Research on Predicting China's Wildlife Conservation Index Based on ARIMA Model

Wenxuan Xie *, Jiangyi Zeng

Faculty of Mechanical and Electrical Engineering, Southwest Petroleum University, Nanchong, China, 637001

* Corresponding Author Email: 13696001374@163.com

Abstract. In recent years, time series analysis has emerged as a vital tool for understanding complex trends and patterns in various fields, including wildlife conservation. Given the dynamic nature of ecological data, developing models capable of capturing both short-term fluctuations and long-term trends is essential for reliable forecasting. This study focuses on predicting the China Wildlife Conservation Society Index through a structured approach to time series modeling. After data collection and preprocessing steps—such as removing seasonality and trends—modeling began with assessments of stationarity using the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests. A first-order differencing operation was applied to make the series stationary, followed by model order selection based on AIC and BIC criteria. An optimal ARIMA model was then established and fitted, with residual diagnostics confirming model adequacy. Using this ARIMA model, predictions were made for the subsequent five-time points, with a calculated 95% confidence interval. The prediction results, well-aligned with the confidence interval, demonstrate the model's reliability. This approach supports strategic decision-making in wildlife conservation by offering accurate forecasting capabilities for key ecological indicators.

Keywords: Time Series Analysis, Wildlife Conservation, Stationarity, ARIMA Model.

1. Introduction

As global environmental issues become increasingly severe, biodiversity conservation has gained significant attention from the international community [1, 2]. China, as one of the most biodiverse countries, plays a critical role in the conservation and management of its wildlife resources [3, 4]. However, questions remain regarding the effectiveness of current conservation measures, how to assess their impact, and how to design more effective conservation strategies for the future. The China Wildlife Conservation Index (CPI) serves as a valuable metric for evaluating the effectiveness of China's biodiversity conservation efforts [5], providing essential data for policymakers and researchers. Existing studies have widely used the CPI as a reference for understanding trends in biodiversity protection, yet gaps remain in establishing reliable predictive models that can inform future strategies.

Time series analysis, known for its ability to reveal data trends, seasonality, and variability, has been widely applied in environmental and ecological studies [6]. This study utilizes time series analysis techniques, particularly the ARIMA model, to analyze and predict CPI data. By employing preprocessing steps such as detrending and differencing for stationarity, and combining model diagnostics and residual analysis, this research aims to establish a reliable model for CPI prediction. Such a model is intended to provide a scientific basis for policymakers, supporting informed decisions in wildlife conservation policy.

While time series analysis has broad applications in biodiversity conservation, many existing studies fall short in performing rigorous stationarity testing, a crucial step to ensure prediction accuracy and model reliability. Addressing this gap, this study implements a comprehensive assessment of stationarity using unit root tests, including the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests, to confirm that the CPI data achieves the necessary stability for robust modeling. By establishing a reliable time series framework, this research enhances predictive precision for future CPI values, thereby providing a scientifically validated tool for evaluating conservation effectiveness. The resulting model offers valuable theoretical support for policymakers and researchers, enabling data-driven planning and decision-making to improve biodiversity conservation strategies across China. Through these efforts, this study not only strengthens the analytical foundation for conservation assessments but also contributes to developing more adaptive and forward-looking conservation policies.

2. Time Series Analysis for Wildlife Conservation

In recent years, time series analysis has become an essential tool for understanding complex trends and patterns in various fields, including wildlife conservation [7, 8]. Given the dynamic nature of ecological data, it is crucial to develop models that can capture both short-term fluctuations and long-term trends to provide reliable forecasts [9]. This study focuses on predicting the China Wildlife Conservation Society index using a structured approach to time series modeling. By carefully preparing the data, testing for stationarity, and optimizing the model, this study aim to generate accurate predictions that can support strategic decision-making in wildlife conservation.

Before making time series predictions, past data was collected and preprocessed by removing seasonality, trend, and other components. The prepared data was then ready for model fitting and subsequent prediction [10]. As shown in Figure 1:

