A Comparative Analysis of ARIMA Models for Forecasting China's GDP

Dongjie Wu

School of Mathematics and Physics, Xiamen University Malaysia, Sepang, Malaysia

MAT2209457@xmu.edu.my

Abstract. Gross Domestic Product (GDP) is the total market value of final goods and services produced by a country in a year. This study attempted to find the best-fit Autoregressive Integrated Moving Average (ARIMA) model for forecasting China's GDP over the next five years (2025 to 2029). In this study, we collected historical GDP data for China from 1960 to 2024 from the World Bank. Using the Box-Jenkins approach, we examined the ACF and PACF plots, performed stationarity tests, and tested several models using the AIC criterion. We determined ARIMA(1,2,1) would be the best model to fit the data. We then used the fitted model to forecast the following five years for GDP in China, demonstrating the capabilities of ARIMA as an effective forecasting model.This study provides valuable insights for policymakers and economists in planning sustainable economic strategies for China's future development.

Keywords: ARIMA, Box-Jenkins approach, GDP forecasting, ACF, PACF, stationarity tests, AIC criterion, predictive modeling

1. Introduction

As one of the most important aspects of a country's overall economic performance, Gross Domestic Product (GDP) represents the total value of all goods a services produced in a given area over a given period. GDP is a key component in measuring a country's overall economic growth, assessing the actions of policy decisions, and determining economic development strategies. Given the significance of GDP, the accurate forecasting of GDP has become one of the most important aspects of a macroeconomic forecast. An accurate GDP forecast is essential in analyzing future trends for makers, investors, and government policy, as it plays a role in allocation of resources, forecasting, and planning.

In recent decades, a variety of statistical and econometric methods have emerged to model and predict GDP. One of the methods used was the Autoregressive Integrated Moving Average (ARIMA) model, proposed by Box and Jenkins, which has emerged as quite popular due to its ease of use and performance in forecasting (in the univariate context) [1]. A substantial body of empirical research has also been conducted, demonstrating that ARIMA models are effective for forecasting GDP. These studies have been conducted for Egypt [2], Kenya [3], India [4], Bangladesh [5], and the African continent [6].

In addition to GDP, ARIMA models have also been successfully applied to forecast various economic and financial variables such as CPI [7], commodity prices [8], and stock prices [9,10]. Some studies have gone further to compare ARIMA with more advanced machine learning models like LSTM, highlighting both the strengths and limitations of ARIMA in complex forecasting tasks [11,12].

This paper proposes to find the best ARIMA model to forecast China's GDP from 2025 to 2029 using annual GDP data from 1960 to 2024. We will take the Box-Jenkins approach to examine a range of ARIMA models and compare them to find the most suitable model for our data.

2. Literature review

2.1. Empirical studies on GDP forecasting using ARIMA

Numerous studies have applied ARIMA models to forecast national GDP, demonstrating their effectiveness across different economic settings. For example, Abonazel and Abd-Elftah [2] used an ARIMA(1,2,1) model to predict Egypt's GDP from 1965 to 2016, selecting the best model based on AIC and residual diagnostics. Mungai [3] studied Kenya's GDP from 1960 to 2012 using the Box-Jenkins approach and selected ARIMA(2, 2, 2) as the best-fitting model based on AIC and stationarity tests. The

model showed good predictive accuracy, with in-sample forecasts deviating less than 5%, and was used to forecast GDP for the next five years. In the case of Bangladesh, Bhuiyan et al. [5] modeled Bangladesh's manufacturing GDP and confirmed that ARIMA models can effectively capture economic patterns. Despite these contributions, ARIMA-based studies focusing on China remain limited. Ning et al. [13] analyzed a regional GDP series for Shaanxi province using a shorter time frame. In contrast, this study has a longer national series from 1960 to 2024 and forecasts China's GDP through 2029, thus providing a more comprehensive perspective.

2.2. ARIMA versus alternative forecasting methods

Although ARIMA models are commonly used for forecasting in economics due to their interpretability and good theoretical foundation, they have limitations due to the nature of economic data, which can exhibit complex nonlinear correlations. More modern machine learning models, such as Long Short-Term Memory (LSTM) neural networks, on the other hand, performed well at reflecting the historical nonlinear dependencies within time series data.

Siami-Namini and Namin [11] conducted a number of large-scale empirical comparisons of LSTM and ARIMA models over multiple economic and financial datasets and found consistently better accuracy performance from LSTM models than ARIMA models. Similarly, Sirisha et al. [12] concluded that LSTM models produced superior performance compared to ARIMA and SARIMA models in forecasting profit levels.

Nonetheless, deep learning models such as LSTM models typically require significantly more training data, have higher computational costs, and provide less interpretability than a traditional statistical model. Thus, in small-sample, single-variable economies such as GDP models, ARIMA provides a more practical and interpretable solution.

