JSR Publications

Available Online

J. Sci. Res. **17** (3), 755-766 (2025)

JOURNAL OF SCIENTIFIC RESEARCH

banglajol.info/index.php/JSR

Forecasting Carbon Dioxide Emissions in Sundarban Islands, India: An ARIMA Model

K. Mondal[1*], A. K. Lal[2], U. S. Singh[3]

[1]University Department of Mathematics, Ranchi University, Jharkhand, India

[2]Department of Mathematics, Simdega College, Simdega, Jharkhand, India

[3]Indian Forest Service (Rtd), Lucknow, Uttar Pradesh, India

Received 11 December 2024, accepted in final revised form 24 May 2025

Abstract

The Sundarban Islands have suffered greatly due to climate change brought on by greenhouse gas emissions. Carbon dioxide (CO_2) is one of the leading emitters that increases the temperature in the atmosphere. Sundarban is one of the largest biospheres in the world facing challenges due to climate change. CO_2 plays a crucial role in climate change and that's why the study on the gas emission in Sundarban is a vital concern. In this article, the CO_2 emissions in Indian Sundarban are studied and analyzed from 1980 to 2008 by using ARIMA model. The stationarity test reveals that the nature of the data is I (1) and our study prescribed that the ARIMA (1, 1, 1) is suitable to capture the underlying trend of the emission data. Model's accuracy tested by the key metrices like Akaike Information Criterion (AIC), RMSE, MAE etc. The proposed model is also utilized for forecasting and its reliable result suggests that the model may be useful in environmental analysis. The findings demonstrate that the CO_2 emissions in Sundarban are rising rapidly and putting the region at risk for climatic threats.

Keywords: Sundarban Island; Climate change; ARIMA model; Carbon dioxide; Stationary; Greenhouse gas.

doi: https://dx.doi.org/10.3329/jsr.v17i3.78318 J. Sci. Res. **17** (3), 755-766 (2025)

1. Introduction

Greenhouse gas (GHG) emissions have badly impacted the Sundarban islands. The effects of climate change are a significant threat to the environment of this unique ecosystem. The GHGs, including CO_2, CH_4 (methane) and N_2O (nitrous oxide) soak up heat within the Earth's atmosphere and finally, it is leading to global warming. Due to the rising temperature, there is an alteration in precipitation patterns, frequency of severe weather incidents like floods, cyclones, storms etc. and rising sea level in Sundarban. Which have a detrimental impact on the largest ecosystem. The polar ice is melting for the rising of global mean temperature and causing the expansion of ocean water. This results in the elevation of sea levels and submerges the low-lying island. Ghoramara Island is one of them situated

[*] *Corresponding author:* kartickmathbhu@gmail.com

in West Bengal, India. The most contributed GHG released into the environment through the deforestation and burning of fossil fuels is CO_2 and it is responsible for the rising of sea-level. As a consequence of rising sea-levels, the Sundarbans are experiencing increased salinity in the water and soil. Ultimately that contributes to habitat loss for mangrove species. Similarly, the survival of a variety of plants and animal species are becoming under threat. A significant portion of the gas is absorbed by the ocean water when the CO_2 level is increased in the atmosphere and it leads to ocean acidification. It negatively impacted marine life. The base of the food chain (plankton and fish etc.) in the Sundarban are affected. In a word, the entire ecosystem of Sundarban is affected due to uncontrolled GHG emissions. Since CO_2 is linked to both carbon sequestration in mangroves and deforestation, it is the most important GHG associated with the Sundarbans. Compared to other wetland ecosystems, CH_4 is less prevalent but still plays a part because of wetland conditions. Although their direct emissions from Sundarbans is negligible, other GHGs such as N_2O and other fluorinated gases contribute to the region's overall climate change impacts. In this crucial situation, reducing greenhouse gas emissions through mitigation efforts is essential. Strategic policy formation and its implementation is necessary for protecting the wetlands and forests to safeguard the Sundarbans.

