Time Series Forecast Analysis of PM2.5 in Beijing

Yidan Chen, Yixin Sun, Wenyi Li

ECON 515 Time Series Analysis

Dr. Ratika Narag

May 7, 2023

Abstract

Our research examined the trend in the PM 2.5 concentration in Beijing from December 2013 to December 2021 and found a downward trend in the PM 2.5 concentration over time. Additionally, we used different models to predict the future PM 2.5 concentration in Beijing in 2023. We found that the SARIMAX model provides the best prediction result, which is aligned with the true PM 2.5 concentration at the beginning of 2023. Our findings may provide support for the success of the Beijing Municipal Government in achieving the sustainable development goal.

I. Introduction

In 2013, Beijing was facing a severe air pollution crisis that posed significant threats to public health and the environment. The city was experiencing extremely high levels of air pollution, with some days reaching hazardous levels, causing visibility to be severely reduced and posing serious health risks to the population. The primary pollutant was particulate matter, with levels of PM2.5 often exceeding 500 micrograms per cubic meter – more than 20 times the recommended limit set by the World Health Organization. Also, the high levels of particulate matter in the air posed a serious health risk to the city's residents, especially to children and the elderly. The pollutants in the air could cause respiratory problems, such as coughing, wheezing, and shortness of breath. Long-term exposure to air pollution could also lead to lung disease, heart disease, and even cancer.

According to our findings, some potential causes lead to the severity of air pollution in Beijing. One of the primary causes of air pollution in Beijing is industrial emissions. Steel production and coal-fired power plants are major contributors to air pollution, releasing large amounts of particulate matter, sulfur dioxide, and nitrogen oxide. Transportation is another significant source of air pollution in Beijing. The increasing number of cars on the roads has led to a significant increase in emissions of carbon monoxide, nitrogen oxides, and volatile organic compounds. Construction activities are also a major source of air pollution in Beijing. Dust and other particles are constantly being kicked up from construction sites, contributing to particulate matter pollution. Local government competition for economic growth is another factor contributing to air pollution in Beijing. Many local governments prioritize economic growth over environmental protection, leading to lax enforcement of environmental regulations. Agricultural practices also contribute to air pollution in Beijing. Burning of crop residue and the use of fertilizer and pesticides release large amounts of particulate matter and other pollutants into the air. Lastly, the geography and climate of the Beijing region also play a role in air pollution. The city is surrounded by mountains, which trap pollutants in the air. In addition, weather conditions such as temperature inversions can worsen air pollution by preventing the dispersal of pollutants.

To deal with the air pollution in Beijing, the Chinese government has also implemented a series of policies. To limit emissions, the government has implemented stricter emission standards for both vehicles and industrial activities. They have also limited the number of cars on the road through a license plate lottery and auction system. In addition, the government is encouraging the use of clean energy sources like natural gas, wind, and solar power. To reduce reliance on personal vehicles, the government is promoting public transportation by expanding subway lines and bus routes. To further reduce pollution levels, highly polluting industries such as steel and cement factories and coal-fired power plants have been closed or relocated. Companies that violate pollution regulations may face fines. The Chinese government has also launched a "war on pollution" campaign and established a ministry dedicated to environmental protection. Increasing public awareness about the health risks of air pollution and the importance of reducing personal exposure is also a priority. Finally, the government is coordinating with neighboring provinces to reduce pollution from agricultural practices like crop burning and pesticide use. These measures are aimed at reducing air pollution levels in Beijing and promoting sustainable development.

The severity of air pollution in Beijing in 2013 prompted us to investigate the forecast of air pollution. The motivation behind this research is to provide a warning of high levels of air pollution, which can help individuals make informed decisions such as wearing masks or limiting outdoor activities. Additionally, this information can be used by the government and policymakers to make informed decisions, such as limiting industrial emissions, restricting car usage, and improving public transportation. The research also aims to provide data and insights that can inform public awareness campaigns and educational efforts. Moreover, this research can promote sustainability in the long run, contributing to economic growth and well-being. The concept of ESG responsibility (Environmental, Social, Governmental) has become increasingly important, and this research can help meet these responsibilities. By investigating the forecast of air pollution, we hope to contribute to the ongoing efforts to reduce air pollution levels in Beijing and promote sustainable development.

Thus, we proposed the following research question: How has the PM 2.5 concentration in Beijing changed over time and how can we forecast the future air pollution level using historical data?

II. Literature Review

Air pollution poses a threat to people's health. Yin et al. (2020) found that the overall age-standardized death rate caused by the air-pollution decreased with declining air pollution exposure. Zhou et al. (2017) examined the relationship between particulate air pollution and mortality by running generalized linear models with different lag structures on PM10 data from the 38 largest cities in China. They found that the short-run effect of PM10 on mortality is statistically significant.

Air pollution also harms economic growth. Richardson and Willis (2013) identified three main ways in which local economic growth could be affected by air quality. The first path is through health and corporate costs. The health of the local population and their productivity will be negatively affected by the worsening air quality. Therefore, the business costs will increase. The second path is through business and residential location decisions. Since the quality of life is greatly determined by the local air quality, people may choose to live and work in an environment with better air quality. Therefore, the local economic growth will be affected. The third path is through regulations. Business costs and operations are affected by the stringency of regulations of local air quality control.

Many previous studies examined the potential causes of air pollution in Beijing. Jiang et al. (2022) pointed out that the local government's competition for economic growth aggravates air pollution. Moreover, the spillover effects of economic growth and air pollution lead to the problem of cross-boundary welfare damage. Feng et al. (2015) used the IPAT model to investigate the effect of Beijing's population, GDP per capita, share of industrial GDP, and energy emission intensity on Beijing's air pollution. They found that emission intensity is the most influential factor in the improvement of air pollution in Beijing.

The consequences of air pollution arose great attention from the international community and local government. Rafaj et al. (2018) analyzed the effect of a Sustainable Development Scenario

(SDS), the aim of which is to address the link between three sustainability goals, namely, achieving universal energy access, limiting climate change and reducing air pollution. Under SDS, CO2 emission shows a steep decline through 2040 after improving energy efficiency, exploiting up-scaled renewable energy, and reducing CO2 emissions. He et al. (2019) summarized the major economic incentives and financial support policies of the Beijing municipal government to fight air pollution from 2001 to 2017. This includes measurements in terms of regulation of industrial firms, promotion of electric vehicles, transformation of vehicles, elimination of old and used vehicles, promotion of new energy, control of bulk coal, and renovation of boilers. However, the deficiency of some policies is also non-negligible. Hu et al. (2018) wrote that Total Emission Control Policy is one of the major pollution control policies in China. Allocating TEC quotas to each provincial government based on industrialization, historical emission data, and reduction capabilities is a creative instrument for implementing TEC. However, the methods for accurately measuring these are not available. Also, this policy lacks continuity in targeting the range of pollutants and the removal of targeted pollutants from the initial list over time.

