Analyzing Temperature Patterns of Dhaka city with SARIMA model

Raisa Rahman Rodela
School of Data and Sciences
Brac University
Dhaka, Bangladesh
raisa.rahman.rodela@g.bracu.ac.bd

Joyanta Jyoti Mondal
School of Data and Sciences
Brac University
Dhaka, Bangladesh
joyanta.jyoti.mondal@g.bracu.ac.bd

Meem Arafat Manab
School of Data and Sciences
Brac University
Dhaka, Bangladesh
meem.arafat@bracu.ac.bd

***Abstract*—The research uses the seasonal autoregressive integrated moving average (SARIMA) model to examine the temperature pattern of Dhaka city. Weather data from 1948 to 2013 was collected from the various cities of Bangladesh, and the SARIMA model was used to identify Dhaka's seasonal and long-term temperature trends. The analysis shows that the temperature of Dhaka had a rising trend in the early years of the studied period, followed by a peak in the middle period, and then a decreasing trend in recent years. The SARIMA model could accurately capture the seasonal patterns in the data and forecast future temperature changes. The RMSE function was used as a base to calculate the error. In our research the predicted value fit very well on the current values. The research offers insightful information on Dhaka's temperature trends and may be utilized to guide choices about climate change and urban planning policy. The SARIMA model can be used in additional regions to examine and forecast temperature trends.**

***Index Terms*—Temperature of Dhaka, SARIMA, RMSE, Time series, Forecasting, Bangladesh weather**

I. INTRODUCTION

The Seasonal Autoregressive Integrated Moving Average (SARIMA) model is a common time series forecasting method that is also known as the Seasonal ARIMA. SARIMA model includes seasonal differencing, autoregression, moving average, and non-seasonal differencing. The SARIMA model can be used in predicting future time series data values and forecasting [1]. The autocorrelation and partial autocorrelation functions are used to discover seasonal and non-seasonal trends in time series data before building a SARIMA model. After identifying patterns, model parameters areb chosen like the order of differencing, autoregression, moving average, and seasonal period.

Due to the urbanisation, industry, and climate change, Dhaka, the capital of Bangladesh, has undergone considerable temperature variations in past years [2]. Researchers, politicians, and environmentalists are all interested in Dhaka's temperature trends because they may provide light on the effects of urbanisation and climate change on the city's environment and inhabitants. The Seasonal Autoregressive Integrated Moving Average (SARIMA) model is an effective method for predicting time series data, and time series analysis is a valuable tool for investigating temperature trends.

Other ways to make predictions about time series include exponential smoothing, neural networks, and regression analysis. Each of these methods has its pros and cons. For example, exponential smoothing is a simple and easy-to-understand method, but it doesn't work as well with seasonal data as SARIMA does. Neural networks are good at figuring out complex, unpredictable connections in data, but they take a lot of work to prepare the data and can use a lot of computing power. Regression analysis is a good way to figure out how independent and dependent factors are related, but it may not work well for time series data with multiple connections and trends.So, the SARIMA model was chosen as the best method for this research because it is made for analysing time series data with seasonal patterns, it is a widely used and accepted method for forecasting time series, and it gives accurate and reliable predictions of future values.

The purpose of this study is to use the SARIMA model to determine the seasonal and long-term temperature patterns in Dhaka, evaluate the model's ability to predict future temperatures, and shed light on the causes of these trends. The research aims to enhance the potential effects of climate change on Dhaka's environment as well. In addition, the SARIMA model used here might be used to examine regional temperature trends elsewhere and provide light on the effects of climate change on populated areas.

II. RELATED WORKS

Dabral and Murry (2017) [3]researched rainfall modeling and forecasting and used time series for various applications, including water resources management, irrigation scheduling, agricultural management, and reservoir operation. The authors compared soft computing techniques, empirically based models, and stochastic time series models for modeling hydrological variables. They emphasize the benefits of time series models such as ARIMA and SARIMA, which offer greater detail on the temporal variation of data and are transparent in their methodology. They also mentioned data-driven techniques such as ANN and fuzzy data interpretation techniques. They then analyzed the model thoroughly. Their primary concentration is on research, which employs SARIMA modeling and forecasting of monthly precipitation series in northeast India. The study discovered that the SARIMA method is an effective procedure for modeling time series. The predictive accuracy of SARIMA was superior to other methods. Using the SARIMA models, the study could analyze monthly and weekly forecasts and daily monsoon precipitation from 2014 to 2027.

