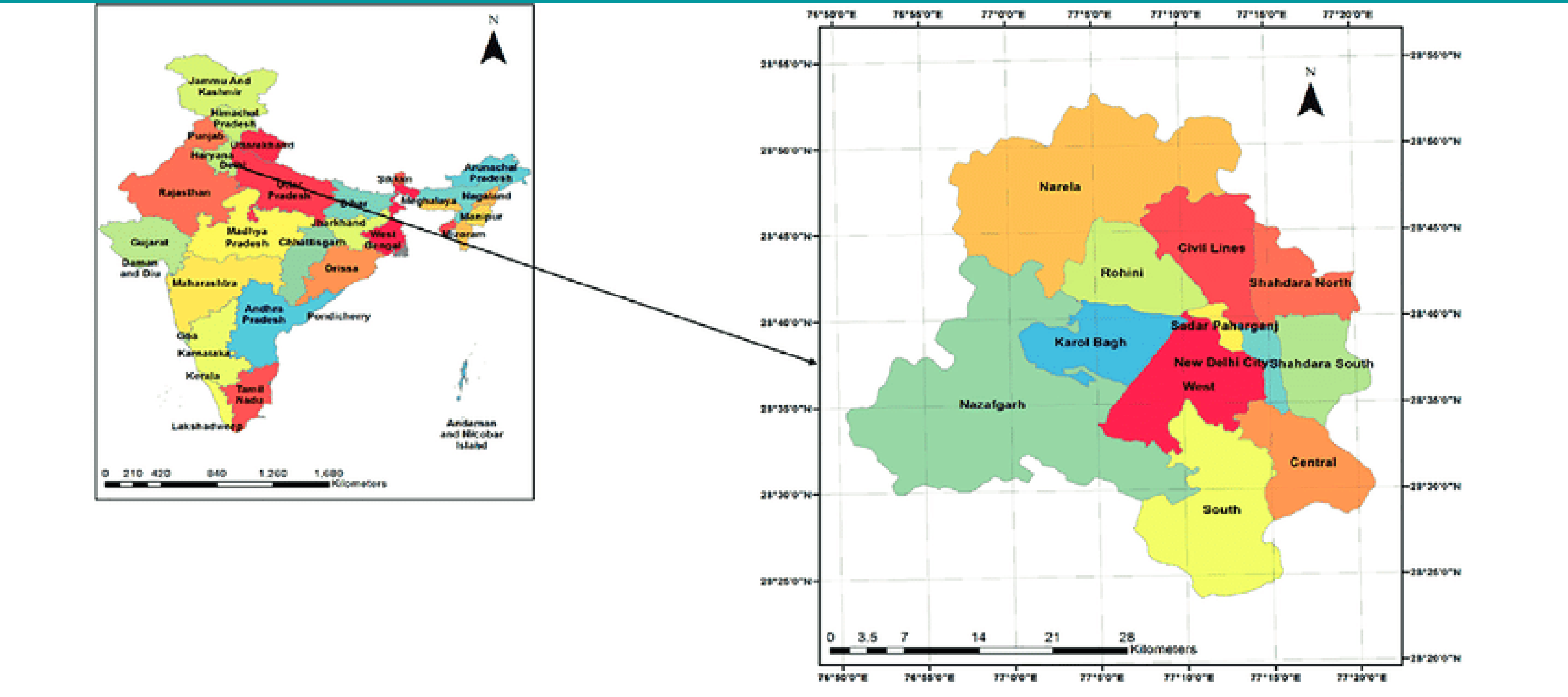
SYMBIOSIS
INSTITUTE OF GEOINFORMATICS
॥वसुधैव कुटुम्बकम्॥

CRACKING THE CODE OF DELHI'S AIR QUALITY: PATTERNS AND PROGRESS

STUDY AREA : DELHI



DTU

Shadipur

NSIR Dwarka

Dilshad Garden

Sirifort

PROBLEM STATEMENT

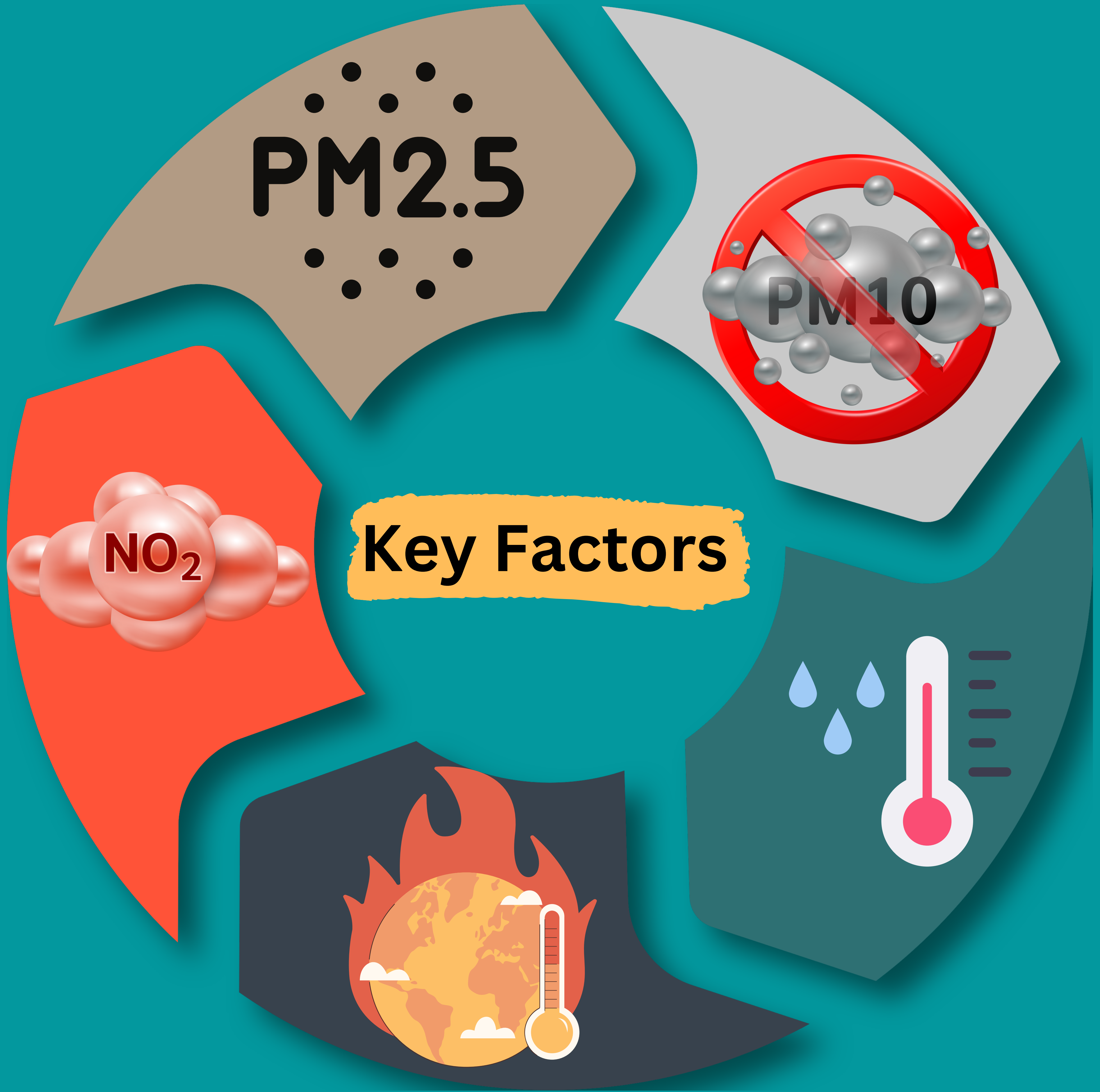
The challenge is to effectively analyze, predict, and visualize air quality parameters (PM2.5, PM10, NO2) over the past 5 years of a specific study area using geostatistical methods like IDW and Kriging, meteorological variables (wind speed) and any other of your choice with correlation, trend analysis, and forecasting models while creating a single legend for Multiple Maps (IDW, Kriging) for the air quality parameter for the study area.

DTU | Shadipur | NSIR Dwarka | Dilshad Garden | Sirifort

OBJECTIVE

The objective of this study is to comprehensively analyze, predict, and visualize the variations in air quality parameters, including PM2.5, PM10, and NO2, within the designated study area over the course of the past 5 years. Leveraging geostatistical methods such as Inverse Distance Weighting (IDW) and Kriging, as well as incorporating meteorological variables Temp and Rainfall

DTU Shadipur NSIR Dwarka Dilshad Garden Sirifort


PM2.5
PM10
NO_2
Key Factors

SPATIAL INTERPOLATION

IDW

KRIGING

IDW

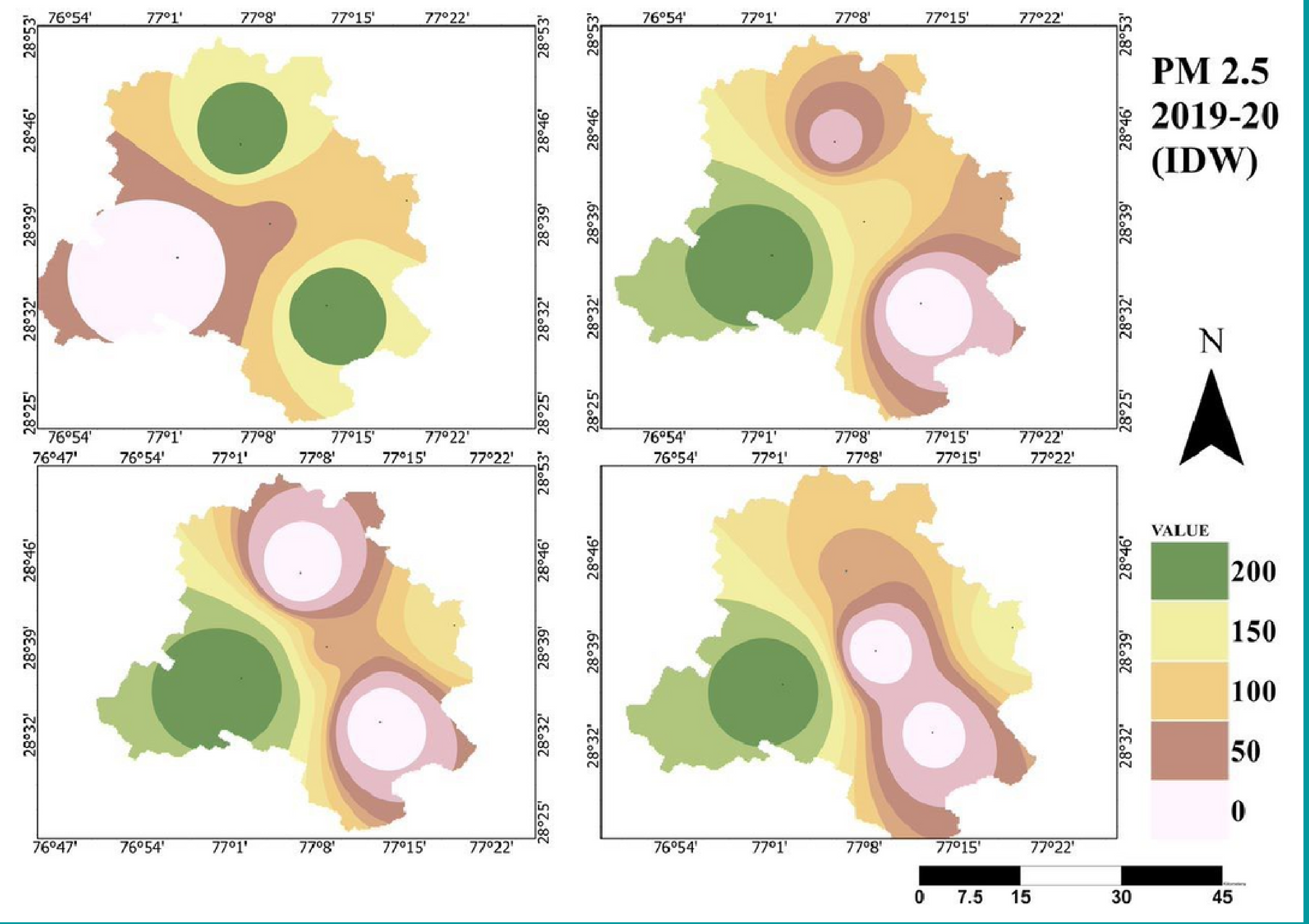
Inverse Distance Weighting (IDW) is a geostatistical interpolation method that plays a pivotal role in understanding the distribution of air quality parameters across Delhi. In this approach, the values of a particular location are estimated based on the weighted average of its neighboring data points. The weights assigned to these points are inversely proportional to their distances from the target location.

DTU Shadipur NSIR Dwarka Dilshad Garden Sirifort

KRIGING

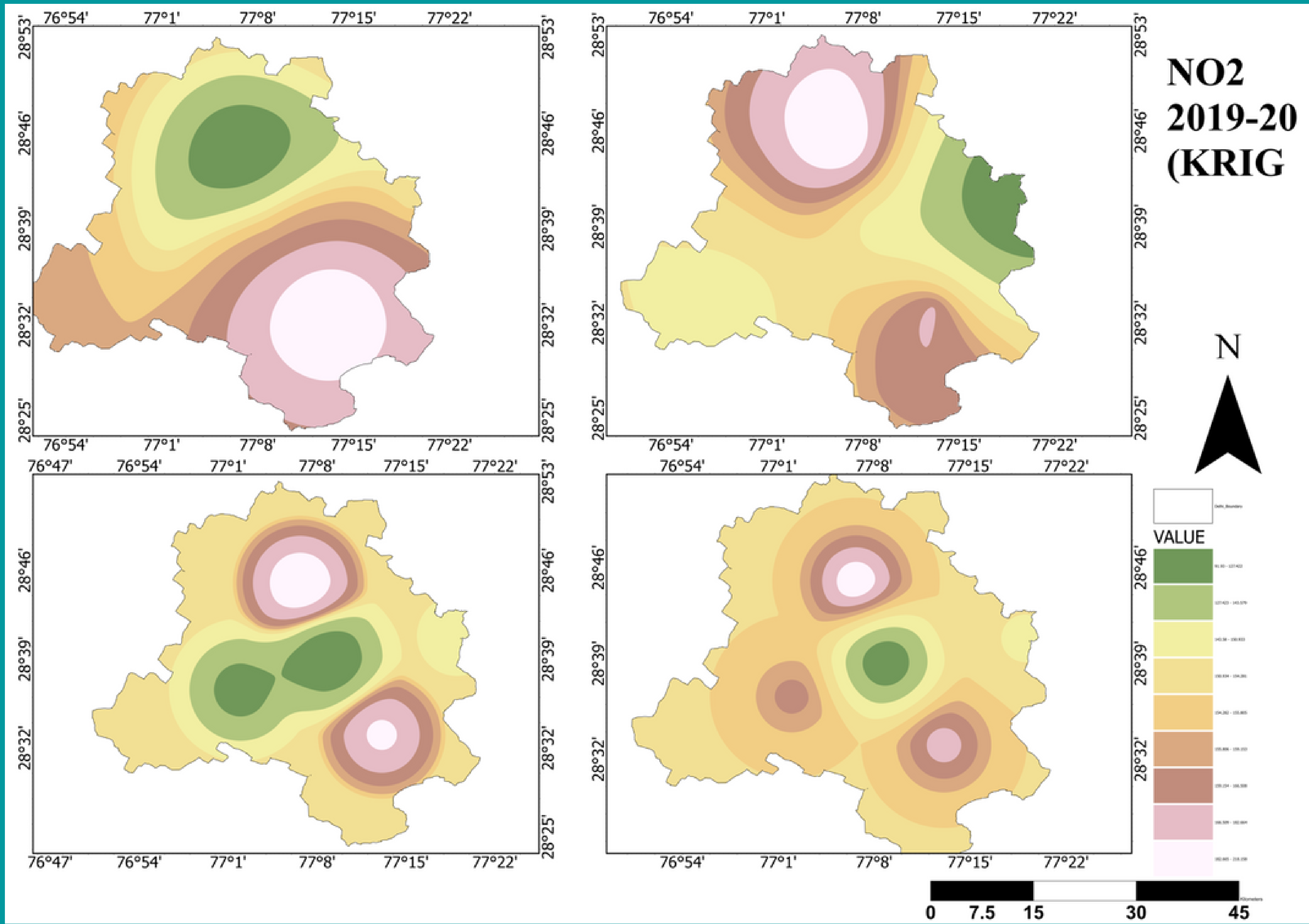
Kriging is a powerful geostatistical interpolation method used to predict air quality values across Delhi by incorporating both spatial correlation and variability information from monitoring stations. It provides a more sophisticated approach compared to simple methods like IDW.

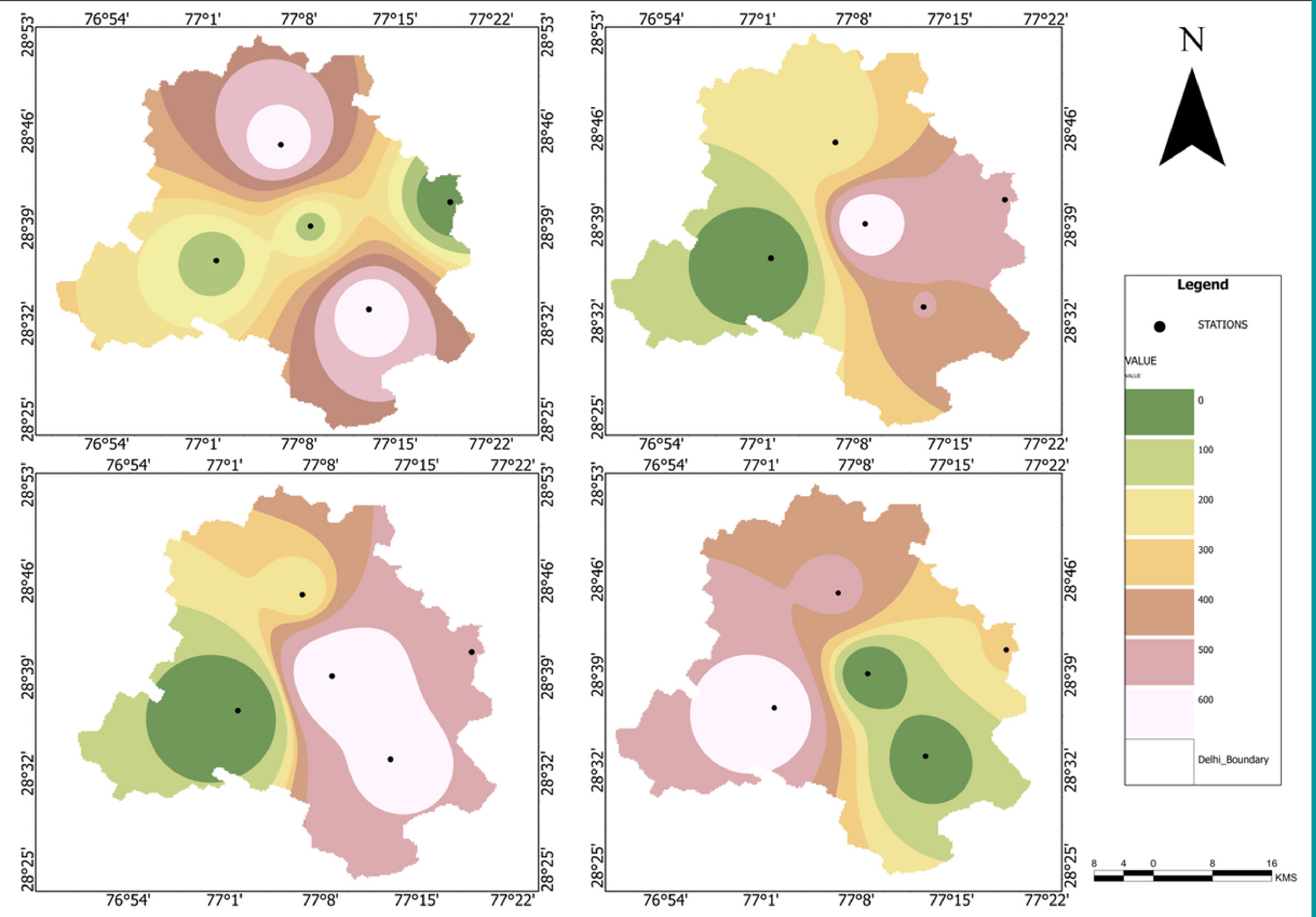
DTU　Shadipur　NSIR Dwarka　Dilshad Garden　Sirifort



KRIGING(PM2.5)

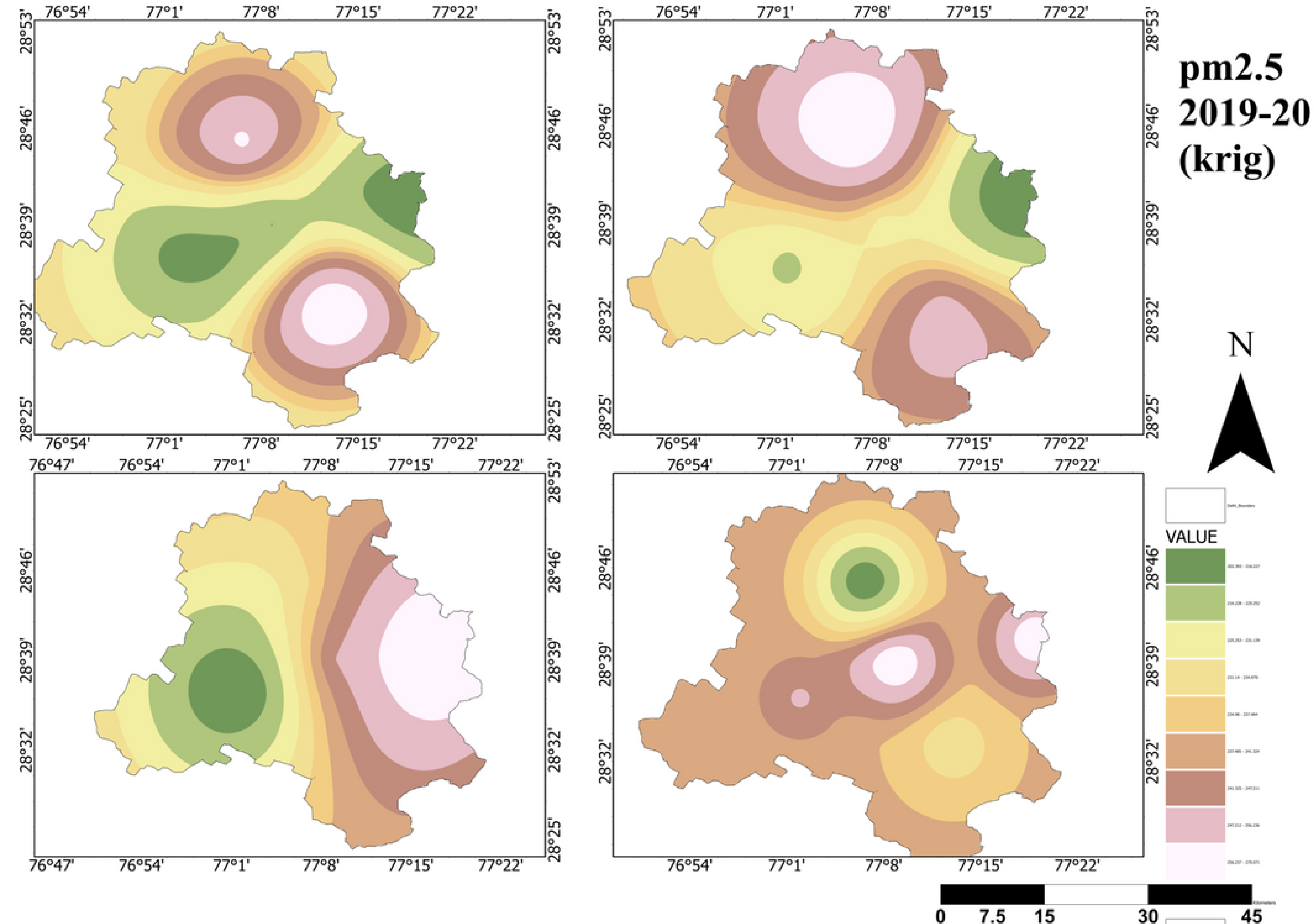
IDW(PM2.5)