There are a lot of empirical studies on air pollution time series forecasting analysis. Some scholars reviewed the existing models and techniques for air pollution forecasts. Chakraborty and Mondal (2019) reviewed various models and techniques used in forecasting air pollutants, such as time series analysis, regression analysis, artificial neural networks (ANN), and machine learning. They emphasized that the accuracy of the forecast depends on the selection of input variables, model selection, and training data. Guo, Wang, and Feng (2020) provided a comprehensive review of different models and methods used for air quality time series forecasting. They discussed the use of statistical models, machine learning models, hybrid models, and ensemble models for air quality forecasting. They also highlighted the importance of selecting appropriate features, dealing with missing data, and model interpretability. Cui and Hao (2018) focused on short-term air quality forecasting. They reviewed statistical models, such as ARIMA, and machine learning models, such as ANN, support vector regression (SVR), and random forest (RF). They also discussed the importance of feature selection, data preprocessing, and model evaluation. Yu et al. (2021) provided a review of air quality forecasting using machine learning. They discussed the advantages and limitations of different machine learning algorithms, such as SVR, RF, decision tree (DT), and deep learning (DL). Liu et al. (2019) reviewed the potential of machine learning in predicting air pollution trends. They discussed the use of machine learning models for forecasting air pollutants, such as ozone, PM2.5, and nitrogen dioxide (NO2). They highlighted the importance of data quality, feature engineering, and model interpretability. Some scholars also did some time series analysis on air quality data in Beijing. Zhang et al. (2017) compared seasonal ARIMA and ARIMA models for forecasting the air quality index (AQI) in Beijing. They found that the SARIMA model outperformed the ARIMA model in terms of accuracy. However, Xu et al. (2016) compared different time series models for forecasting fine particulate matter concentration (PM2.5) in Beijing. They found that the ARIMA model performed better than other models, such as the exponential smoothing model and the autoregressive integrated moving average model with exogenous variables (ARIMAX). Hao and Cui (2019) proposed a hybrid model for air quality forecasting in Beijing. They found that the hybrid model outperformed both the ARIMA and ANN models in terms of accuracy. Chen et al. (2020) used machine learning algorithms to predict PM2.5 concentrations in Beijing. They compared the performance of different algorithms, such as RF, extreme gradient boosting (XGBoost), and support vector machine (SVM). They found that the RF model performed better than other models in terms of accuracy. Wang, Gu, and He (2019) used machine learning and meteorological data to predict daily PM2.5 concentrations in Beijing. They compared the performance of different machine learning models, such as ANN, SVR, and RF. They found that the RF model outperformed other models. Freeman et al. (2018) trained a deep learning model, consisting of a recurrent neural network (RNN) with long short-term memory (LSTM), to accurately predict local 8-hr averaged O3 concentrations based on hourly air monitoring station measurements as a tool to forecast air pollution. Zheng et al. (2015) evaluated their predictive method with data from 43 cities, presenting the results of four major cities: Beijing, Tianjin, Guangzhou, and Shenzhen. Their method outperforms four sets of baselines significantly by combining four major components, including temporal predictor, spatial predictor, prediction aggregator, and inflection predictor.

However, no existing literature carefully predicts the PM 2.5 concentration using the seasonal ARIMA model, and seasonal ARIMAX model, because we assume that there’s a strong seasonality in the PM 2.5 concentration. Additionally, we attempt to use the machine learning model, such as the LSTM model, to predict the PM 2.5 concentration in Beijing.

III. Theoretical Analysis

This paper uses the time series forecasting model ARIMA (Autoregressive Integrated Moving Average) model and extensions of the ARIMA model: The seasonal ARIMA model and the seasonal ARIMA model with exogenous variables. Other than time series models, this paper also uses a type of Recurrent Neural Network (RNN) model long short-term memory (LSTM) to predict the concentration of PM2.5 in the air. The main goal of this paper is to fit an appropriate model to the residuals of PM2.5 data. By using different forecasting models according to the specific characteristics of the data, we successfully provided a comprehensive comparison of these forecasting models and found the best model to predict the PM2.5 concentration in Beijing.

The standard ARIMA model referred to as Autoregressive Integrated Moving Average, is a commonly used univariate time series forecasting model that combines three essential components: autoregressive (AR) component, differencing component (I), and moving average component. For this paper, the AR term represents the current PM 2.5 concentration's correlation to its past concentration. The I term is to make the time series data of PM 2.5 concentration stationary by applying different levels of differencing; the MA term in the model captures how PM 2.5 concentration depends on the past errors in the series. The standard ARIMA model is denoted as ARIMA (p,d,q) where p represents the orders of the AR components, d represents the I components and q represents the MA components. These components make the ARIMA model suitable for capturing complex relationships in the time series data.

The seasonal ARIMA model is an extension of the standard ARIMA model which can specifically address the seasonality in our PM 2.5 data. It incorporates a seasonal part into the standard ARIMA model which can better fit time series data with seasonal patterns. As we can see in Figure 1, PM 2.5 concentration fluctuates with time recurrently. The seasonal ARIMA model is denoted as ARIMA (p, d, q) (P, D, Q) [s]. Non-seasonal AR, I, and MA components are represented by p, d, q, and seasonal components are represented by P, D, Q, and s.

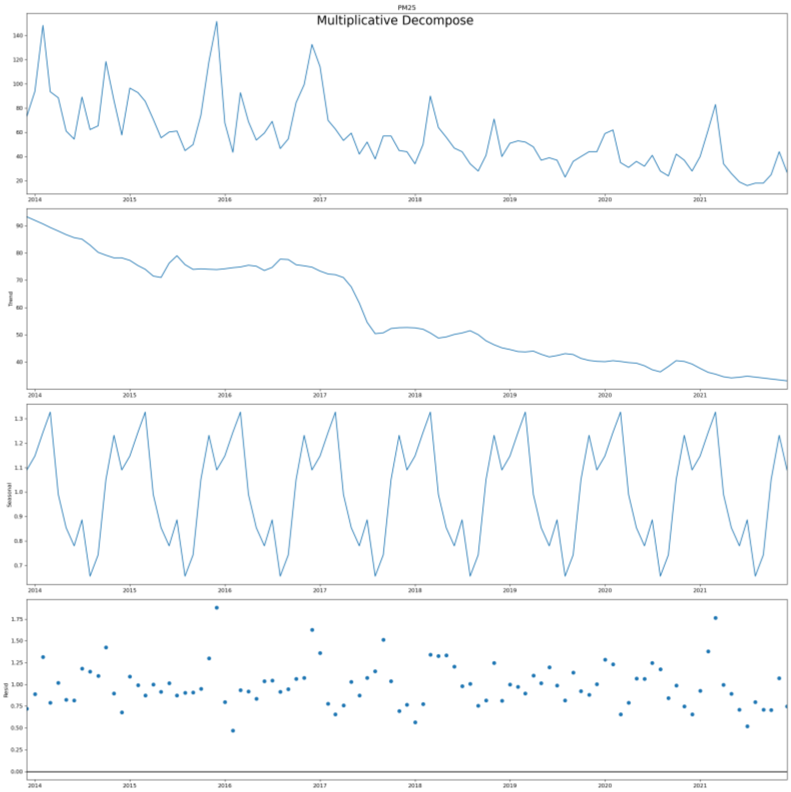


Figure 1: Multiplicative decompose of PM 2.5 concentration

The seasonal ARIMA with exogenous variables (SARIMAX) is a further extension of the standard ARIMA model by adding external variables that could influence the time series data. In our case, exogenous variables such as temperature and humidity could affect the concentration of PM 2.5 in the air. Adding the exogenous variable to the model eliminates the effects of the exogenous variables on the time series while handling the seasonality at the same time. Therefore, the SARIMAX model theoretically could obtain better predictions of time series data which has seasonal patterns and is influenced by exogenous variables.

Long Short-Term Memory (LSTM) is a Recurrent Neural Network (RNN) model that can predict sequences in time series data. The major difference between ARIMA models and LSTM is that LSTM is better at capturing complex non-linear relationships in the time series data while ARIMA models are better at handling linear relationships. This is because the LSTM model is capable of remembering information and it can learn about how the variables in the time series are correlated to its past values and errors because of backpropagation through time (BPTT) training. These characteristics of the model enable it to increase its prediction accuracy.