A significant number of research work using ARIMA model was done to predict CO_2 emissions in India. For example, Nyoni and Bonga [1] observed that the emission trend is increasing and Singh and Kumari [2] suggested that the ARIMA model is one of the most appropriate tools for this study and used ARIMA model to forecast emissions in India. Few researchers predicted CO_2 emissions in Bangladesh using Auto-ARIMA (Autoregressive Integrated Moving Average) models. Rahaman and Hasan [3] tried to fit an ARIMA model to predict the emissions during 2016-2018. According to him ARIMA (0, 2, 1) is the best fitting model and it was used for prediction. After considering various statistical metrics Hossain *et al.* [4] also used ARIMA models. Fatima *et al.* [5] and Abhilash *et al.* [6] worked on CO_2 emissions and air pollution in Bangladesh and other Asian countries by utilizing Simple Exponential Smoothing models and ARIMA models. All the studies recommended strategic planning and continuous monitoring for air quality management. Some researchers have employed forecasting economic indicators by using ARIMA models in East African countries. ARIMA models were applied by Wabomba [7] and Awel [8] for forecasting the GDP of Kenya and Ethiopia respectively. Where Wabomba suggested that ARIMA (2, 2, 2) is the most effective model. Similar work was done by Gikungu [9] using the Seasonal ARIMA (SARIMA) model to predict inflation in Kenya. In 2022, Muma presented a comprehensive study on GDP modelling by applying ARIMA models. Author emphasized the importance of model identification, diagnosis, estimation and forecasting. All the studies of this region collectively highlighted the potentiality of the ARIMA modelling in the future prediction of economic indicators. In 2022, Deepa and Vijayasree [10] worked on forecasting carbon dioxide equivalent emission in different key sectors of India utilizing ARIMA model. Similarly, Jamii *et al.* [11] followed the Box-Jenkins ARIMA approach for modelling and forecasting CO_2 emissions in Morocco. Where authors were specifically focused on long-term future prediction and observed a significant increasing trend in the emissions. The studies highlighted valuable insights regarding the future CO_2 emissions patterns in Morocco. The same approach was followed by Kumar and Jain [12] to study the prediction of air pollutants in urban Delhi, India. Few researchers applied this model in economic sectors. For example, Muma and Karoki [13] applied in GDP modelling and they recommended that the ARIMA (2, 2, 2) can be applied in studies 182 days treasury bills and ARIMA (3, 1, 1) for 91 days respectively. A sufficient number of research works continued on greenhouse gas emissions (mainly CO_2 or its equivalent emission) in India and abroad. In India, some researchers worked on state-wise emissions [14, 15, 16] of the gas using the least square method but very few works were done on the Indian Sundarban region in this particular direction of research. Sundarban is a UNESCO world heritage site and has international importance but this site is facing various challenges and threats now due to global warming and climate change. Though the Sundarbans are a local hotspot for climate sensitivity but the global trends in CO_2 emissions from industrialized countries are directly responsible for the crisis of the region. Both local conservation policies like mangrove restoration, climate adaptation regulations etc. and global emission reductions through international climate agreements like the Paris Agreement are necessary to protect this special environment. Being CO_2 emissions are one of the most responsible factors for this, it is important to study the gas emissions in the insight of a mathematical point of view. This study is an attempt to model carbon dioxide gas emissions so that some suitable policies and important measures may be taken in advance.

This paper explores how ARIMA, a modern time series modelling technique, can be used to analyze and forecast CO_2 emissions. With environmental concerns on the rise, understanding past emission patterns and predicting future trends is crucial. ARIMA (1,1,1) is an efficient tool for modelling complex time series. We apply this model to forecast CO_2 emissions and evaluate its accuracy in predicting future trends. So, in brief, our study finds a model and forecasts the gas emission in Indian Sundarban using ARIMA models.

2. Materials and Methods

2.1. *ARIMA models*

ARIMA represents Autoregressive Integrated Moving Average. It is employed to forecast future series points in time series studies. The three components of ARIMA are moving average (MA), integrated (I) and autoregressive (AR). For the evolving variable, the AR component means that is regressed on its own lag values. According to the MA component, the regression error is a linear sum of errors that happened at various points in the past. The I component shows that the difference between the data values and previous values has been used in place of the data values.