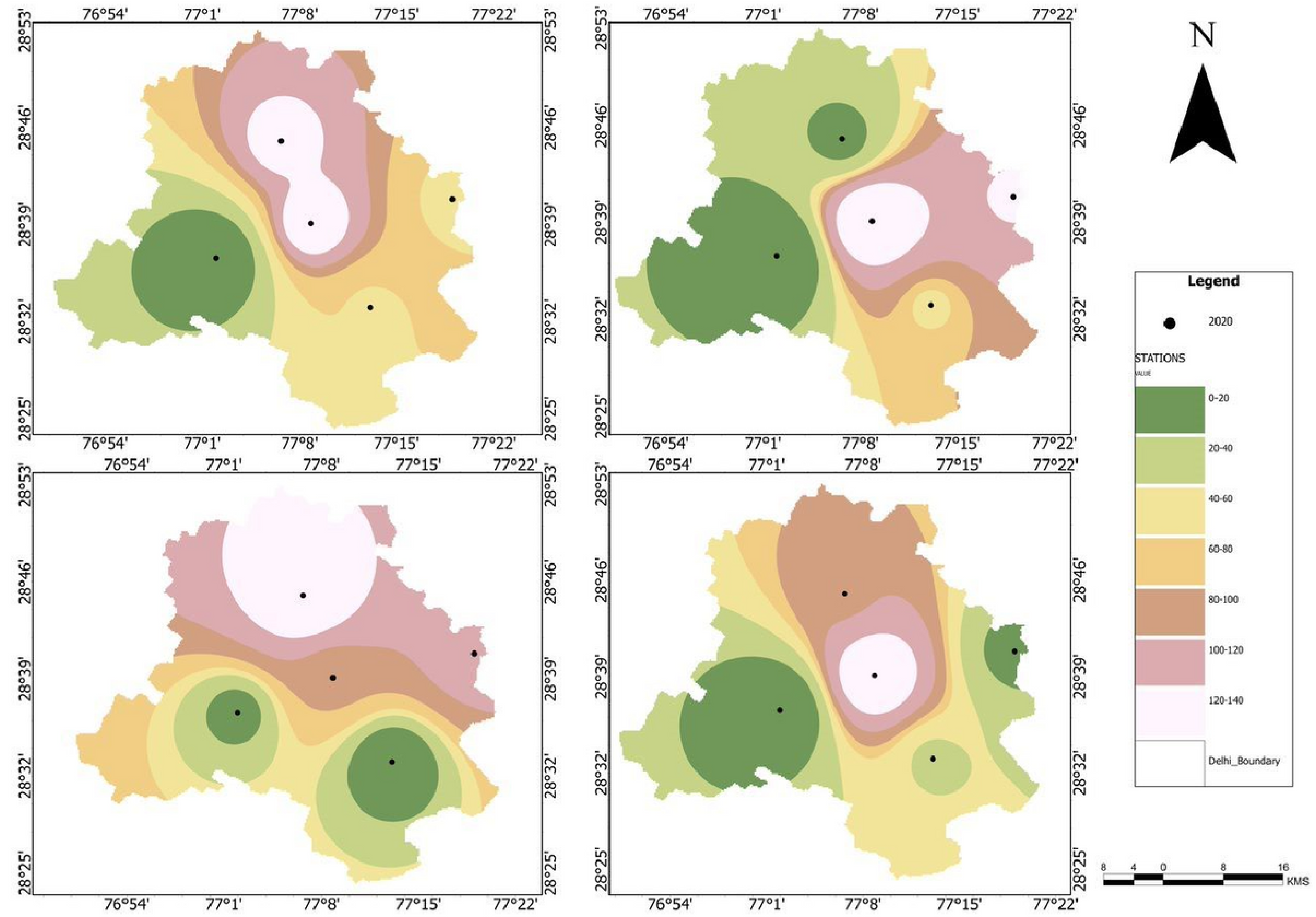




KRIGING(PM10)

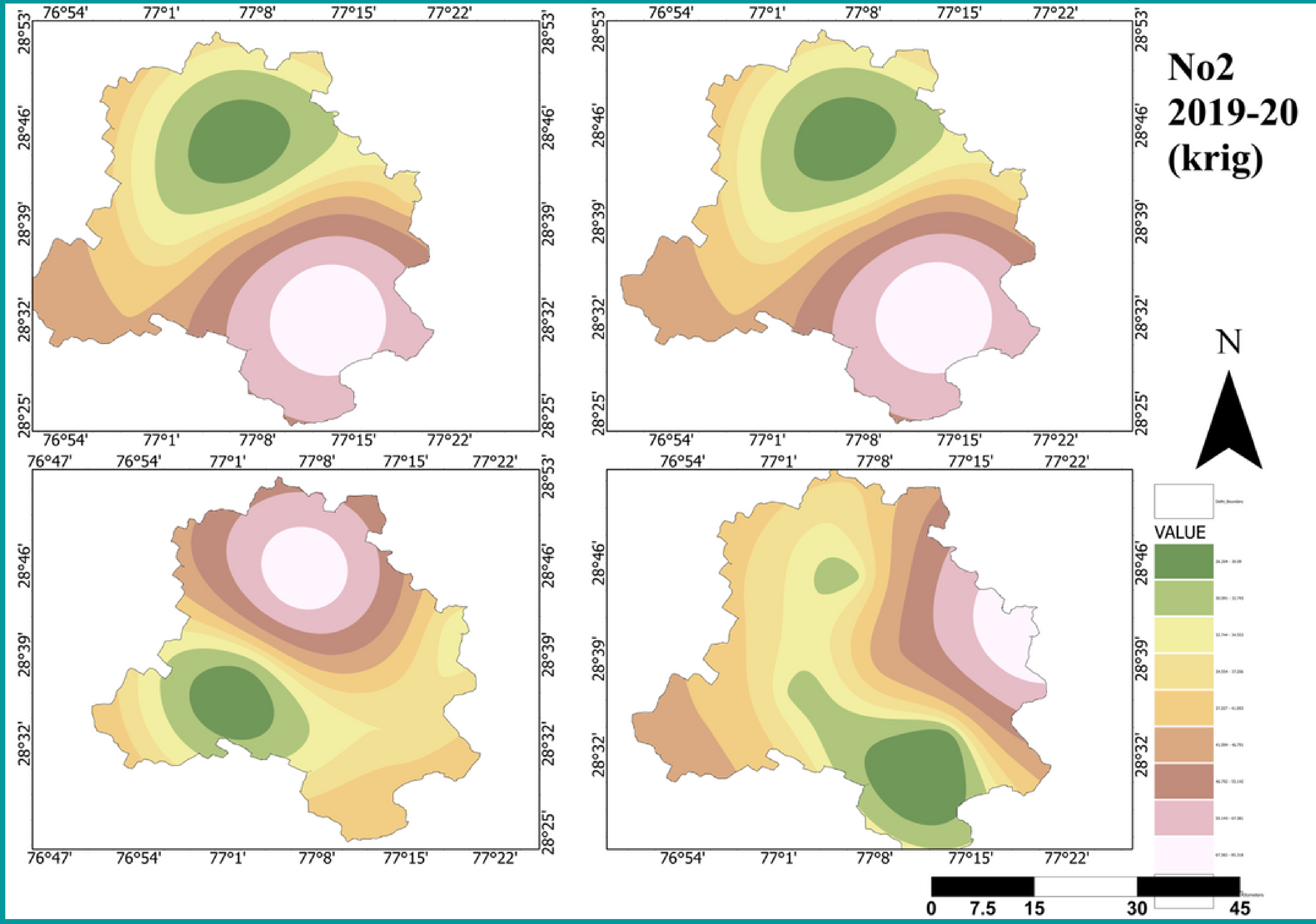
IDW(PM10)





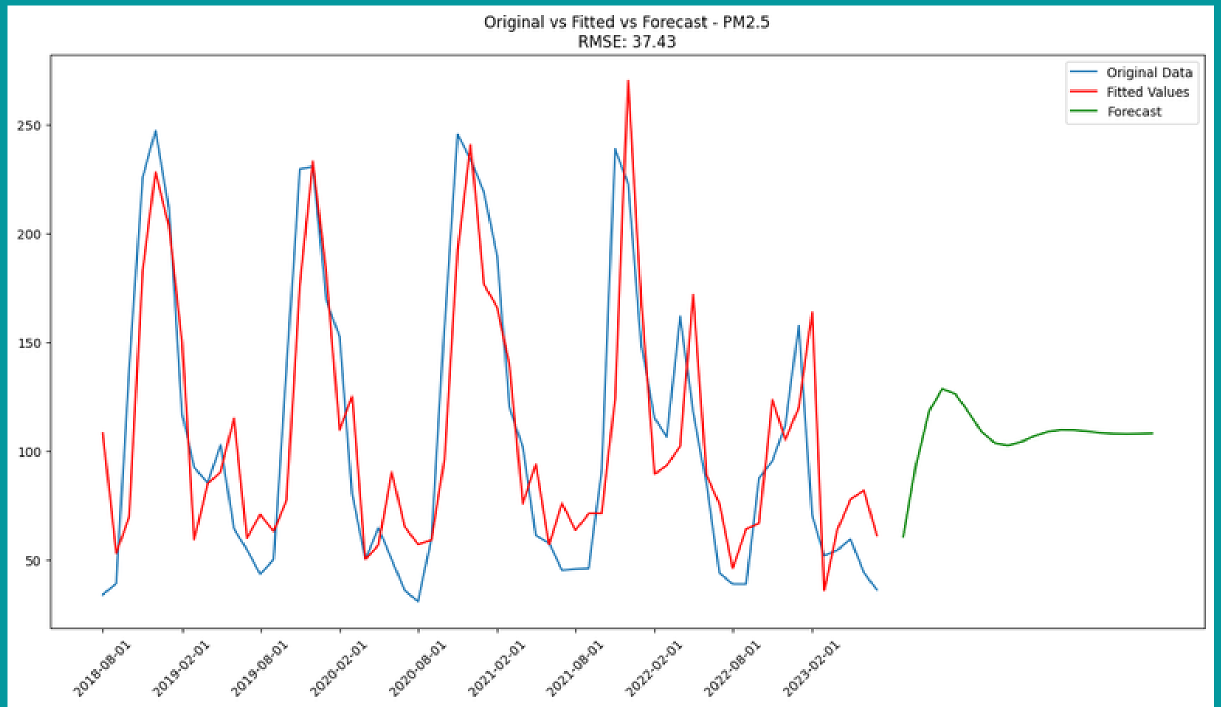
IDW(NO2)

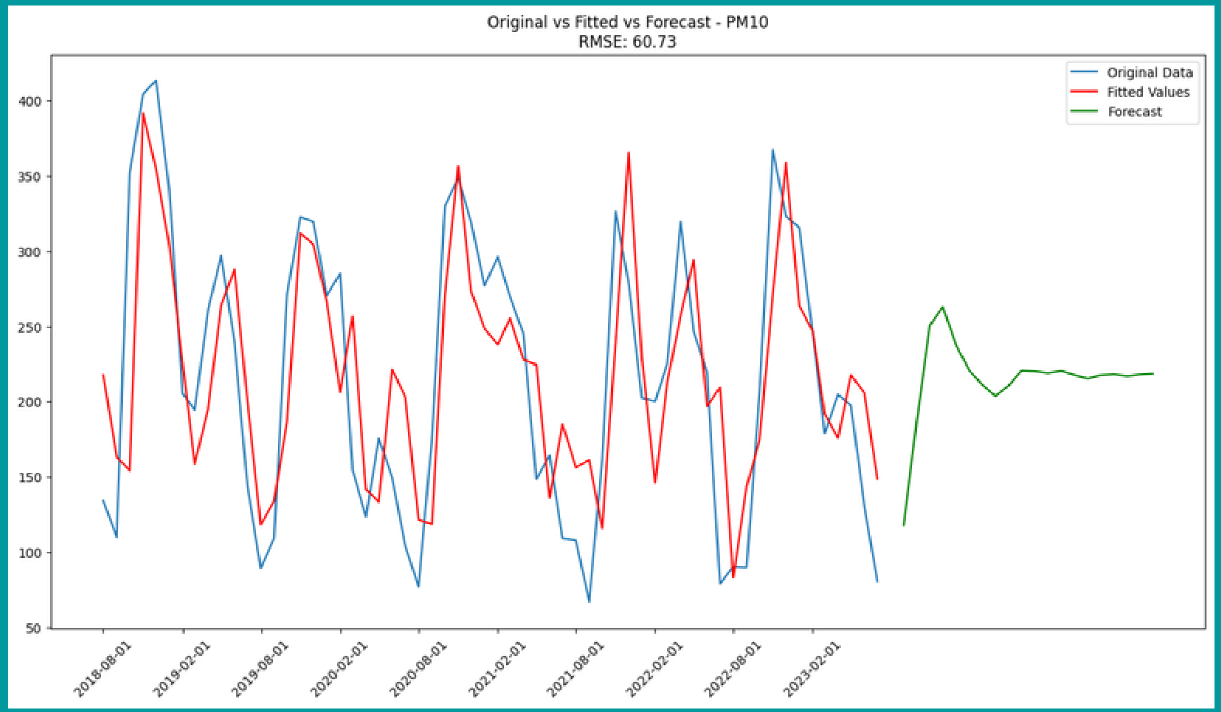
KRIGING(NO2)



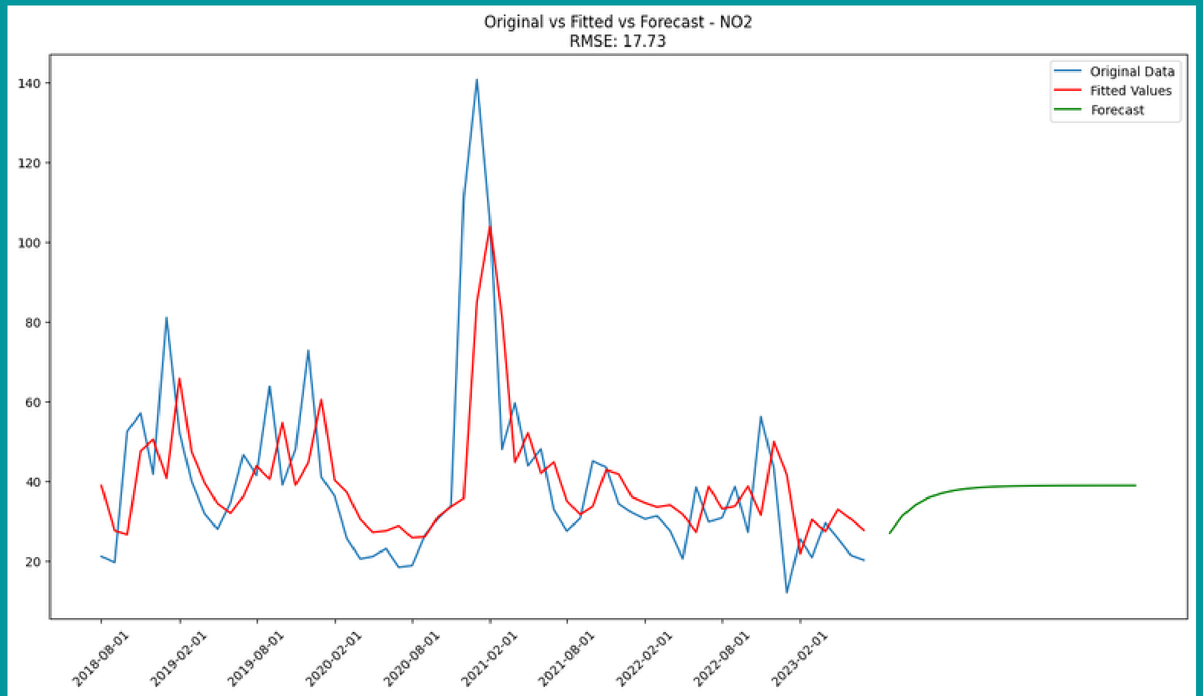
ERROR VALUES

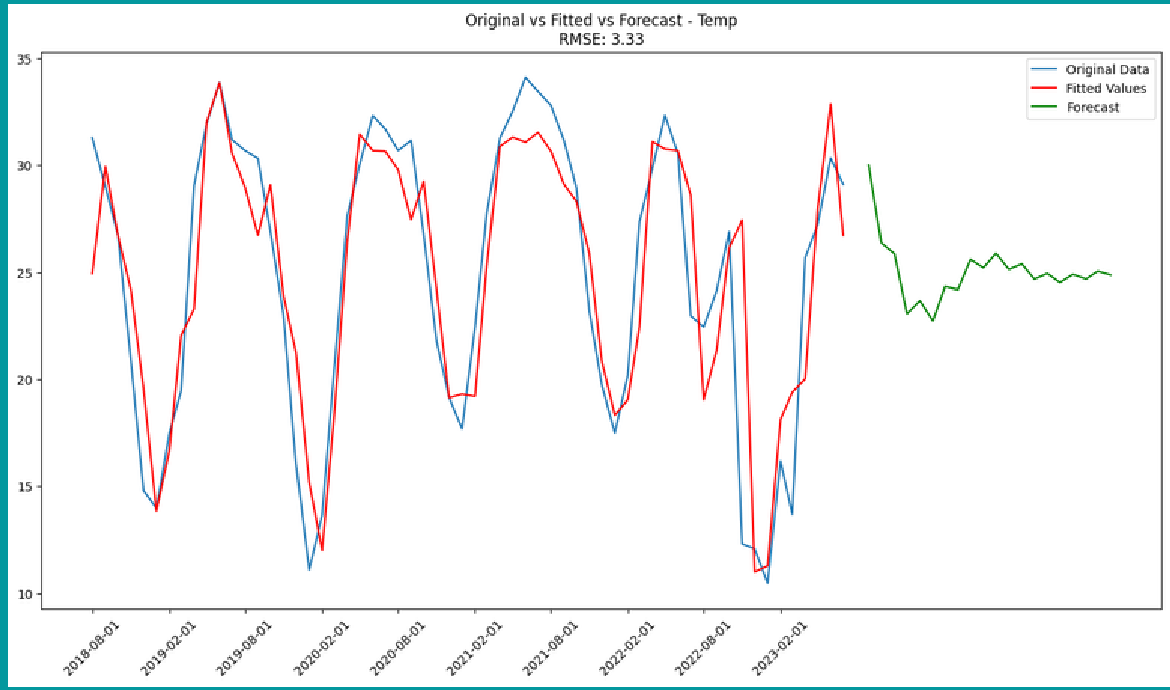
		RMSE VALUES		
		PM 2.5	PM 10	No2
2019	IDW	4.09	34.22	15.32
	KRIGING	9.37	25.86	11.81
2020	IDW	20.37	21.21	14.72
	KRIGING	14.51	16.17	11.81
2021	IDW	20.88	17.52	11.35
	KRIGING	14.42	11.94	9.36
2022	IDW	11.78	13.53	14.95
	KRIGING	10.56	11.91	11.35

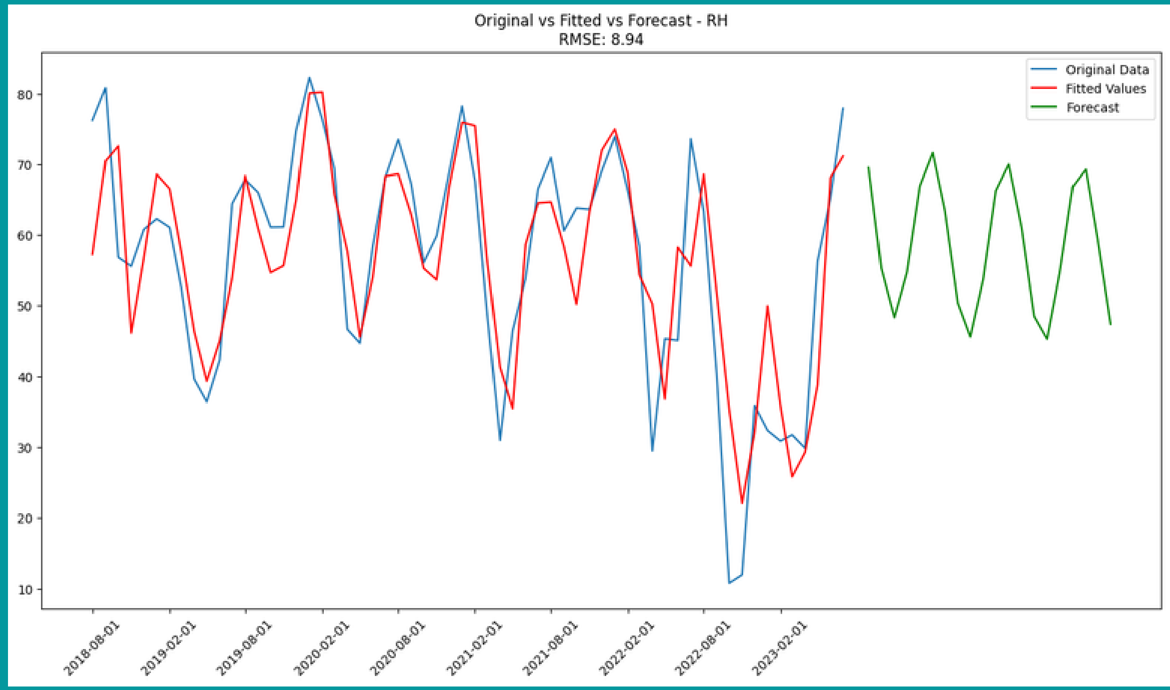

Original vs Fitted vs Forecast - PM2.5
RMSE: 37.43
Original Data
Fitted Values
Forecast