2.3. Autoregressive Integrated Moving Average (ARIMA) models

2.3.1. Moving-average (MA) model

The moving-average (MA) model is one of the fundamental components in time series analysis. It represents the current value of a time series as a linear combination of current and past white noise error terms. Specifically, a time series $\{x_t\}$ is said to follow an MA(q) process if it can be written as:

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \tag{1}$$

where εt denotes a white noise error term at time t, and $\theta 1$, $\theta 2$,...,θq are the model parameters. This model captures short-term dependencies in the data by expressing x_t in terms of a finite number of past forecast errors. Since only a limited number of lags are involved, the MA model is typically classified as a short-memory process [1].

2.3.2. Autoregressive (AR) model

The autoregressive (AR) model describes a time series where the current value depends linearly on its previous values. An AR(p) process of order p is defined as:

$$X_t = \phi_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \varepsilon_t \tag{2}$$

where x_t is the value of the time series at time t, $\phi 0$ is a constant term, $\phi 1, \phi 2, \ldots, \phi p$ are the autoregressive coefficients, and εt is a white noise error term with zero mean and constant variance.

Unlike the MA model, the AR model allows past values of the series to influence the current observation, often resulting in longer memory and persistence in the data. The autocorrelation function (ACF) of an AR process typically decays gradually, which is a key indicator in identifying its structure [1].

2.3.3. Autoregressive Moving Average (ARMA) model

The ARMA model combines the properties of both the autoregressive (AR) and moving average (MA) models. A time series $\{x_t\}$ is said to follow an ARMA(p, q) process when it incorporates both p lagged values of the series and q lagged error terms. The general form of the model is:

$$X_t = \phi_0 + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q} \tag{3}$$

In this structure, ϕi represents the autoregressive parameters and θ_j denotes the moving average coefficients. The error term εt is assumed to be white noise. By incorporating both past values and past errors, the ARMA model is capable of capturing more complex time series dynamics compared to the AR or MA models individually [1].

2.3.4. Autoregressive Integrated Moving Average (ARIMA) model

The ARIMA model generalizes the ARMA framework to accommodate non- stationary time series by introducing a differencing component. In an ARIMA(p,d,q) model, p denotes the order of the autoregressive part, d represents the number of times the series is differenced to achieve stationarity, and q is the order of the moving average part.

When a time series $\{ x_t \}$ exhibits non-stationary behavior, differencing is applied d times to obtain a stationary series $\{\nabla d x_t\}$. A first-order differenced ARIMA(p,1,q) model can be written as:

$$\nabla X_t = \phi_0 + \phi_1 \nabla X_{t-1} + \cdots + \phi_p \nabla X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q} \tag{4}$$

where $\nabla x_t = x_t - x_{t-1}$ represents the first difference of the original series. If both p and q are set to zero, the model reduces to a simple random walk, which corresponds to an ARIMA(0,1,0) process. This extension makes ARIMA models particularly suitable for real-world economic and financial data, which often exhibit trends or evolving variance over time [1].

2.4. Box-Jenkins methodology

The Box–Jenkins approach, introduced by George Box and Gwilym Jenkins in the early 1970s, is a systematic approach for identifying, estimating, and checking ARIMA models in time series modelling. The Box–Jenkins procedure is intended to provide the best statistical fit of a model to past observations of a univariate time series. The Box–Jenkins procedure iterates on diagnostics and modifications to ensure that the model captures the basic data structure with minimal residual autocorrelation [1]. The Box–Jenkins process involves four steps, as shown in Figure 1.
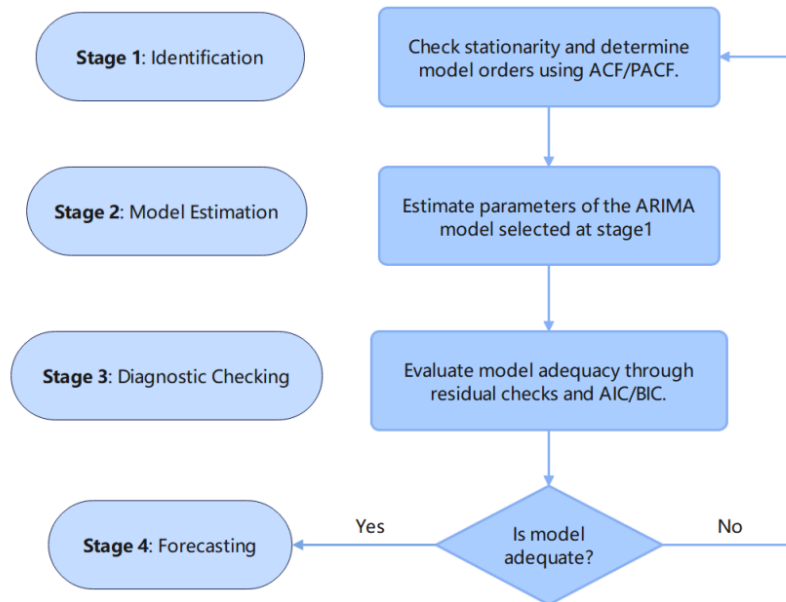
Figure 1. The Box–Jenkins iterative modeling process