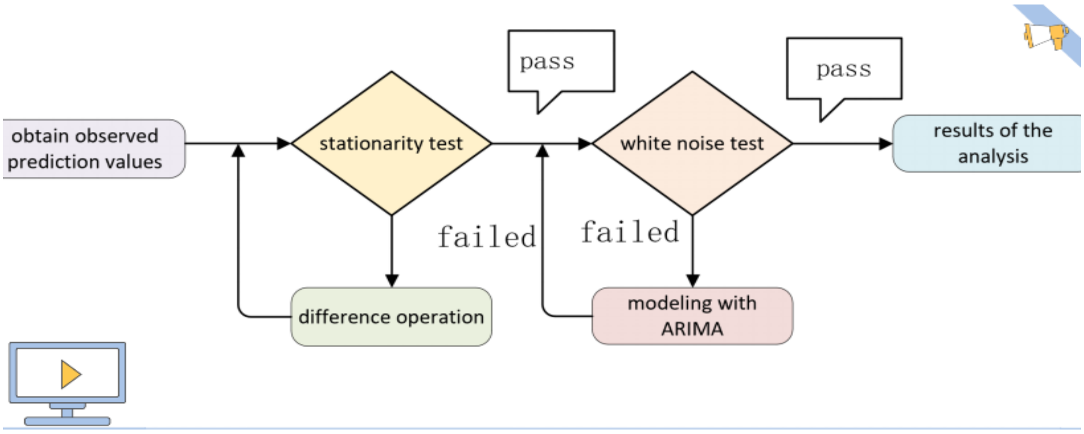

Figure 1. The flowchart of time series model construction

Firstly, the data of the China Wildlife Conservation Society index was loaded and visually displayed. As shown in the figure 2:

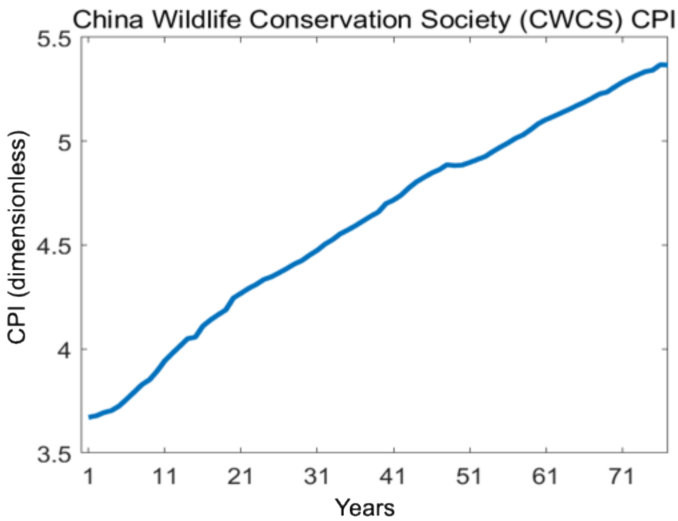

Figure 2. China Wildlife Conservation Society CPI.

To assess stationarity, unit root tests were conducted on the original time series data using the Augmented Dickey-Fuller (ADF) test and the Phillips-Perron (PP) test. The results showed a test statistic (h-value) of 0 and a p-value of 0.990, indicating non-stationarity in the series. Consequently, a first-order differencing operation was applied to make the series stationary.

A first-order differencing operation was applied to the original time series data to obtain the first-order differenced series. Following this, a unit root test was conducted on the differenced series to assess stationarity. Figure 3 illustrates the results of the differencing operation.

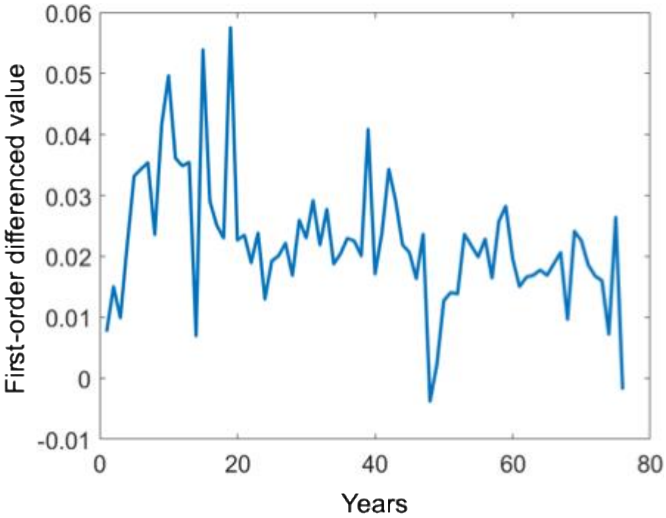

Figure 3. Differential arithmetic results

The original time series data underwent unit root tests with the ADF and PP tests to determine stationarity, and the results showed an h-value of 1 with a p-value of 0.244, suggesting that the series is stationary.

The order of the model was determined using AIC and BIC criteria by iteratively testing different AR and MA orders and recording the corresponding AIC and BIC values, which were then visualized, as shown in the figure 4.