Original vs Fitted vs Forecast - PM10
RMSE: 60.73
Original Data
Fitted Values
Forecast

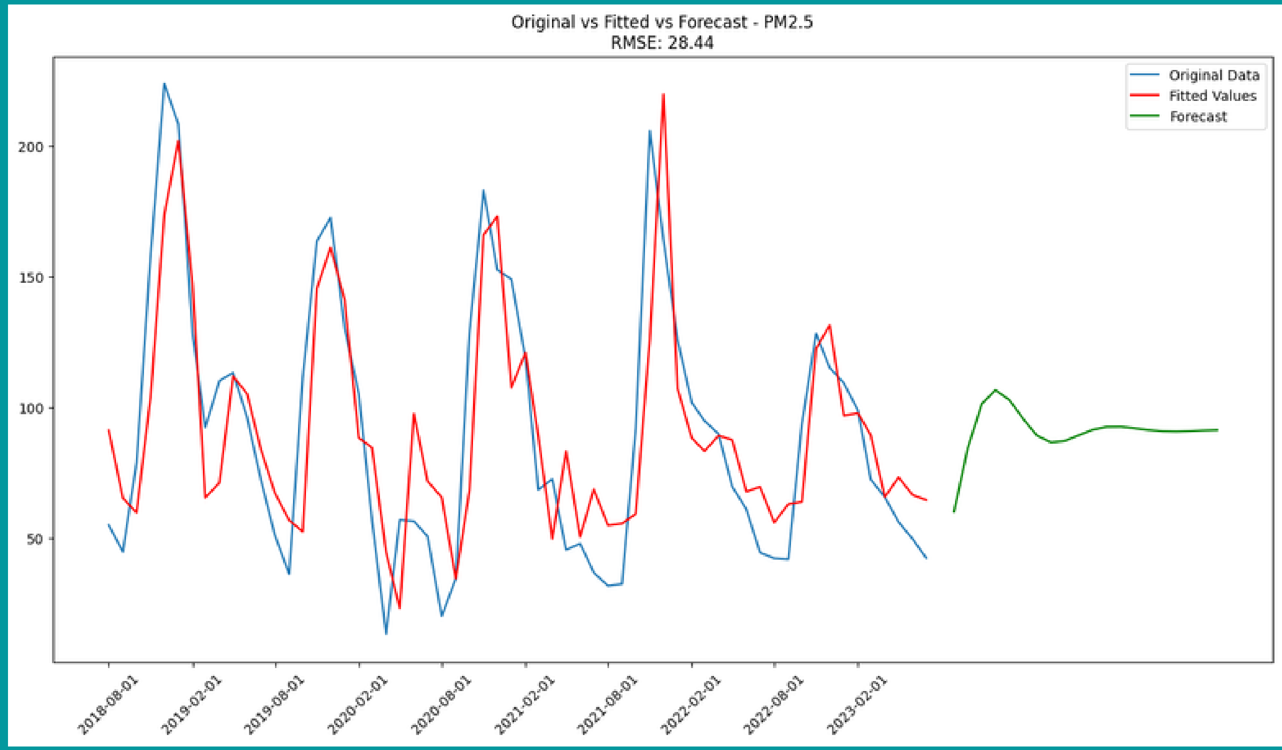
DTU

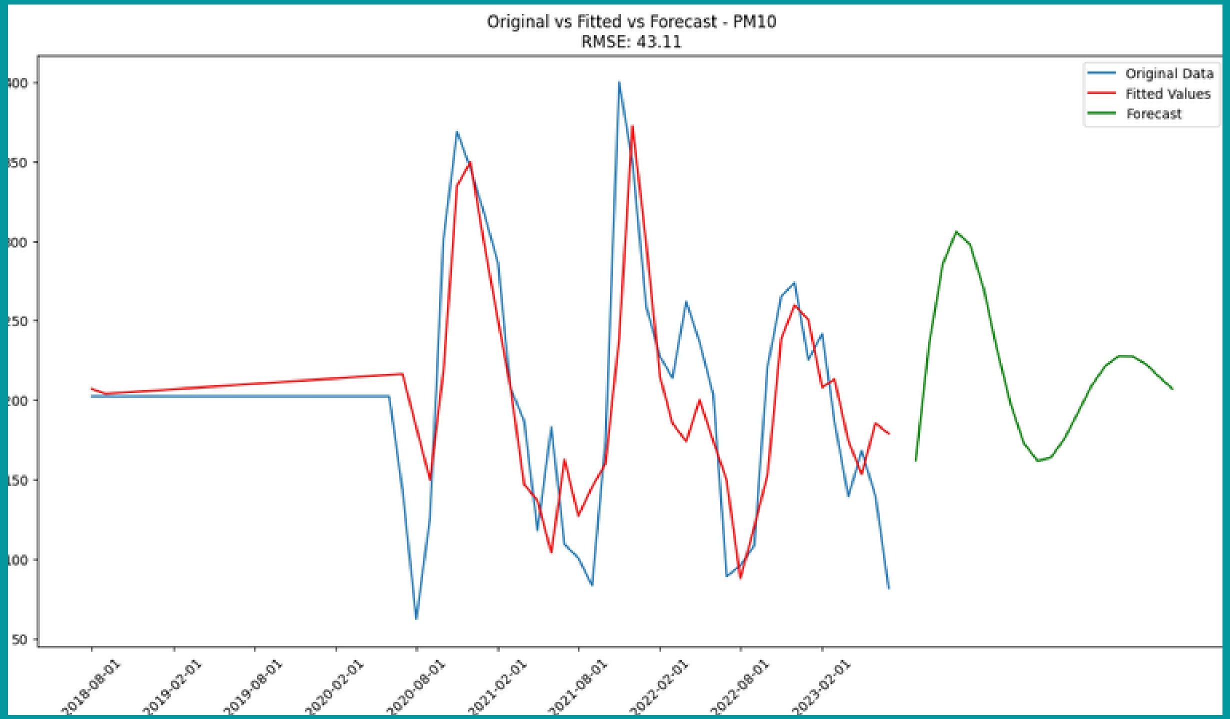

Original vs Fitted vs Forecast - NO2
RMSE: 17.73
Original Data
Fitted Values
Forecast

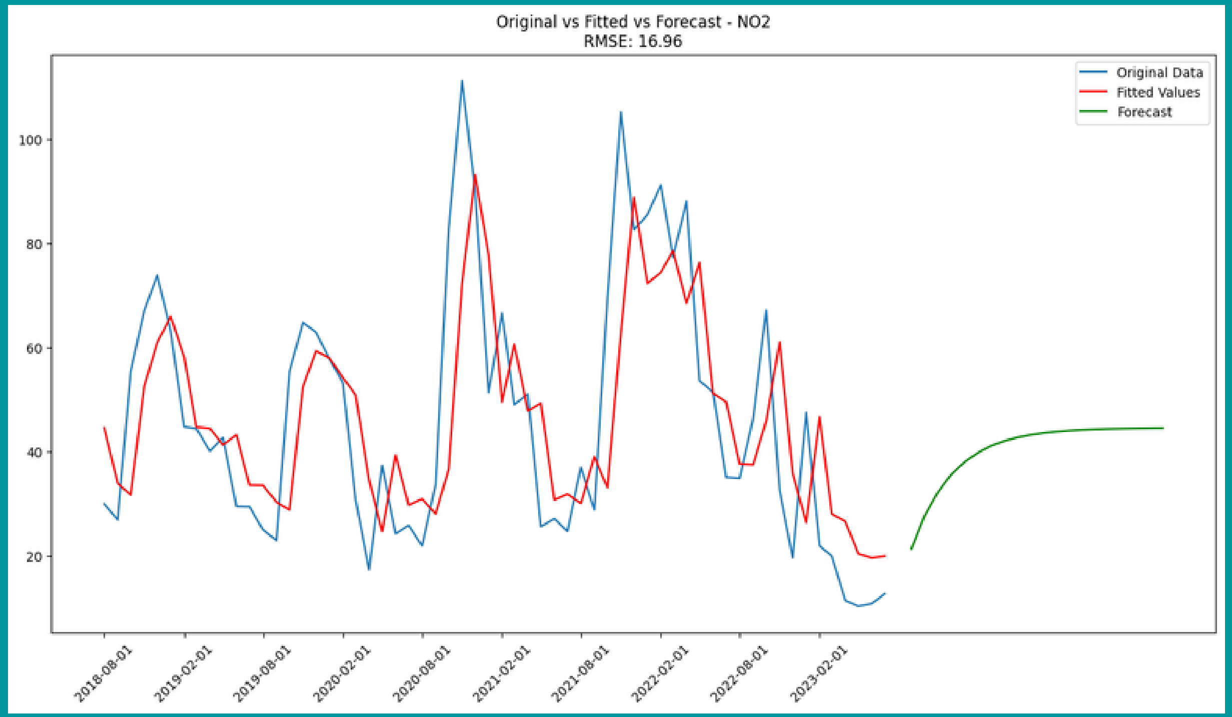

Original vs Fitted vs Forecast - Temp
RMSE: 3.33
Original Data
Fitted Values
Forecast

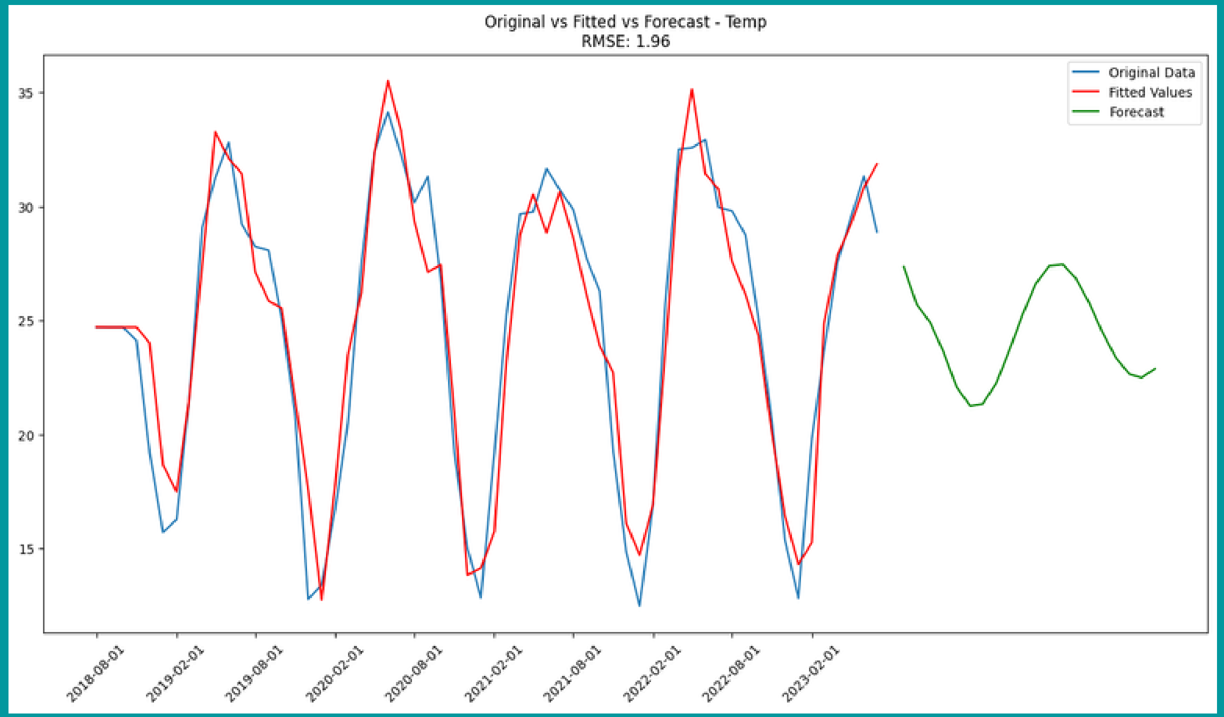

Original vs Fitted vs Forecast - RH
RMSE: 8.94
Original Data
Fitted Values
Forecast

