Data

Lamyaa Mohammed Ali Hameed*

Department of Statistics, College of Administration and Economics, University of Baghdad, Baghdad, Iraq.
* **Corresponding Author Email**: lamiaa.mohammed@coadec.uobaghdad.edu.iq - **ORCID:** 0000-0001-7213-2115

Article Info:

DOI: 10.22399/ijcesen.3446
Received : 22 May 2025
Accepted : 17 July 2025

Keywords

Non-Gaussian
Gamma-ARIMA
EGWO algorithm
Bayesian inference
Tuberculosis
Iraq

Abstract:

Conventional forecasting models require time series that are stationary over time in terms of mean and variance. However, we often encounter data that rarely meet this condition. The data may have Non-Gaussian (N-G) distribution or contain heavy tails or extreme values. In order to improve and strengthen the predictive performance, various (N-G) models have been used, each of which has a different property from the other models. The combined formulas of discrete distributions such as Poisson or Negative – Binomial (NB) distribution with Autoregressive Integrated Moving Average (ARIMA) models provide an interpretable methodology when modeling time series data by following the characteristics of count data because it relies on the distributional properties represented by the general linear model based on count data and the time dependence represented by the ARIMA model of the residuals. Predicting time-dependent patterns of count data involves complexities resulting from the discrete and positive nature of the data, which is not compatible with the classical ARIMA methodology. To address this shortcoming, models combining the two were used as an alternative solution. These models are Gamma-ARIMA, Poisson-ARIMA, and NB-ARIMA. To fit discrete data to a continuous gamma distribution, a new framework, the transformed Gamma-ARIMA model, was proposed. By applying a mathematical transformation to discrete data, the series formation becomes more consistent, and the Gamma-ARIMA technique is successful on non-Gaussian discrete data sets.. Four different mathematical formulations were used, and the Enhanced Grey Wolf Optimizer (EGWO) algorithm was used to compare them. The results show that the square root transformation is the best using the No-U-Turn Sampler (NUTS) algorithm, and that the Bayesian estimation performance is robust and suitable for reliable inference and future predictions. Using an annual time series of the number of pulmonary Tuberculosis (TB) cases in Iraq, the results showed that the Poisson-ARIMA model outperformed the other models using Mean Square Error (MSE)and Mean Absolute Percentage Error (MAPE).

1. Introduction

The majority of time series that describe real events and situations in various fields are Non-Gaussian (N-G) series ,[1] and predicting these time series is a prominent statistical challenge, especially since traditional prediction models such as Autoregressive Integrated Moving Average (ARIMA) models require conditions, foremost among which is that the distribution of the error series be normal, [2] but this condition is not met with time series in which the variable is of the type of countable variable such as people, students, employees, crimes, traffic accidents, patients, units sold, cars, customers, calls, messages, and many others. All of these examples take integer and non-negative values and therefore cannot be accurately represented with a continuous distribution such as the normal distribution. If the series contains zeros or large values, this means that the distribution is severely skewed to the right, and also Kurtosis when it is more than three confirms that the distribution is non-Gaussian. [3] Deviations from a symmetric Gaussian distribution may arise due to the integer number of the variable or the binary nature of the variable, such as such as whether a person has or does not have Tuberculosis (TB). Another reason is that the values are spread far from the mean in an uneven manner [4]. The Gaussian process does not apply to a skewed

distribution. [5] New studies have indicated that the random variable component may have (N-G) distribution in the presence of time dependent parameter. [6] Therefore, the main objective of this research is to predict the (N-G) time series of annual pulmonary tuberculosis cases in Iraq for the period from 1985 to 2023, which is a (N-G) series, using four different (N-G) models and compare them using the statistical measures of Mean Squared Error (MSE) and Mean Absolute Percentage Errors (MAPE).

1.1 Tuberculosis (TB)

Tuberculosis is a contagious bacterial disease caused by a bacterium called Cobacterium tuberculosis [7]. It is widespread throughout Iraq without clear regional boundaries. Tuberculosis fluctuates between ups and downs [8], and over the years, it has been considered endemic in Iraq [9]. There are three categories of the disease [7]. The first category is sputum-positive cases, the second category is reversible or failed cases, and the third category is extrapulmonary tuberculosis.

1.2 Related Work

In Terengganu -Malaysia (2009) [10] they assumed that Holt's trend modified with exponential smoothing is the best model for forecasting the yearly TB cases, in Iran [11] They justified that SARIMA is the best model for forecasting the monthly TB cases, in China - Jiangsu province [12] They illustrated that ARIMA-NAR is the best model for forecasting the monthly TB cases, in Beijing [13] They confirmed that SARIMA is the best model for forecasting the monthly TB cases , and found the incidence of PTB is associated with seasonal weather factors, in China [14] They considered SARIMA-GRNN is the best model for forecasting the monthly TB CASES , in Japan [15] They showed that the degree of seasonality of active TB cases was significantly related with population mass by using spectral analysis for the monthly TB cases, in Shaanxi- China [16]They asserted that SARIMA is the best model for forecasting the monthly TB cases, in Niger State [17] They considered that ARIMA (2, 1, 3) is the best model for forecasting the monthly TB cases, in Xinjiang-China [18] They confirmed that the best method for predicting annual tuberculosis cases is AR-Elman model, in Colombia [19] They emphasized that traditional methods in time series are the best for forecasting TB, [20] in Antioquia - Colombia They confirmed that Kalman filter is the best model for forecasting the weekly TB cases, in Kenya [21] they use ARIMA -ANN to forecast TB cases among children in some in some countries, in Malaysia [22] They use SARIMA model for forecasting the monthly TB cases, in Kazakhstan [23] They use SARIMA model for forecasting the monthly TB cases, in India [24] They showed that NNAR is the best model for forecasting the monthly TB cases, in Indonesia [25] They confirmed that ARMA (1, 1) is the best model for forecasting the monthly TB cases, in USA [26] He clarified that the model of decision tree (DT) is more accurate than ANN for forecasting the monthly TB cases, in Brazil [27] They used SARIMA model for forecasting the monthly TB cases, in Changde – China [28] They affirmed that self-attention model is the best model for forecasting the monthly TB cases, In a review [29] They confirmed that ARIMA; SARIMA; ETS; GRNN; BPNN; NARNN; NNAR; and RNN are common time series models for TB incidence prediction, in Yingjisha County- China (2025) [30] They asserted that SARIMA is the best model for forecasting annual TB.