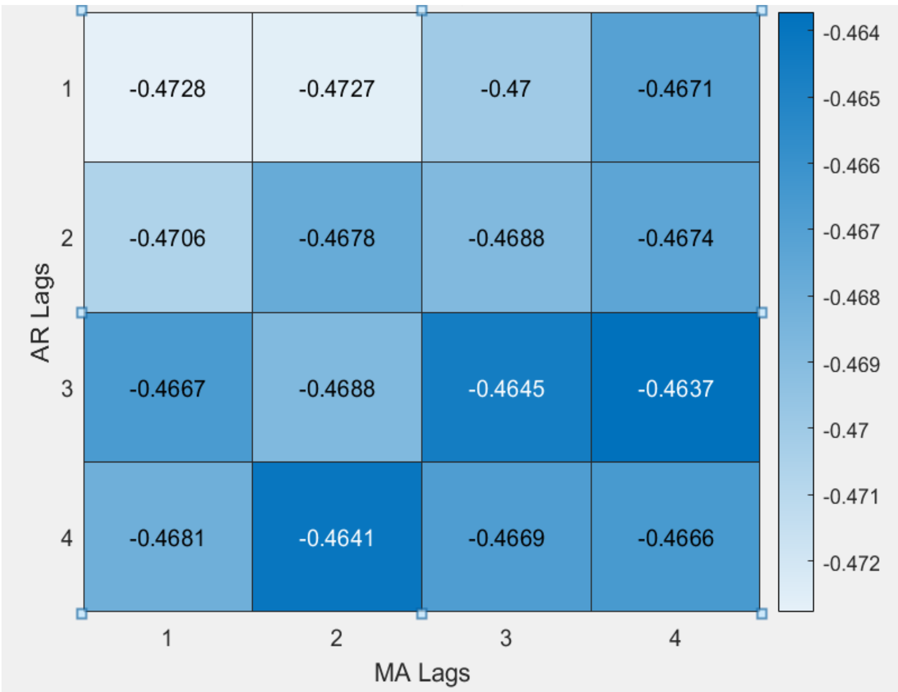

Figure 4. BIC guideline diagram

To establish an optimal ARIMA model, the model parameters were carefully selected, and the model was fitted using the estimation function. Following model fitting, residual diagnostics were conducted, including the Ljung-Box Q test on the residuals, standardization of residuals, and examination through QQ plots, autocorrelation plots, and partial autocorrelation plots. As shown in

Figure 5. These diagnostics, shown in the figure below, helped assess the adequacy of the model and ensured that the residuals met the assumptions required for a well-fitted ARIMA model.

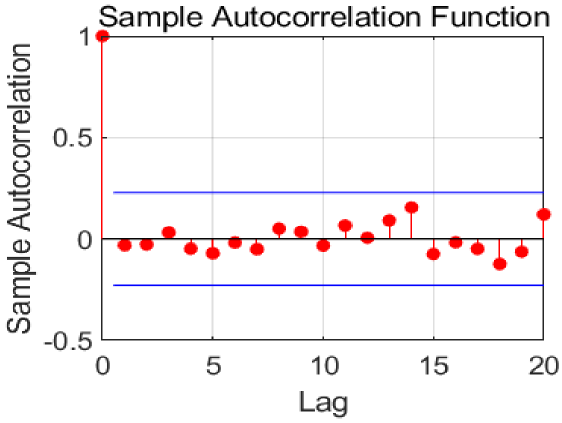

Figure 5. Residual distribution diagram

Using the established ARIMA model, predictions were made for the next five time points, along with the calculation of the 95% confidence interval's upper and lower limits. The prediction results were then visualized alongside the original data for comparison. As shown in Figure 6.