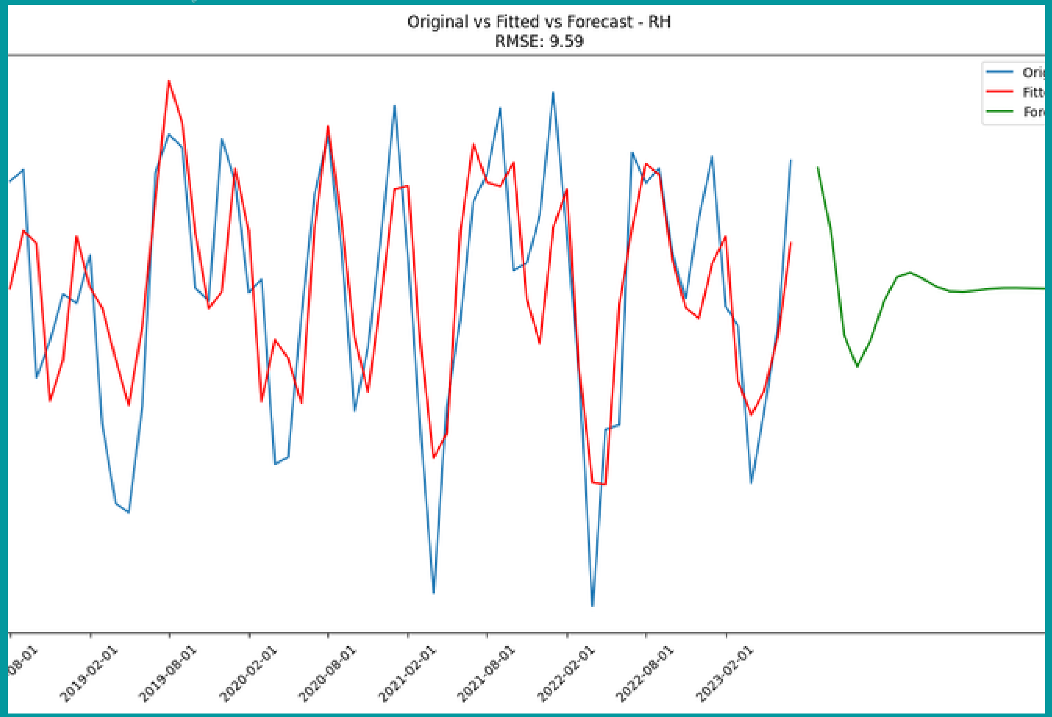
DILSHAD GARDEN



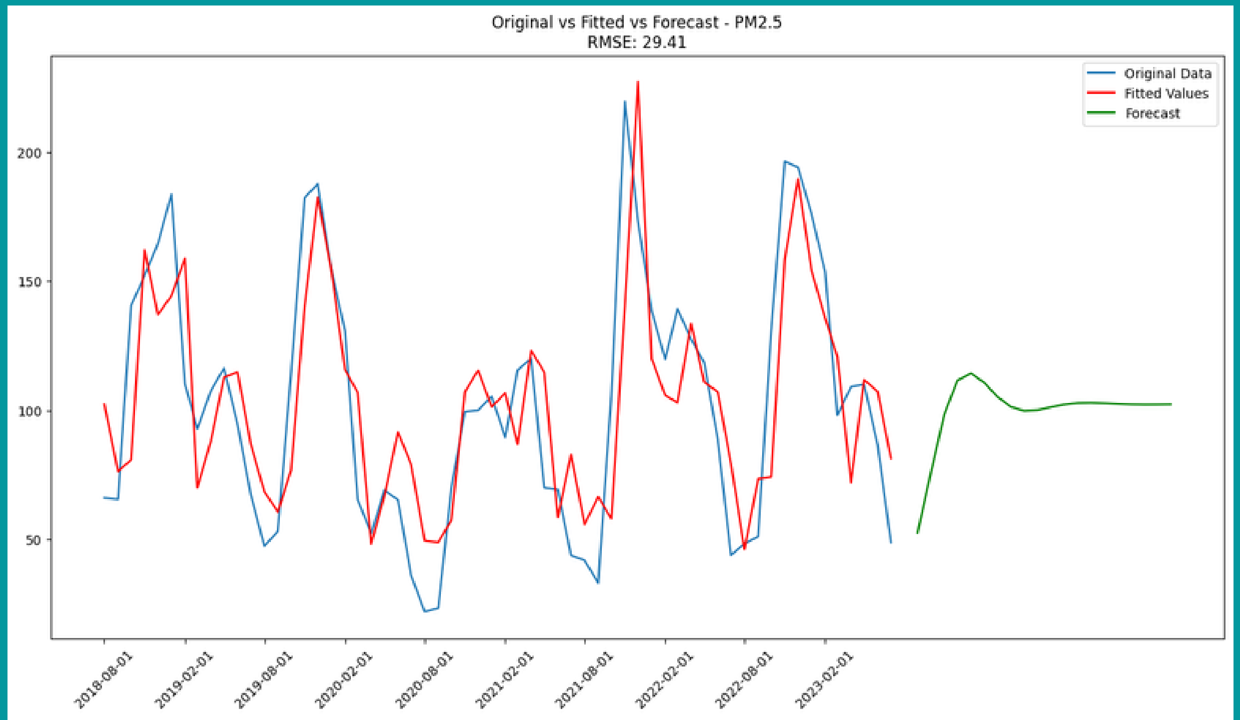


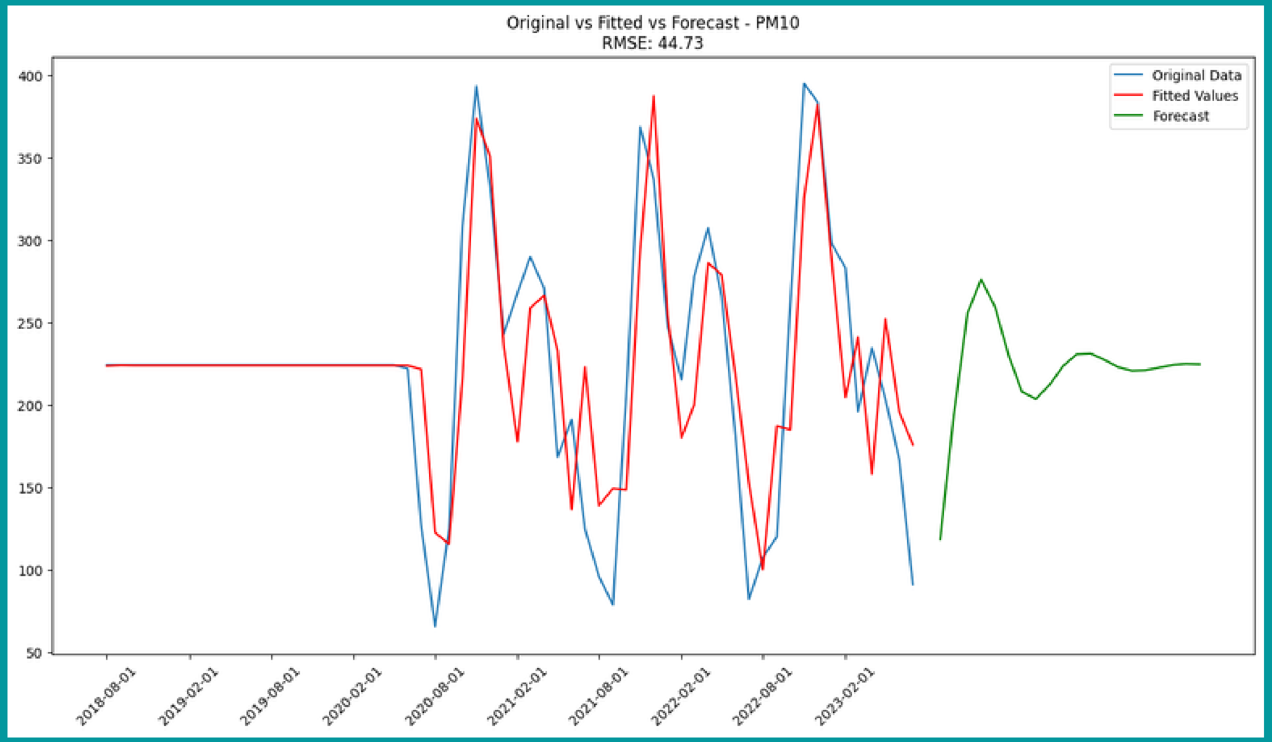


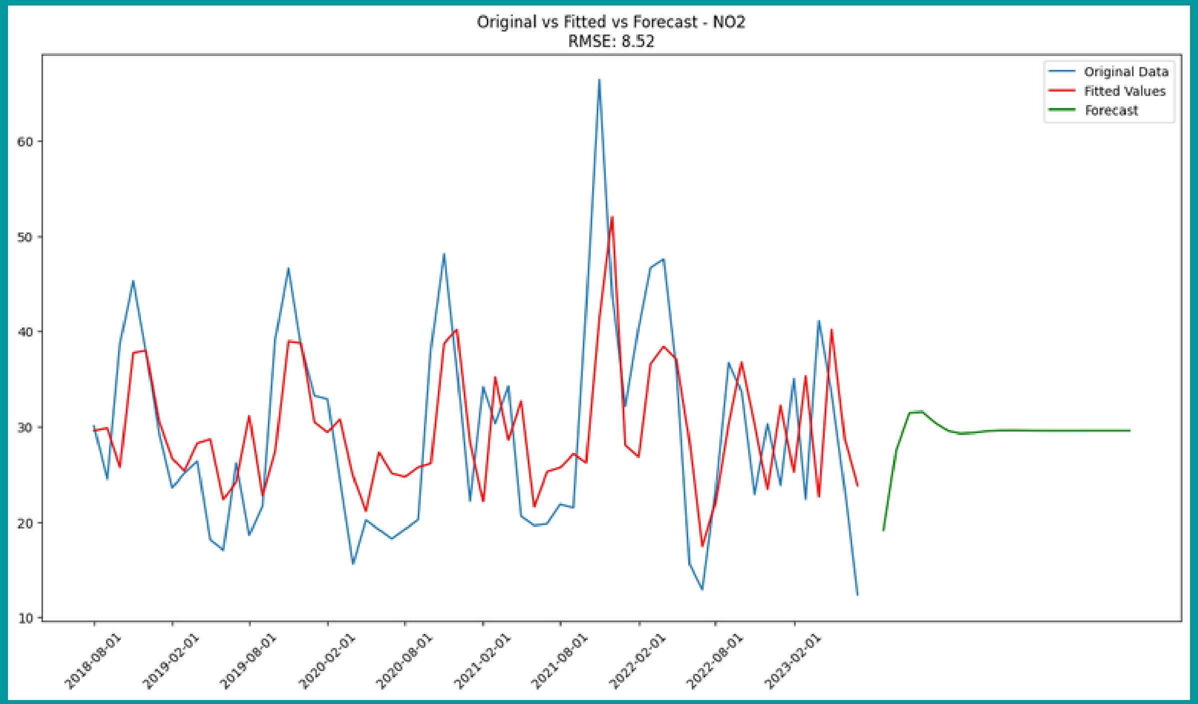


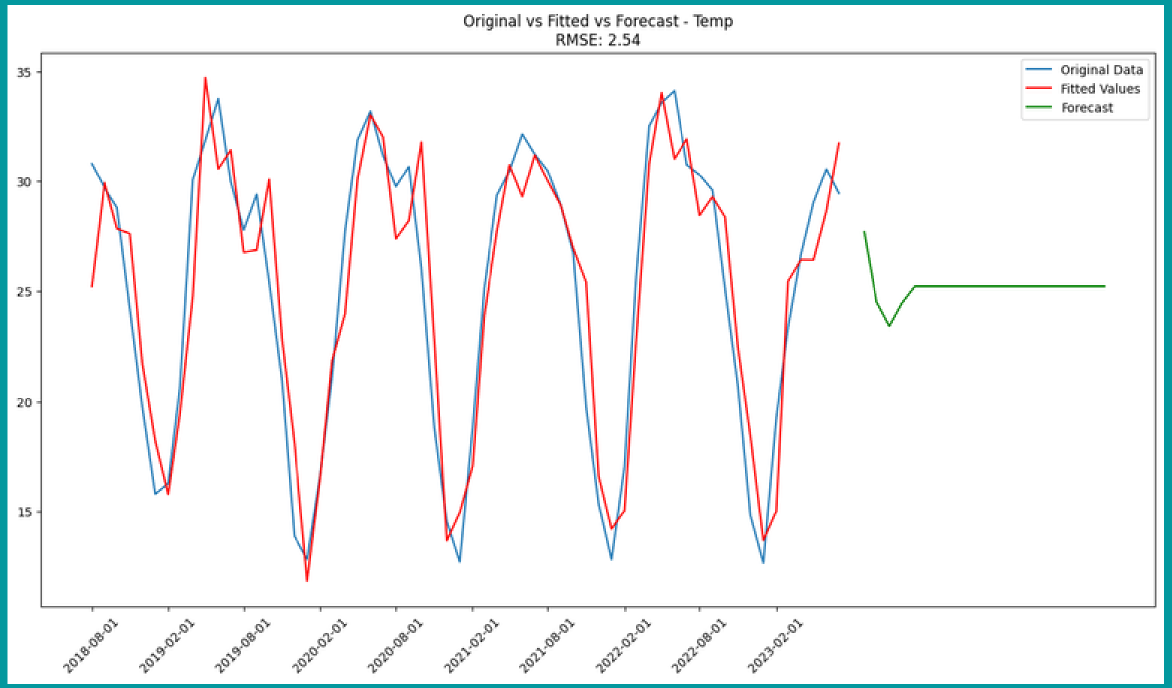


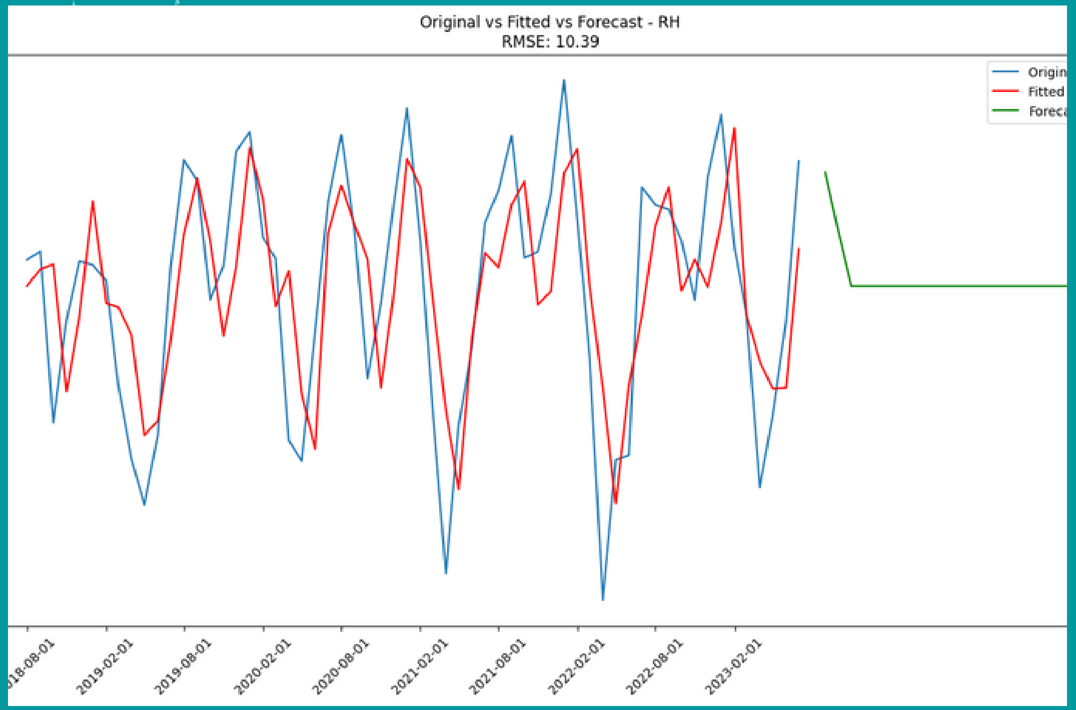
DWARKA


Original vs Fitted vs Forecast - PM2.5
RMSE: 29.41
Original Data
Fitted Values
Forecast


Original vs Fitted vs Forecast - PM10
RMSE: 44.73
Original Data
Fitted Values
Forecast

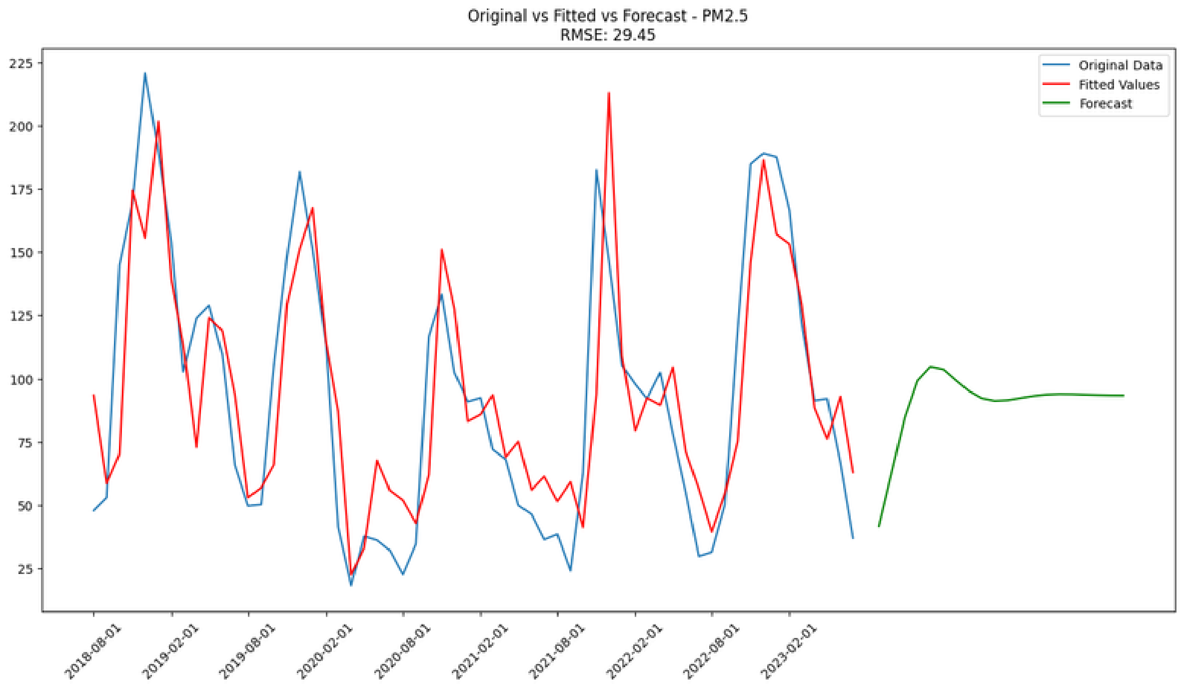

Original vs Fitted vs Forecast - NO2
RMSE: 8.52
Original Data
Fitted Values
Forecast


Original vs Fitted vs Forecast - Temp
RMSE: 2.54
Original Data
Fitted Values
Forecast