A related publication addresses the problem of long-term average forecasting for developing solid environmental policies [4]. They researched the monthly temperatures of Iran's various climate zones using time series models and neural networks, such as SARIMA and Support Vector Regression (SVR). To alleviate the problem of local optimality, the authors used a combination of SVR and FA, SVR-FA. The study contrasts the effectiveness of SARIMA, SVR, and SVR-FA models in predicting temperature in distinct climatic environments. The SARIMA model performed the best in long-term temperature forecasting, especially for Iran's arid and warm/cold regions.

Earlier work on forecasting evaluated two models: the multiple linear regression (MLR) models and the seasonal autoregressive integrated moving average (SARIMA) model [5]. The critical issue addressed in this paper is the accurate forecasting of heat demand in district heating systems. The authors observed that forecasting errors can produce costly excess energy and are environmentally destructive. So, accurate forecasting is crucial for enhancing the efficiency of district heating systems. To address the issue, they assess the efficacy of the MLR and SARIMA models using heat demand data from the past. The analysis indicates that both models can accurately predict heat demand from a short-term perspective. The SARIMA model, however, outperformed in longer-term forecasting. The study demonstrates that the SARIMA could be enhanced by including additional variables, such as information regarding economic and demographic factors that may influence heat demand.

Ho and Xie (1998b) used ARIMA model in their study for time series forecasting. They discussed that time series approach is incredibly versatile and requires minimal assumptions [6]. The method to repairable system reliability predictions based on Autoregressive Integrated Moving Average (ARIMA) models was implemented effectively in the study. A comparison with the standard Duane model is also presented. It is determined that the ARIMA model is a feasible option that provides good prediction performance.

In a particular work, [7]the ARIMA model was used for forecasting and time series modelling. The most recent advancements in this field were reviewed, and the methodology was connected to many other forecasting techniques.

An earlier article concentrated on accurate, multi-step advanced wind power forecasting for the efficient operation of intermittent wind power systems [8]. The proposed model was based on a dual-attention mechanism applied to an encoder-decoder-based sequence-to-sequence model composed of LSTM blocks. The Bayesian optimization algorithm was used to obtain the optimal hyperparameters. The proposed model outperformed advanced models like LSTM. N-BEATS used in the research ensembled the competence scores for forecasting: MAE, RMSE, and RMSE. They also analyzed input features such as wind speed and air pressure for multi-horizon forecasting. The average score for forecasting was 0.6625.

Another study examines numerous bias correction techniques and evaluates their efficacy using a variety of metrics [9]. Quantile mapping, distribution-based scaling, and local intensity scaling are parametric and non-parametric methodologies the authors consider evaluating. Then they compare the uncorrected and corrected RCM simulations to the methods, observed climate data, and the impact on hydrological variables, such as precipitation and runoff. They also considered the effects of various bias correction techniques for future hydrological variable projections. The research demonstrates that bias correction is essential for detailed hydrological impact studies. RCMs can generate significant biases. The publication offers a thorough analysis and evaluation of numerous bias correction techniques. RCM simulations emphasize the significance of contemplating various methods.

Maraun (2016) addresses the issue of quantifying climate models [10]. His research focuses on regional climate change impacts in the future. Regional climate models (RCMs) are gaining popularity for bridging the scale divide. The study was done to inherit substantial errors from the primary global climate models (GCMs). Therefore, climate bias correction has become a viral influence investigation. The research introduces bias correction, including its historical context, conceptual aspects, ongoing debates, and methodological limitations. It also examines the evaluation and effectiveness of bias correction techniques. The approach estimated that the climate model could not solve fundamental flaws despite its skillful input.

Another publication examines the implications of time series (TS) forecasting in modern business and its financial impact [11]. Despite the preponderance of machine learning (ML) and deep learning (DL) techniques in other areas, classical statistical TS forecasting systems often surpass them. The article inquires whether hybrid methods combining classical statistical techniques with DL methods are required to enhance forecasting precision. Instead, it explores the potential of refined DL architectures for TS forecasting. They used N-BEATS to decompose the time series into a set of basis functions. Then they analyzed the essential functions to determine the driving factors contributing to the time series and forecasted values. They also used them to develop an accurate and interpretable model for time series forecasting. The experiment shows that N-BEATS outperforms other deep learning and classical time series forecasting methods on benchmark datasets, achieving state-of-the-art accuracy.