In Stage 1, we will assess the stationarity of the time series via a series of tests and differencing non-stationary data until stationarity is achieved. The plots of the autocorrelation function (ACF) and partial auto- correlation function (PACF) will be examined to suggest appropriate orders for the autoregressive (p) and moving-average (q) components of our model. The outcome of this step will provide an initial candidate ARIMA(p, d, q) model for validation.

After selecting the model, the parameters of our model will be estimated using either maximum likelihood estimation (MLE) or non- linear least squares methods, which seek to provide the best model by optimally estimating the relationship between the actual and predicted values while ensuring that these estimates are statistically consistent.

Then in Stage 3, the adequacy of the fitted model is evaluated. We want to ensure the residuals are representative of white noise (i.e., they should be uncorrelated, normally distributed, and homoscedastic). We will use diagnostic tests like the Ljung–

Box test, and ACF/PACF plots ofthe residuals. We will also compare the model performance using information criteria like the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), defined as follows:

$$\text{AIC} = 2m - 2\ln\left(\widehat{L}\right), \quad \text{BIC} = \ln\left(n\right)m - 2\ln\left(\widehat{L}\right) \tag{5}$$

where denotes the maximum value of the likelihood function, m is the number of estimated parameters, and n is the sample size. Lower values of AIC or BIC indicate a more optimal balance between model complexity and goodness of fit.

Once a model passes the diagnostic checks, it can be used to produce forecasts. A validated ARIMA model provides reliable short-term projections by effectively modeling the underlying dynamics of the time series in a statistically sound manner.

3. ARIMA model estimation and GDP forecasting

3.1. Data description

This analysis uses annual GDP data for China from the World Bank over the period 1960-2024, which consists of 65 observations. This is well above the generally accepted minimum of 50 observations for ARIMA data in the Box–Jenkins paradigm, and the measurements are in current US dollars (USD).