In summary, this paper aims to compare the forecasting performance of the above four models by using Akaike Information Criterion (AIC) and RMSE （Root Mean Square Error) as evaluation criteria and choose the best model for the prediction of PM 2.5 concentration. AIC is used to assess the goodness of fit of the model and RMSE is used to assess the accuracy of the prediction. Combined, they can provide us with a thorough evaluation of our models. The theoretical procedures of fitting an ARIMA model and calculating the forecast follow the map of what Figure 2 shows: We first plot the data to understand its patterns and select the according models. Then we would perform an ADF test to determine if the time series data is stationary or not and do differencing to make the data stationary, if possible, for the ARIMA process. Then we will determine the orders of the ARIMA models using ACF and PACF plots. After we find the best model according to our evaluation criteria, we will do residual analysis and calculate the forecast if the residuals appear to be white noise.

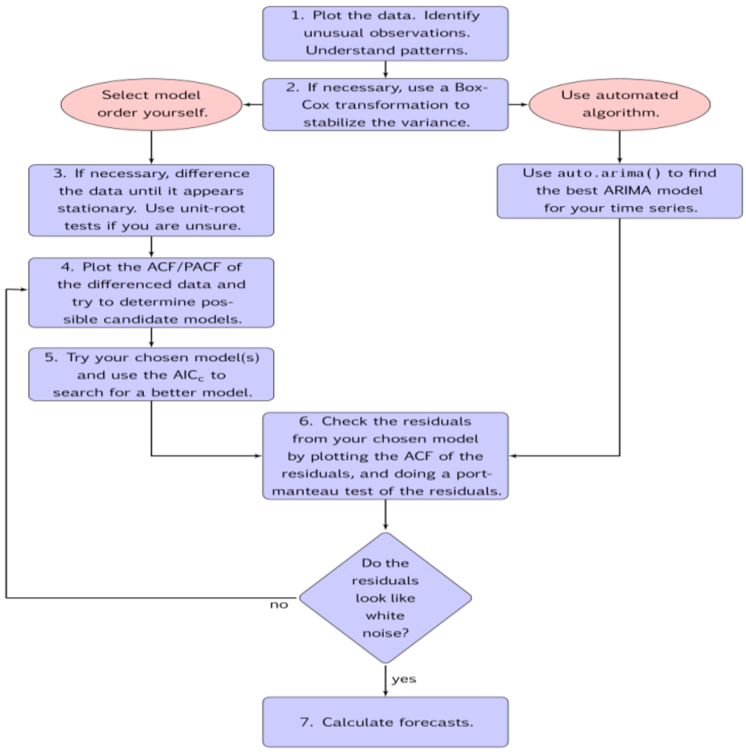


Figure 2: Analysis process

IV. Empirical Analysis

i. Raw Data

The raw data is obtained from a data collection by Dalhousie University Atmospheric Composition Analysis Group and contains the air concentration of 6 major pollutants and the AQI index in Beijing City from December 2013 to December 2021. As we can see from Table 1, the data has 97 rows and 8 columns and is displayed visually here.

	CityName	AQI	PM25	PM10	SO2	CO	NO	O3
date								
2013-12-01	北京市	100	73.3	97.0	37.1	1.73	56.0	37.7
2014-01-01	北京市	125	93.9	122.8	51.4	1.95	65.0	36.6
2014-02-01	北京市	184	148.4	155.4	56.4	2.17	68.6	41.5
2014-03-01	北京市	130	93.6	137.8	33.7	1.39	62.3	82.9
2014-04-01	北京市	127	88.6	144.7	16.3	0.93	57.4	129.0
...	...	...	...	...	...	...	...	...
2021-08-01	北京市	72	18.0	36.0	3.0	0.62	15.0	124.0
2021-09-01	北京市	56	18.0	35.0	3.0	0.63	20.0	98.0
2021-10-01	北京市	50	25.0	43.0	3.0	0.47	30.0	57.0
2021-11-01	北京市	75	44.0	65.0	3.0	0.64	36.0	42.0
2021-12-01	北京市	55	27.0	52.0	3.0	0.58	36.0	40.0

97 rows × 8 columns

Table 1: PM 2.5 data

The concentration of these pollutants in the air fluctuates with time as shown in Figure 3. Although there are 6 major pollutants, in this paper, we are only going to focus on predicting the air concentration of PM 2.5 in Beijing as PM 2.5 pollution has been a major concern in Beijing's air condition since 2013.

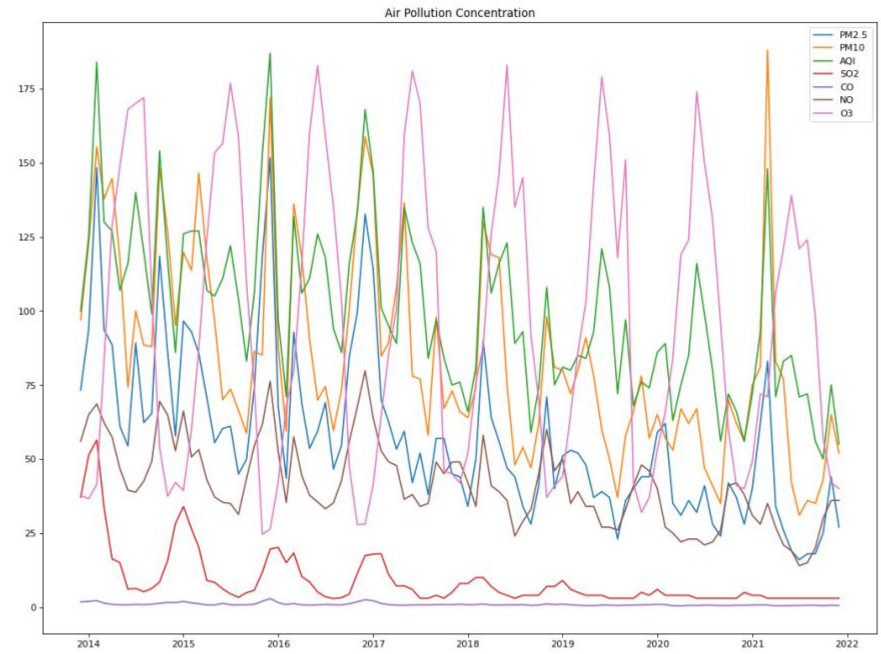
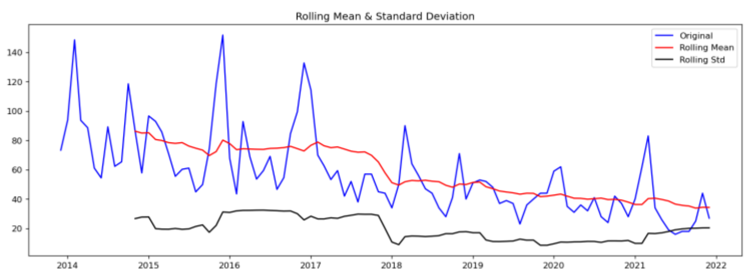


Figure 3: Trend of air pollutants concentration

Therefore, we plot PM 2.5 time series data individually in Figure 4 and we can see that PM 2.5 concentration fluctuates with time from 2013 to 2022. We apply ADF (Augmented Dickey-Fuller) test to the original series and obtain a p-value of 0.883571 and we fail to reject the null hypothesis that the data is stationary. Therefore, we do a first-order differencing to the original series and test the stationarity again using the ADF test. We obtain a P-value of 0 and we can conclude that the first-order differencing of the data is stationary.



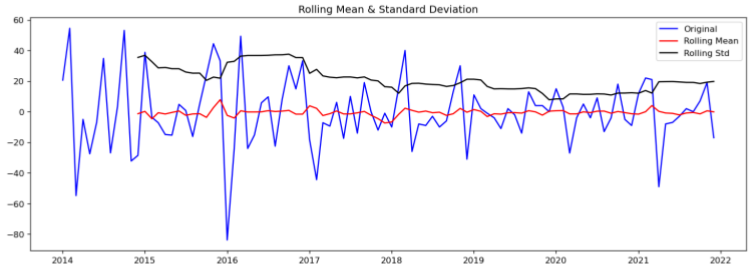


Figure 4: Rolling mean & rolling standard deviation

Also, we plot the first-order differencing and autocorrelation graphs of the data to see if and how current PM 2.5 data is related to the past values (Figure 5).