If p is the number of auto regressive orders, d is the order of the differentiation applied to the time series and q is the number of the moving average orders in the data series then ARIMA models are denoted by ARIMA (p, d, q). The non-negative numbers p, d and q are the parameters [1,17].

ARIMA (p, d, q) can be expressed as:

$\bar{z}_t = \emptyset_1\bar{z}_{t-1} + \emptyset_2\bar{z}_{t-2} + \cdots + \emptyset_p\bar{z}_{t-p} + a_t - \theta_1\bar{z}_{t-1} - \theta_2\bar{z}_{t-2} - \cdots \theta_q\bar{z}_{t-q}$

Where $\bar{z}_t = z_t - \mu$ and a_t is shock.

Once the backward shift operator B has been determined, equation (1) [17,18] can be used as follows:

$$\emptyset(B)(1-B)^d z_t = \theta(B)a_t \tag{1}$$

2.2. *Data collection*

Historical data on CO_2 emissions from the year of 1980 to 2008 was collected from the article 'Carbon Dioxide Emissions of Indian States: An Update' authored by Tapas Ghoshal, Bureau of Applied Economics and Statistics, Government of West Bengal, India and Ranajoy Bhattacharyya, Indian Institute of Foreign Trade, Kolkata, India [19].

2.3. *Methodology*

In this study, the used data set containing annual CO_2 emissions data from the year of 1980 to 2008 having two columns 'CO_2 emissions' and 'Year.' The data was loaded from an Excel file, and the year was properly formatted to fit the time series requirements. Utilizing Python, the pandas library first imported the data set and visualized by using matplotlib function treating as 'CO_2 Emission' as dependent variable (Y) and 'Year' as independent (X) variable. For stationarity checking we used Augmented Dickey-Fuller (ADF) test, Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test, autocorrelation (ACF), partial autocorrelation (PACF) function from the statsmodel library over the dependent variable 'CO_2 emission'. Diagnosis test indicates that the emission data was non-stationary and it becomes stationary after first difference I (1). Now to get the value of the parameters p and q, numpy and statsmodel library was used from where the ARIMA imported. Error metrics were imported from the library sklearn. Analysis shows the lower Akaike Information Criterion (AIC) value as p, q equals to one. So ARIMA (1, 1, 1) model selected. For examining the model performance, the data was divided into two parts: training data (80 %) to develop the model, and test data (20 %) to evaluate its performance. Train Test analysis assures the model performance.

3. Diagnostic Test and Moled Selection

3.1. *Stationarity test*

A stationary time series is defined as one whose statistical properties such as variance and mean remain unchanged over time. Simply put, the data should look consistent throughout its duration. Stationarity is crucial because the forecasting model (ARIMA model) works best with stationary data. Using both the KPSS and ADF tests allows us to establish the non-stationarity of the CO_2 emissions data. While the ADF test indicates non-stationarity, the KPSS test verifies it, allowing for a more solid conclusion before applying the ARIMA model. Using both, it can be confirmed the nature of the data and ensure that suitable preprocessing processes such as differencing are implemented.

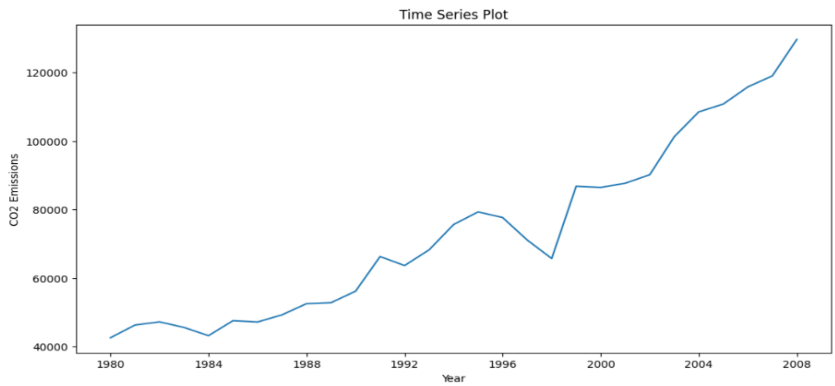

Fig. 1. Plot of CO_2 emission trend over time.