### 3.2. Stationarity and preliminary analysis

A simple, direct look at the original GDP series (Figure 2) shows a clear exponential upward trend over time (particularly after the early 2000s), which leads to the assumption that there may be non-stationarity in the series, which would violate the assumptions necessary for ARIMA processes. To determine whether the data were non-stationary, the autocorrelation function (ACF) and partial autocorrelation function (PACF) were reviewed (Figure 3 and Figure 4). The ACF shows evidence of slowly decaying while the PACF noticeably spikes at lag 1 and drops off quickly, which indicates the existence of non-stationarity.
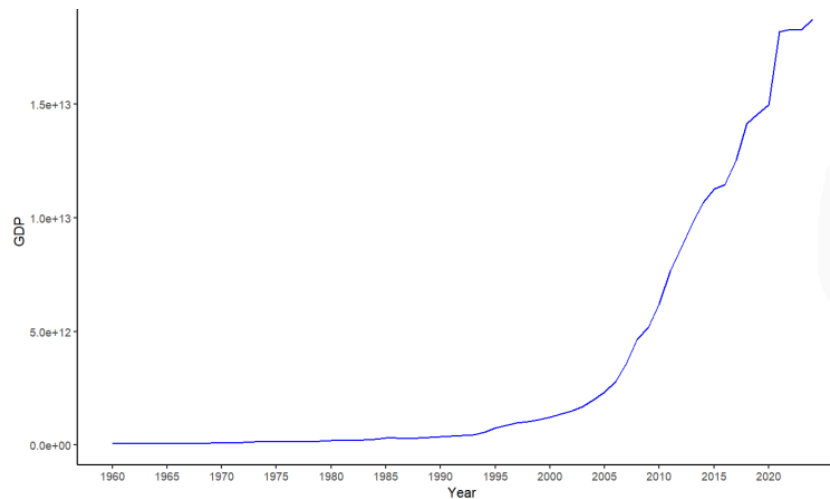
**Figure 2.** Time series plot for China GDP
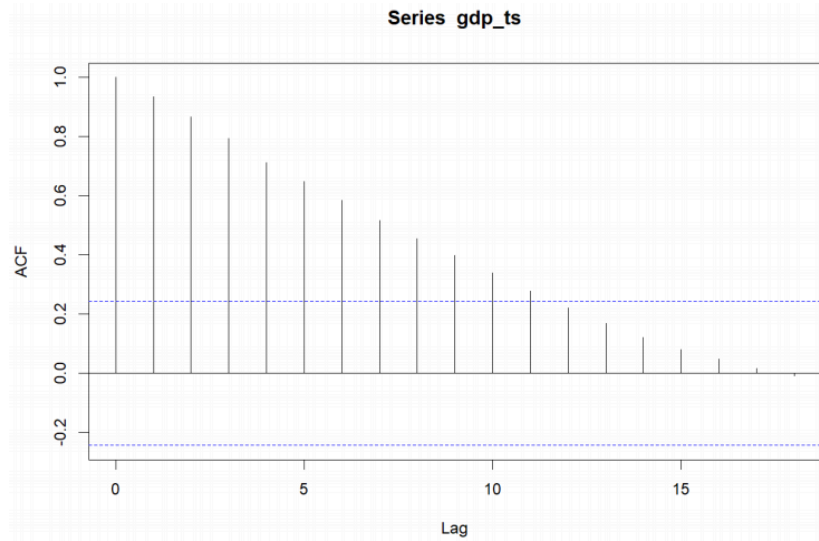
Figure 3. ACF plot for China GDP data
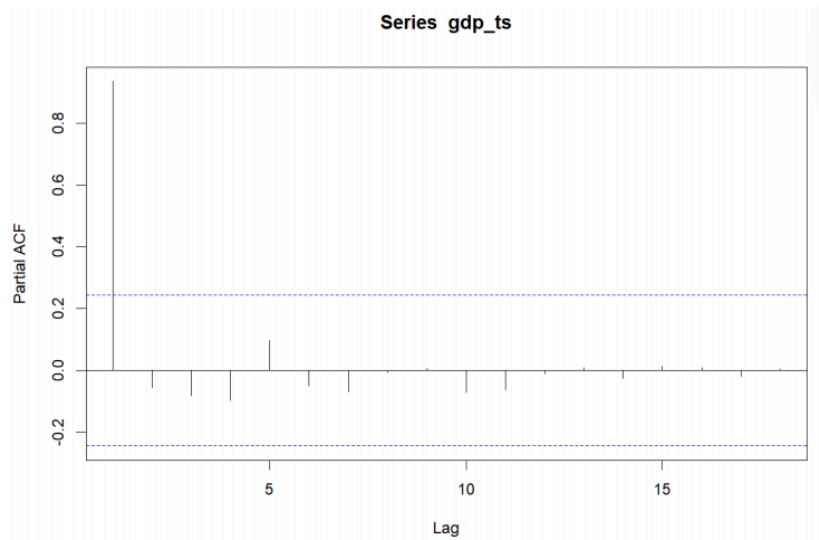
Figure 4. PACF plot for China GDP data

The Augmented Dickey–Fuller (ADF) test was used to verify this statistically. The ADF test returned a p-value of 0.978, which is much larger than the common threshold of 0.05, meaning that the unit root null hypothesis can not be rejected. Thus, the series is nonstationary.

To create a stationary series, we logarithmically transformed the original series and performed first-order differencing. The resulting series is much less volatile, as demonstrated in Figure 5, but it still shows some signs of non-stationarity. The ADF test gives a p-value of approximately 0.167, indicating that the series is still non-stationary at the 5% significance level.
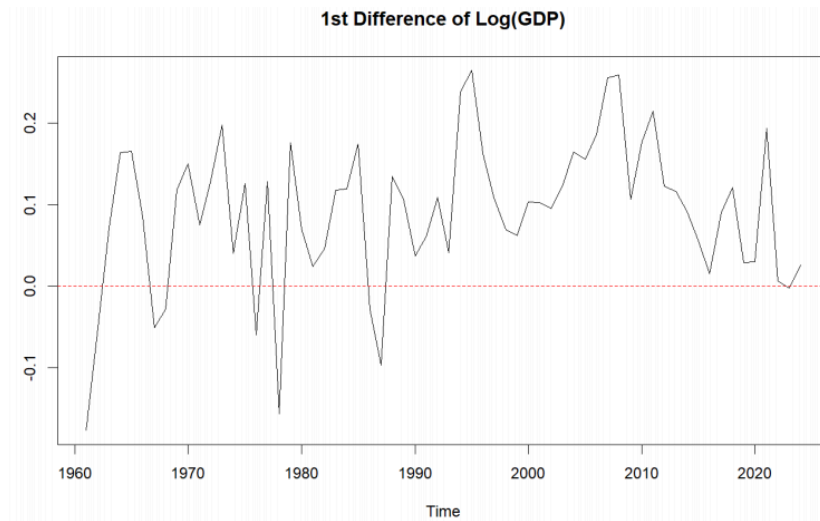
Figure 5. Time series plot for the first-order difference of ln(GDP)

Therefore, additional differencing is required before fitting an ARIMA model.