Another source emphasized the difficulty of converting climate projections into practical information for agricultural decision-making [12]. The study addresses farmers' need to adjust to changing climatic circumstances to secure sustainable agricultural production. However, converting climate projections into usable information that farmers can use and comprehend remains challenging. The authors analyzed recent progress in translating climate projections into agricultural terms to address the issue. They highlight several methodologies, such as statistical and machine learning models, ensemble forecasting, and participatory forecasting approaches that include farmers in the forecasting process. They also highlight the issues involved with these techniques, including data limits, uncertainty in climate predictions, and the necessity for good communication strategies to guarantee that forecast information is transmitted to farmers successfully. The study's findings show the potential advantages of converting climate projections into agricultural terms, such as better crop yields, lower water consumption, and enhanced resistance to climatic unpredictability. However, the authors point out that considerable hurdles remain in converting climate projections into practical information that farmers can use and comprehend.

Shourav et al. (2016) examined Dhaka's climate's past and projected changes in their research [13]. Statistical downscaling model (SDSM) projected future changes in daily rainfall and temperature, and non-parametric trend analysis assessed changes in rainfall, temperature, and associated extremes. The research found that Dhaka's nighttime temperature has climbed 0.22oC/decade over the last 50 years and is expected to continue rising. Dhaka also has more severe temperature incidents. However, rainfall and extremes remain unchanged. Thus, climate change consequences will be caused by rising temperatures and extremes.

It is anticipated that climate change, as well as variations in the climate, would have a negative impact on Bangladesh. Many people, particularly those living in the coastal zone, low-lying regions, and the north-western section of the nation, are already being put in danger by a variety of climate-related dangers. So, research on the climate change should be continued for the betterment of the country. [14]

III. Research Methodology

This study utilised a quantitative methodology to analyse Dhaka's temperature patterns. The study used an existing dataset that has Monthly average of Maximum Temperature, Minimum Temperature, Rainfall, Relative Humidity, Wind Speed, Cloud Coverage and Bright Sunshine of Bangladesh from the period 1948 to 2013 of specific region. The data is fetched from Bangladesh Meteorological Department (BMD) that represents 65 years of weather data of Bangladesh .
Using the SARIMA model, seasonal and long-term temperature trends were identified from the temperature data. The SARIMA model was chosen due to its ability to capture seasonal and non-seasonal trends in time series data, as well as its precision in predicting future values.

The steps involved in applying the SARIMA model included Data Preprocessing, Model Identification, Model Estimation and checking and Forecasting.

A. Data Preprocessing

In the preprocessing part, we droped the unnecessary columns which were not related to our project. We also merged two columns which are year and month for better calculation.

B. Model Identification

The autocorrelation and partial autocorrelation functions were used to identify the seasonal and non-seasonal patterns in the temperature data. The appropriate SARIMA model parameters, including the order of differencing, the order of autoregression, the order of moving average, and the seasonal period, were selected based on the PACF and ACF. For forecasting SARIMAX was used.

C. Model Estimation and Checking

The selected SARIMA model was fitted to the temperature data using maximum likelihood estimation. The fitted model was checked for the adequacy of the model assumptions, including the normality of the residuals, the absence of serial correlation, and the absence of heteroscedasticity. To RMSE function helped to calculate the error.

D. Forecasting

For the forecasting and validation, a function is to use one-step-forecast in the whole validation set and measure the error. And then predicted value is plotted. After that, predicted value was compared to current value.

IV. Results

From the research we found out that, temperature in Dhaka city was increasing mostly and was highest in 1980 era. After that, the temperature had a decreasing trend till 2013.

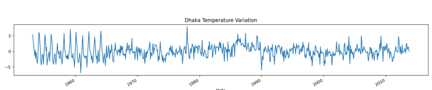

Fig. 1. Temperature variation in Dhaka

In our finding, the Error vs Predicted values has a linear distribution (the errors are between -1.0 and +1.0 while the temperature increases). SO, it can be said that, the predictions fitted well with the current values. The baseline RMSE for the test baseline score was 1.02 celsius degrees.