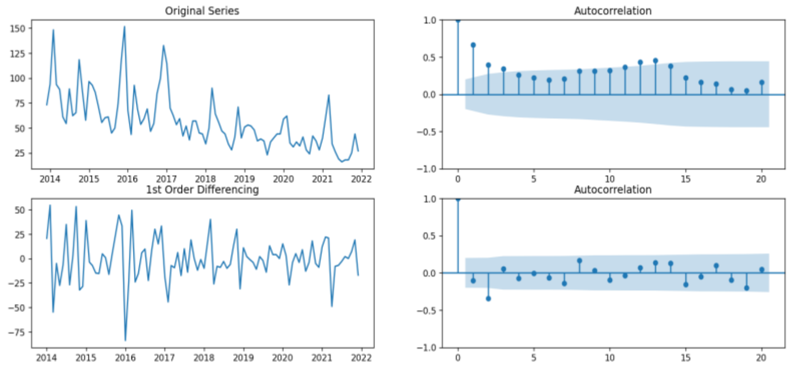


Figure 5: 1st difference & ACF

ii. Model and Results

Having successfully transformed the time series into a stationary process through the application of first differencing, we can now proceed to examine the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) of the differenced series (Figure 6). This will allow us to gain further insights into the underlying patterns and dependencies in the data, which can in turn inform our choice of appropriate modeling techniques and parameters.

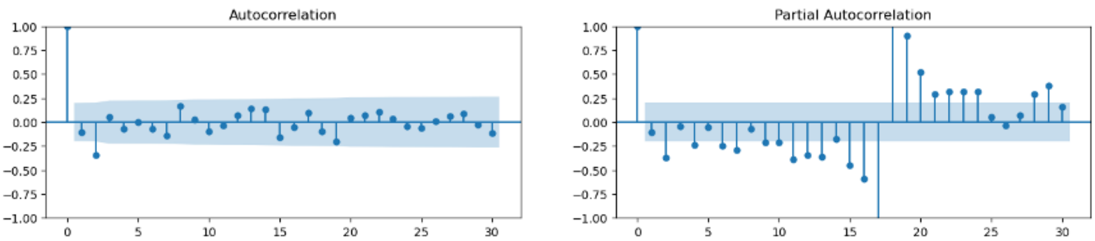


Figure 6: ACF & PACF plot

Based on our analysis of the Autocorrelation Function (ACF) and Partial Autocorrelation

Function (PACF) plots of the differenced time series, we can make some initial observations regarding the presence of various components in the data.

Specifically, the ACF plot reveals that the autocorrelation values for each lag period generally fall within the bounds of the confidence interval. This suggests that there is no significant Autoregressive (AR) component in the time series after the first differencing. On the other hand, the PACF plot exhibits a few significant peaks that extend beyond the confidence interval during the lag period. This indicates that there may be a strong seasonal component in the data, which could potentially be modeled using a Seasonal Autoregressive Integrated Moving Average (SARIMA) model or a similar approach.

Overall, these findings provide important insights into the underlying patterns and dependencies in the differenced time series, which can help guide the selection of appropriate modeling techniques and parameter values for our analysis.

After examining the ACF and PACF plots of the differenced time series, we proceeded to test for the presence of seasonality. To do so, we first redraw the PM2.5 concentration plot on a monthly basis in Figure 7, then we generated two box plots displaying the PM2.5 concentration data on a yearly and monthly basis in Figure 8, respectively.

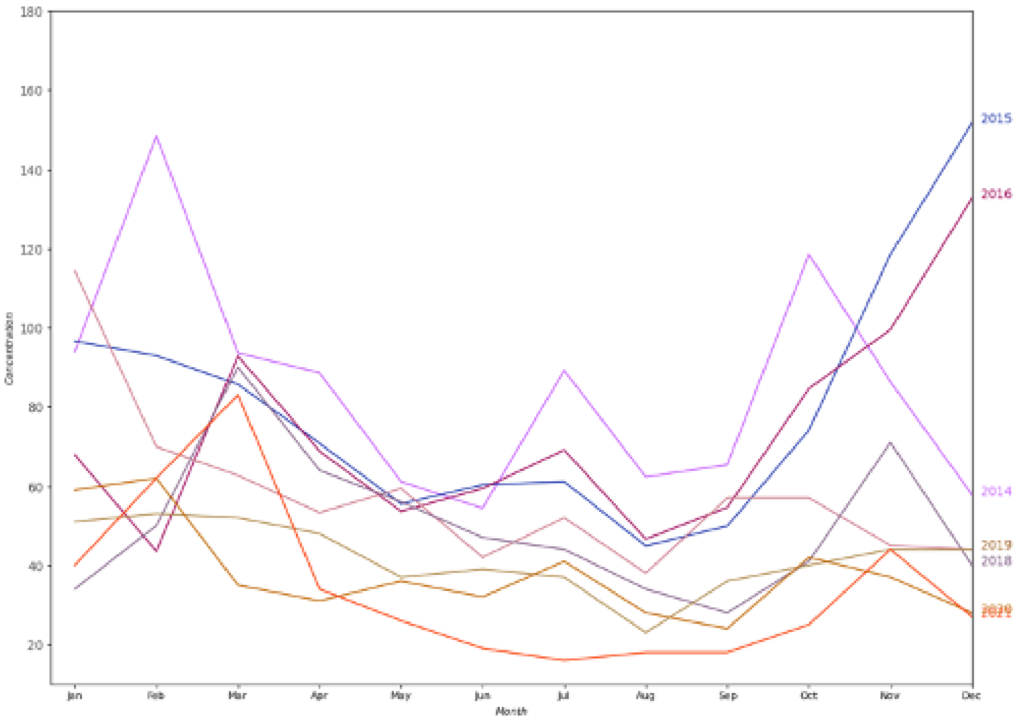


Figure 7: PM2.5 concentration on monthly basis

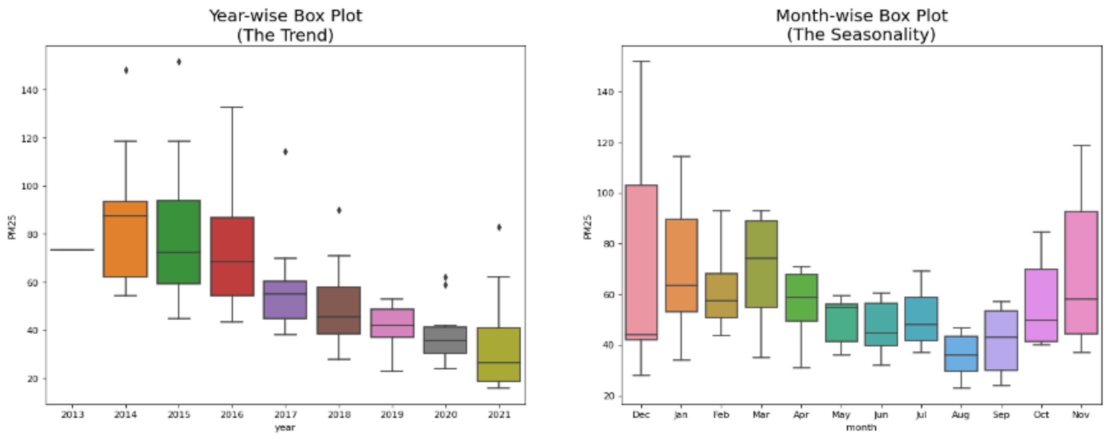


Figure 8: Box plot on yearly basis & monthly basis

In the monthly box plot, we included a median line within each box to represent the median value of PM2.5 concentration for each month. This allowed us to visually identify any recurring patterns or seasonal components in the data. As expected, the monthly box plot clearly shows evidence of seasonality, with concentration levels varying systematically across the months.