2. Non-Gaussian (N-G) time series

Most time series that represent real-world situations are actually series that follow N-G time series distributions. N-G time series models were first introduced over 40 years ago in order to capture characteristics that the Gaussian distribution cannot capture.[1], while Gaussian models were introduced nearly a century ago. What distinguishes ARIMA models from other models is the presence of time-dependent factors within them. It is known that Gaussian ARIMA models have an error distributed with a mean zero and a variance of one, while in the case of N-G models, where the error distributed N-G in other forms. The research is based on the idea of generating a model of the values of random variables of the real series from a heavy-tailed or skewed distribution when time series models with Gaussian marginal distributions fail to provide a true description of the actual situation.[1], The classical approach assumes that the time series has both deterministic and random components. Recent studies have confirmed that there is a not insignificant probability that the random component has a (N-G) distribution with time-dependent parameter values [31]. the N-G models used in this paper are: Poisson ARIMA, Negative- Binomial (NB) ARIMA and Gamma ARIMA, this section explores ARIMA models with innovations is members of different types of N-G distributions compares the predictive performance of the three models along with the traditional ARIMA model, for TB forecasting in Iraq.

2.1 Bayesian N-G- ARIMA Models

In the Bayesian method, prior information is entered and the full distribution of parameters is estimated by choosing initial distributions, calculating the posterior distribution and implementing it using MCMC.

Bayesian Inference

The posterior conditional distribution is calculated according to Bayes' law[32]:

The general form of Bayesian inference is given by:

$$P(\theta \mid y) \propto p(y \mid \theta) \cdot p(\theta)$$

Where:

p (θ | y) is the posterior conditional distribution of the parameters θ given the data y.

p (y | θ) is the likelihood function of observing the data y given the parameters θ.

p (θ) The prior distribution of the parameters θ.

To ensure the efficiency and effectiveness of the model, the convergence of the two-sampling series is confirmed. The effectiveness of Monte Carlo (MC) performance is also confirmed by observing the MC error, which is assumed to be less than 5% of the standard deviation from the posterior mean estimate. Otherwise, an alternative distribution is tested and the iterations is increased [32].

Numerical estimation is performed using MCMC (such as Metropolis-Hastings or NUTS), Variational Bayes / INLA for speed, Indices used for comparison, RMSE, MAPE [33], WAIC, LOO, Bayes Factor [33]

2.2 Gamma–ARIMA Model

This model is used with data that exhibits heterogeneous variation over time. The data must also be positive continuous and **(N-G)** in distribution. This model combines the precision of the temporal structure of ARIMA with the suitability of the Gamma distribution for data, making it ideal for application. Bayesian estimation in this model is an effective tool for incorporating prior knowledge and analyzing uncertainty. Given the presence of high values in the data, heavy tails, and extreme values, and thus a non-normal distribution, it is suitable for application. [34]

$$Y_t = \mu_t + \epsilon_t \quad (1)$$

$$\epsilon_t \sim Gamma(\alpha, \beta_t)$$

Where's

μ_t is the expected value at time t calculated from ARIMA model.

ϵ_t is the noise (Random innovation) distributed Gamma.

α the shape parameter.

β is the average parameter associated with μ as follows:

$$\mu_t = c + \sum_{i=1}^{p} \emptyset_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} \quad (2)$$

2.3 Bayesian Estimation [34]

The numerical Bayesian estimation method is used through the MCMC algorithms. The NUTS (No-U-Turn Sampler) algorithm is adopted for its efficiency in exploring the probability space.

Prior assumptions:

Prior Distributions:

$\emptyset_i \sim N(0,1)$, $\theta_j \sim N(0,1)$,

$\alpha \sim Half\ Normal(5), c \sim N(0,10)$

Likelihood Function:

$$p(Y_t \mid \mu_t, \alpha) = (\beta_t{}^{\alpha} / \Gamma(\alpha)) * Y_t{}^{(\alpha-1)}\ e^{(-\beta_t Y_t)}$$

Where $\beta_t = \alpha / \mu_t$

Posterior conditional distribution is performed according to Bayes' law:

$$p(\varphi, \theta, \alpha, c \mid Y) \propto \prod p(Y_t \mid \mu_t, \alpha) p(\varphi)\ p(\theta) p(\alpha)\ p(c)$$

The implementation is by Markov Chain Monte Carlo (MCMC) using samplers such as the No-U-Turn Sampler (NUTS)[33].

Proposed Transformed Gamma–ARIMA for Discrete Time Seri

Ordinary ARIMA models are built on the assumption that the residuals are normally distributed. However, this assumption is often violated when real-world data consists of numerical values. In contrast, Gamma-ARIMA models are only suitable for continuous and positive data, which leaves us puzzled as to how these models fit discrete count series. However, many practical time series—such as daily

hospital visits or statistics for specific events—are not only discrete and non-negative, but also exhibit properties such as heteroscedasticity and asymmetry. These features make the Gamma distribution attractive, but the challenge remains in finding a way to transform the data from discrete to continuous. Therefore, a new framework, the transformed Gamma-ARIMA model, has been proposed. By applying a mathematical transformation to the discrete data, the series can be reshaped into a more consistent form of Gamma modeling, making the Gamma-ARIMA technique successful on discrete, non-Gaussian data sets. Four different mathematical formulas have been used, and these are the formulas that have been applied:

$$yt = yt + \big(log(yt) - log(yt-1)\big) \quad (7)$$

$$yt = yt + 0.05 \quad (8)$$

$$yt = yt + \big(ln(yt) - log(yt)\big) \quad (9)$$

$$yt = yt + log(yt) \quad (10)$$

$$yt = \text{sqrt}(yt) \quad (11)$$

The Enhanced Grey Wolf Optimizer (EGWO) algorithm is used to select the best transformation from among the proposed transformation.

Enhanced Grey Wolf Optimizer (EGWO) [35]

The Enhanced Grey Wolf Optimizer (EGWO) represents an improved variant of the Grey Wolf Optimizer (GWO), originally developed by Mirjalili et al.(2014) [36]. This algorithm draws its conceptual foundation from the natural hunting patterns and hierarchical organization of grey wolves. In EGWO, the mechanisms of locating, surrounding, and attacking prey are mathematically modeled to guide the optimization process, reflecting the cooperative behavior seen in wild wolf packs.

In the GWO algorithm, wolves are classified into four categories:
- Alpha (α): Denotes the best candidate solution.
- Beta (β): The second best solution.
- Delta (δ): The third best solution.
- Omega (ω): The remaining wolves that follow α, β, and δ.

The GWO algorithm uses the (12),(13) equations for updating the positions of the wolves:

$$D = |C * Xp(t) - X(t)| \quad (12)$$

$$X(t+1) = Xp(t) - A * D \quad (13)$$

$Where$:

$Xp(t)$: $Position\ of\ the\ prey$ ($best\ solution$).
- The coefficients are dynamically adjusted using stochastic components that gradually diminish over iterations, aiming to achieve a trade-off between exploration of the search space and exploitation of promising areas.

2.4 Poisson–ARIMA Model

This model is used with data that exhibits Positive-valued outcomes, Skewed distributions and time-varying means with flexible dynamics.