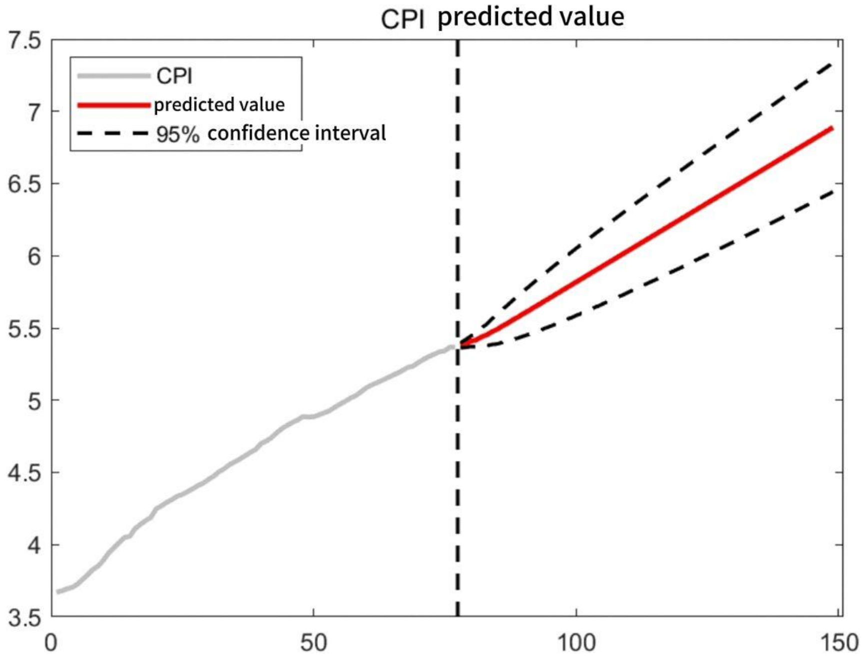

Figure 6. CPI predicted value

The prediction results falling within the confidence interval indicate that they are quite reliable.

3. Conclusion

This study successfully developed a time series prediction model for the China Wildlife Conservation Society index by employing systematic preprocessing, model selection, and diagnostic evaluations. Initial stationarity testing using the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests revealed non-stationarity in the original data, prompting a first-order differencing transformation. Following the differencing operation, stationarity was confirmed, allowing for optimal ARIMA model selection through AIC and BIC criteria. Residual diagnostics, including the Ljung-Box Q test, standardization of residuals, and visual inspections via QQ plots, autocorrelation, and partial autocorrelation plots, validated the model's adequacy and fit. Using the finalized ARIMA model, predictions were generated, and the forecasted values demonstrated a reliable fit within the calculated 95% confidence intervals. This comprehensive approach to time series modeling provides a robust framework for understanding and forecasting conservation indices, offering valuable insights for strategic planning and resource allocation in wildlife conservation efforts.

Future research could enhance this model by incorporating environmental variables (e.g., temperature, precipitation) through multivariate time series or machine learning approaches, such as VAR or LSTM, to capture complex relationships influencing CPI trends. Additionally, integrating spatial-temporal analysis could create a geographically sensitive conservation index, supporting more localized and adaptive wildlife conservation policies.

References

[1] CLÉMENÇON R. Is sustainable development bad for global biodiversity conservation? [J]. Global sustainability, 2021, 4: e16.

[2] RANDS M R, ADAMS W M, BENNUN L, et al. Biodiversity conservation: challenges beyond 2010 [J]. science, 2010, 329 (5997): 1298 - 1303.

[3] MI X, FENG G, HU Y, et al. The global significance of biodiversity science in China: an overview [J]. National Science Review, 2021, 8 (7): nwab32.

[4] ZHANG L, LUO Z, MALLON D, et al. Biodiversity conservation status in China's growing protected areas [J]. Biological Conservation, 2017, 210: 89 - 100.

[5] WANI Z A, BHAT J A, NEGI V S, et al. Conservation Priority Index of species, communities, and habitats for biodiversity conservation and their management planning: A case study in Gulmarg Wildlife Sanctuary, Kashmir Himalaya [J]. Frontiers in Forests and Global Change, 2022, 5: 995427.

[6] SHAUKAT K, ALAM T M, LUO S, et al. A review of time-series anomaly detection techniques: A step to future perspectives [C]//. Advances in Information and Communication: Proceedings of the 2021 Future of Information and Communication Conference (FICC), Volume 1: Springer, 2021: 865 - 877.

[7] CHOI K, YI J, PARK C, et al. Deep learning for anomaly detection in time-series data: Review, analysis, and guidelines [J]. IEEE Access, 2021,9: 120043 - 120065.

[8] LIANG Y, WEN H, NIE Y, et al. Foundation models for time series analysis: A tutorial and survey[C]//. Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2024: 6555 - 6565.

[9] MA Y, HU Y, MONCRIEFF G R, et al. Forecasting vegetation dynamics in an open ecosystem by integrating deep learning and environmental variables [J]. International Journal of Applied Earth Observation and Geoinformation, 2022,114: 103060.

[10] ATWAN T A. Time Series Analysis with Python Cookbook: Practical recipes for exploratory data analysis, data preparation, forecasting, and model evaluation [M]. Packt Publishing Ltd, 2022.