The graphical presentation depicted in Fig. 1 shows the CO_2 emission trend. The plot describes that the curve is trending upwards and implies that the variable is not stationary as mean and variance is changing over time. The plot of Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) of emission data (treated as Y) is depicted in Figs. 2a,b.

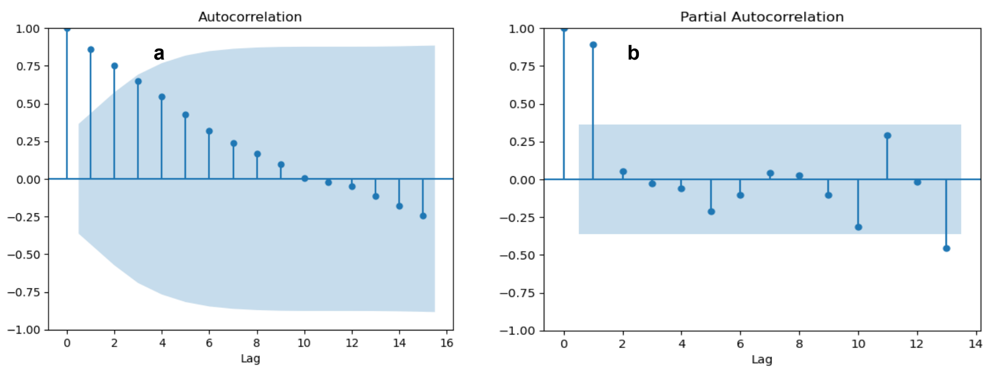

Fig. 2. (a) ACF plot of emission data (Y); (b) PACF plot of emission data (Y).

The ADF and KPSS test statistics are presented in the following Tables 1 and 2 respectively.

Table 1. Augmented Dicky-Fuller test statistic.

Variable	ADF Statistic	P value	Critical Values	Confidence Interval (%)	Conclusion
Y	1.005265	0.994321	-2.625296	90	Not Stationary
			-2.971989	95	Not Stationary
			-3.688926	99	Not Stationary

A negative valued ADF statistics with the p-value below 0.05 generally suggest that the data seems to be stationary. The CO_2 emissions data in this analysis appear to be non-stationary, as indicated here the p-value is above 0.05 and compared to critical values the ADF statistic is less negative.

Table 2. KPSS test statistic.

Variable	KPSS Statistic	P value	Critical Values	Confidence Interval (%)	Conclusion
Y	0.798360	0.010000	0.347000	90	Not Stationary
			0.463000	95	Not Stationary
			0.739000	99	Not Stationary

The data is most likely non-stationary as the p-values is below 0.05 and KPSS statistic is small. The series is non-stationary according to our study since the KPSS statistic is more than the critical values and p-value is smaller than 0.05.

As per the ACF and PACF plot in Figs. 2a,b and ADF test and KPSS test statistics (Tables 1 and 2) the given data is not stationary and therefore we need to move forward to the first difference.

3.2. ***First difference I(1):*** The first difference (d=1) of the non-stationary emission data and its plot is deployed in Fig. 3.

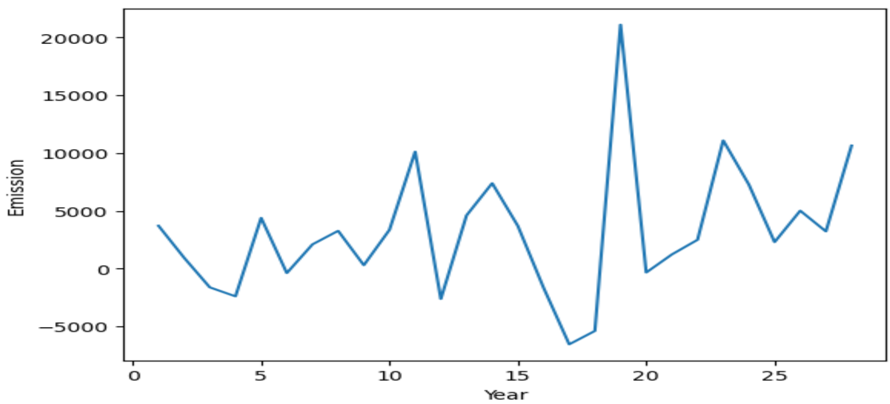

Fig. 3. Transforms: difference (1).