According to Formula (1)

Let $\{y_t\}$ be a time series, with density:

$$p(y_t|\lambda_t) = \exp(-\lambda_t)\frac{\lambda_t^{y_t}}{y_t!}, \quad y_t = 0,1,2,\dots, \quad \lambda > 0, \quad (14)$$

$$\epsilon_t \sim Poisson(\lambda_t)$$

The model supposes that E (y_t) = Var (y_t)

2.5 Negative Binomial ARIMA (NB – ARIMA) Model [37]

The Negative Binomial–ARIMA model is an extension of the typical ARIMA model used to analyze time series containing count data that suffer from overdispersion that is: E(y_t) < Var (y_t), This model assumes that the residuals do not follow a normal distribution, as in traditional ARIMA, but rather follow a more flexible negative binomial distribution. The negative binomial distribution is a probability distribution used to model the number of failures before reaching a certain number of successes in a series of trials with two outcomes (success or failure). The probability of success in each trial is pp.

Probability Density Function (PMF) Formula [38]

If: y is the Number of failures (or observed events) before reaching r success.

p: Probability of success in each trial

The probability function is given by the following equation:

$$P(Y = y) = C(y + r - 1, r - 1)\, p^r\, (1 - p)^y \quad (16)$$

According to Formula (1)

$$\epsilon_t \sim NB(\mu_t, \alpha)$$

The estimation is performed using the MCMC method to extract the dimensional distribution of all parameters.

3. Methodology

- Preprocessing: Perform Normality tests (Shapiro–Wilk, Jarque–Bera, Anderson–Darling) [39].
  - Stationarity test: Conduct the (Augmented Dickey-Fuller) ADF) test [40] before and after taking appropriate differences or transformations.
  - After Determine the required degree of difference, Determine the appropriate p and q degrees based on the graphs of the autocorrelation function (ACF) and the partial autocorrelation function (PACF).
- Apply ARIMA and **(N-G)** ARIMA which are: Poisson- ARIMA; NB -ARIMA; and Gamma-ARIMA By Bayesian inference
  - Plot the results and calculate RMSE
- Select the best model according to the RMSE and MAPE
- Prediction.

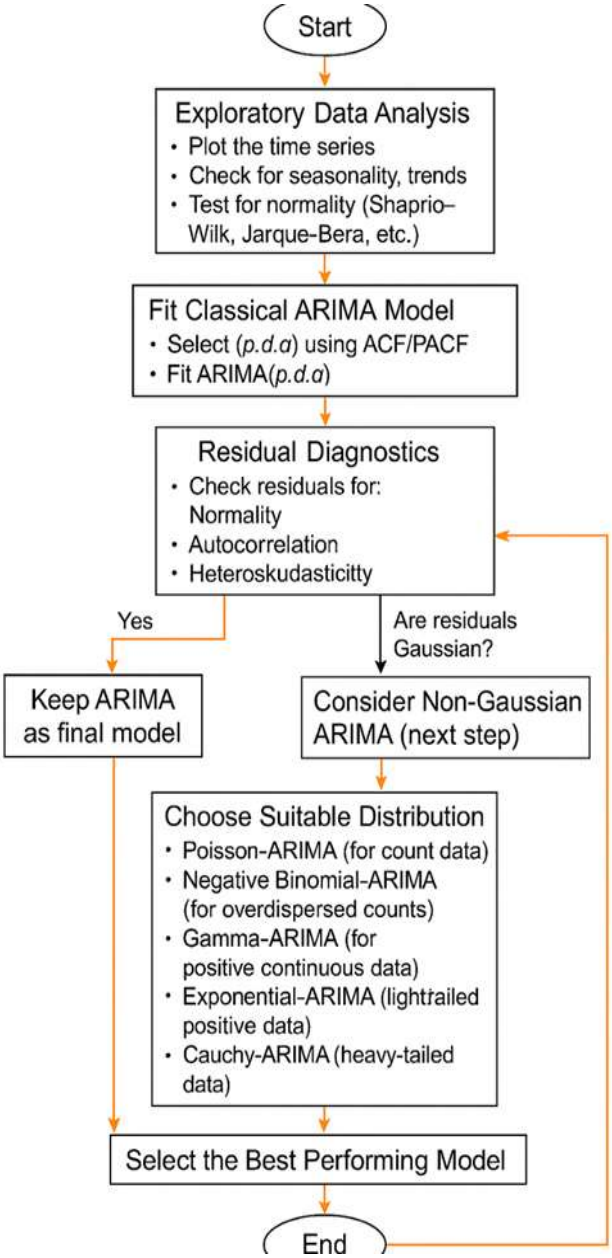

Figure 1. *Proposed algorithm flowchart for selecting the best forecasting model*

4. Results and Discussion

4.1 Annual Tuberculosis cases Data

Data on pulmonary tuberculosis annual count data across Iraq were obtained from the Ministry of Health / Department of Health and Life Statistics for the years (1985 –2023.) The cases are divided into three categories: active tuberculosis, inactive tuberculosis, and extrapulmonary tuberculosis. This research relied on the total number of cases for the three categories. All analyses were performed using Python and MATLAB. The Descriptive Statistics were calculated, the values in table (1):

Table 1. the Descriptive Statistics

The statistic	Value	The statistic	Value	The statistic	Value
Count	39	Min	4843	Q3	14523.5
Mean	12496.15	Max	29906	IQR	6374.5
Median	9668	Range	25063	Skewness	1.351392
Standard Deviation	6939.062	Q1	8149	Kurtosis	0.832763
Variance	48,150,580				

From table 1. We note that the data has a significant deviation from the mean of approximately 6,939 units and a large variance of 6,939, indicating a clear disparity in the data. The difference between the largest and smallest values is 25,062, indicating the presence of extreme data that has a significant impact on the dispersion. (50%) of the data falls between 8,147 and 14,523.6, with a skewness of 1.351. This means that the values are skewed to the right side of the distribution. A positive skewness indicates a long tail on the right. Platykurtic Distribution (Kurtosis < 3)With very high values affecting the distribution, such as the maximum value of 29,905.

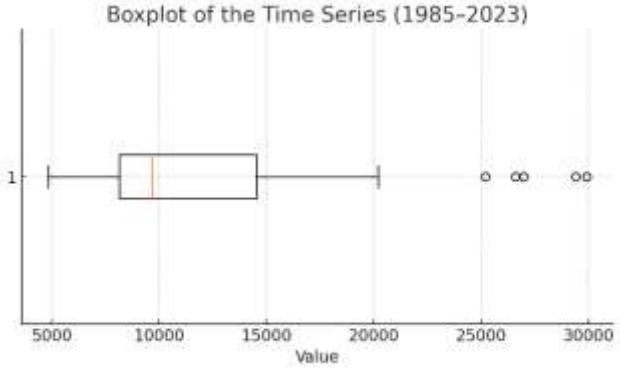

Figure 2. *Boxplot of Original Time Series*

From figure 2. we notice the median is close to the first quartile Q1, which is closer to the minimum acceptable limit, indicating the presence of positive skewness with the presence of 5 extreme values that exceed the acceptable upper limit, while there are no values less than the acceptable lower limit.