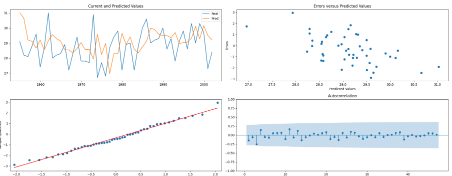

Fig. 2. Graph of Prediction and Error

A. Discussion

The analysis revealed that the temperature in Dhaka had a fluctuation in degrees Celsius per year from 1948 to 2013. The seasonal component of the temperature pattern also showed a clear fluctuation, with the highest temperatures occurring in May-June and the lowest temperatures occurring in December-January. However, after visualizing the average yearly temperature, We can confirm that there is a constant increasing trend and that the average temperature increased from 33.0º to 34.0º, that's 4.25% in over 65 years. Although the prediction in our study fitted well with the current value, extrapolated prediction did not fit well with the real value.

V. Conclusion

This research analyzed the temperature patterns of Dhaka using the SARIMA model. The analysis revealed a fluctuating trend in temperature from 1948 to 2013 temperature patterns. The SARIMA model was found to be a reliable and accurate method for forecasting future temperature values in Dhaka, and its superiority was demonstrated by its lower values of RMSE compared to other time series forecasting methods. The results of this study provide valuable insights into the impacts of urbanization and climate change on the temperature patterns of Dhaka. Further research is needed to understand the specific causes of the observed temperature patterns and to develop effective strategies for mitigating their impacts on the environment and human health. Future research could explore the use of more advanced modeling techniques, such as deep learning or hybrid models, to improve the accuracy of temperature forecasting in Dhaka. This could involve incorporating additional data sources, such as satellite data, into the analysis to improve the accuracy of the models.

References

[1] S. I. Vagropoulos, G. Chouliaras, E. G. Kardakos, C. K. Simoglou, and A. G. Bakirtzis, "Comparison of sarimax, sarima, modified sarima and ann-based models for short-term pv generation forecasting," in *2016 IEEE international energy conference (ENERGYCON)*. IEEE, 2016, pp. 1–6.

[2] S. Banu, W. Hu, Y. Guo, C. Hurst, and S. Tong, "Projecting the impact of climate change on dengue transmission in dhaka, bangladesh," *Environment international*, vol. 63, pp. 137–142, 2014.

[3] P. Dabral and M. Z. Murry, "Modelling and forecasting of rainfall time series using sarima," *Environmental Processes*, vol. 4, no. 2, pp. 399–419, 2017.

[4] P. Aghelpour, B. Mohammadi, and S. M. Biazar, "Long-term monthly average temperature forecasting in some climate types of iran, using the models sarima, svr, and svr-fa," *Theoretical and Applied Climatology*, vol. 138, no. 3-4, pp. 1471–1480, 2019.

[5] T. Fang and R. Lahdelma, "Evaluation of a multiple linear regression model and sarima model in forecasting heat demand for district heating system," *Applied energy*, vol. 179, pp. 544–552, 2016.

[6] S. L. Ho and M. Xie, "The use of arima models for reliability forecasting and analysis," *Computers & industrial engineering*, vol. 35, no. 1-2, pp. 213–216, 1998.

[7] P. Newbold, "Arima model building and the time series analysis approach to forecasting," *Journal of forecasting*, vol. 2, no. 1, pp. 23–35, 1983.

[8] M. Aslam, J.-S. Kim, and J. Jung, "Multi-step ahead wind power forecasting based on dual-attention mechanism," *Energy Reports*, vol. 9, pp. 239–251, 2023.

[9] C. Teutschbein and J. Seibert, "Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods," *Journal of hydrology*, vol. 456, pp. 12–29, 2012.

[10] D. Maraun, "Bias correcting climate change simulations-a critical review," *Current Climate Change Reports*, vol. 2, no. 4, pp. 211–220, 2016.

[11] B. N. Oreshkin, D. Carpov, N. Chapados, and Y. Bengio, "N-beats: Neural basis expansion analysis for interpretable time series forecasting," *arXiv preprint arXiv:1905.10437*, 2019.

[12] J. W. Hansen, A. Challinor, A. Ines, T. Wheeler, and V. Moron, "Translating climate forecasts into agricultural terms: advances and challenges," *Climate research*, vol. 33, no. 1, pp. 27–41, 2006.

[13] M. S. A. Shourav, M. Mohsenipour, M. Alamgir, S. H. Pour, and T. Ismail, "Historical trends and future projection of climate at dhaka city of bangladesh," *Jurnal Teknologi*, vol. 78, no. 6-12, 2016.

[14] S. Huq, "Climate change and bangladesh," pp. 1617–1617, 2001.