To further explore the seasonal patterns in the data, we performed a decomposition of the time series in Figure 9, separating it into its trend and seasonal components. The resulting plot revealed the underlying trend and seasonality of the PM2.5 concentration data. This information can be used to inform our choice of appropriate modeling techniques and parameters for further analysis of the data.

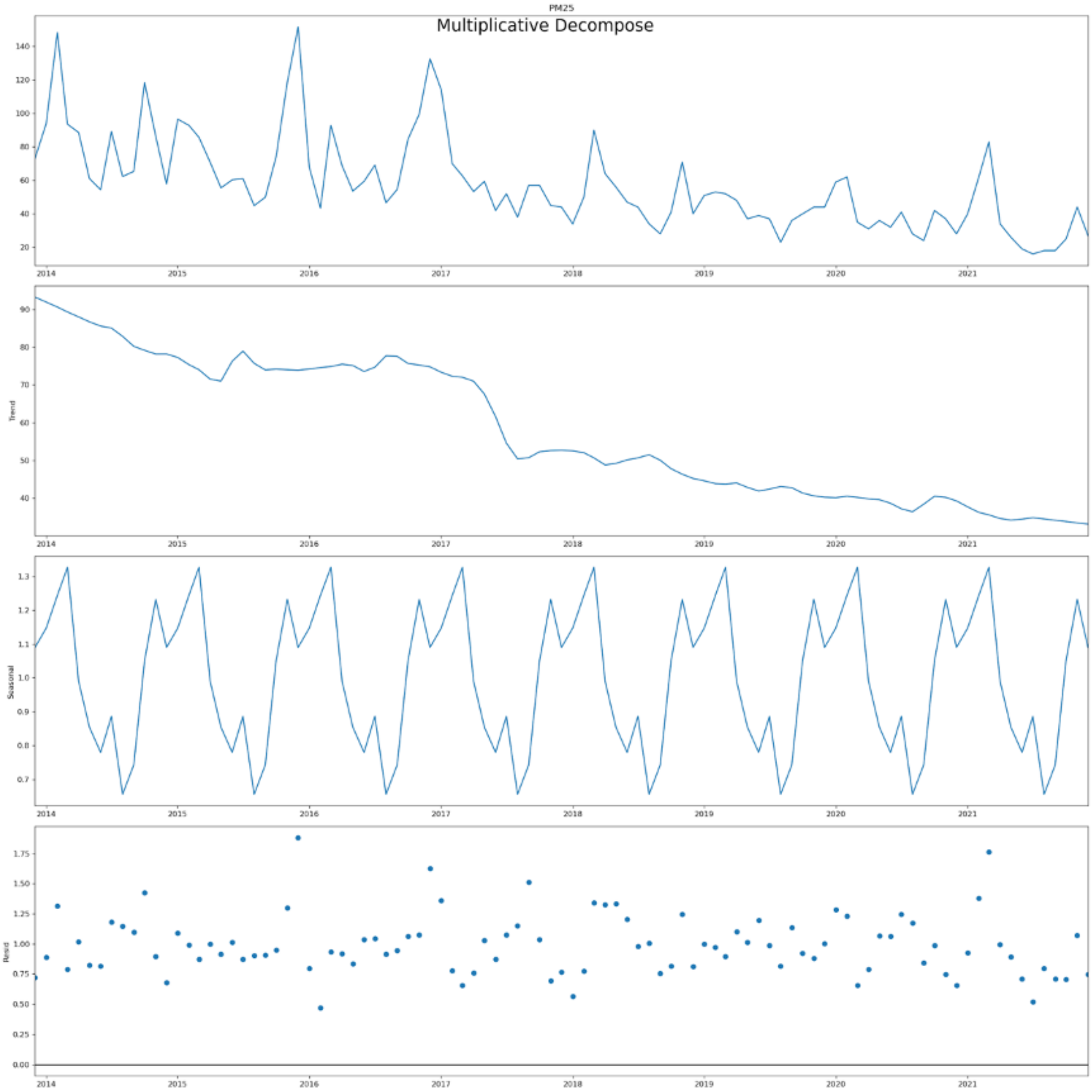


Figure 9: Multiplicative decompose of PM2.5 concentration

We deseasonalize the time series to obtain a more accurate trend representation and generate new ACF and PACF plots in Figure 10. Most values fell within the confidence interval, indicating that the deseasonalized time series is now stationary with no significant trends or patterns remaining in the data.

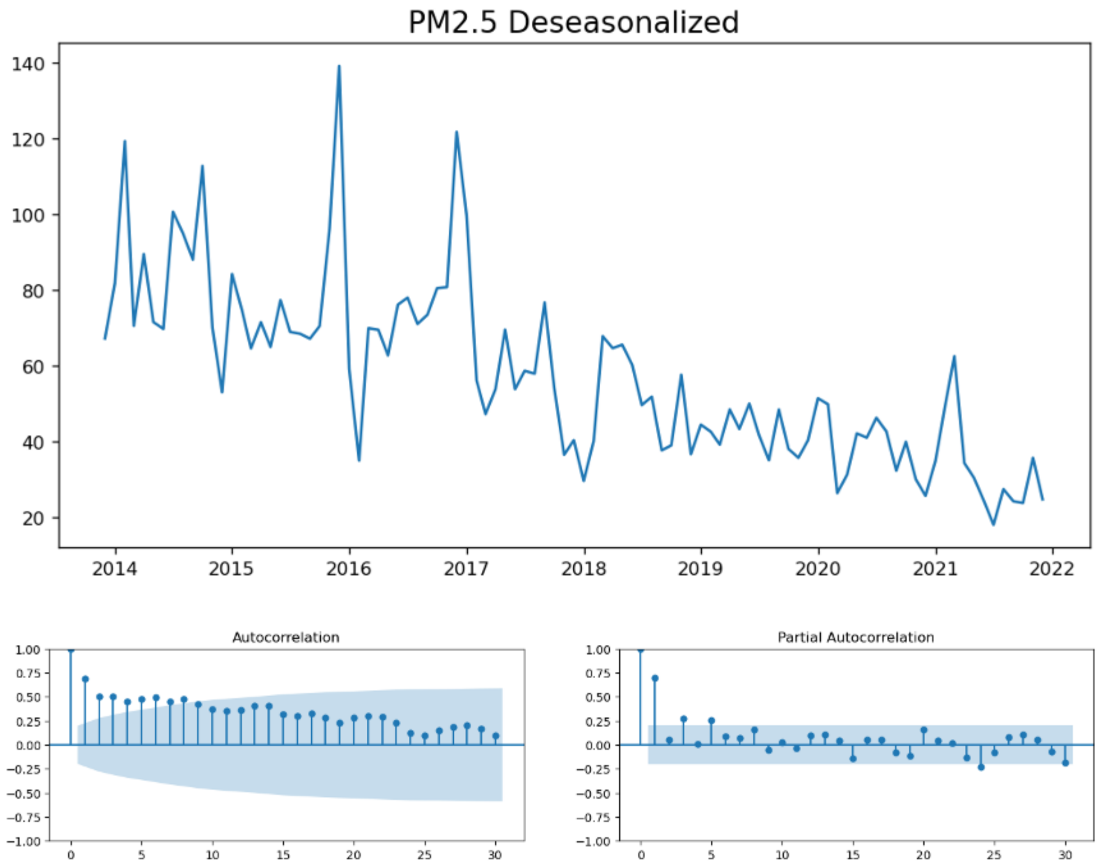


Figure 10: Deseasonalized PM2.5 concentration & ACF & PACF

1) ARIMA Model

In our analysis, we used ARIMA modeling as the first approach to forecasting PM2.5 concentration. The deseasonalized PM2.5 concentration data was used as input, and AIC was used as the criterion for selecting the optimal AR and MA terms for the model.