Then, a second-order difference to the natural logarithm of the GDP series was applied. As shown in Figure 6, the series becomes more stable and fluctuates around a constant mean (0). The ADF test returns a p-value of 0.01, confirming stationarity at the 5% level. Based on this, we determine that the differencing order is d=2.
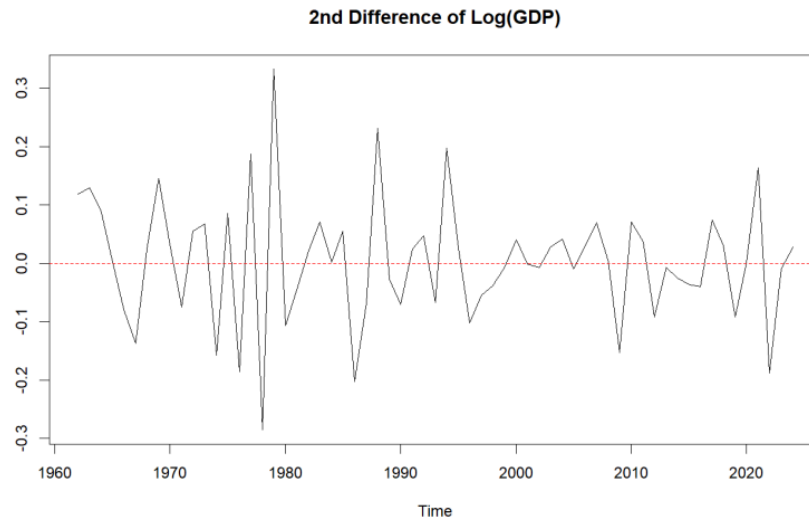
Figure 6. Time series plot for the second-order difference of ln(GDP)

The transformation is defined as:

$$X_t = \ln(GDP_t) - \ln(GDP_{t-2}) \tag{6}$$

where ln(GDPt) denotes the natural logarithm of the GDP at time t.

3.3. Model identification and selection

After performing second-order differencing on the log-transformed GDP series, we examined the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots with a view to determining the appropriate ARIMA model structure.

In the ACF plot (Figure 7), we observe a significant negative spike at lag 1 with a rapid decay. The PACF plot (Figure 8) shows a significant value at lag 1, then diminishes rapidly at subsequent lags. This pattern suggests the data has both an order one autoregressive (AR) and a moving average (MA) component.
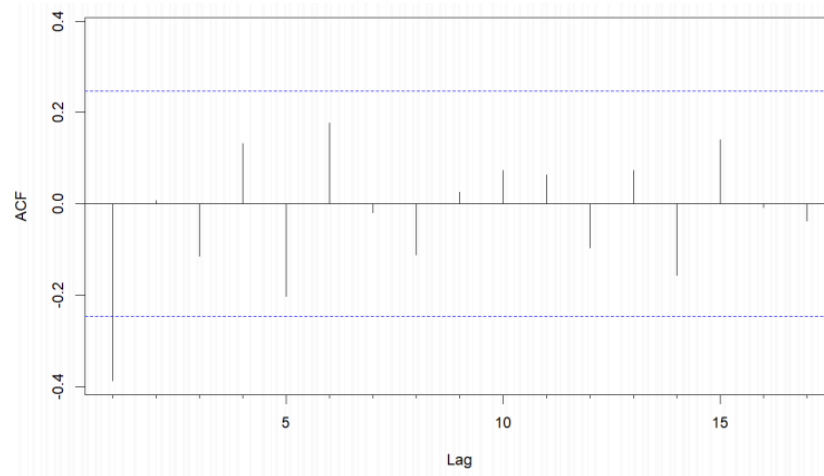
Figure 7. ACF plot for the second-order difference of ln(GDP)
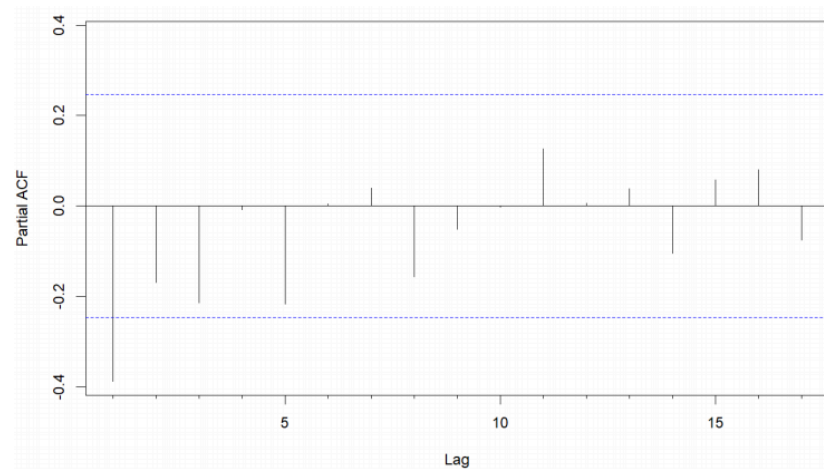
Figure 8. PACF plot for the second-order difference of ln(GDP)

Based on the analysis, the ARIMA(1,2,1) model seems to be an acceptable choice for modeling the transformed GDP. This model will be further validated during the parameter estimation and diagnostic checking stages.