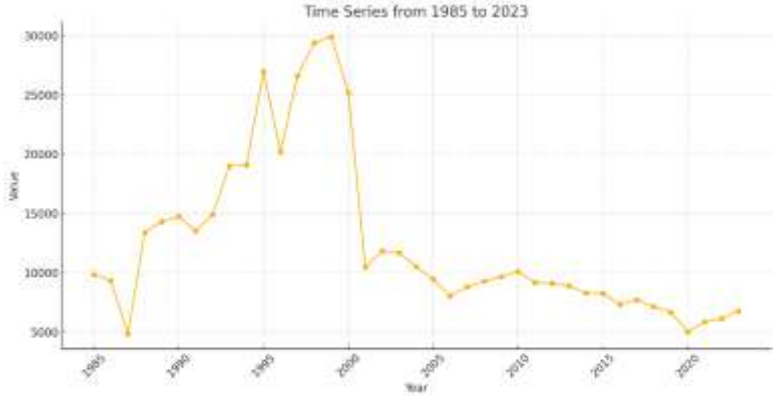

Figure 3. *Graph of the original time series of TB infections in Iraq for the years 1985-2023.*

From Figure 3. we notice that the series does not contain a general trend, with a clear peak at the end of the nineties and a gradual decline in recent years.

4.2 Normality , stationarity and diagnostic test results

Table 2. *Gaussian Test Results*

Test / Metric	p- Value	Interpretation
Shapiro-Wilk Test (p-value)	**p = 0.000023**	Not Gaussian
Jarque-Bera Test (p-value)	0.0032	Not Gaussian
Anderson-Darling Test Statistic	2.67 > Anderson Critical Value (5%) = 0.725	Not Gaussian
Skewness	1.3	Positive skew (right-tailed)
Kurtosis	0.58	Lower than normal kurtosis

table 2. shows the TB time series **follows a (N-G) model, This confirms the inevitability of using (N-G) models. Such as Poisson-ARIMA , NB-ARIMA and Gamma-ARIMA**

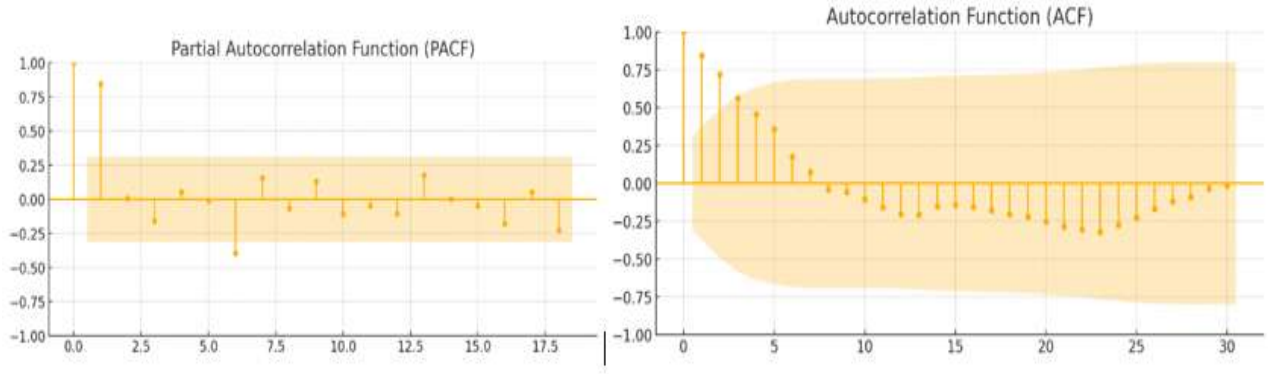

Figure 4. *plots of the Autocorrelation function (ACF) (sample) and the Partial Autocorrelation Function (PACF) (sample) for the original series.*

From Figure 3. The ACF plot indicates that the time series is not stationary because it is decreasing slowly and may also indicate the presence of a moving average (MA) component. From the PACF plot, we find that the values approach zero after lag 1, compared to the theoretical properties, this indicates the possibility of the presence of an AR component (1).

Table 3. *ADF test*

Metric	The original series	after one difference	After Second Differencing	After Log Transformation
ADF Statistic	-1.594	-2.32	-3.116	-1.629
p-value	0.487	0.166	0.025	0.468
Critical Value (1%)	-3.616	-3.654	-3.661	-3.616
Critical Value (5%)	-2.941	-2.957	-2.961	-2.941
Critical Value (10%)	-2.609	-2.618	-2.619	-2.609

Table 3. Confirms that the series becomes stationary after performing the second difference and that performing the logarithmic transformation does not make the series stationary.

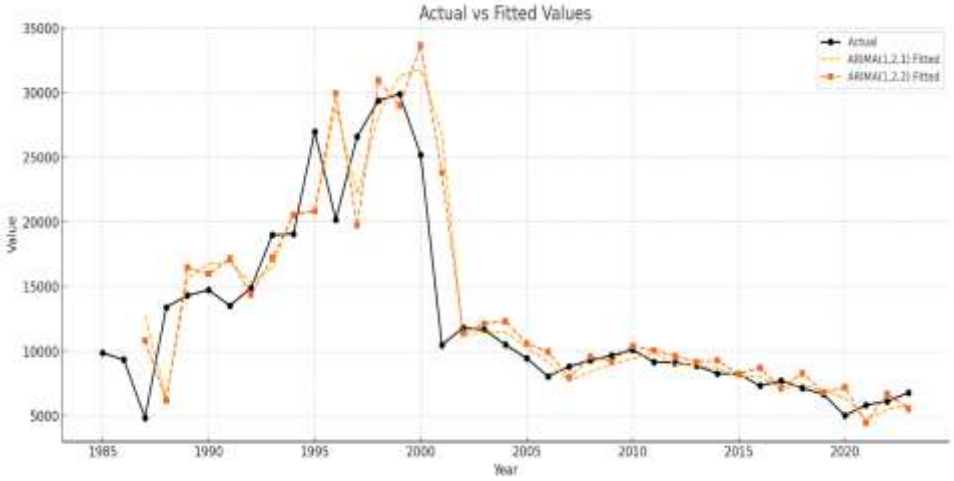

Figure 5. *plot of ARIMA (1,2,1) and ARIMA (1,2,2)*

Table 4. *Comparative analysis and Residual Diagnostics for ARIMA (1,2,1) and ARIMA (1,2,2)*

The Model	RMSE	MAPE	AIC	BIC	Residuals mean	Residuals (Std)	Ljung-Box	ARCH	Shapiro-Wilk
ARIMA (1,2,1)	4060.10	2390.44	724.672	729.505	-583	4267.58	0.2791238	0.953	0.0001325
ARIMA (1,2,2)	4004.85	2536.48	724.200	730.643	-764	4187.28	0.2704236	0.320	0.0010435