Based on the resulting plot (Figure 11), we determined that the model was best fitted with an AR term of 0 and an MA term of 3. Therefore, we chose to implement the ARIMA(0,1,3) model for our initial forecast. Figure 12 displays the fitted values against the actual PM2.5 concentration data.

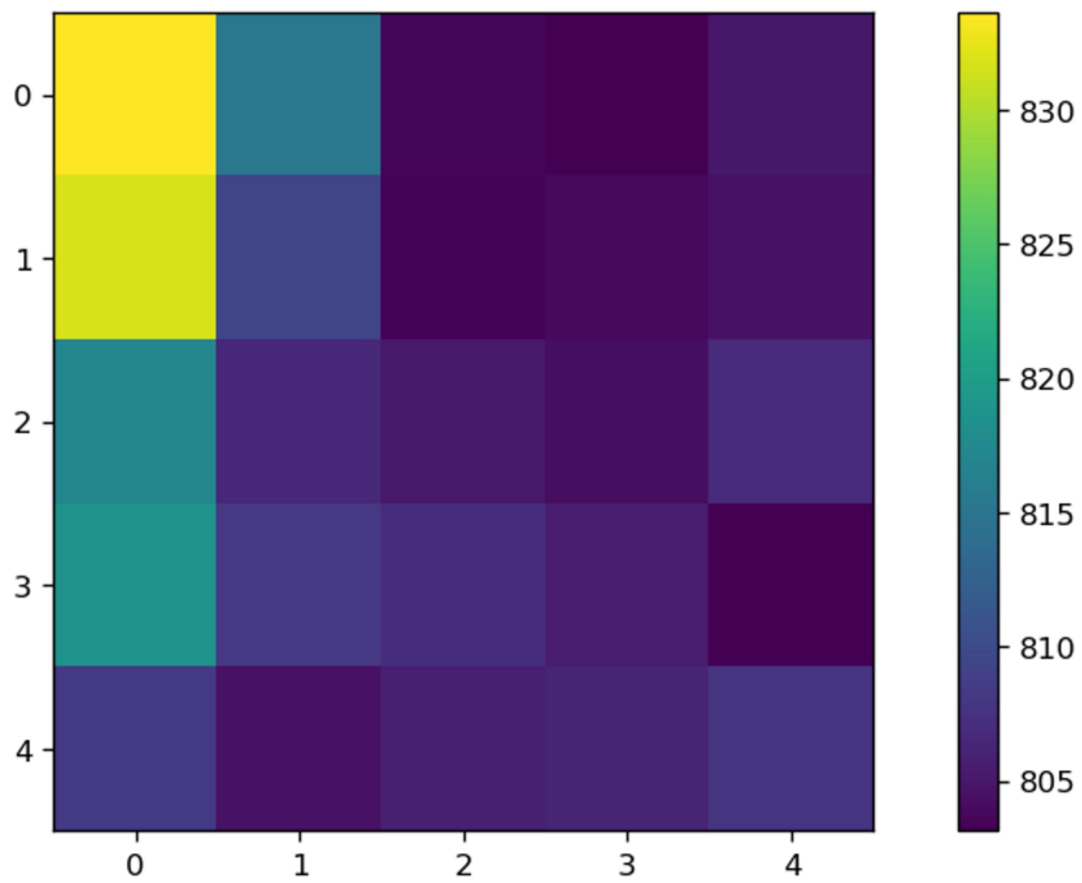


Figure 11: AIC plot

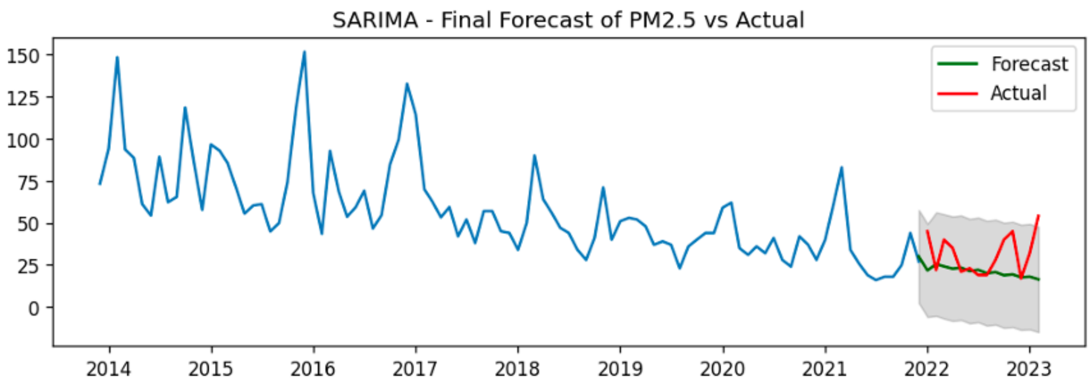


Figure 12: Forecast PM2.5 concentration v.s. actual

2) SARIMA Model

For our second forecasting model, we opted for SARIMA due to the strong seasonal component in our time series. We began by applying seasonal differencing to the data (Figure 13) and then used the auto-ARIMA package to determine the optimal AR and MA parameters. The resulting best-fit model was found to be SARIMA (0,0,2) (2,1,0) [12], where the parameters in the parentheses denote the non-seasonal AR and MA orders, while the parameters in the square brackets denote the seasonal AR and MA orders with the seasonality of 12 (monthly basis). The AIC value for this model was 771.74.

Figure 13: Seasonal difference

Using this model, we forecasted the PM2.5 concentration for the period between January to December of 2022. Figure 14 shows the actual concentration values represented by the red line and our predicted values represented by the green line.

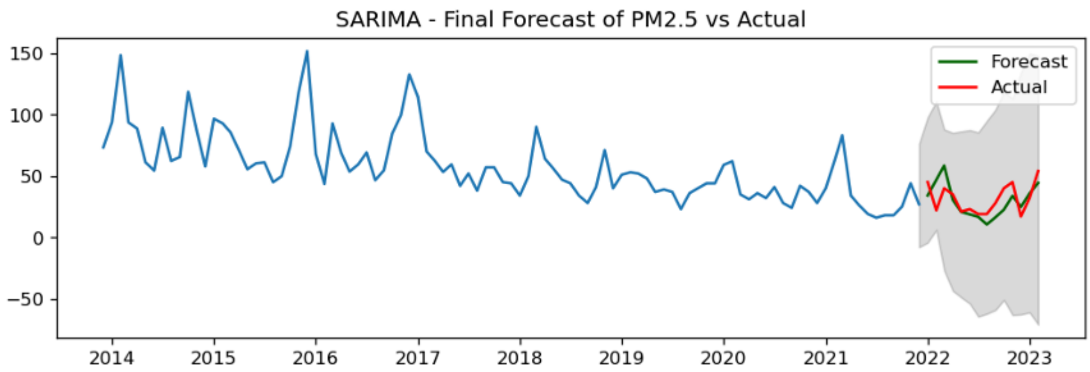


Figure 14: Forecast PM2.5 concentration v.s. actual

3) SARIMAX Model

The third model utilized in this analysis is SARIMAX, which incorporates exogenous variables to account for external factors that could potentially impact PM2.5 concentration levels. One such variable that was included in our model is temperature. The AIC value for this model was calculated to be 503.562, which is indicative of a relatively good fit to the data. And Table 2 is our SARIMAX regression result.

	coef	std err	z	P>|z|	[0.025, 0.975]	
ar.L1	-0.1960	0.106	-1.844	0.065	-0.404	0.012
ar.L2	-0.2483	0.145	-1.713	0.087	-0.532	0.036
ar.S.L1	-0.6187	0.098	-6.346	0.000	-0.810	-0.428
ar.S.L2	-0.3903	0.082	-4.734	0.000	-0.552	-0.229
sigma2	290.5605	47.395	6.131	0.000	197.669	383.452
Ljung-Box (L1) (Q):		0.27		Jarque-Bera (JB):		2.56
Prob(Q):		0.600		Prob(JB):		0.28
Heteroskedasticity (H):		0.36		Skew:		0.14
Prob(H) (two-sided):		0.03		Kurtosis:		3.99

Table 2: SARIMAX result

The model generated a forecast of PM2.5 concentration levels, which was compared to the actual values. Figure 15 depicts the forecasted values alongside the observed concentrations. Overall, the SARIMAX model with temperature as an exogenous variable appears to provide a reasonable estimate of PM2.5 concentrations.