The results of the estimation of the ARIMA(1,2,1) model are presented in Table 1. The AR(1) coefficient is not statistically significant at the level 5% (p=0.1364), which indicates a weak autoregressive effect. In contrast, the MA(1) coefficient is highly significant (p<0.001), suggesting a strong moving average component in the model. Despite the insignificance of AR(1), the overall model is still considered reasonable, as the residuals show minimal autocorrelation and the model achieves low forecast errors. Therefore, the ARIMA(1,2,1) model can be deemed a suitable choice for capturing the dynamics of the GDP time series.

Table 1. Parameter estimates of ARIMA(1,2,1) model

Variable	Estimate	Std. Error	z-value	p-value
AR(1)	0.2410	0.1619	1.4892	0.1364
MA(1)	-0.9074	0.0991	-9.1589	$< 2 \times 10^{-16}$

To determine the optimal ARIMA model for GDP forecasting, multiple candidates were evaluated using AIC, BIC, and RMSE. Table 2 summarizes the results.

Table 2: Evaluation of Various ARIMA Models

Table 2. Parameter estimates of ARIMA(1,2,1) model

Model	AIC	BIC	RMSE
ARIMA(1,2,1)	-118.64	-112.21	0.0878
ARIMA(0,2,1)	-118.56	-111.27	0.0895
ARIMA(0,2,2)	-118.35	-111.92	0.0881
ARIMA(2,2,1)	-116.67	-108.09	0.0878
ARIMA(1,2,2)	-116.65	-108.08	0.0878
ARIMA(2,2,2)	-115.51	-104.80	0.0881
ARIMA(2,2,0)	-110.15	-103.72	0.0947
ARIMA(1,2,0)	-109.91	-105.63	0.0964
ARIMA(0,2,0)	-101.62	-99.48	0.1048

Among all models, ARIMA(1,2,1) shows the lowest AIC, BIC, and RMSE, making it the most suitable for forecasting. Its balance of simplicity and performance justifies its selection.

The estimated ARIMA(1,2,1) model can be expressed as:

$$X_t = 0.2410X_{t-1} - 0.9074\varepsilon_{t-1} + \varepsilon_t \tag{7}$$

As suggested by the Box-Jenkins approach, diagnostic testing is a critical part of confirming the adequacy of the chosen model. Figure 9 provides some diagnostic plots for the residuals of an ARIMA(1,2,1) model, including the residuals vs fitted, ACF, and PACF of the residuals, residual time series plot, and the normal Q-Q plot. From these plots, the residuals appear randomly scattered around zero with no apparent pattern, indicating homoscedasticity. The ACF and PACF plots indicate that most autocorrelations fall within the confidence bounds, indicating no significant serial correlation. Additionally, the residuals are normally distributed, confirmed by the Jarque-Bera test with a p-value of 0.789 greater than 0.05, meaning we fail to reject the null hypothesis of normality. Overall, this implies that the model is well specified and suitable for forecasting.
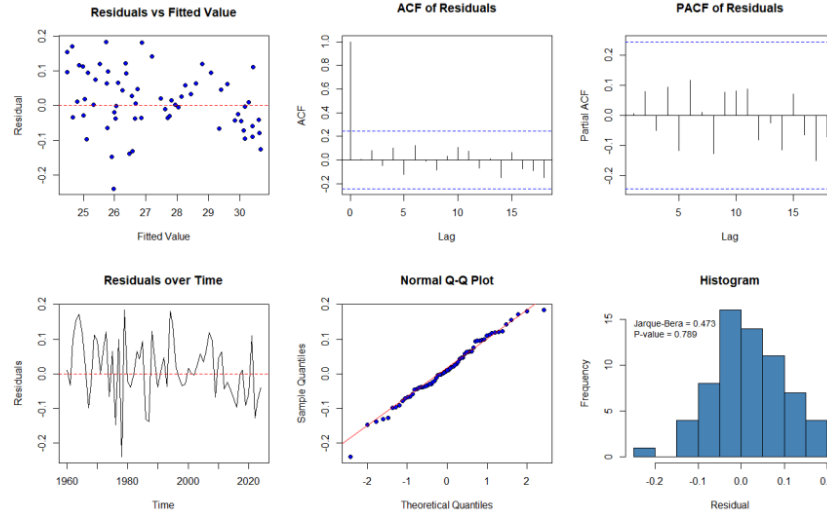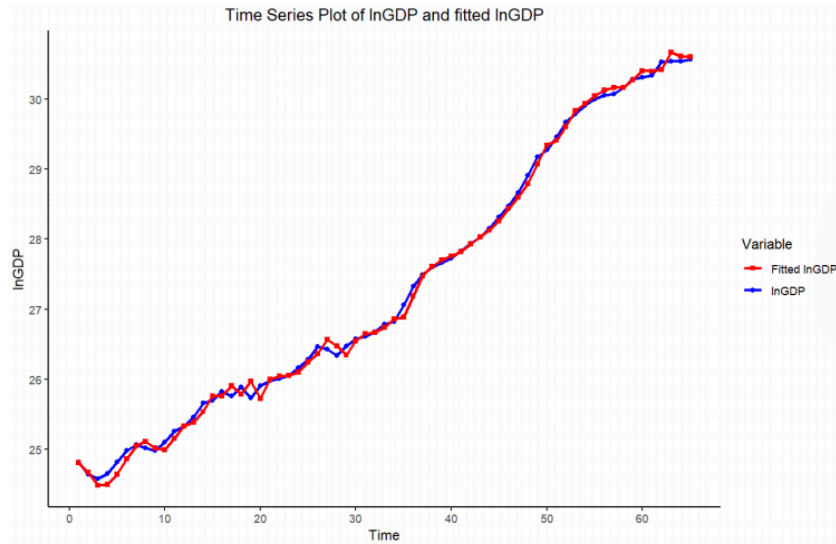