To check the stationarity of the CO_2 emission data after first difference it is necessary to check the ACF and PACF plot and ADF and KPSS test statistic. ACF and PACF plots are presented in Figs. 4a,b respectively.

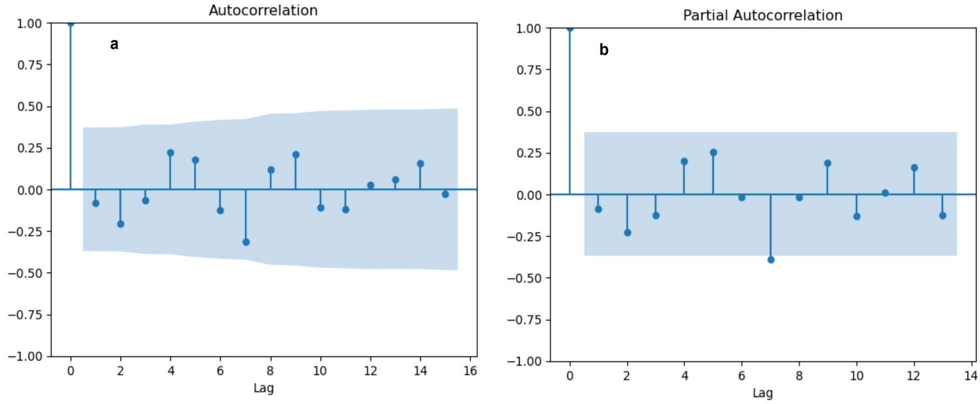

Fig. 4. (a) ACF plot of I (1); (b) PACF plot of I (1).

ADF and KPSS test statistics of the first differenced emission data are presented in the following Tables 3 and 4, respectively.

Table 3. Augmented Dicky-Fuller test statistic.

Variable	ADF Statistic	P value	Critical Values	Confidence Interval (%)	Conclusion
Y	-5.271492	0.000006	-2.627601	90	Stationary
			-2.976430	95	Stationary
			-3.699608	99	Stationary

Table 4. KPSS test statistic.

Variable	KPSS Statistic	P value	Critical Values	Confidence Interval (%)	Conclusion
Y	0.341647	0.100000	0.347000	90	Stationary
			0.463000	95	Stationary
			0.739000	99	Stationary

Figs. 3, 4a,b and Tables 3, 4 assure that the emissions data are stationary after first difference and therefore is I (1) variable.

3.3. *Evaluation of ARIMA models (p, d, q)*

The following models were fitted with various combinations of p, d =1 and q and evaluated them by utilizing BIC (Bayesian information Criterion) and AIC (Akaike Information Criterion). Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) are calculated as error metrics and all these are presented in Table 5.

Table 5. Parameters selection of ARIMA model.

Models	AIC	BIC	RMSE	MAE
ARIMA (1,1,1)	**569.087811**	573.084424	9637.422964	5300.884306
ARIMA (2,1,1)	571.357780	576.686598	9639.819401	5348.223541

ARIMA (0,1,0)	571.725640	573.057844	10060.663243	5920.241138
ARIMA (2,1,2)	571.831407	578.492430	9563.913172	5147.735496
ARIMA (1,1,2)	571.923770	577.252589	9638.400896	5311.284645
ARIMA (1,1,0)	572.777207	575.441616	9996.376817	5774.985273
ARIMA (0,1,1)	573.095207	575.759616	9999.496130	5807.368089
ARIMA (2,1,0)	574.798663	578.795277	9994.231462	5732.342554
ARIMA (0,1,2)	575.971753	579.968367	9995.904577	5733.600903

A lower AIC (and BIC) value indicates better fitting model. Considering the lower value of RMSE and MAE our diagnosis test suggests that the ARIMA (1, 1, 1) model is suitable to represent the CO_2 emissions.