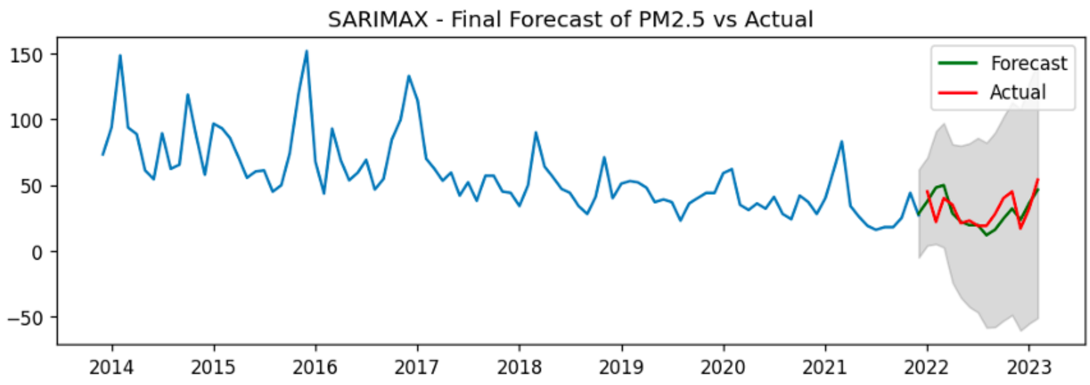


Figure 15: Forecast PM2.5 concentration v.s. actual

4) LSTM Model

The final model we employed in our analysis is the LSTM model, which is a type of machine-learning algorithm. Unfortunately, the results obtained from this model were rather poor, we can see clearly in Figure 16. This is primarily due to the limited amount of data available for training the model, which hindered its ability to effectively learn the underlying patterns and relationships within the data.

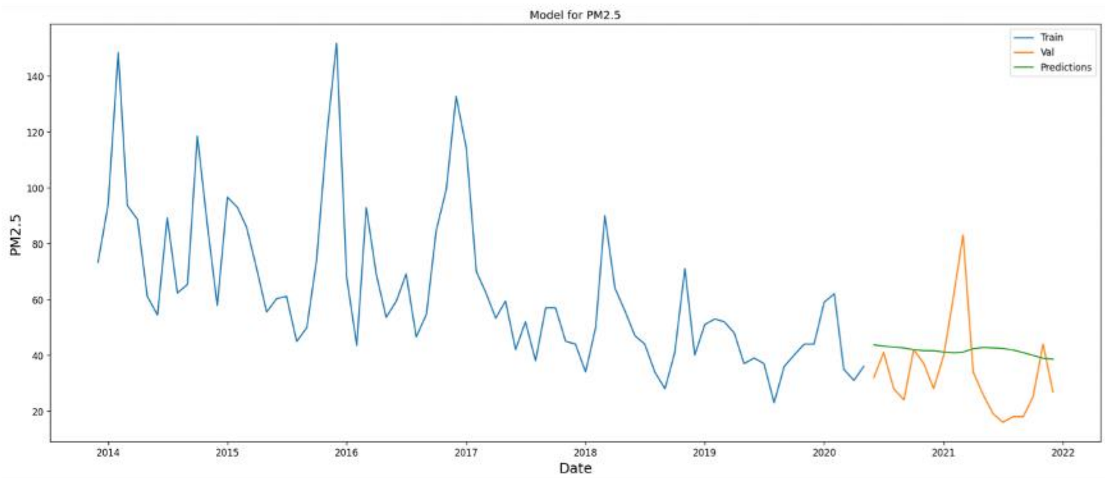


Figure 16: Forecast PM2.5 concentration v.s. actual

As a result, the LSTM model was unable to generate accurate forecasts of PM2.5 concentrations. While this outcome is disappointing, it highlights the importance of having sufficient data for training machine learning models, particularly for complex and sophisticated algorithms such as LSTM.

5) Model Comparison

To compare the prediction accuracy of the four models we have utilized, we can use the Root Mean Square Error (RMSE) metric. The RMSE values obtained for the ARIMA, SARIMA, and SARIMAX models were 16.338, 11.724, and 10.724, respectively.

Model	**AIC**	**BIC**	**RMSE**
ARIMA	803.133	813.391	16.338
SARIMA	771.740	783.894	11.724
SARIMAX	503.562 *	513.864 *	10.724 *

Table 3: Model comparison

From Table 3, we can see that the SARIMAX model produced the lowest RMSE, indicating that it had the highest level of prediction accuracy among the four models. This finding is further supported by the AIC values and BIC values obtained for the models, which also indicated that the SARIMAX model was the best fit for the data.

Therefore, we can confidently conclude that the SARIMAX model is the most effective of the four models we employed for predicting PM2.5 concentrations.

6) Final Model

After analyzing the SARIMAX model, we examined the residual plot (Figure 17) and found that the residuals appeared to be random and resembled white noise. This indicates that the model was well-fitted to the data and effectively captured the underlying patterns and trends.

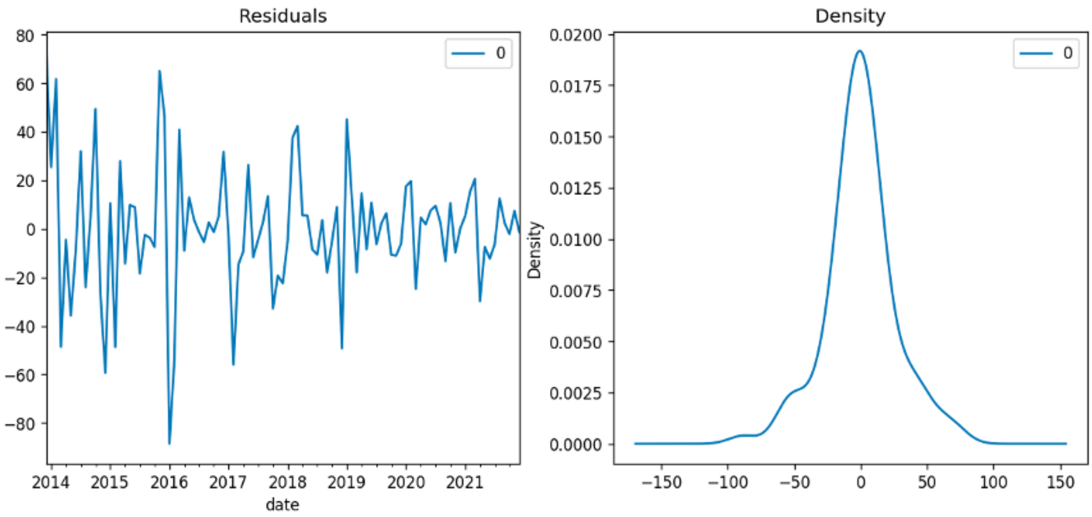


Figure 17: Residual & density plot

Additionally, we generated a forecast of PM2.5 concentrations using the SARIMAX model and compared the forecasted values to the actual concentrations. Table 4 depicts the forecasted values alongside the observed concentrations, and Figure 18 shows our final PM2.5 concentration prediction versus actual PM2.5 concentration.