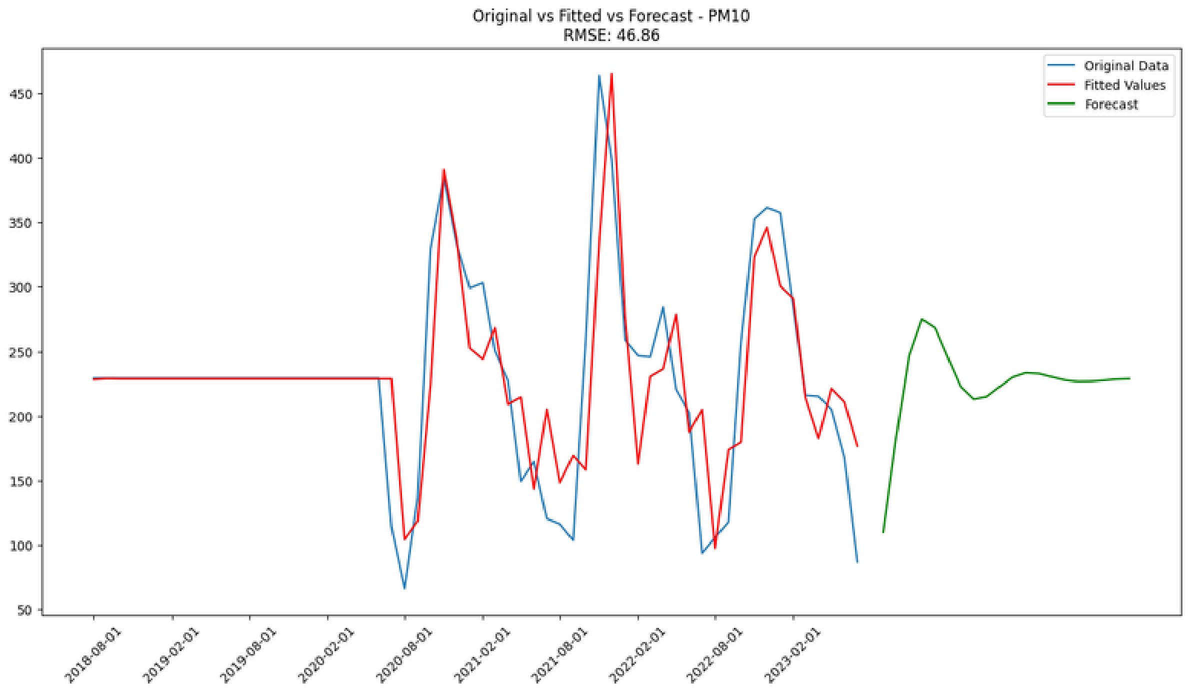

Original vs Fitted vs Forecast - RH
RMSE: 10.39

SHADIPUR

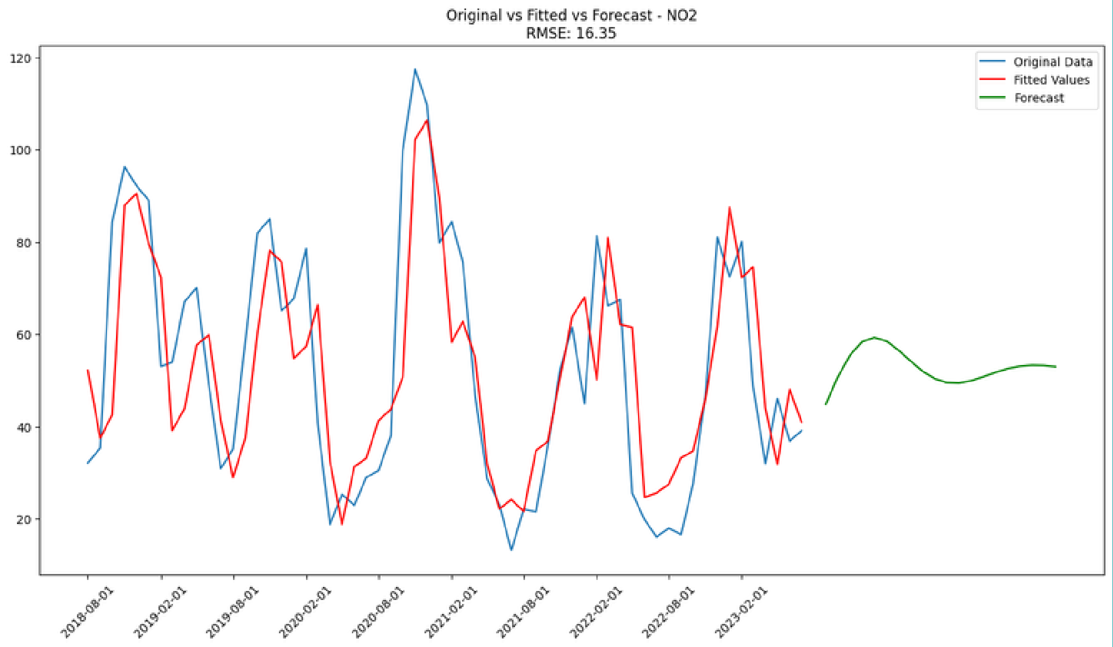
PM2.5

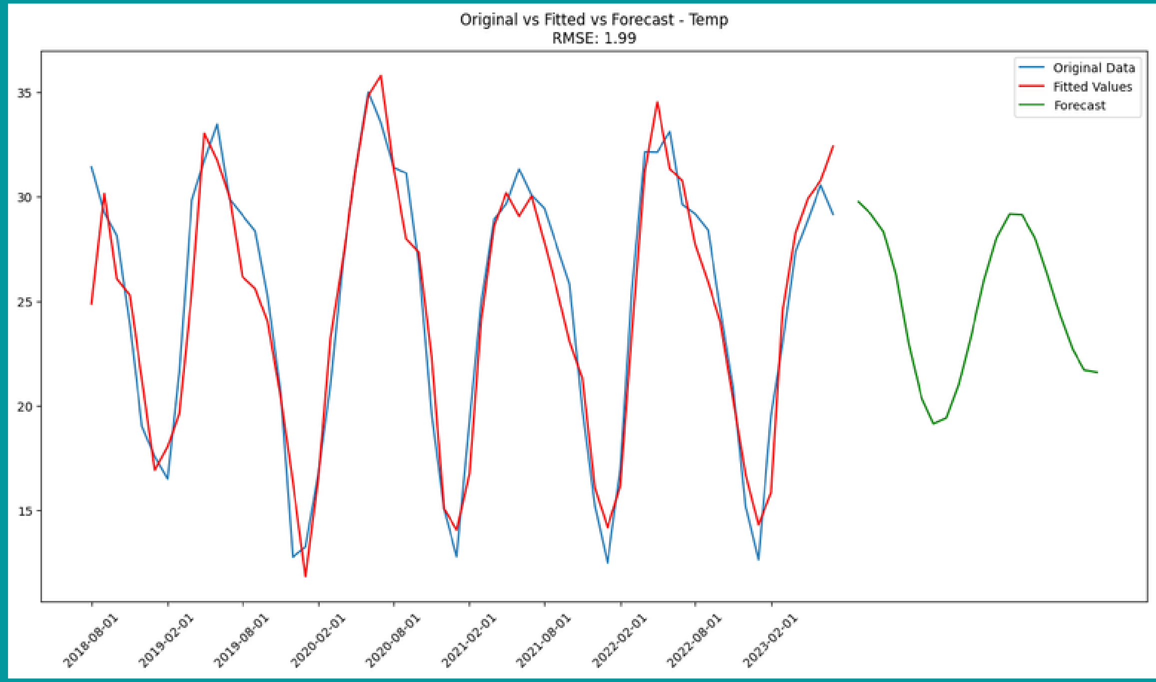


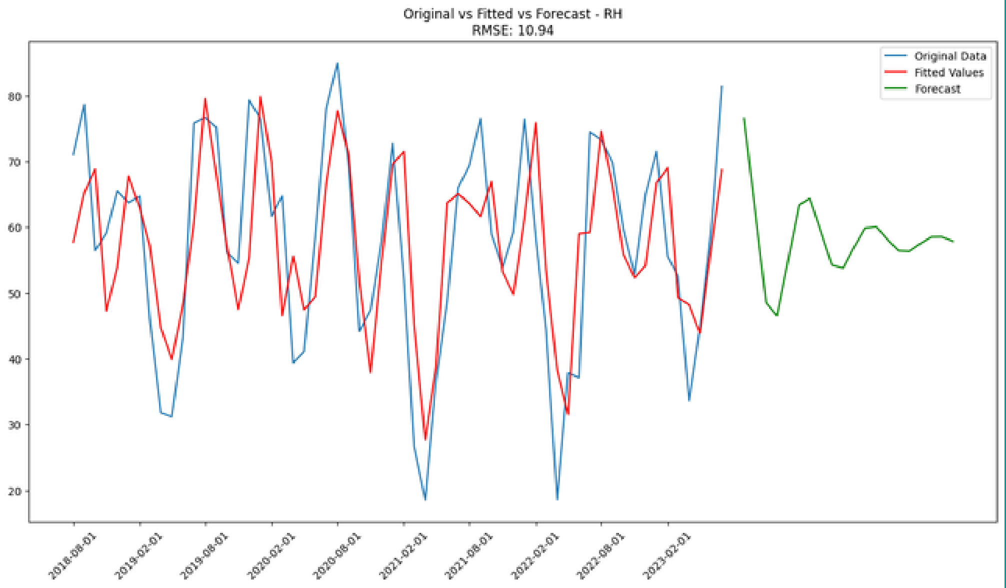
PM10



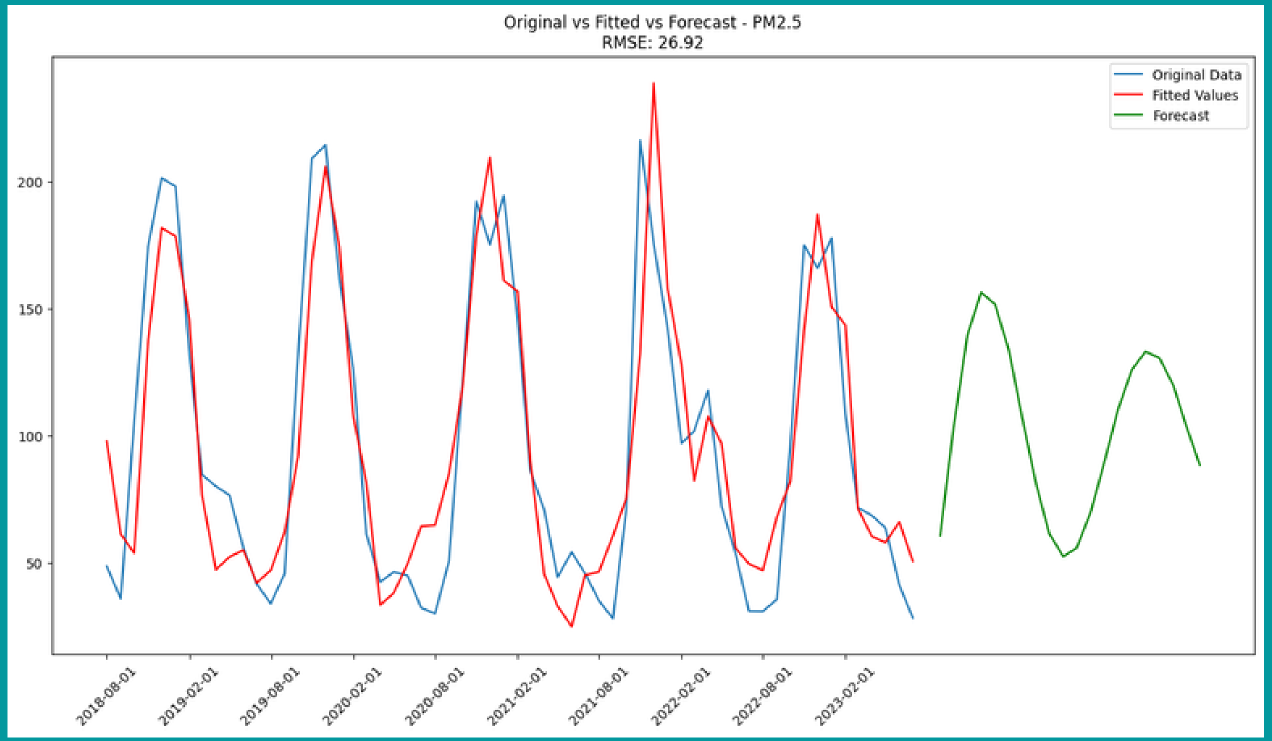
NO_2

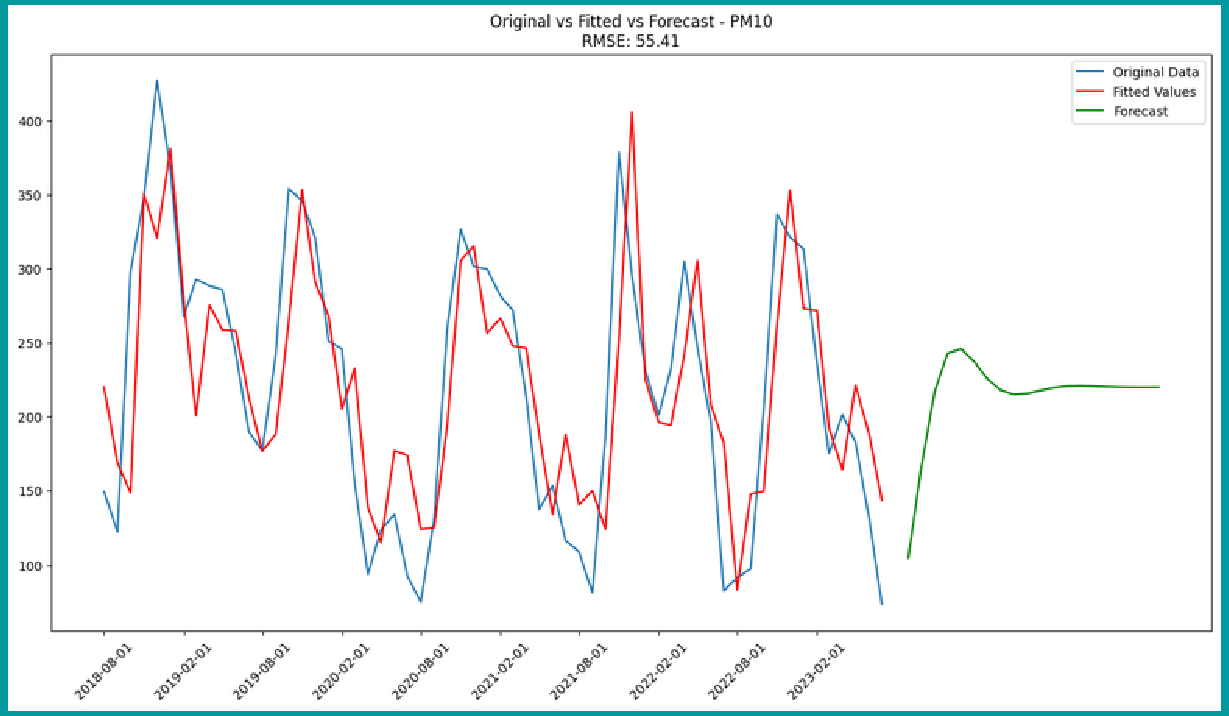


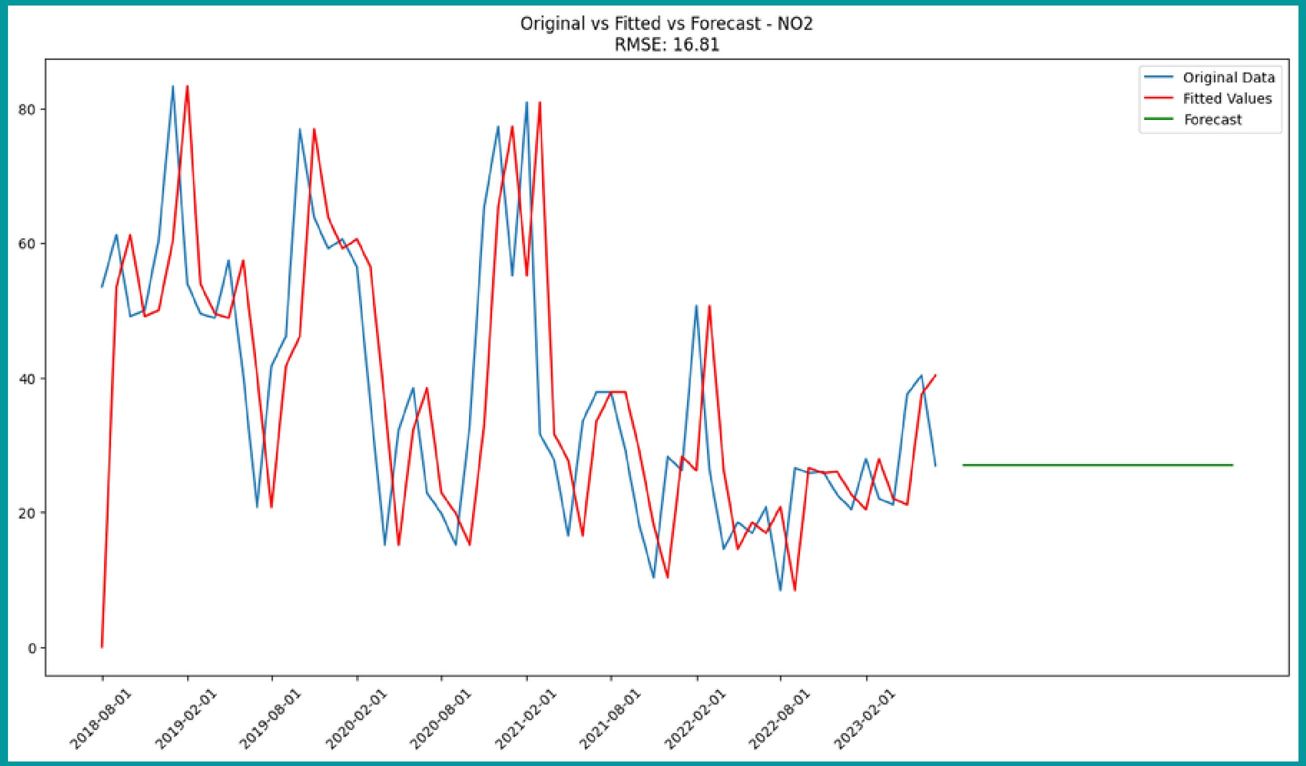


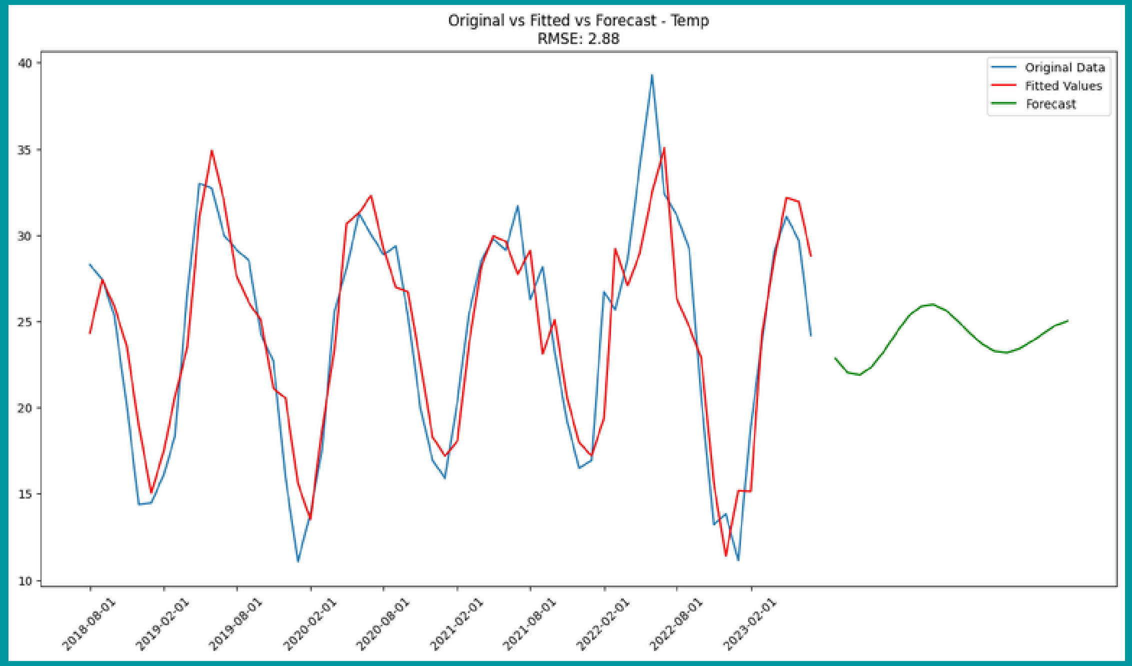


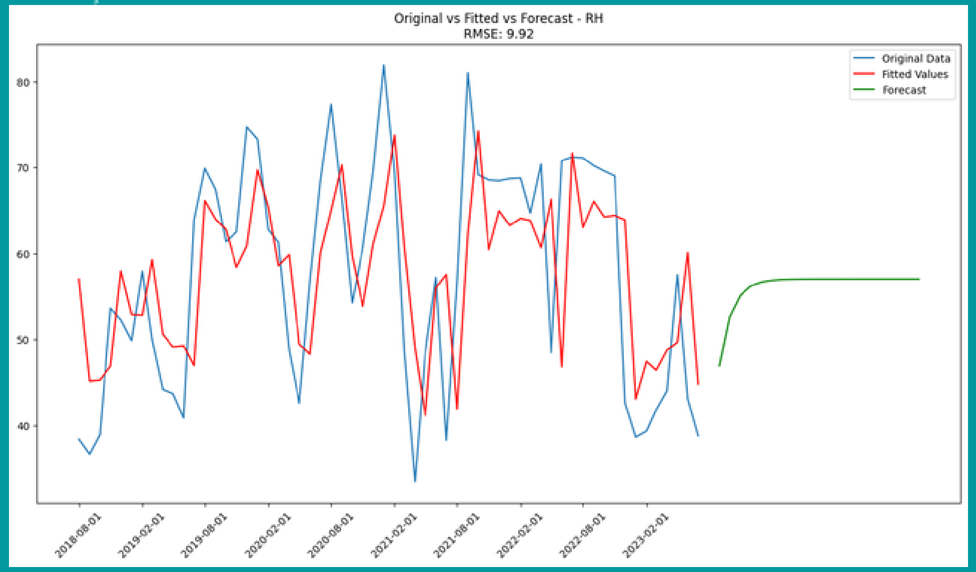
SIRIFORT

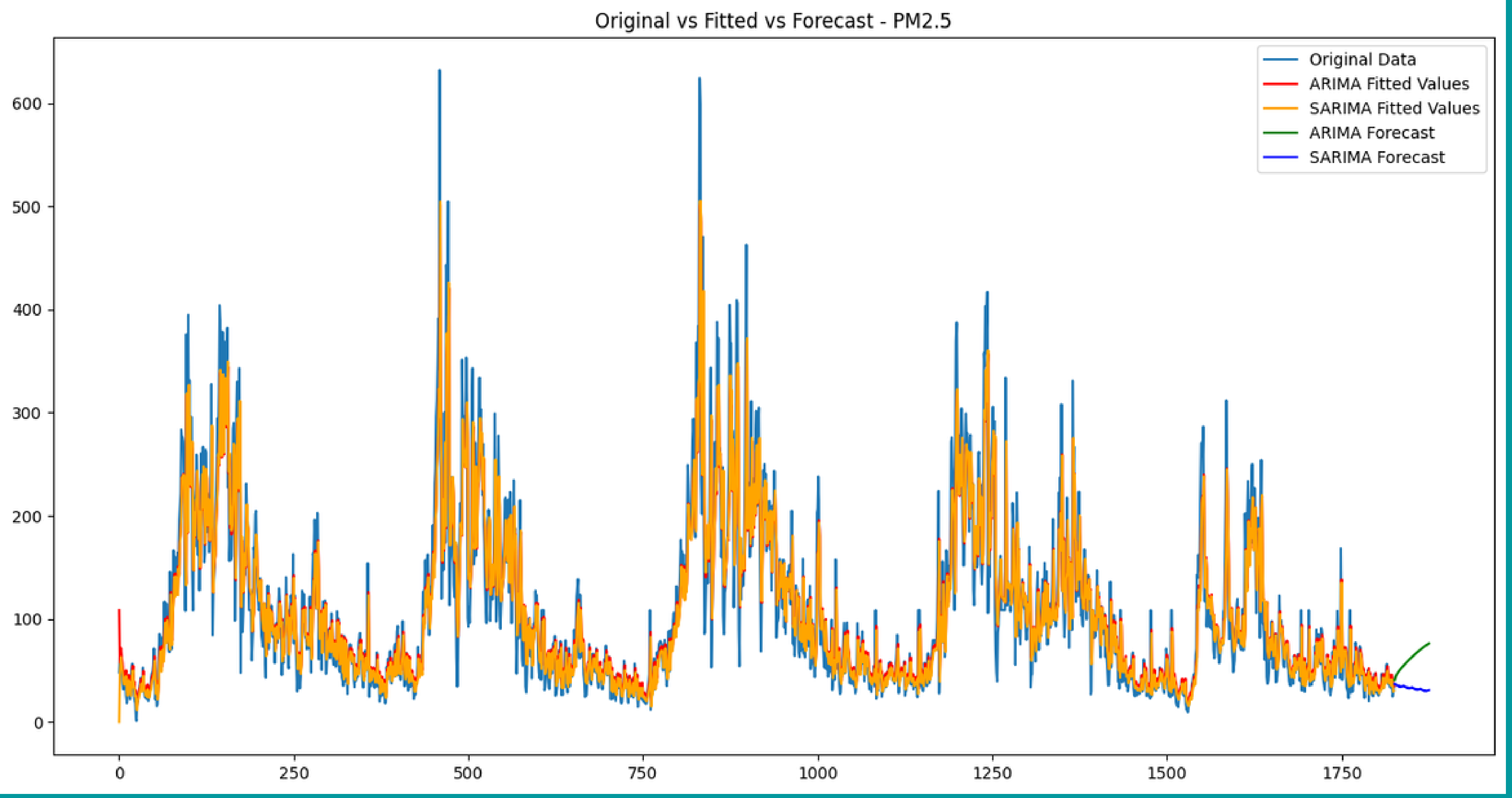

Original vs Fitted vs Forecast - PM2.5
RMSE: 26.92
Original Data
Fitted Values
Forecast


Original vs Fitted vs Forecast - PM10
RMSE: 55.41
Original Data
Fitted Values
Forecast


Original vs Fitted vs Forecast - NO2
RMSE: 16.81
Original Data
Fitted Values
Forecast

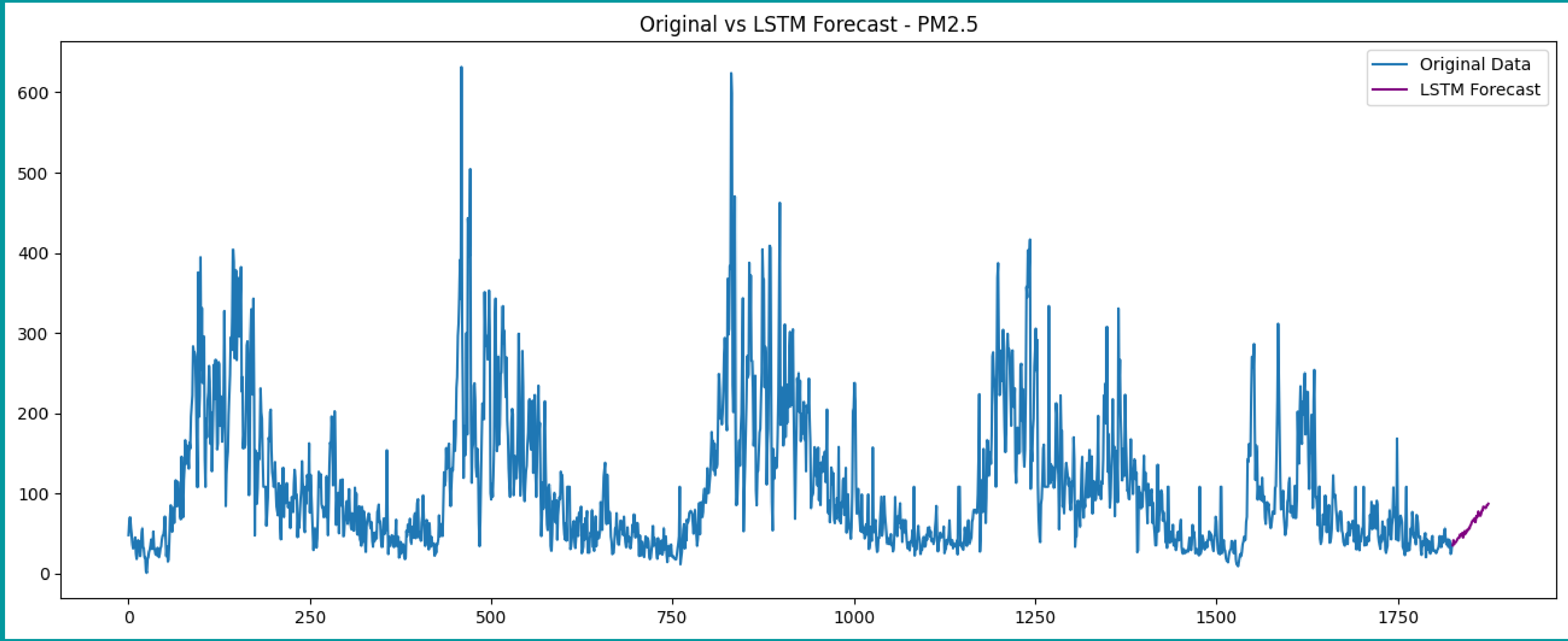

Original vs Fitted vs Forecast - Temp
RMSE: 2.88
Original Data
Fitted Values
Forecast

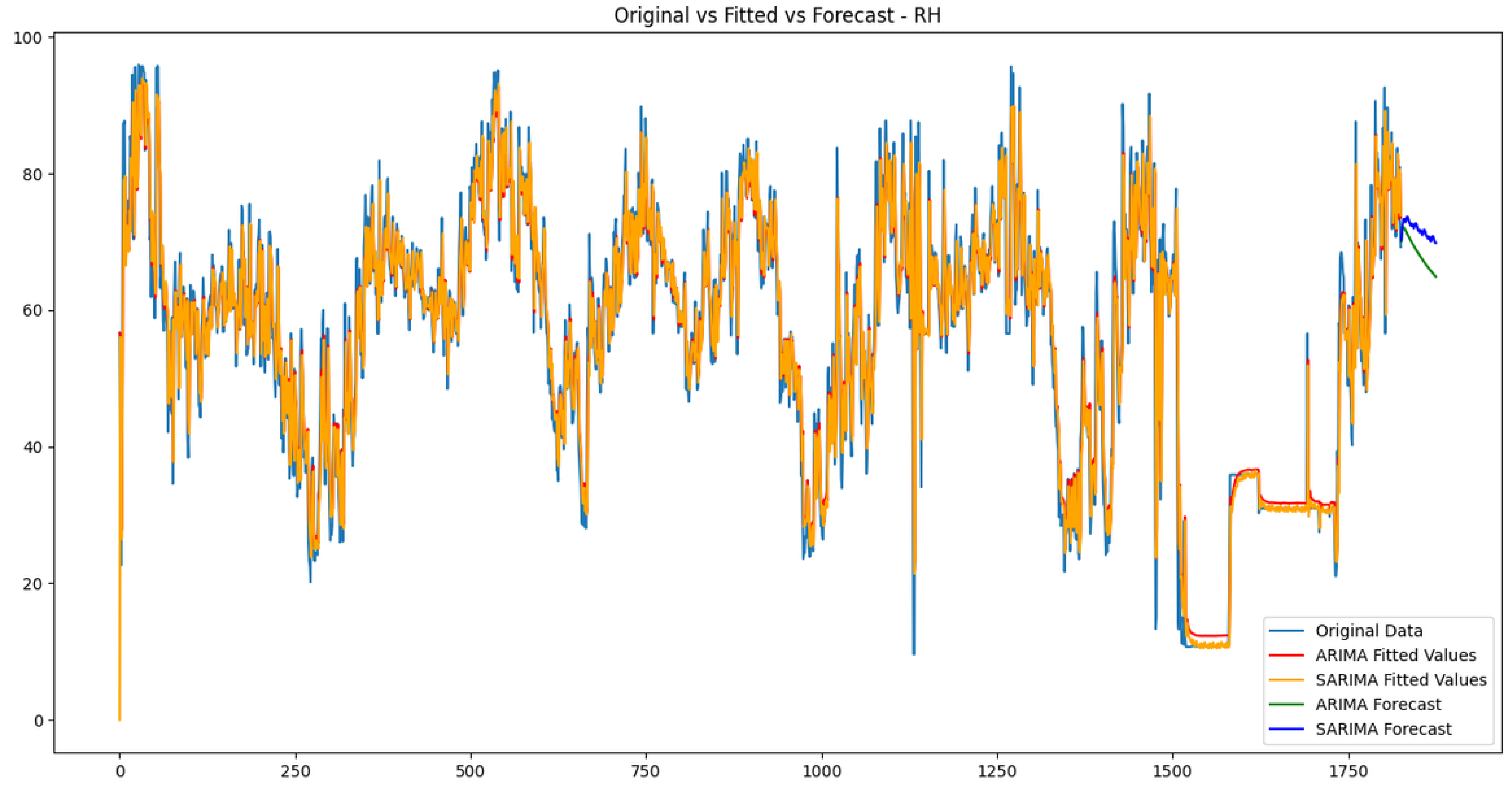

Original vs Fitted vs Forecast - RH
RMSE: 9.92
Original Data
Fitted Values
Forecast