Table 4. shows that the RMSE and MAPE values do not agree on one model because the first model is the best according to the MAPE measure, while the second model is the best according to RMSE. As for Model Selection Criteria AIC and BIC there is no agreement between the two measures on a specific model, as the AIC measure shows the superiority of the second model, but the BIC measure shows the superiority of ARIMA (1,2,1) because it has the lower value. For the mean of residuals, ARIMA (1,2,1) equals -5.83 has a slight downward skew, while ARIMA (1,2,2) equals -7.64 has a slightly larger downward skew, i.e., the ARIMA (1,2,1) model predicts better. For residuals Standard Deviation (Std) The most accurate model is the one that has the lowest standard deviation of the residuals. It is clear from the table that ARIMA (1,2,2) has a lower standard deviation of the residuals, meaning that it generates predictions that are closer to the actual values. For the Ljung-Box test, since the p-values for both models are greater than 5%, which means that the residuals are random and there is no autocorrelation. For the ARCH test, since the p-value is > 0.05 for both models, we accept the null hypothesis, there is no statistical evidence of homoscedasticity, the residuals exhibit homoscedasticity. As for the Shapiro-Wilk test, this test confirms that the residuals do not follow the normal distribution in both models because the p-values are very small, less than 0.05, indicating the possibility of thick tails or bias in extreme values. Therefore, for more accurate predictive purposes, it is preferable to use a **(N-G)** ARIMA model.

Top row: Original data The histogram with KDE shows that the data is still non-normal (skewed and with long tails). Q–Q plot: Clear deviation from a straight line, especially at the tails, indicates that the original data does not follow a normal distribution. Bottom row: Residuals generated by ARIMA(1,2,1) Histogram of residuals: Reasonable symmetry around zero is evident, but there are still some deviations. Q–Q plot of residuals: The residuals are closer to a normal distribution than the original data, but there are still deviations in the tails.

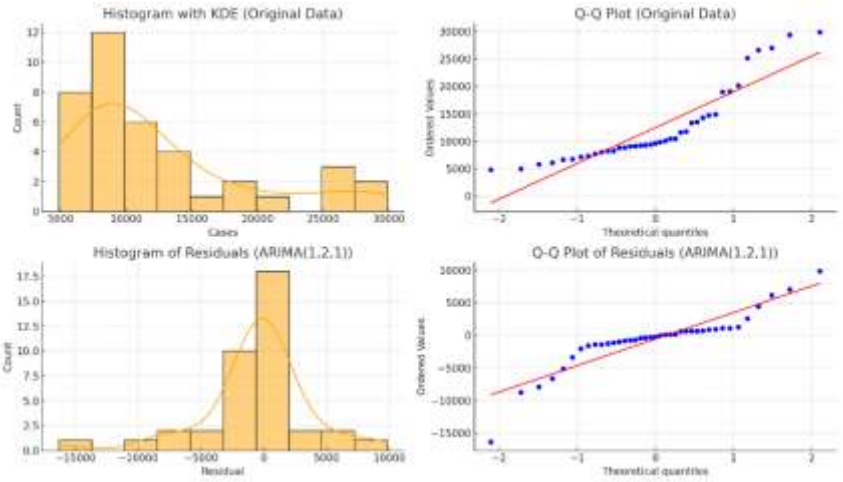

Figure 6. *The Histogram with KDE and Q–Q Plot for the original series and for ARIMA(1,2,1) model*

4.3 Modeling Discrete Time Series via a Continuous Transformation Using a Bayesian Gamma-ARIMA(1,2,1) Model

Figure 7. *The best transformation by EGWO: T11 for Gamma ARIMA (1,2,1)*

The EGWO algorithm, inspired by the behavior of wolves, has proven that the square root transformation is the best and computationally optimal among all proposed transformations.

The Bayesian Gamma–ARIMA (1,2,1) framework was calibrated using the No-U-Turn Sampler (NUTS) algorithm obtainable in the PyMC library. The sampling process contained of a total of 2,000 drawdowns, divided into 1,000 adaptation iterations and 1,000 drawdowns used for subsequent analysis. Throughout the sampling process, no divergent transitions were encountered, and the Markov chains displayed satisfactory mixing performance. The adaptive stage size stabilized inside the range of 0.04 to 0.05, indicative of effective exploration of the posterior distribution. Convergence valuations, including trace visualizations and diagnostic metrics, confirmed that the posterior estimations were robust and appropriate for dependable inference and future predictions.

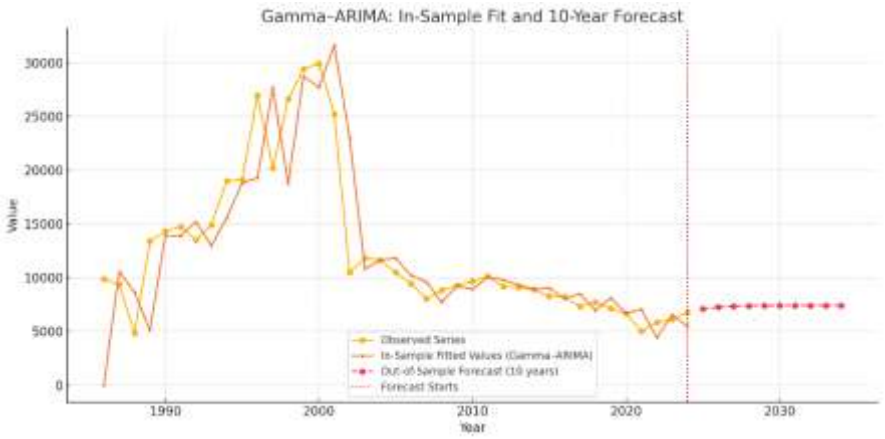

Figure 8. *Bayesian Gamma ARMA (1,2,1)*

	mean	sd	hdi_3%	hdi_97%	mcse_mean	mcse_sd	ess_bulk	ess_tail	r_hat
alpha	1.42	0.89	-0.16	3.08	0.03	0.02	733.43	763.13	1.01
phi	0.85	0.10	0.66	1.01	0.00	0.00	724.43	854.44	1.01
sigma	0.28	0.03	0.23	0.35	0.00	0.00	1128.02	1057.91	1.00
theta	-0.04	1.00	-2.02	1.69	0.03	0.02	1308.89	1146.15	1.00