	Actual	Forecast
Jan-22	45	37.5
Feb-22	22	48.1
Mar-22	40	49.8
Apr-22	35	28.4
May-22	21	22.2
Jun-22	23	19.5
Jul-22	19	19.6
Aug-22	19	11.9
Sep-22	28	16.1
Oct-22	40	24.4
Nov-22	45	32.1
Dec-22	17	23.6
Jan-23	32	36.2
Feb-23	54	46.4

Table 4: Actual PM2.5 concentration value v.s. Forecast

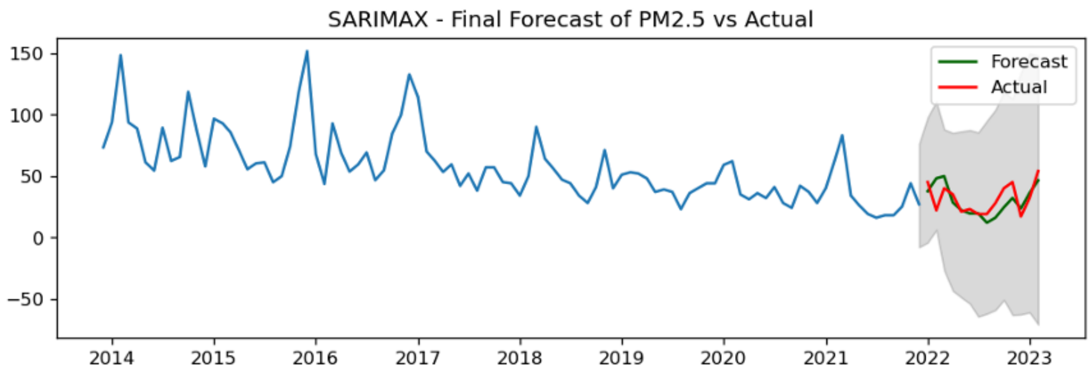


Figure 18: Forecast PM2.5 concentration v.s. actual

V. Conclusion

Our research aims at investigating the trend of the PM 2.5 concentration in Beijing and making predictions on the future PM 2.5 value in Beijing in 2023. As we can see from the de-seasonalized data, there is a downward trend of PM2.5 in Beijing over time. There are a lot of scholars have examined air pollutants using different statistical models and machine learning models. We use the ARIMA model, the seasonal ARIMA model, the seasonal ARIMAX model, and the LSTM model to predict the PM 2.5 concentration in Beijing in 2023 and find that the SARIMAX model provides the best fit for the data. The forecasted PM 2.5 value quite aligns with the true PM 2.5 value in the previous two months of 2023. Our forecast suggests that this trend will continue in the future. This suggests that the implemented policy is quite effective and supportive of achieving sustainable development goals.

However, there are two major limitations in our research:

The first one is that we only include temperate as an exogenous variable in the SARIMAX model. However, there are other exogenous variables, such as wind and humidity, that may affect PM2.5 emissions. In the future, we can include them as additional variables in our SARIMAX model to investigate their impact.

The second one is that we use monthly data in our model for the forecast. However, the dataset is not large enough to use machine learning methods, such as artificial neural networks (ANN), and is not large enough to better train the data in the LSTM model. In the future, we can find daily PM2.5 data to increase our data size and use more models for forecasting to examine whether other models that could give us better forecast results.

VI. Reference

Chakraborty, S. and Mondal, A. (2019). A review on time series forecasting of air pollutants. *Journal of Environmental Management, 233*, pp. 610-620.

Chen, Y., Cui, L., Chen, H., and Chen, J. (2020). Spatiotemporal prediction of PM2.5 concentrations in Beijing using machine learning algorithms. *Atmospheric Pollution Research, 11(2)*, pp. 166-173.

Cui, X and Hao, Y. (2018). Short-term air quality forecasting: An overview of methods and applications. *Journal of Environmental Sciences, 71*, pp. 6-17.

Freeman, B. S., Taylor, G., Gharabaghi, B., & Thé, J. (2018). Forecasting air quality time series using deep learning. *Journal of the Air & Waste Management Association, 68(8)*, 866-886.

Guo, X., Wang, L. and Feng, L. (2020). A review of time series forecasting for air quality: models, methods, and applications. *Journal of Cleaner Production, 246*, 118913.

Hao, Y., & Cui, X. (2019). A hybrid model for air quality forecasting: A case study in Beijing. *Atmospheric Pollution Research, 10(2)*, pp. 527-535.

He, K.; Zhang, Qiang; Ming, D.; Wu, Ye; Catherine, W.; Valentin, F.; Han, Y.; Cheng, J. and Qu, Y. (2019). A Review of 20 Years' Air Pollution Control in Beijing. *UN Environment*, 978-92-807-3743-1.

Hu, Q., Li, X., Lin, A., Qi, W., Li, X., & Yang, X. J. (2018). Total emission control policy in China. *Environmental Development, 25*, 126-129.

Jiang, S., Tan, X., Hu, P., Wang, Y., Shi, L., Ma, Z., & Lu, G. (2022). Air pollution and economic growth under local government competition: Evidence from China, 2007–2016. *Journal of Cleaner Production, 334*, 130231.

Liu, J., Lin, X., Zhou, Y., Zhang, J., and Li, M. (2019). The potential of machine learning in predicting air pollution trend: A review. *Journal of Cleaner Production, 220*, pp. 202-209.

Li, S., Feng, K., & Li, M. (2017). Identifying the main contributors of air pollution in Beijing. *Journal of Cleaner Production, 163*, S359-S365.

Nataraj, S., Chari, R., Richardson, A., & Willis, H. H. (2013). *Links Between Air Quality and Economic Growth: Implications for Pittsburgh.* Rand Corporation.

Rafaj, P., Kiesewetter, G., Gül, T., Schöpp, W., Cofala, J., Klimont, Z., ... & Cozzi, L. (2018). Outlook for clean air in the context of sustainable development goals. *Global Environmental Change*, *53*, 1-11.

Wang, Y., Gu, X., and He, J. (2019). Predicting the daily PM2.5 concentration in Beijing using machine learning and meteorological data. *Journal of Environmental Sciences, 76*, pp. 320-329.

Xu, X., Ma, Z., Zhang, L., and Chen, L. (2016). A comparison of time series models for forecasting fine particulate matter concentration in Beijing. *Environmental Science and Pollution Research, 23(2)*, pp. 1598-1607.

Yin, P., Brauer, M., Cohen, A. J., Wang, H., Li, J., Burnett, R. T., ... & Murray, C. J. (2020). The effect of air pollution on deaths, disease burden, and life expectancy across China and its provinces, 1990–2017: an analysis for the Global Burden of Disease Study 2017. *The Lancet Planetary Health*, 4(9), e386-e398.

Yin, P., He, G., Fan, M., Chiu, K. Y., Fan, M., Liu, C., ... & Zhou, M. (2017). Particulate air pollution and mortality in 38 of China's largest cities: time series analysis. *bmj, 356*.

Yu, L., Wu, S., Zhang, Y. & Li, M. (2021). Air quality forecasting using machine learning: A review. *Journal of Cleaner Production*, *279*, 123600.

Zhang, Y., Zhang, X., Shi, J. and Song, X. (2017) Forecasting the air quality index in Beijing: A comparison of seasonal ARIMA and SARIMA models. *Atmospheric Environment, 150*, pp. 142-153.

Zheng, Y., Yi, X., Li, M., Li, R., Shan, Z., Chang, E., & Li, T. (2015, August). Forecasting fine-grained air quality based on big data. In *Proceedings of the 21st ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 2267-2276).