ARIMA VS SARIMA

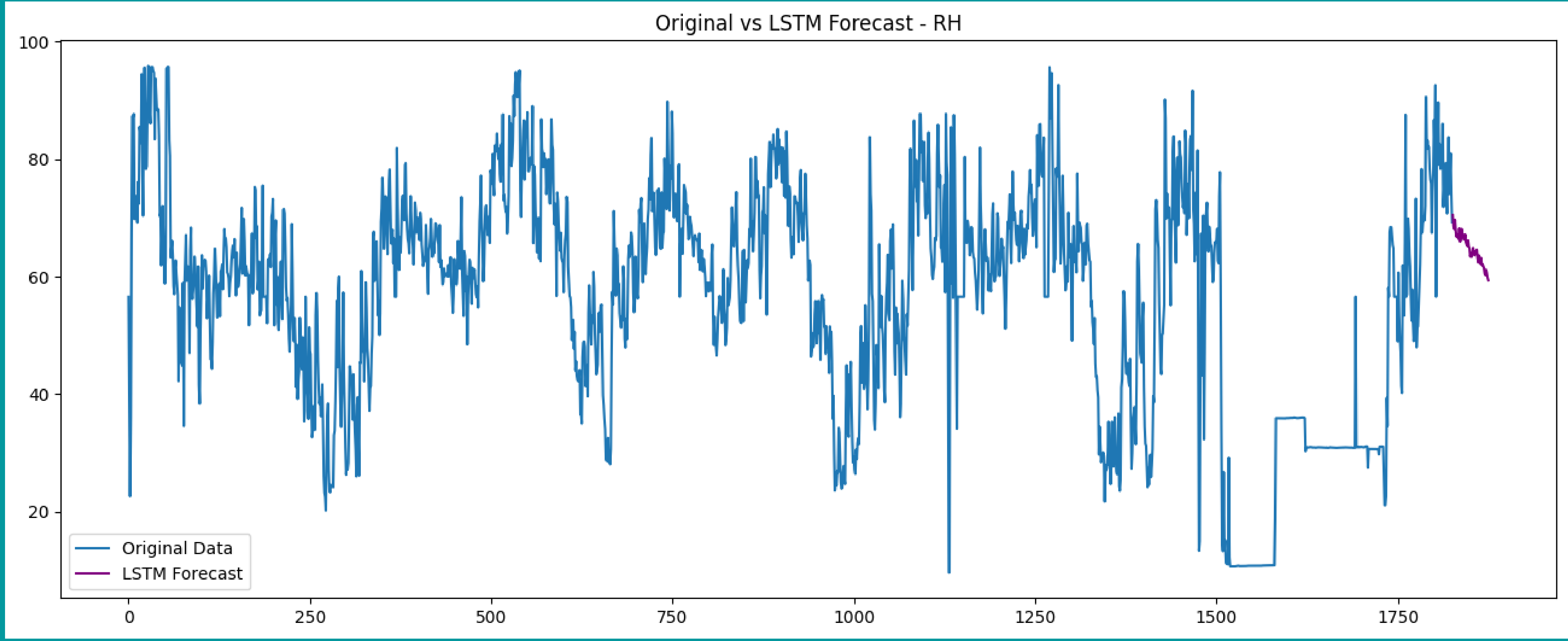
LSTM



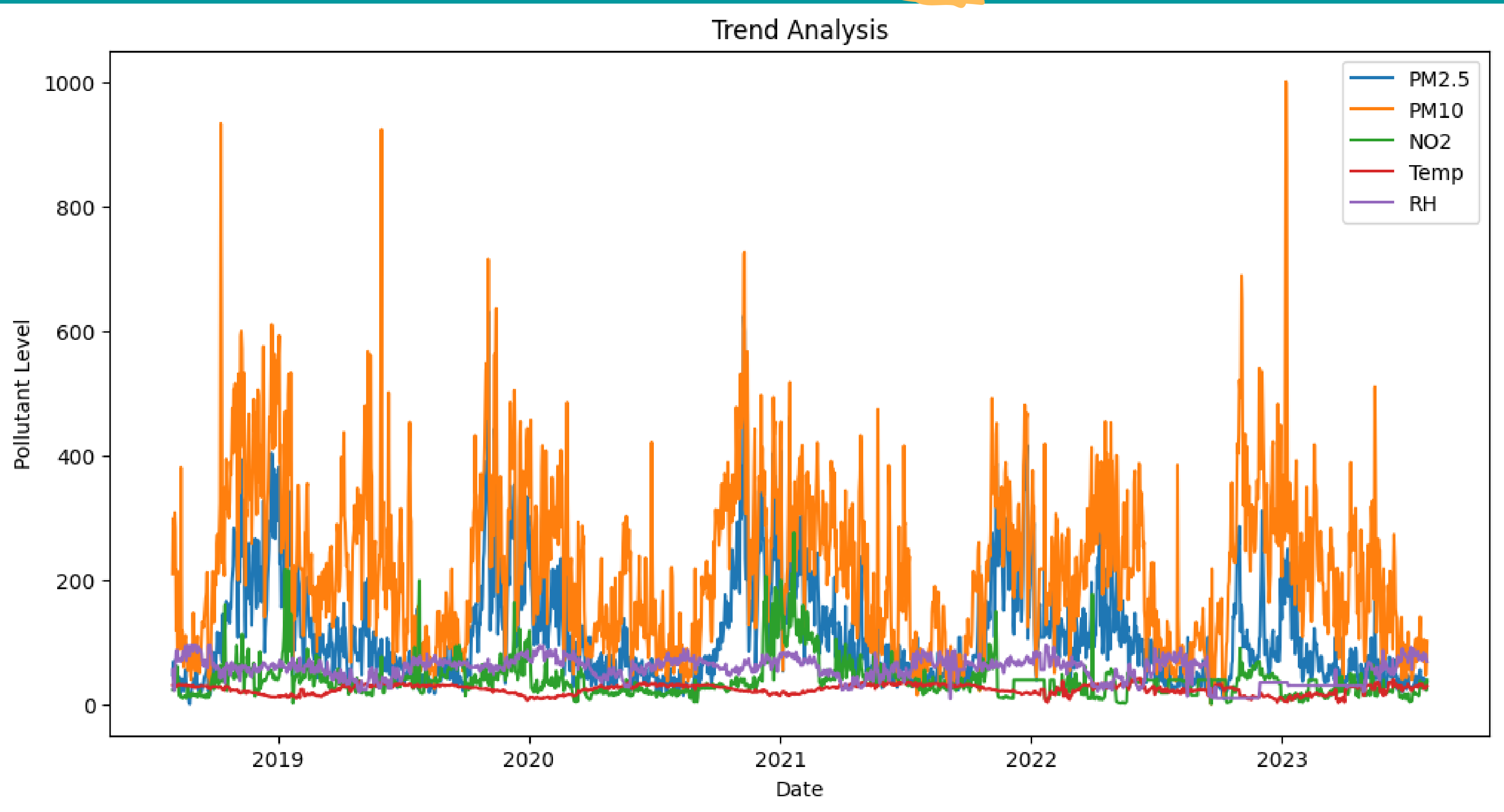


ARIMA VS SARIMA

LSTM



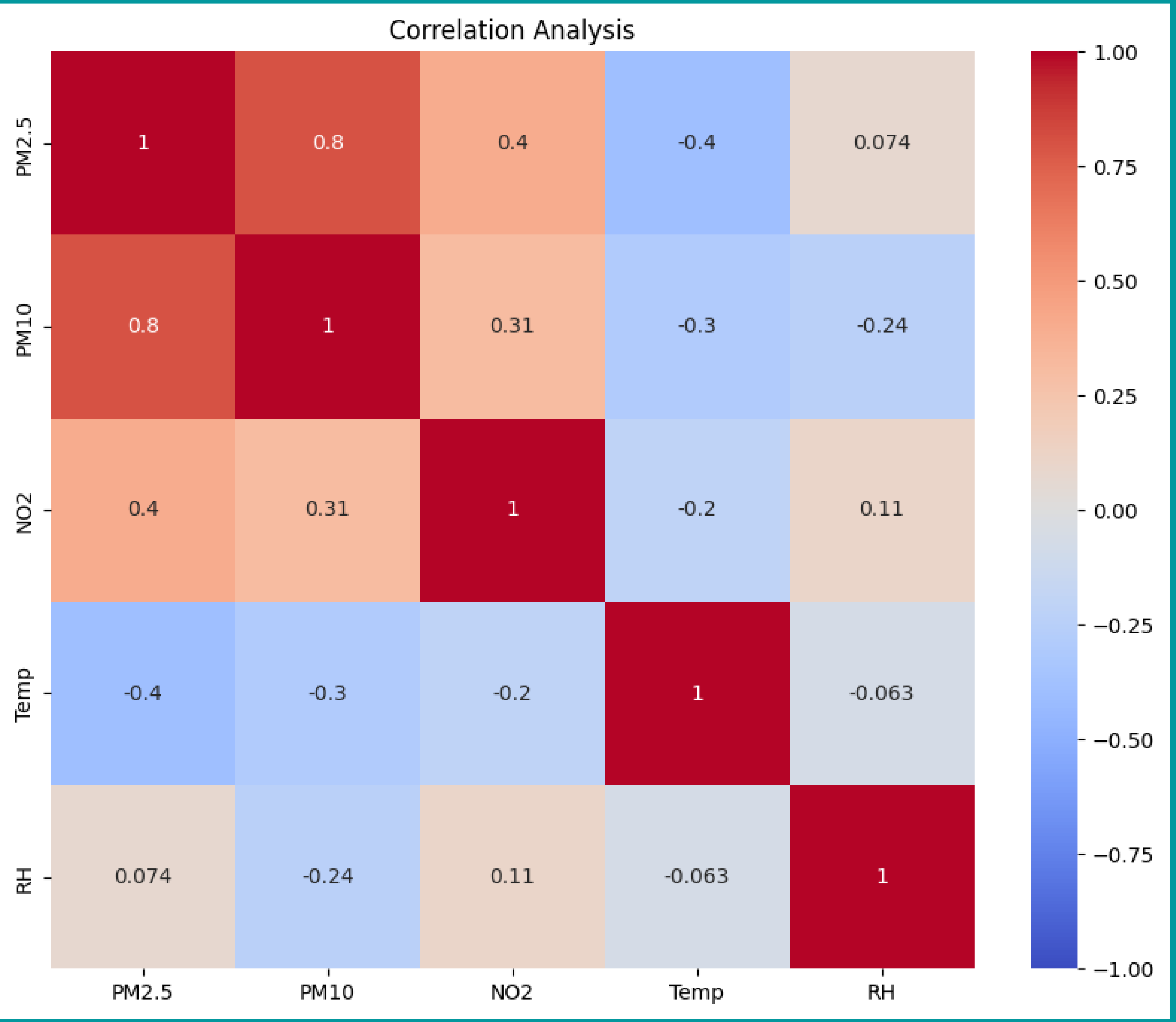
TREND ANALYSIS



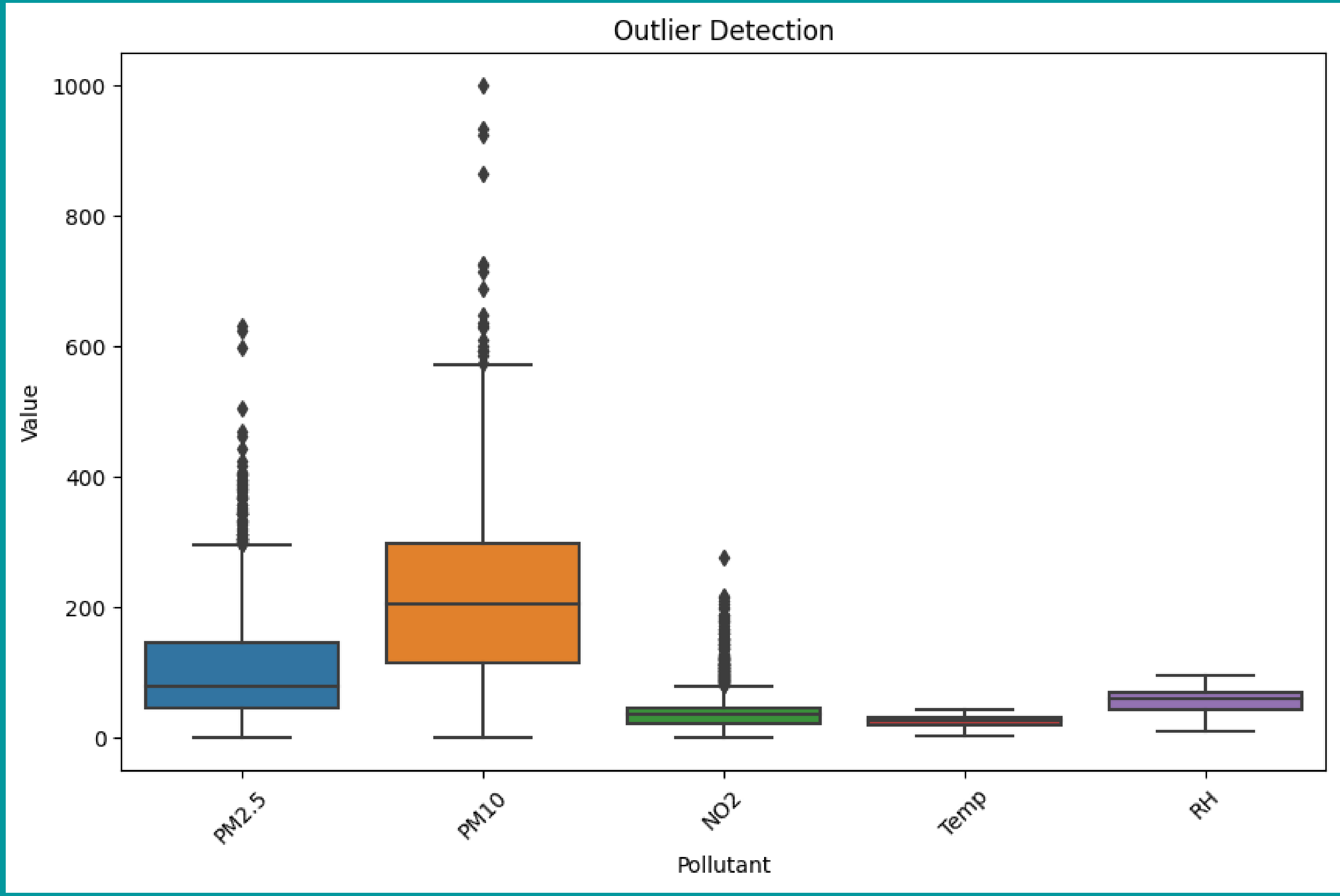
DESCRIPTIVE STATS

	PM2.5	PM10	NO2	Temp	RH
count	1826.000000	1826.000000	1826.000000	1826.000000	1826.000000
mean	108.416294	218.014497	39.943173	24.845647	56.568067
std	85.492241	129.188313	29.372160	7.507347	18.929223
min	1.000000	1.000000	1.060000	3.060000	9.620000
25%	45.190000	114.307500	21.792500	19.340000	43.962500
50%	79.465000	204.325000	36.240000	26.250000	60.245000
75%	145.800000	297.922500	45.680000	31.007500	69.602500
max	631.850000	1000.000000	276.460000	42.830000	95.880000