3.4. *Training and test data analysis*

The time series was split up into training and testing datasets, with the train data being used to fit the model and the test data being used to assess the forecast performance. Utilised the training dataset to apply the ARIMA (1,1,1) model and to predict the upcoming observations in the test set, we applied the model. The plot presented in Fig. 5 highlights the train and test data sets with the prediction curve.

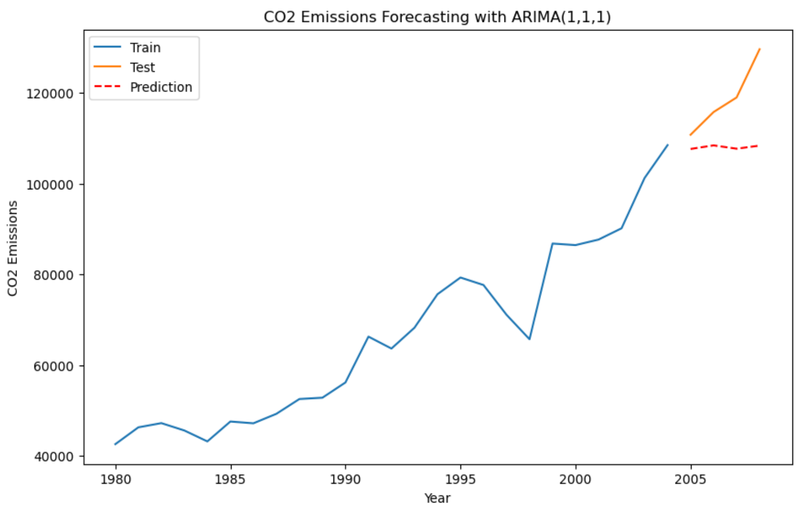

Fig. 5. Plot of training and test data set.

The model summary presents that ARIMA (1,1,1) fitted over training data with the AIC value of 493.464 and BIC value of 496.999 respectively. Test performance metric RMSE equals to 12687.00 and it suggests that the proposed model describes well to the unseen emissions data. The outcomes validate the model's precision in predicting CO_2 emissions.

4. Results and Discussion

Present article is an attempt to study the CO_2 emissions pattern in Indian Sundarban by utilizing ARIMA approach. The proposed model well represents the real situation regarding the emissions in the region. Our diagnostic test suggests that the ARIMA (1, 1 ,1) model is suitable for the time series data, and its performance was evaluated on the test dataset. The model showed a good fit for the training data. The chosen model configuration was appropriate for the emissions data since the differencing component (d = 1) removed any non-stationarity and the AR and MA terms (with p =1, q = 1) effectively captured the auto correlations present in the series. Key metrics such as the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were relatively low, indicating that the model was well-suited to the data. Visual inspection of forecasted emissions versus actual emissions confirmed that the model is able to predict the movement of the time series in major cases, especially for soother periods in the emissions data. A plot of the forecasted values vs. the actual test values presented in Fig. 5 shows that the ARIMA (1, 1, 1) model successfully captured the trend and dynamics of the time series. The forecasted values closely follow the actual test values with minor deviations, indicating strong predictive power.