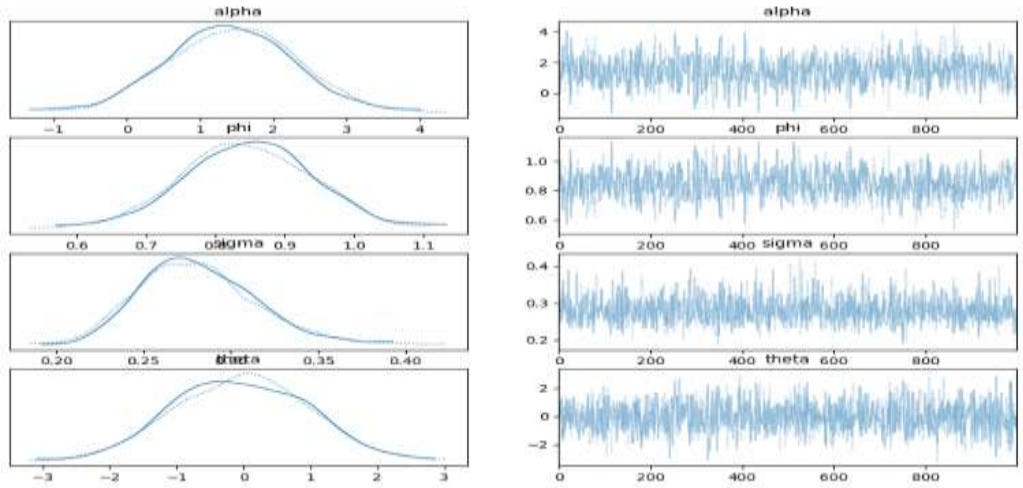

Figure 9. *Visualization of the Bayesian Gamma–ARIMA model parameter estimates through marginal density plots and sampling trace paths.*

The figure shows that the residual deviation is small, indicating relative stability in the model. Graphically speaking, the probability density plot shows that each approximate parameter distribution appears single-peaked and relatively symmetric (with the exception of the theta distribution). Most distributions do not exhibit multiple peaks, indicating sample stability. As for the trace plots, all series exhibit variance around the mean without any obvious skewness. There are no signs of undermixing or convergence problems for MCMC chains, which enhances the reliability of the inference. Therefore, we conclude that R ≈ 1.00 is excellent, while almost all parameters converge well. The sigma mass > 700 is good, and the actual sample size is sufficient for accurate analysis. The low sigma indicates low residuals, indicating a good fit to the series. From all of this, we conclude that the model is effective in describing the series after square root transformation.

5.1 NB- ARIMA Model using Bayesian NB- ARIMA (1,2,1) Model

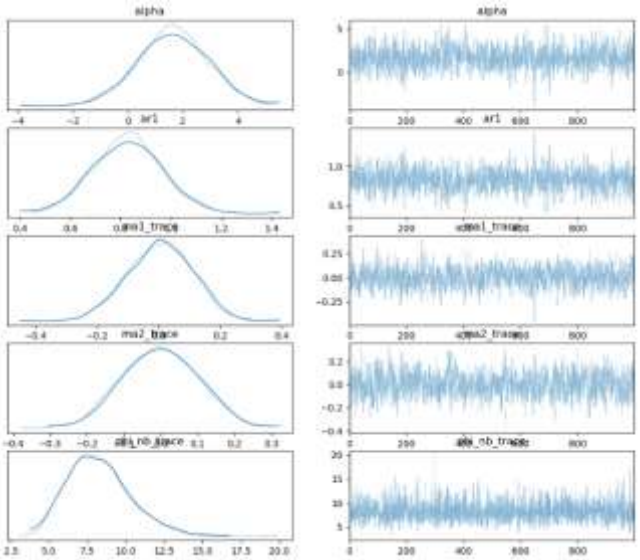

Figure 10. *Visualization of the Bayesian NB–ARIMA (1,2,1) model parameter estimates through marginal density plots and sampling trace paths.*

From Figure 10. The density plots show All parameters show well-shaped posterior distributions, indicating that MCMC is working correctly. The trace plots indicate the chains are stable, with no apparent drift or mixing failure. All chains have normal migrations without bottlenecks or suspicious clusters.

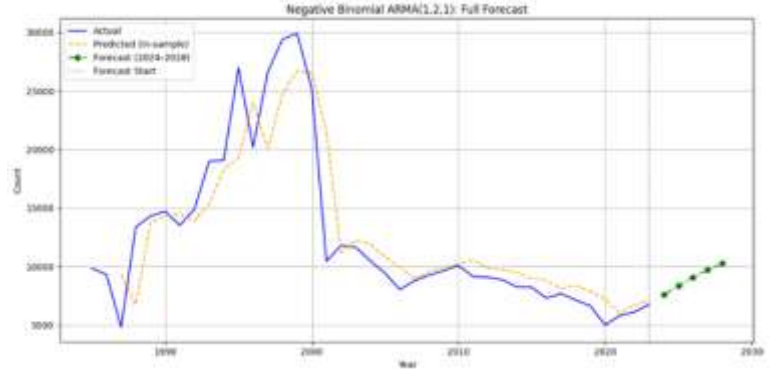

Figure 11. *Bayesian NB - ARMA (1,2,1)*

There are no convergence problems, the central values of the coefficients are clear and consistent, the posterior distributions are smooth and homogeneous, and the RMSE and MAPE are within acceptable limits, especially for data with high variance

5.2 Poisson ARIMA Model using Bayesian Poisson ARMA (1,2,1)

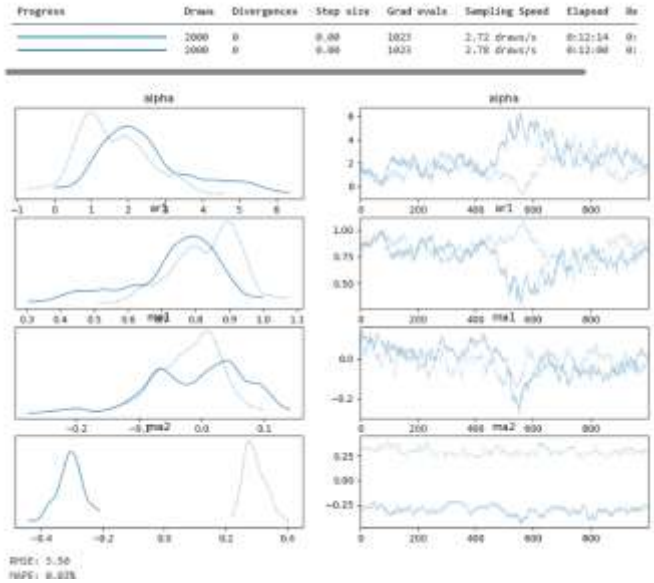

Figure 12. *Visualization of the Bayesian Poisson–ARIMA model parameter estimates through marginal density plots and sampling trace paths.*

The model is stable and appropriate, with no indications of poor convergence or sampling issues, and the prediction quality is high (low MAPE, good RMSE).