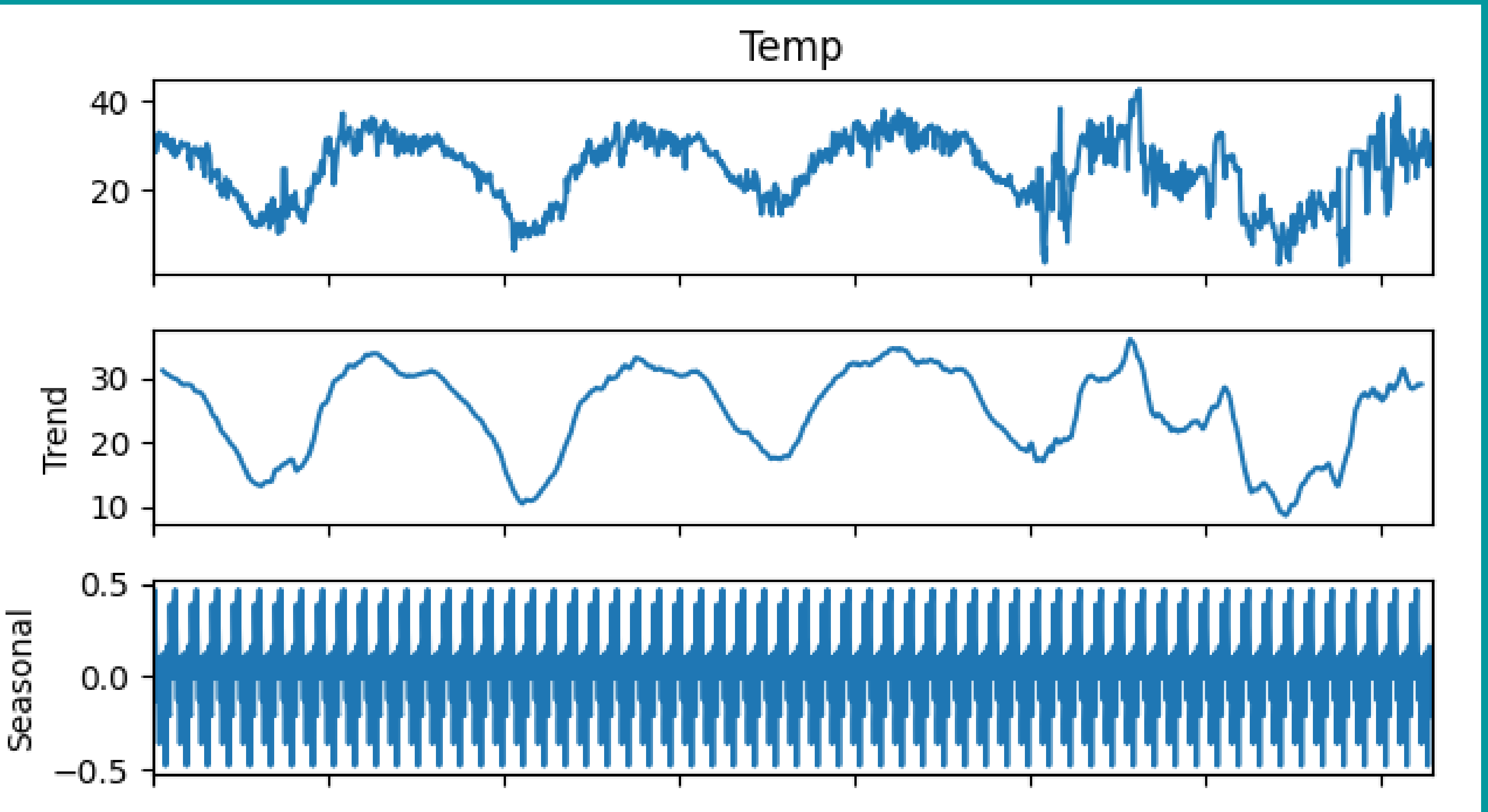
CORRELATION

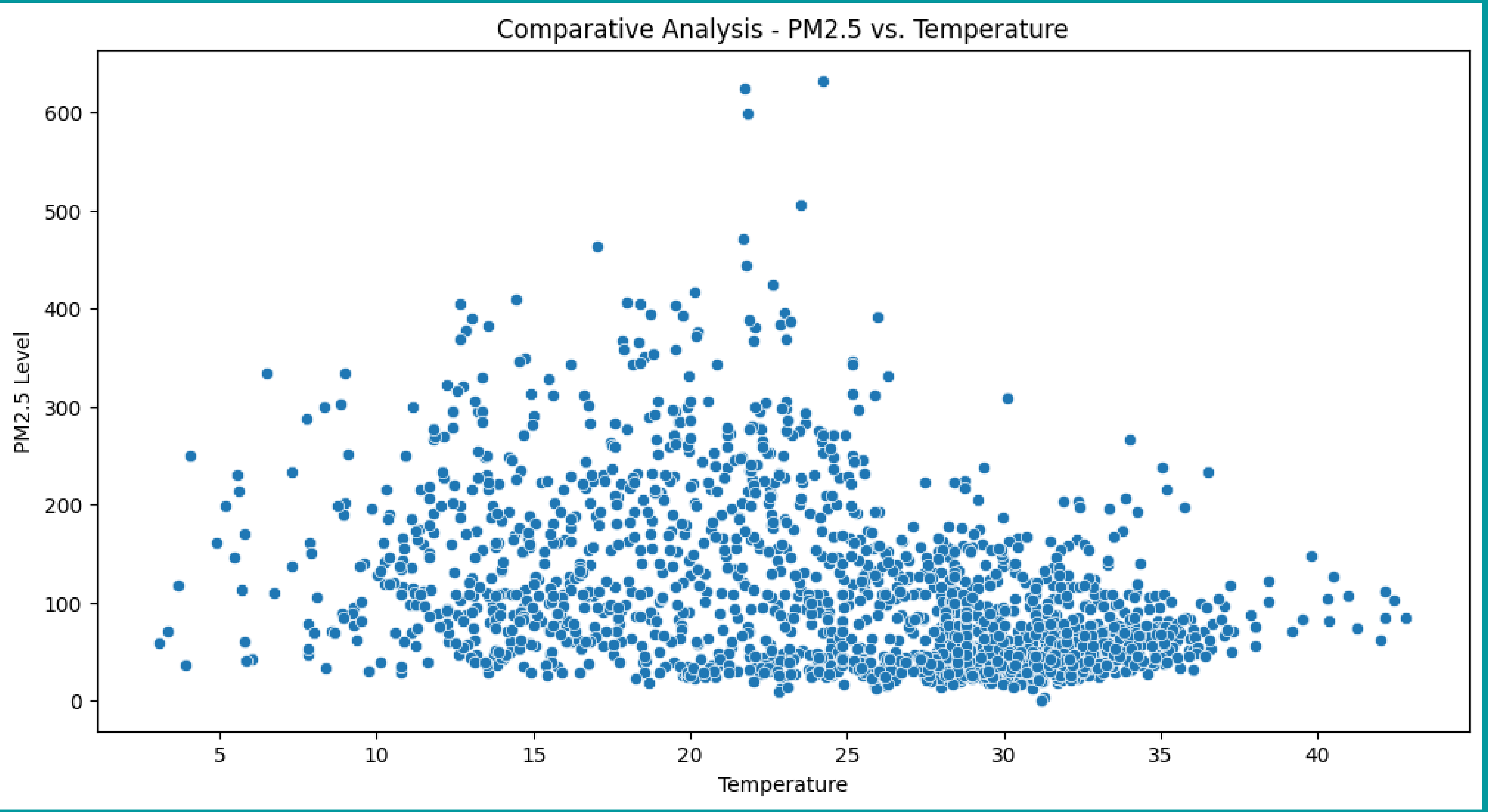


OUTLIER DETECTION

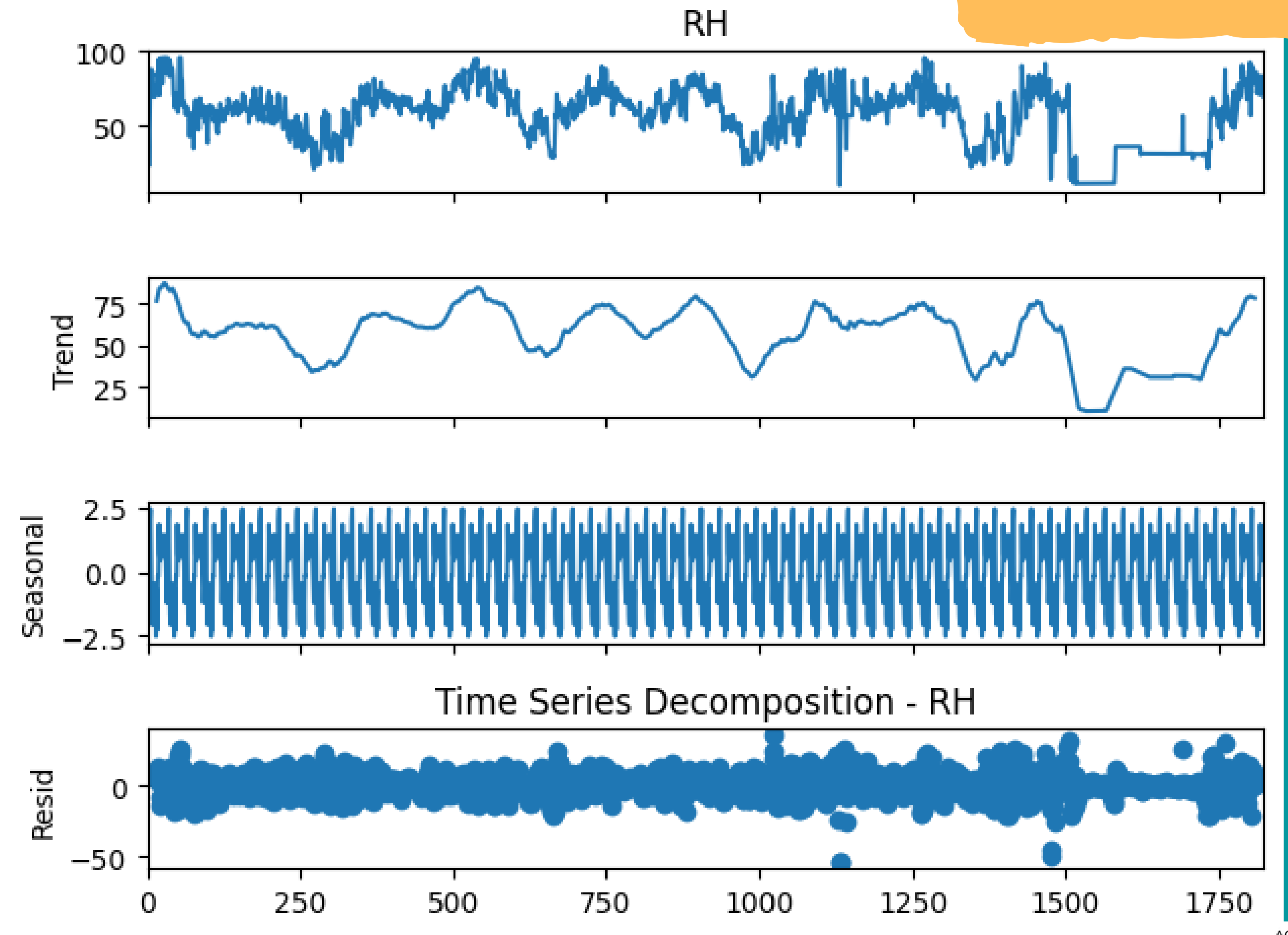


TEMPERATURE

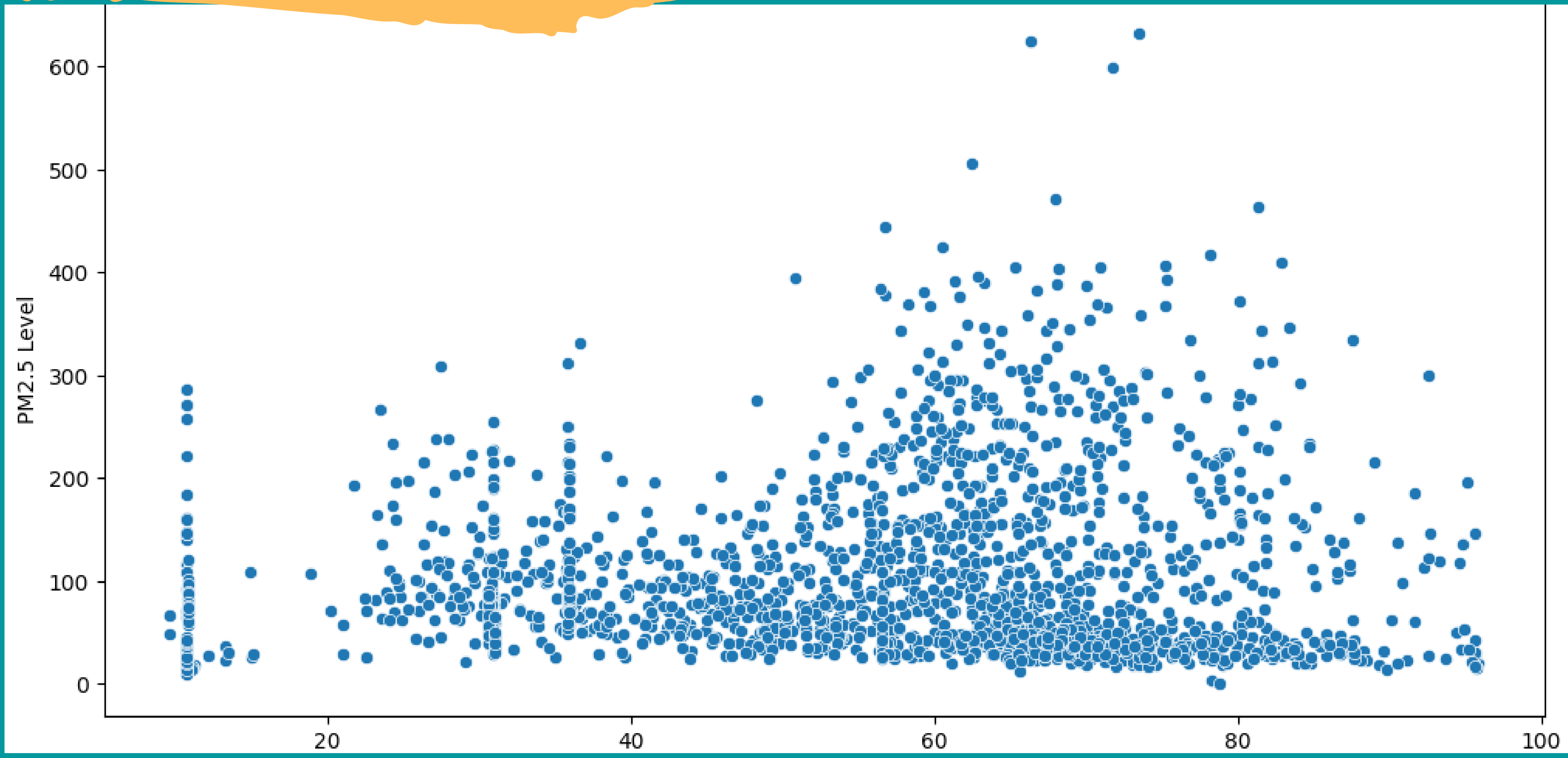




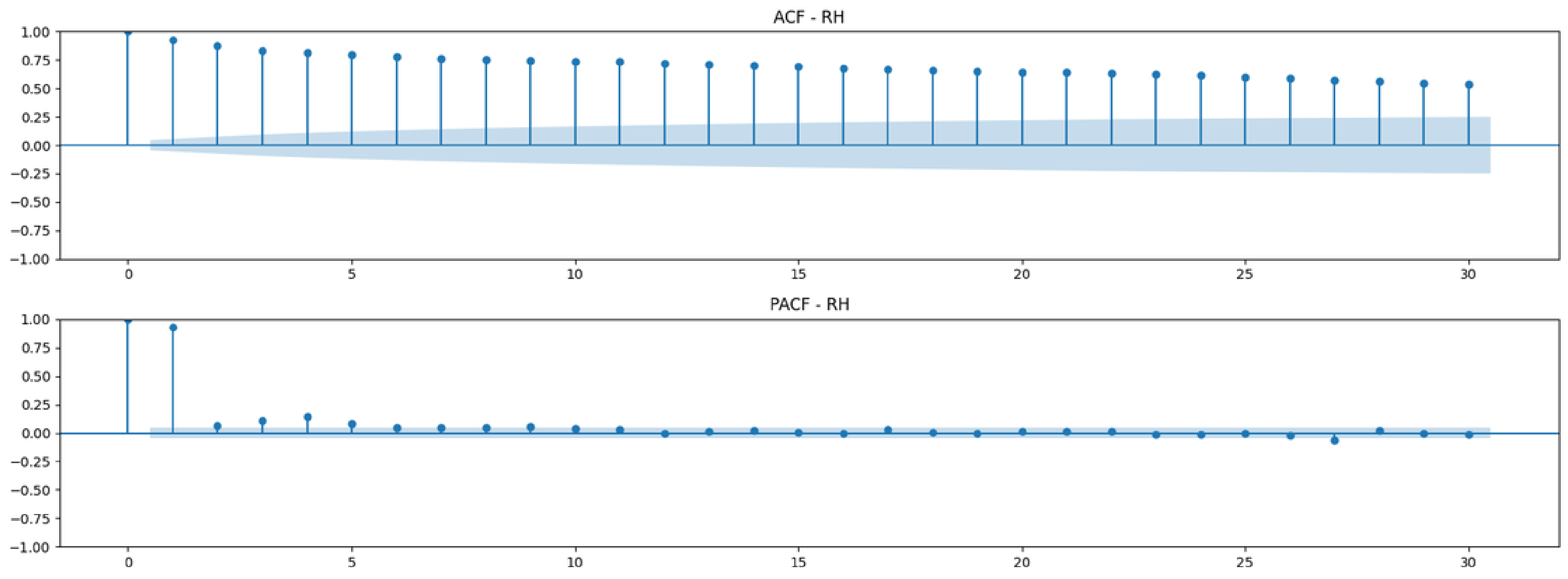
RH



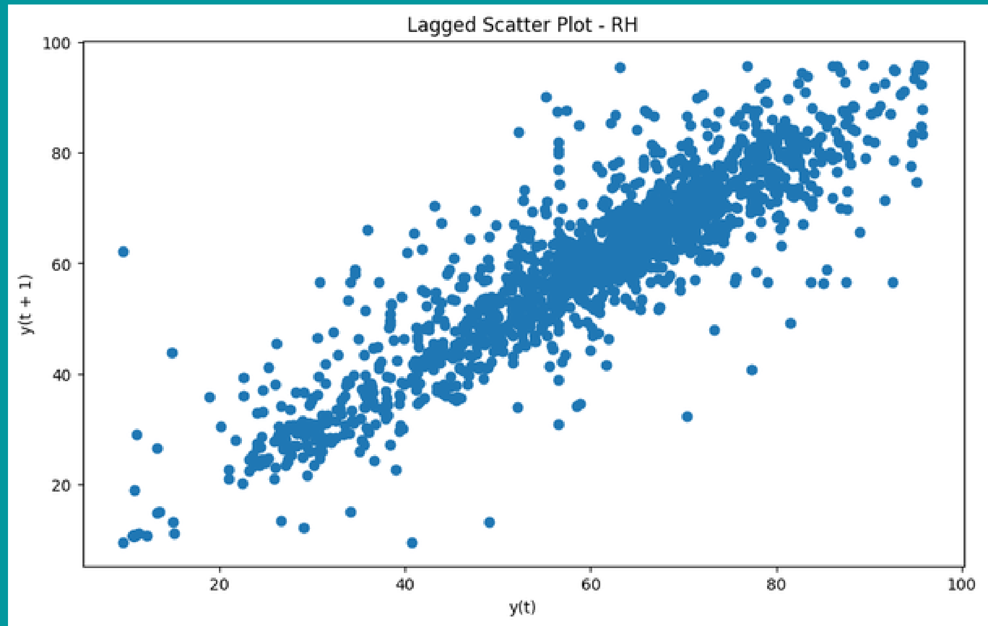
RH
Trend
Seasonal
Time Series Decomposition - RH
Resid



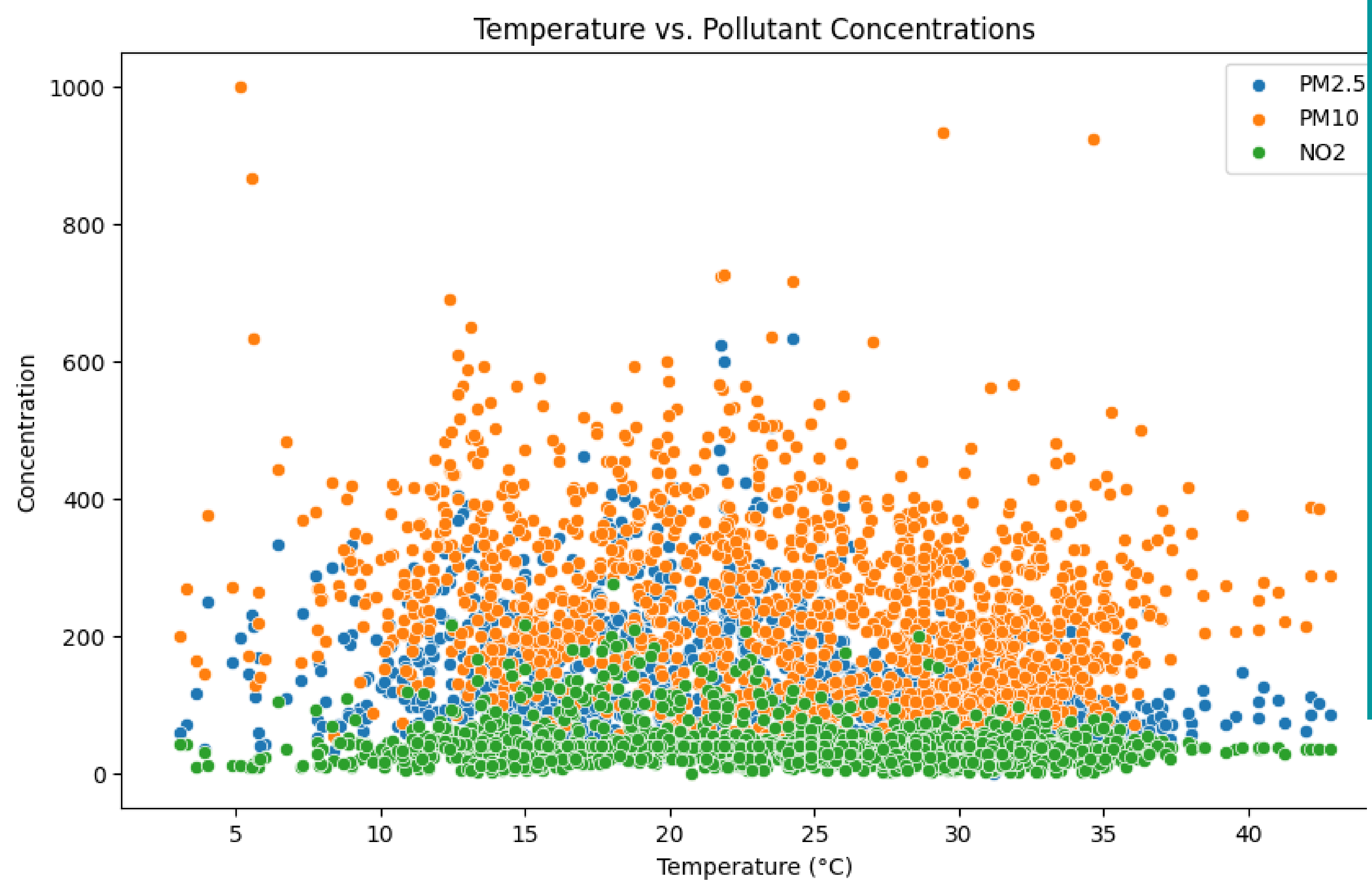
PM2.5 Level



ACF - RH
PACF - RH

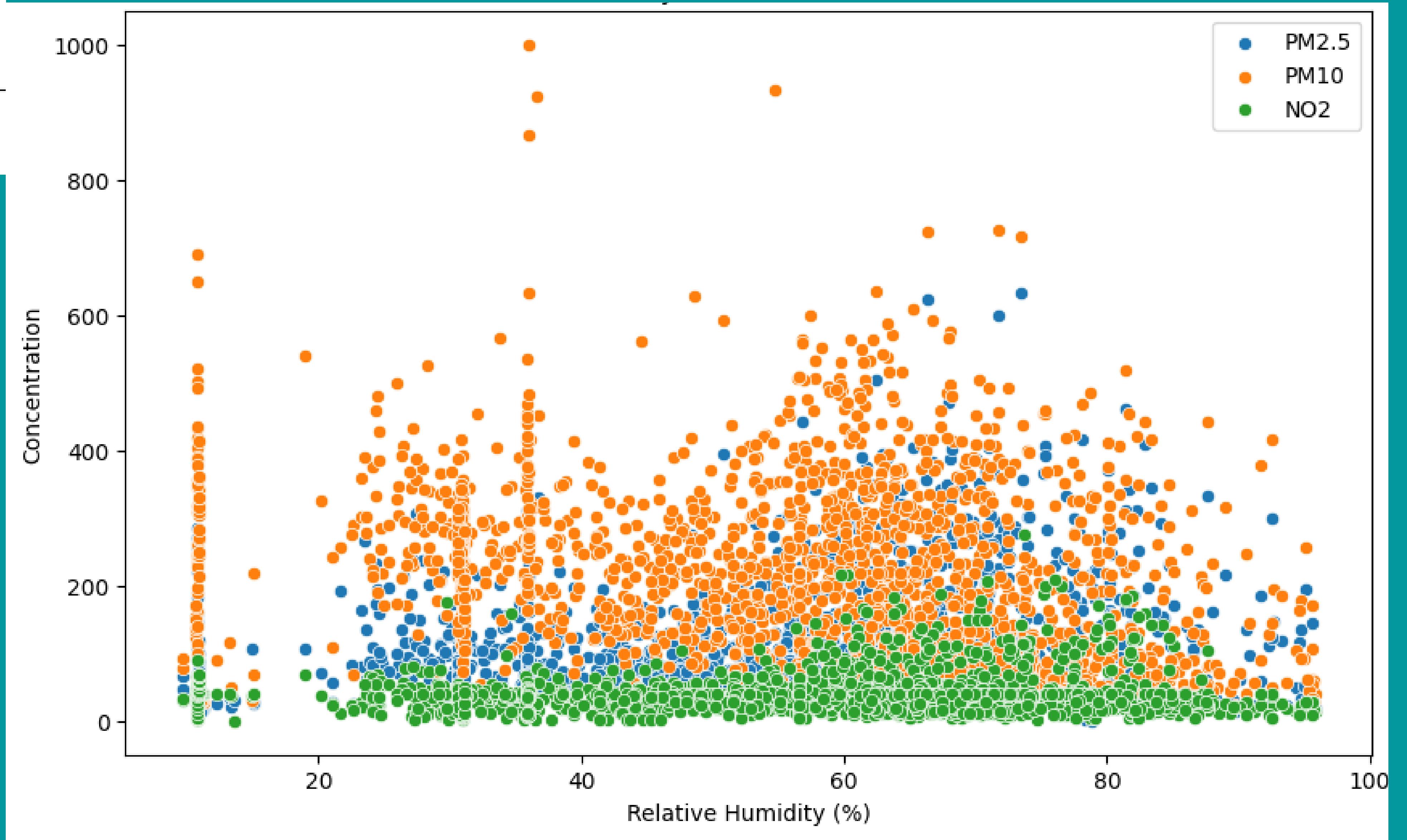


Lagged Scatter Plot - RH
y(t + 1)
y(t)



RH VS OTHER

TEMP VS OTHER



ACF (Autocorrelation Function)

The ACF measures the correlation between a time series and its own lagged values at various time lags.

It helps identify the overall degree of autocorrelation in a time series at different lags.

The ACF plot shows how correlated each observation is with its past observations. It provides information about the general dependence structure and the presence of seasonality.

A significant autocorrelation in the ACF plot can suggest the need for an autoregressive (AR) component in a time series model. If the ACF tails off gradually, it can indicate a non-stationary process. If there are spikes at certain lags, it might imply the presence of a seasonal pattern.