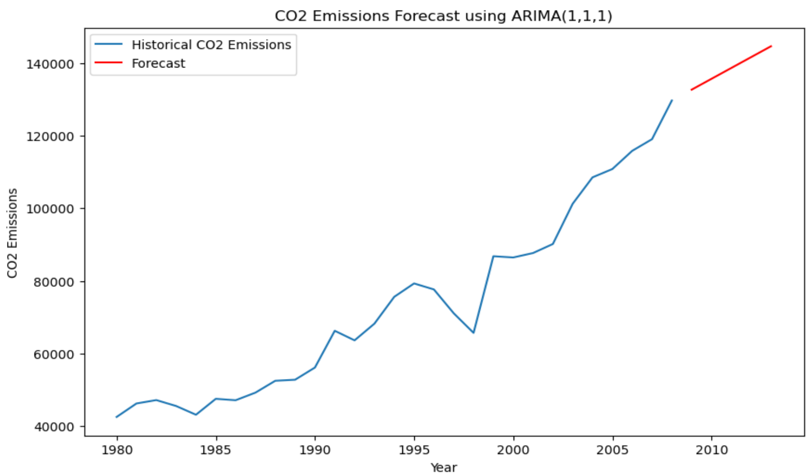

Fig. 6. Forecast graph of CO_2 emissions.

Attempt was made to fit the model over the entire dataset and predicted CO_2 emissions from 2009 -2013. The numeric values of the emissions data and the diagram deployed in Fig. 6, clearly highlight the fact that there is an increasing trend of emissions observed in Sundarban Island, India. The predicted emission data is presented in Table 6.

Table 6. Forecasting CO_2 emission.

Year	2009	2010	2011	2012	2013
Emission	132642.33	135628.96	138615.56	141602.11	144588.62

5. Limitations of the Study

There are some limitations of the study. Such as:

- *Data Constraints:* The study is based on historical CO_2 emissions data from 1980 to 2008. More recent data could improve model accuracy and relevance to current environmental conditions.
- *Simplified Model Assumptions:* The ARIMA model captures the trend and seasonality but does not incorporate exogenous variables such as economic activities, industrial expansion, or regulatory changes that may influence CO_2 emissions.
- *Limited Forecasting Horizon:* The model forecasts emissions up to 2013, and extending the predictions further may reduce accuracy due to evolving environmental policies and unforeseen global climate events.
- *Regional Focus:* The study is specific to the Sundarbans, and findings may not be directly generalizable to other ecological regions with different climate and industrial conditions.
- *Unaccounted Feedback Mechanisms:* The study does not consider feedback loops such as the impact of increased emissions on mangrove health and subsequent effects on CO_2 absorption capacity.

6. Conclusion

This study attempts to model the CO_2 emissions data in Indian Sundarban utilizing the ARIMA approach. Findings of the study suggest that ARIMA (1, 1, 1) is suitable for capturing the underlying trend in the emission data and it provides accurate predictions over the test period. In our analysis, physical relevancy and robustness of the model is reliable and it is applied as a tool for forecasting CO_2 emissions. Due to the minimality of the forecasting errors the model highlights its effectiveness in presenting the underlying structure of the emission data series. The low RMSE value on the test set shows that the model can generate accurate forecasts. These measures prove that there was no overfitting and the model successfully identified the underlying patterns in the datasets. For real-world applications (such as financial forecasting and inventory management) where precise forecasting is required to support decision-making, this level of precision is essential. The forecasting periods of 5 years from 2009 to 2013 clearly highlight that CO_2 emissions in Indian Sundarban are gradually increasing and as an immediate effect of this, the island will face more challenges of climatic hazards and global warming. Sustained increases in carbon dioxide (CO_2) emissions in the Sundarbans pose serious risks to public health, the ecosystem, and the climate. It is imperative to take the necessary measures to lessen these negative impacts.

Some actionable recommendations for policymakers based on the forecasted trends are as follows:

- *Strengthen Carbon Monitoring Systems*: Establish continuous CO_2 monitoring stations in the Sundarbans to track emissions in real-time and improve predictive modeling.

- *Enhance Mangrove Conservation*: Since mangroves act as natural carbon sinks, policies should prioritize afforestation, restoration, and protection against deforestation.
- *Promote Renewable Energy*: Encourage solar and wind energy adoption in local communities to reduce dependence on fossil fuels, which contribute to rising emissions.
- *Implement Sustainable Development Policies*: Introduce regulations limiting industrial expansion in ecologically sensitive areas and promote eco-friendly tourism practices.
- *Cross-Border Climate Cooperation*: Since the Sundarbans extend into Bangladesh, policymakers should collaborate on transboundary environmental strategies to address CO_2 emissions collectively.
- *Encourage Community Participation*: Involve local communities in conservation efforts through incentives for sustainable livelihoods such as eco-tourism, afforestation projects, and alternative fuel sources.
- *Leverage Predictive Analytics*: Use ARIMA-based forecasting to design adaptive policies that preemptively mitigate rising CO_2 emissions rather than reactively addressing climate consequences.