Figure 9. Diagnostic plots for ARIMA(1,2,1) model

Figure 10 illustrates the comparison of the actual GDP values vs. the fitted values from the ARIMA(1,2,1) model. The shapes of these two curves are very similar and close, indicating that the model fits the trend and patterns of China's GDP as intended. This evidence further supports the validity of the model.

**Figure 10.** Time series plot of actual and fitted lnGDP using ARIMA(1,2,1)

3.4. Out-of-sample forecasts

In the preceding section, we determined that the ARIMA (1,2,1) model was appropriate, and we now use it to predict China's future GDP over the next five years (2025–2029). We used the model equation (6) to predict outside the sample and a 95% confidence interval to help indicate uncertainty in the prediction. Table 3 presents our GDP forecasts for the period 2025–2029.

Figure 11 displays the time series plot of the actual ln(GDP) and the predicted values. The red lines indicate the project path and display the limits of the 95% confidence limits. The forecast suggests GDP is projected to continue to increase in the future, assuming the historical case continues. It is worth noting that these forecasts are model-based forecasts and thus bear the limitations of the ARIMA models. The economic decisions of the future might be affected by unpredictable policy, structural, or external shocks. Thus, while the model serves as a useful reference point, subsequent adjustments will be constantly made as real-world economic developments unfold.

Table 3. Forecasted values of China's GDP from 2025 to 2029

Year	lnGDP Forecast	GDP Forecast	95% CI Lower	95% CI Upper
2025	30.6313	2.0092×10^{13}	1.6822×10^{13}	2.3997×10^{13}
2026	30.7114	2.1766×10^{13}	1.6188×10^{13}	2.9266×10^{13}
2027	30.7940	2.3640×10^{13}	1.5861×10^{13}	3.5234×10^{13}
2028	30.8772	2.5691×10^{13}	1.5674×10^{13}	4.2110×10^{13}
2029	30.9605	2.7924×10^{13}	1.5552×10^{13}	5.0137×10^{13}
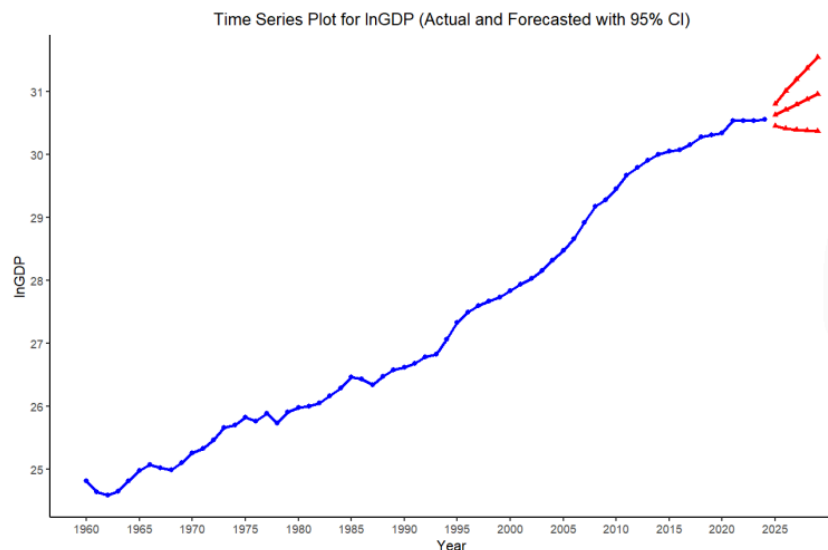
Figure 11. Time series plot for lnGDP (actual and forecasted with 95% CI)