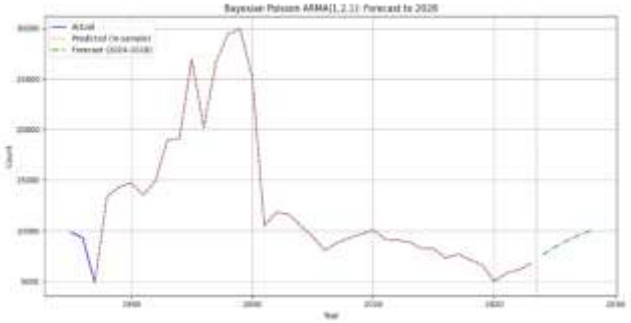

Figure 13. *Bayesian Poisson- ARMA (1,2,1)*

There is no irrational explosion or sharp decline, indicating that the model is Bayesian stable.

Table 5. *Model Evaluation Metrics*

Model	RMSE	MAPE
ARIMA (1,2,1)	4060.10	23.90%
Gamma-ARIMA (1,2,1) with sqr-transformation	3884.95	19.31%
Gamma-ARIMA (1,2,1) with log-transformation	3753.81	18.31%
Poisson-ARIMA (1,2,1)	3.50	0.03%
NB- ARIMA (1,2,1)	3181.47	17.56%

The table 5. indicates that the best model is Poisson-ARIMA (1,2,1) , followed by the N- B-ARIMA, the Gamma -ARIMA model came in third place, the Ordinary ARIMA model came in fourth place, it is clear from the comparison that models with heavy-tailed distributions are more robust to outliers and extreme fluctuations.

Table 6. *Poisson-ARIMA 8-Year Forecast*

Year	Forecast	Year	Forecast
2024	6569.553	2028	5803.205
2025	6368.838	2029	5625.486
2026	6174.773	2030	5452.99
2027	5986.263	2031	5285.594

6. Conclusions

The number of reported tuberculosis cases in Iraq has declined significantly, especially after 1999, reaching its lowest point in 2020. This is due to increased healthcare and health awareness. The overall trend was significantly upward, but in 1999, the overall trend became downward, meaning that the time series is not stationary on average. In this article, a comparison was made between Gamma-ARIMA, Poisson-ARIMA, Negative Binomial-ARIMA, and Classical ARIMA models. The Poisson-ARIMA model outperformed the other models in accurately predicting tuberculosis cases for the next eight years using RMSE and MAPE. Annual predictions of tuberculosis cases in Iraq were made. These predictions are useful for decision-making to maintain public health, control, and prevent the disease. We hope that through this research, we have taken a step forward toward more reliable predictions of tuberculosis cases in Iraq. Furthermore, we hope to conduct future studies, more detailed by governorate and gender, using monthly, quarterly, or semi-annual data, to provide insights into TB cases over short periods. The Bayesian model is excellent for long-tailed count data. It preserves the nature of the original data without transformation. It is well-suited for future predictions of annual pulmonary TB case numbers. ARIMA appears to perform poorly under the assumption of a normal error distribution. The Bayesian Poisson-ARIMA model is an effective model for analyzing time series count data using the ARIMA structure and the Poisson distribution. Bayesian inference provides reliable estimates. Bayesian ARIMA (NG) is an advanced approach for analyzing time series that exhibit NG residuals. It is an ideal choice when the data contain outliers, large variances, and skewed distributions with thick tails. The model provides a better understanding of uncertainty and delivers more realistic predictions compared to the traditional ARIMA model. The predictive ability of the Gamma-ARIMA model was evaluated in the context of continuous, positive, and large-valued time series data exhibiting perfectly skewed distributional properties. The results showed that the NG-ARIMA model improved the prediction accuracy of NG time series with an autocorrelated structure. This model is of particular interest for data such as the annual number of TB cases. The ARIMA-Poisson model performed better than the N-B-ARIMA model, while the discrete non-Gaussian models performed better than the continuous Gamma-ARIMA model with root and log transformations.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.

- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1] Balakrishna, N. (2021). *Non-Gaussian autoregressive-type time series*. Springer.

[2] Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2009). *Time series analysis: Forecasting and control* (3rd ed.). Pearson.

[3] Wang, J., & Chen, H. (2024). A bi-Gamma distribution model for a broadband non-Gaussian random stress rainflow range based on a neural network. *Applied Sciences*, *14*(18), 8376. https://doi.org/10.3390/app14188376

[4] Ng, J., Forbes, C. S., Martin, G. M., & McCabe, B. P. M. (2013). Non-parametric estimation of forecast distributions in non-Gaussian, non-linear state space models. *International Journal of Forecasting*, *29*(3), 411–430. https://doi.org/10.1016/j.ijforecast.2012.10.005

[5] Kaur, S., & Rakshit, M. (2019). Gaussian and non-Gaussian autoregressive time series models with rainfall data. *International Journal of Engineering and Advanced Technology*, *9*(1).

[6] Szarek, D., Jabłoński, I., Zimroz, R., & Wyłomańska, A. (2023). Non-Gaussian feature distribution forecasting based on ConvLSTM neural network and its application to robust machine condition prognosis. *Expert Systems with Applications*, *230*, 120588. https://doi.org/10.1016/j.eswa.2023.120588

[7] Al-Mudhaffar, M. A. S. (1979). *Spatial analysis of endemic diseases in Iraq: A study of the geographical foundations of health planning*. Al-Irshad Press.

[8] Ali, A. H. (2007). *The impact of administrative functions on the DOTS program for the treatment of tuberculosis in Iraq* (Master's thesis). St. Clements International University.

[9] Ministry of Health – Department of Public Health and Primary Health Care, Institute of Chest and Respiratory Diseases. (2023). *Epidemiology of tuberculosis in Iraq*.

[10] Kilicman, A., & Roslan, U. A. M. (2009). Tuberculosis in the Terengganu region: Forecast and data analysis. *ScienceAsia*, *35*, 392–395.

[11] Moosazadeh, M., Nasehi, M., Bahrampour, A., Khanjani, N., Sharafi, S., & Ahmadi, S. (2014). Forecasting tuberculosis incidence in Iran using Box-Jenkins models. *Iran Red Crescent Medical Journal*, *16*(5), e11779.

[12] Wang, K. W., Deng, C., Li, J. P., Zhang, Y. Y., Li, X. Y., & Wu, M. C. (2017). Hybrid methodology for tuberculosis incidence time-series forecasting based on ARIMA and a NAR neural network. *Epidemiology and Infection*, *145*(6), 1118–1129. https://doi.org/10.1017/S0950268816003216

[13] Keerqinfu, Zhang, Q., Yan, L., & He, J. (2018). Time series analysis of correlativity between pulmonary tuberculosis and seasonal meteorological factors based on theory of Human-Environmental Inter Relation. *Journal of Traditional Chinese Medical Sciences*, *5*(2), 119–127. https://doi.org/10.1016/j.jtcms.2018.03.001

[14] Wang, H., Tian, C. W., Wang, W. M., & Luo, X. M. (2018). Time-series analysis of tuberculosis from 2005 to 2017 in China. *Epidemiology and Infection*, *146*, 935–939. https://doi.org/10.1017/S0950268818001115

[15] Sumi, A., & Kobayashi, N. (2019). Time-series analysis of geographically specific monthly number of newly registered cases of active tuberculosis in Japan. *PLoS ONE*, *14*(3), e0213856. https://doi.org/10.1371/journal.pone.0213856

[16] Yang, C., et al. (2020). Analysis and prediction of pulmonary tuberculosis using an ARIMA model in Shaanxi Province, China. *Journal of Physics: Conference Series*, *1624*, 022013.