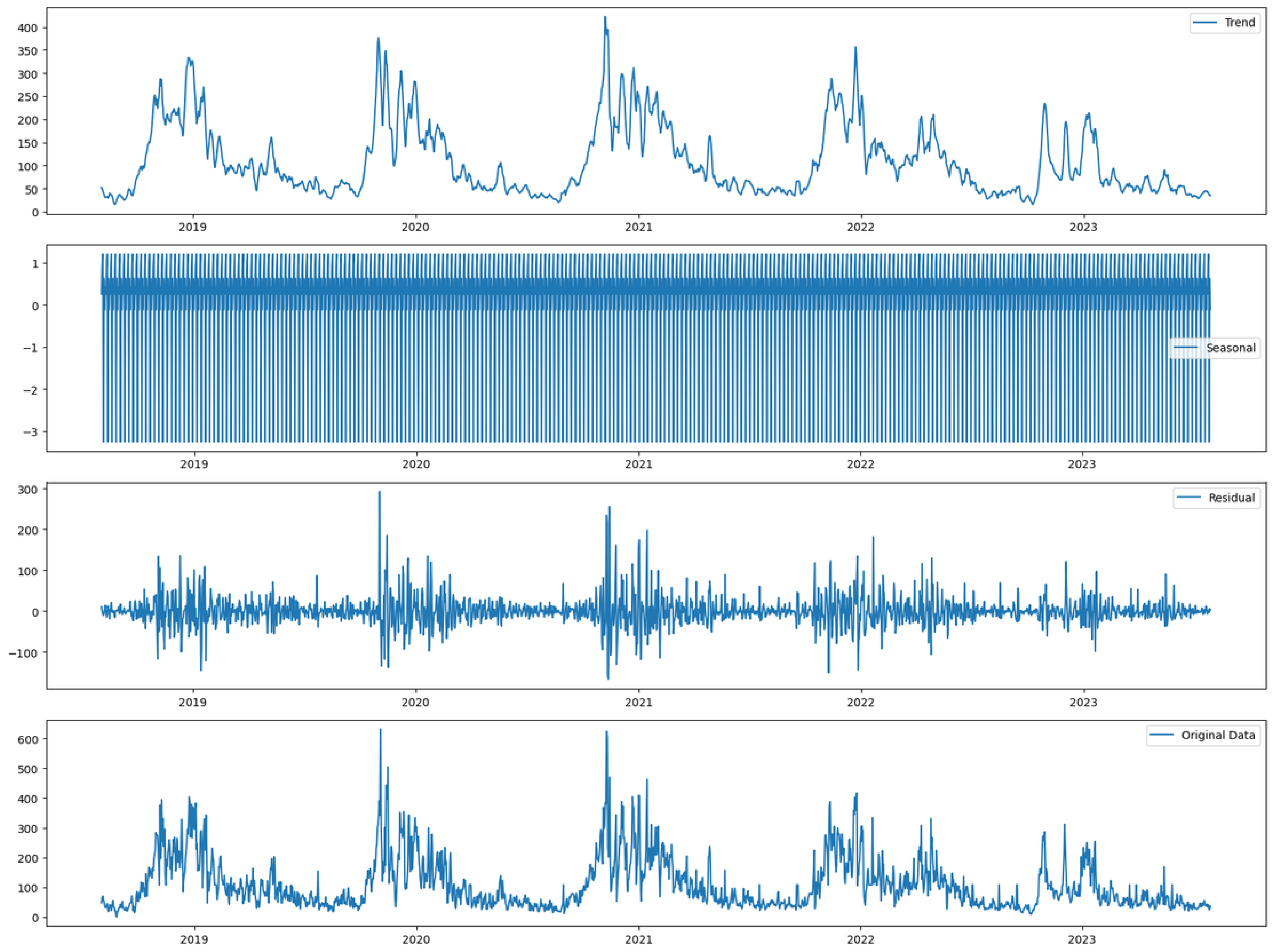
PACF (Partial Autocorrelation Function)

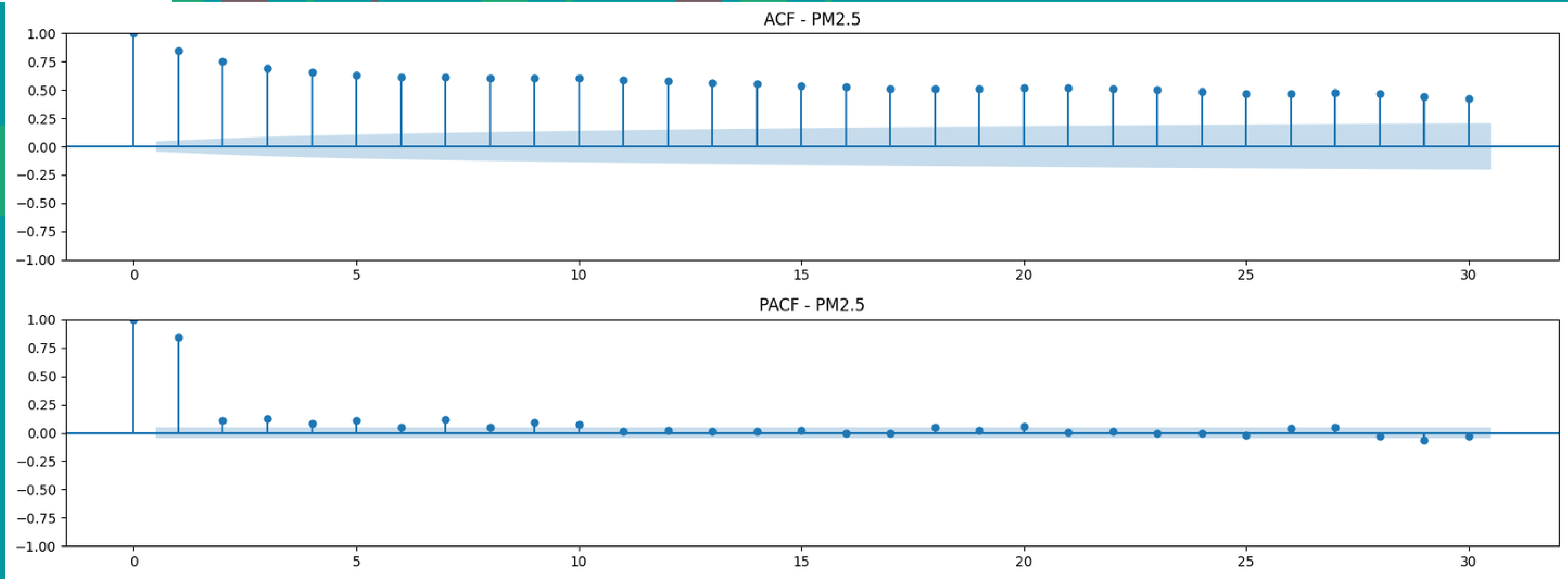
The PACF measures the correlation between a time series and its own lagged values while accounting for the influence of intermediate lags.

It helps identify the direct relationship between an observation and its specific lagged values, removing the effects of intervening lags.

The PACF plot indicates the direct impact of each lag on the current observation, allowing for the identification of the order of autoregressive terms in a model.

Significant spikes in the PACF plot can help determine the appropriate order of the autoregressive (AR) component in an ARIMA model. A significant spike at lag 1 followed by a gradual decrease might suggest an AR(1) term is needed, and so on.


Trend
Seasonal
Residual
Original Data
2019
2020
2021
2022
2023


ACF - PM2.5
PACF - PM2.5

REMOVING TRENDS

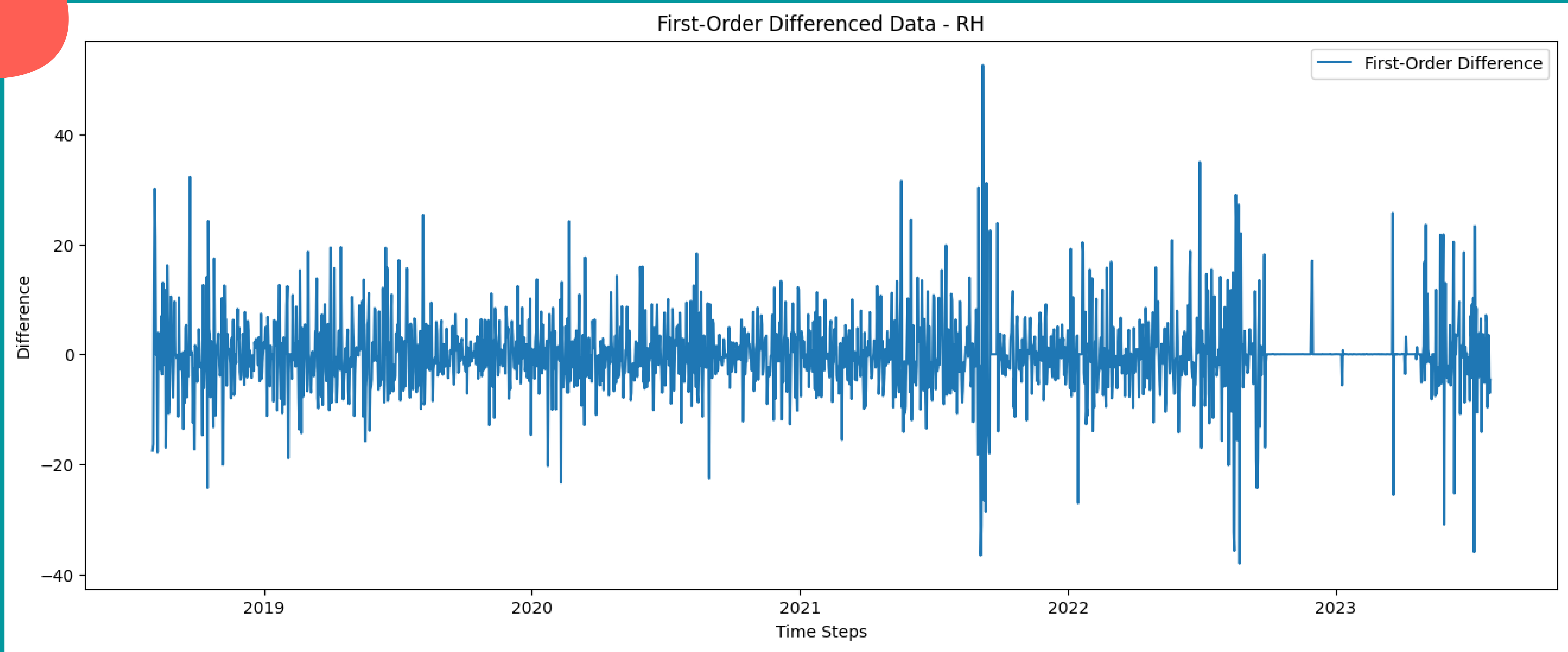
REMOVING SEASONALITY

IDENTIFYING PATTERNS

PREPROCESSING

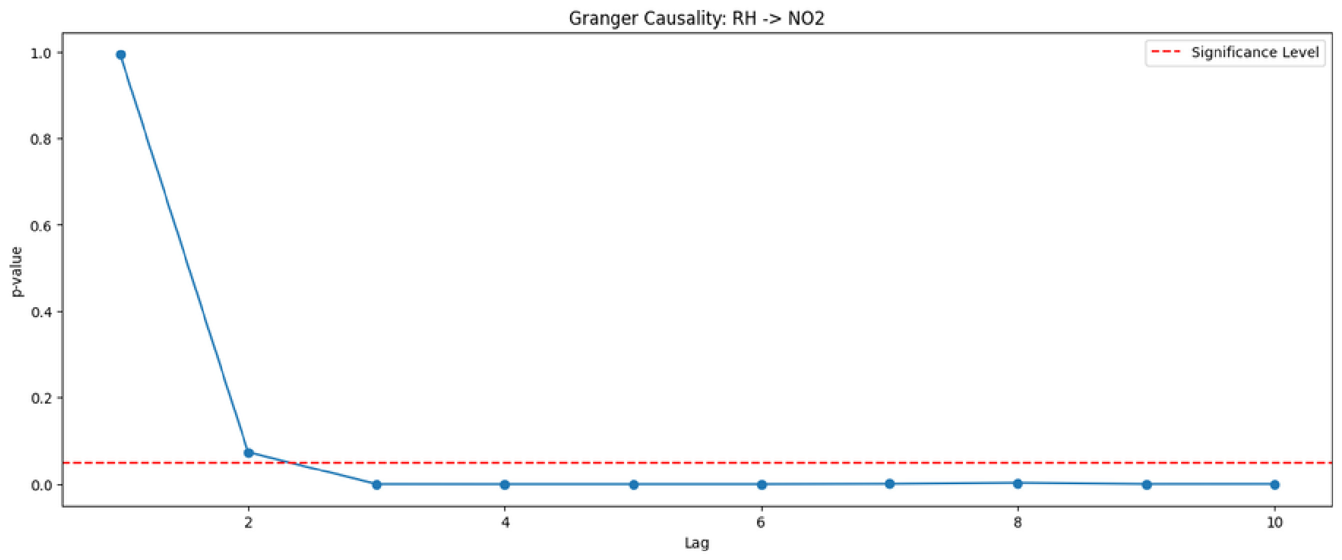
FIRST ORDER DIFFERENCE

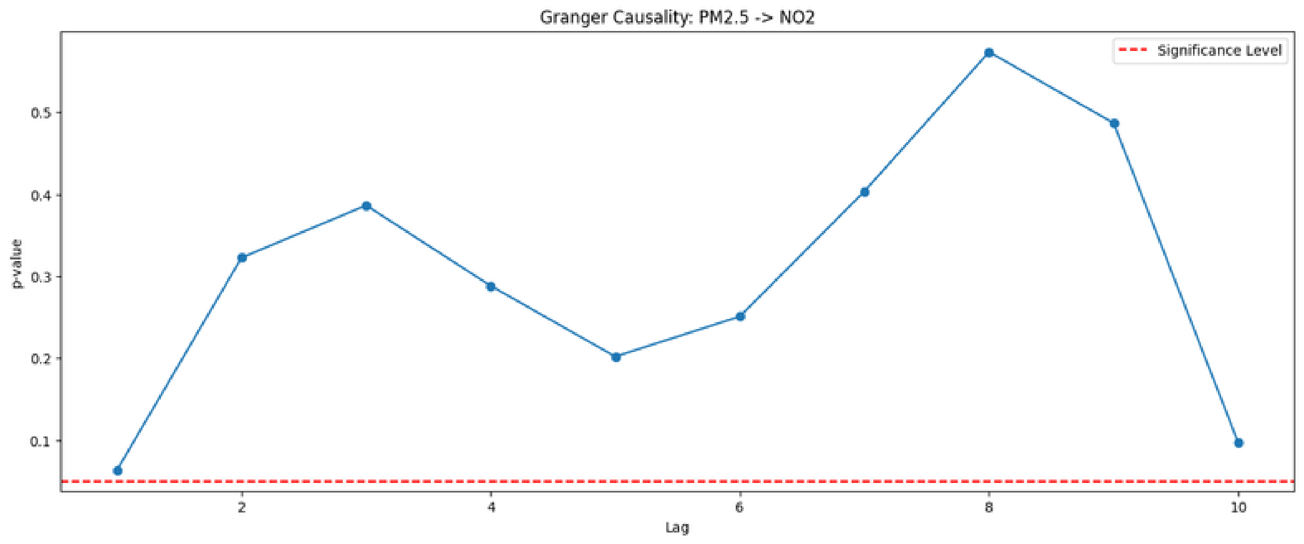
$$D_t = X_t - X_{t-1}$$

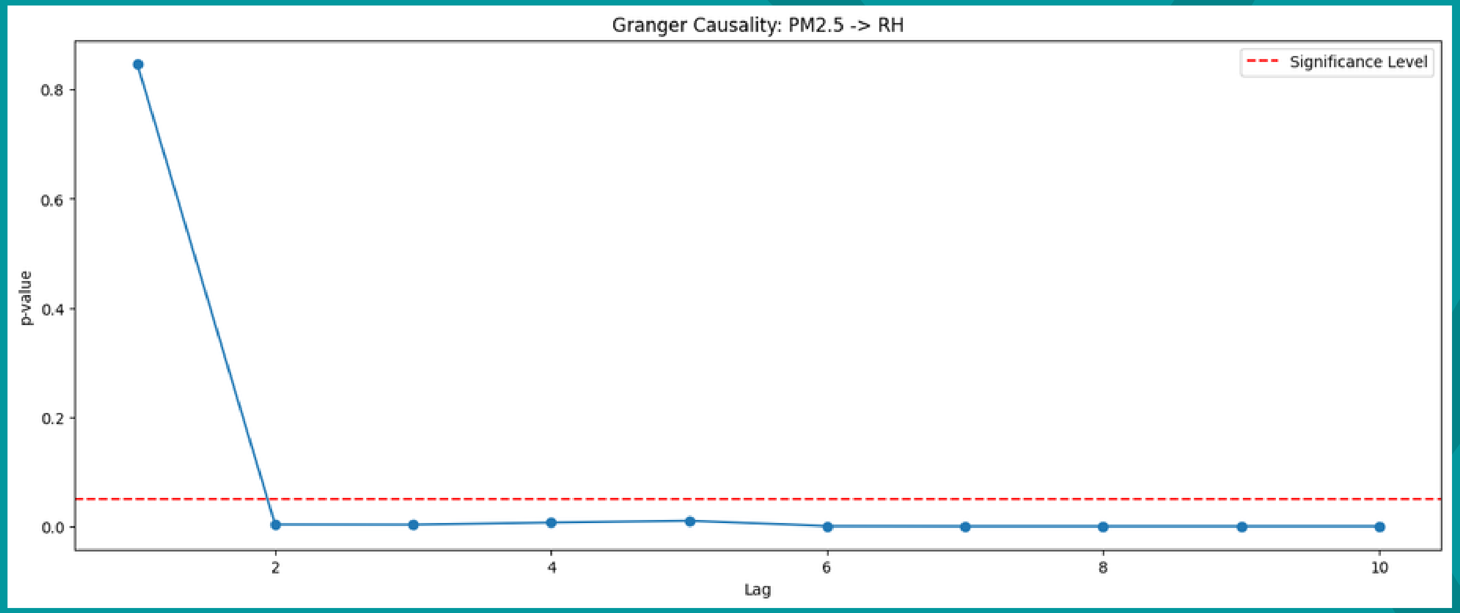


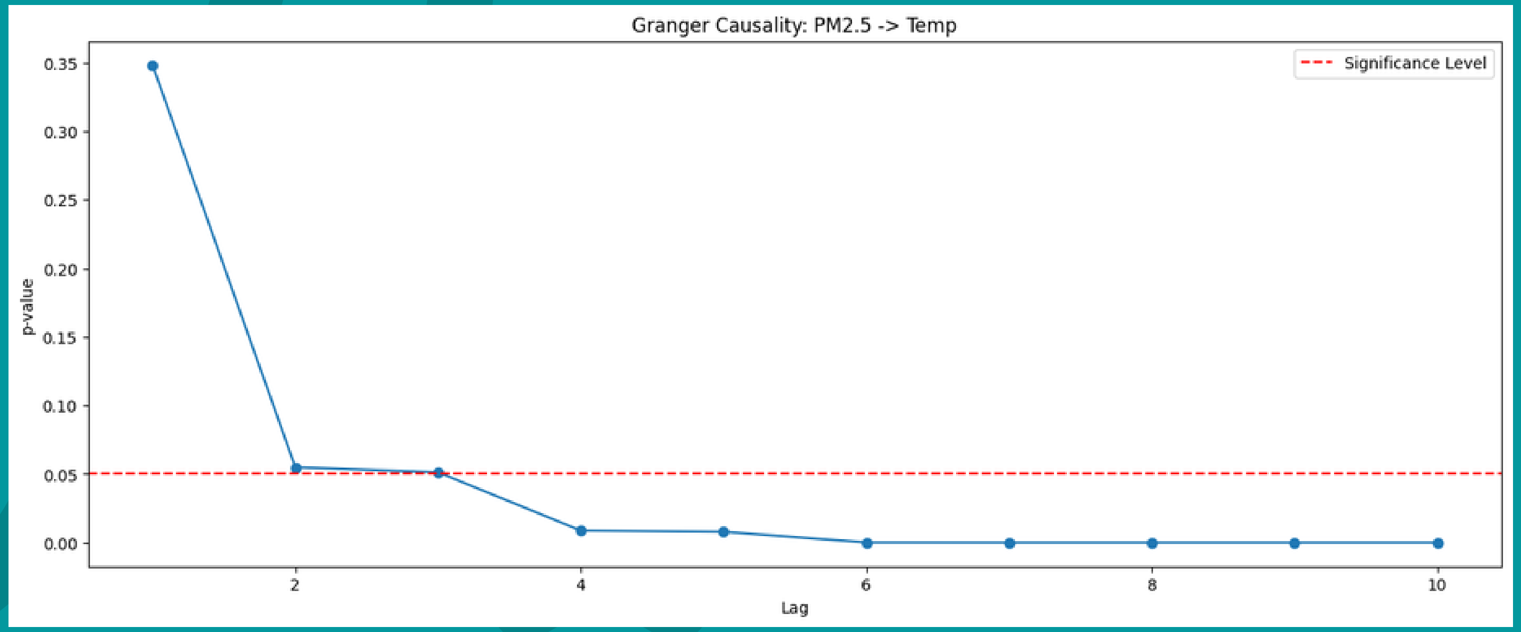
GRANGER CAUSALITY TEST

The Granger causality test helps determine whether one time series can predict another. You can apply this test to understand if the past values of one pollutant can predict the future values of another pollutant.









ANALYSIS

DWARKA

- **Lowest PM2.5**: March 2022 - This could be due to favorable weather conditions, reduced vehicular traffic, or effective pollution control measures during this month.
- **Highest PM2.5**: December 2021 - Winter months often see high PM2.5 levels due to temperature inversions trapping pollutants.
- **Lowest PM10**: Similar trend to PM2.5, with the lowest in March 2022 and highest in December 2021.
- **Highest PM10**: December 2021 - Consistent with high PM2.5, indicating the prevalence of coarse particulate matter during winter.
- **Lowest NO2**: March 2020 - Possibly a result of reduced vehicular and industrial activity during the initial COVID-19 lockdown.
- **Highest NO2**: December 2020 - Increase in vehicular movement and unfavorable meteorological conditions could contribute.

ANALYSIS

DTU

- **Lowest PM2.5**: April 2022 - Improved air quality due to various factors, including meteorology and pollution control.
- **Highest PM2.5**: November 2020 - Winter months experience high PM2.5 due to pollution getting trapped close to the ground.
- **Lowest PM10**: Similar trend to PM2.5, with the lowest in March 2022 and highest in December 2021.
- **Highest PM10**: November 2020 - Consistent with high PM2.5 levels during winter.
- **Lowest NO2**: May 2020 - Lockdown measures and reduced traffic lead to lower NO2 levels.
- **Highest NO2**: December 2020 - Increase in vehicular movement during winter contributes to NO2 buildup.

ANALYSIS

DWARKA

Lowest PM2.5: June 2022 - Possibly due to meteorological conditions dispersing pollutants effectively.

Highest PM2.5: January 2021 - Cold weather and stagnant air lead to high PM2.5 levels.

Lowest PM10: Similar trend as PM2.5, with lowest in June 2022 and highest in January 2021.

Highest PM10: January 2021 - Corroborating with elevated PM2.5 levels in winter.

Lowest NO2: July 2020 - Lockdown and monsoon could contribute to lower NO2 levels.

Highest NO2: December 2020 - Increased vehicular emissions during colder months.

ANALYSIS

SHADIPUR

Lowest PM2.5: March 2022 - Improved air quality during this period, possibly due to mitigation measures.
Highest PM2.5: November 2021 - Winter months show higher PM2.5 due to atmospheric conditions.
Lowest PM10: March 2022 - Consistent with lower PM2.5 levels during the same period.
Highest PM10: November 2021 - Corresponding with high PM2.5 levels in winter.
Lowest NO2: April 2020 - Reduced activity due to lockdown measures.
Highest NO2: November 2020 - Increased NO2 emissions from vehicles and industries during winter.

ANALYSIS

SIRIFORT

Lowest PM2.5: May 2022 - Possibly due to a combination of factors, including weather and emission controls.

Highest PM2.5: November 2021 - Winter months see elevated PM2.5 due to stagnant air.

Lowest PM10: May 2022 - Similar trend as PM2.5, likely reflecting overall lower particulate levels.

Highest PM10: November 2021 - Consistent with high PM2.5 during winter.

Lowest NO2: April 2020 - Lockdown measures and reduced vehicular movement.

Highest NO2: November 2020 - Increased NO2 emissions during winter.

OVERALL ANALYSIS

Winter months, especially December and November, consistently show the highest PM2.5, PM10, and NO2 levels across all stations due to unfavorable weather conditions and increased emissions.

POLICY IMPLICATIONS

Implement stringent pollution control measures during winter months, focusing on reducing vehicular emissions and regulating industrial activity.

Strengthen efforts to control construction and road dust during periods of high pollution.

Promote cleaner transportation alternatives to mitigate vehicular emissions throughout the year.

THANK YOU!