4. Conclusion, limitations and future work

This study utilized the Box-Jenkins method for modeling and forecasting China's GDP using data from 1960 to 2024 on an annual basis. After conducting stationarity tests and differencing the non-stationary GDP series, the ARIMA(1,2,1) was chosen as the most appropriate model. The decision around the ARIMA specification was evaluated using values of AIC, BIC, and RMSE, which all pointed to the same specifications. Residual diagnostics also confirmed that all assumptions were validated, indicating no significant autocorrelation and assuming the residuals were normally distributed. The model was then used to forecast China's GDP from 2025 to 2029. There was a noted close alignment between the actual realized and fitted values, and the increase in GDP values in the forecast was clear, as China has sustained GDP growth. Overall, the statistical forecasting approach using the ARIMA(1,2,1) model appears to have captured the underlying dynamic of China's economic growth, thereby validating the results of this study and demonstrating that ARIMA models can be effectively utilized for short-term economic forecasts.

The ARIMA(1,2,1) model is a good fit for modelling China's GDP; nevertheless, the study has its constraints. The study only uses univariate time series data and does not consider macroeconomic variables or structural breaks. This makes the ARIMA models less effective in representing potential nonlinearities or abrupt economic changes. Future work could use multivariate models or higher-order forecasting models such as LSTM. Future work could also incorporate high-frequency data to help define better economic dynamics. However, the study offers a useful starting point for any future work in economic time series forecasting.

References

[1] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, Time Series Analysis: Forecasting and Control, 5th ed., Wiley, 2015.
[2] M. R. Abonazel and A. I. Abd-Elftah, "Forecasting Egyptian GDP Using ARIMA Models, " *Reports on Economics and Finance*, vol. 5, no. 1, pp. 35–47, 2019.
[3] F. Mungai, "Modeling and Forecasting Kenyan GDP Using Autoregressive Integrated Moving Average (ARIMA) Models, " *Science Journal of Applied Mathematics and Statistics*, 2016.
[4] B. Maity and B. Chatterjee, "Forecasting GDP Growth Rates of India: An Empirical Study, " *International Journal of Economics and Management Sciences*, vol. 1, pp. 52–58, 2012.
[5] M. N. A. Bhuiyan, K. S. Ahmed, and R. Jahan, "Study on Modeling and Forecasting of the GDP of Manufacturing Industries in Bangladesh, " *Chiang Mai University Journal of Social Science and Humanities*, vol. 2, pp. 143–157, 2008.
[6] A. Uwimana, B. Xiuchun, and Z. Shuguang, "Modeling and Forecasting Africa's GDP with Time Series Models, " *International Journal of Scientific and Research Publications*, vol. 8, pp. 41–46, 2018. https: //doi.org/10.29322/ijsrp.8.4.2018.p7608
[7] A. Ghazo, "Applying the ARIMA Model to the Process of Forecasting GDP and CPI in the Jordanian Economy, " *International Journal of Financial Research*, vol. 12, no. 3, pp. 70–77, 2021.
[8] R. W. Divisekara, G. J. M. S. R. Jayasinghe, and K. W. S. N. Kumari, "Forecasting the red lentils commodity market price using SARIMA models, " *SN Business & Economics*, vol. 1, no. 1, p. 20, 2020.
[9] P. Mondal, L. Shit, and S. Goswami, "Study of Effectiveness of Time Series Modeling (ARIMA) in Forecasting Stock Prices, " *International Journal of Computer Science, Engineering and Applications*, vol. 4, no. 2, pp. 13–21, 2014.

[10] S. Khan, "ARIMA Model for Accurate Time Series Stocks Forecasting, " *International Journal of Advanced Computer Science and Applications*, 2020.

[11] S. Siami-Namini and A. S. Namin, "Forecasting Economics and Financial Time Series: ARIMA vs. LSTM, " arXiv preprint, arXiv: 1803.06386, 2018.

[12] U. M. Sirisha, M. C. Belavagi, and G. Attigeri, "Profit Prediction Using ARIMA, SARIMA and LSTM Models in Time Series Forecasting: A Comparison, " IEEE Access, vol. 10, pp. 124715–124727, 2022.

[13] W. Ning, B. Kuan-jiang, and Y. Zhi-fa, "Analysis and Forecast of Shaanxi GDP Based on the ARIMA Model, " *Asian Agricultural Research*, vol. 2, pp.34–41, 2010.