7. Future Scope

There is ample scope to utilize ARIMA models to characterize greenhouse gases in other demographic areas and it may be use for understanding emissions pattern of other GHGs. Further study is needed for understanding the trend of long-term emissions. By analysing different ARIMA configurations, the model might be further refined. However, future analyses could consider more advanced models or incorporate additional exogenous variables to further enhance forecasting accuracy, especially in more complex datasets.

Acknowledgment

We are thankful to Damodar Prasad Goswami for his valuable suggestions and constant support for this research work.

References

1. T. Nyoni and W. G. Bonga, DRJ-J. Econ. Finance **4**, 1 (2019).
2. S. K. Singh and S. Kumari, Preprint (2022). https://doi.org/10.21203/rs.3.rs-1265771/v1
3. A. Rahman and M. M. Hasan, Open J. Stat. **7**, 560 (2017). https://doi.org/10.4236/ojs.2017.74038
4. A. Hossain, M. A. Islam, M. Kamruzzaman, M. A. Khalek, and M. A. Ali, Int. J. Stat. Sci. **16**, 33 (2017).
5. S. Fatima, S. S. Ali, S. S. Zia, E. Hussain, T. R. Fraz, and M. S. Khan, Int. J. Econ. Environ. Geol. **10**, 64 (2019). https://doi.org/10.46660/ojs.v10i1.219
6. M. S. Abhilash, A. Thakur, D. Gupta, and B. Sreevidya, Adv. Intell. Syst. Comput. **696**, 413 (2018). https://doi.org/10.1007/978-981-10-7386-1_36
7. M. S. Wabomba, Sci. J. Appl. Math. Stat. **4**, 64 (2016). https://doi.org/10.11648/j.sjams.20160402.18
8. Y. M. Awel, Empir. Econ. Rev. **1**, 1 (2018). https://doi.org/10.29145/eer/12/010201

9. S. W. Gikungu, Am. J. Theor. Appl. Stat. **4**, 15 (2015). https://doi.org/10.11648/j.ajtas.20150401.13
10. S. Deepa and G. Vijayasree, Ecol. Environ. Conserv. **28**, 459 (2022). https://doi.org/10.53550/eec.2022.v28i02s.070
11. M. Jamii, N. Oumidou, and M. Maaroufi - *Proc. 2nd Int. Conf. Big Data Modelling Machine Learn*ing BML, **1**, 496 (2021). https://doi.org/10.5220/0010737600003101
12. U. Kumar and V. K. Jain, Stoch. Environ. Res. Risk Assess. **24**, 751 (2009). https://doi.org/10.1007/s00477-009-0361-8
13. B. Muma and A. Karoki, OALib **9**, 1 (2022). https://doi.org/10.4236/oalib.1108355
14. P. Basak and S. Nandi, Int. J. Environ. Sci. **4**, 956 (2014).
15. K. Mondal, S. Sinha, P. Basak, and D. P. Goswami, Ensemble **2**, 1 (2021). https://doi.org/10.37948/ensemble-2020-0202-a001
16. O. P. Dubey, A. K. Lal, K. Mondal, and U. Singh, Int. J. Stud. Res. Technol. Manage. **11**, 1 (2023). https://doi.org/10.18510/ijsrtm.2023.1121
17. C. K. Pandit, A. K. Lal, and U. S. Singh, J. Sci. Res. **16**, 723 (2024).
18. M. H. Masum, R. Islam, M. A. Hossen, and A. A. Akhie, J. Sci. Res. **14**, 215 (2022). https://doi.org/10.3329/jsr.v14i1.54973
19. T. Ghoshal and R. Bhattacharyya, SSRN Electron. J. (2012). https://doi.org/10.2139/ssrn.2166900