[17] Olanrewaju, S. O., Ojo, E. O., & Oguntade, E. S. (2020). Time series analysis on reported cases of tuberculosis in Minna Niger State Nigeria. *Open Journal of Statistics*, *10*, 412–430. https://doi.org/10.4236/ojs.2020.103027

[18] Zheng, Y., Zhang, X., Wang, X., et al. (2021). Predictive study of tuberculosis incidence by time series method and Elman neural network in Kashgar, China. *BMJ Open*, *11*, e041040. https://doi.org/10.1136/bmjopen-2020-041040

[19] Orjuela-Cañón, A. D., Jutinico, A. L., Duarte González, M. E., Awad García, C. E., Vergara, E., & Palencia, M. A. (2022). Time series forecasting for tuberculosis incidence employing neural network models. *Heliyon*, *8*(7), e09897. https://doi.org/10.1016/j.heliyon.2022.e09897

[20] Wyld, D. C., et al. (2022). Towards tuberculosis incidence time series forecasting in Colombian regions: The Antioquia case. *Computer Science & Information Technology*, 117–126.

[21] Siamba, S., Otieno, A., & Koech, J. (2023). Application of ARIMA, and hybrid ARIMA models in predicting and forecasting tuberculosis incidences among children in Homa Bay and Turkana Counties, Kenya. *PLOS Digital Health*, *2*(2), e0000084. https://doi.org/10.1371/journal.pdig.0000084

[22] Ab Rashid, M. A., Ahmad Zaki, R., Wan Mahiyuddin, W. R., & Yahya, A. (2023). Forecasting new tuberculosis cases in Malaysia: A time-series study using the autoregressive integrated moving average (ARIMA) model. *Cureus*, *15*(9), e44676. https://doi.org/10.7759/cureus.44676

[23] Yerdessov, S., Arupzhanov, I., Aimyshev, T., Makhammajanov, Z., Kadyrov, S., Kashkynbayev, A., & Gaipov, A. (2023). Time series analysis on tuberculosis cases in Kazakhstan from 2014 to 2025. *European Journal of Public Health*, *33*(Suppl_2). https://doi.org/10.1093/eurpub/ckad160.888

[24] Yadav, B. K., Srivastava, S. K., Thillai Arasu, P., & Singh, P. (2023). Time series modeling of tuberculosis cases in India from 2017 to 2022 based on the SARIMA-NNAR hybrid model. *Canadian Journal of Infectious Diseases and Medical Microbiology*, Article ID 5934552. https://doi.org/10.1155/2023/5934552

[25] Sumanto, Y. D., Hariyanto, S., & Andriyati, A. (2024). Forecasting TB in Indonesia using ARMA(1,1). *International Journal of Software & Hardware Research in Engineering*, *12*(9), 8–13.

[26] Elhag, A. A. (2024). Prediction and classification of tuberculosis using machine learning. *Journal of Statistics Applications & Probability*, *13*(3), 939–946.

[27] Silva, M. T., & Galvão, T. F. (2024). Tuberculosis incidence in Brazil: Time series analysis between 2001 and 2021 and projection until 2030. *Revista Brasileira de Epidemiologia*, *27*, e240027. https://doi.org/10.1590/1980-549720240027

[28] Lv, Z., Sun, R., Liu, X., et al. (2024). Evaluating the effectiveness of self-attention mechanism in tuberculosis time series forecasting. *BMC Infectious Diseases*, *24*, 1377. https://doi.org/10.1186/s12879-024-10183-9

[29] Maipan-Uku, J. Y., & Cavus, N. (2024). Forecasting tuberculosis incidence: A review of time series and machine learning models for prediction and eradication strategies. *International Journal of Environmental Health Research*. Advance online publication. https://doi.org/10.1080/09603123.2024.2368137

[30] Chang, M., Emam, M., Zhang, N., et al. (2025). Analysis of changes in tuberculosis incidence trends before and after COVID-19 based on time series models. *Advances in Continuous and Discrete Models*, *48*. https://doi.org/10.1186/s13662-025-03905-z

[31] Żuławiński, W., Maraj-Zygmąt, K., Shiri, H., Wyłomańska, A., & Zimroz, R. (2023). Framework for stochastic modelling of long-term non-homogeneous data with non-Gaussian characteristics for machine condition prognosis. *Mechanical Systems and Signal Processing*, *184*, 109677.

[32] Azizi, A. (2014). Integration of seasonal autoregressive integrated moving average and Bayesian methods to predict production throughput under random variables. *Journal of Mechanical Engineering and Sciences*, *7*, 1236–1250. https://doi.org/10.15282/jmes.7.2014.23.0121

[33] Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*, 1413–1432. https://doi.org/10.1007/s11222-016-9696-4

[34] Helske, J., & Vihola, M. (2021). bssm: Bayesian inference of non-linear and non-Gaussian state space models in R. *The R Journal*, *13*(2).

[35] Rajakumar, R., Sekaran, K., Hsu, C.-H., & Kadry, S. (2021). Accelerated grey wolf optimization for global optimization problems. *Technological Forecasting and Social Change*, *169*, 120824. https://doi.org/10.1016/j.techfore.2021.120824

[36] Mirjalili, S., Mirjalili, S. M., & Lewis, A. (2014). Grey wolf optimizer. *Advances in Engineering Software*, *69*, 46–61. https://doi.org/10.1016/j.advengsoft.2013.12.007

[37] Pala, L. O. D. E., Carvalho, M. D. M., & Sáfadi, T. (2023). Analysis of count time series: A Bayesian GARMA (p, q) approach. *Austrian Journal of Statistics*, *52*, 131–151. https://doi.org/10.17713/ajs.v52i5.1568

[38] Jia, Y. (2018). Some models for count time series (Doctoral dissertation). Clemson University. https://open.clemson.edu/all_dissertations/2213

[39] Razali, N. M., & Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics*, *2*(1), 21–33.

[40] Wazeer, E. A. A., & Hameed, L. M. A. (2021). Forecasting the exchange rates of the US dollar against the Iraqi dinar using the Box-Jenkins methodology in time series with practical application. *International Journal of Agricultural and Statistical Sciences*, *17*, 2085. https://connectjournals.com/03899.2021.17.2085