The Relationship between Pollution and Economic Activities

Sara Almadani

Introduction

In this study, I aim to replicate and extend the findings of Vandenbroucke and Zhu (2017) from their work Measures of Pollution. Their research explored the relationship between economic activity, measured by GDP per capita, and air pollution, specifically total PM2.5 and PM10.

Using a new dataset—the WMO SDS-WAS Global Dust Reanalysis Ensemble spanning 2003 to 2022—I seek to examine whether the patterns observed in their study hold true with higher spatial and temporal resolution data. This replication allows for a validation of their findings and provides an updated analysis using up-to-date datasets.

My research investigates the relationship between economic activity, measured by ***GDP per capita***, a key indicator of average economic output per person, and ***the National Activity Index (CFNAI)***, a combination of 85 economic indicators assessing U.S. economic growth and inflationary pressure. It also analyzes ***air pollution levels*** through the following variables:

1. **Dust PM2.5:** Fine particulate matter (<2.5 µm) from dust sources (e.g., deserts and arid regions).
2. **Total PM2.5:** Total fine particulate matter (<2.5 µm) from all sources, including natural and anthropogenic sources.
3. **Dust PM10:** Coarse and fine dust particles (<10 µm) originating from natural dust sources, focusing on dust's role in visibility and atmospheric processes.
4. **Total PM10:** Total particulate matter (<10 µm) from all sources, including anthropogenic sources (e.g., construction and road dust), and natural sources.
5. **Dust Aerosol Optical Depth (AOD):** Measures dust particles' scattering and absorption of sunlight, reflecting their transport, climatic effects, and impact on air quality.

The research aims to answer the following question: How does economic activity influence levels of air pollution?

The hypotheses guiding this study are:

• **Null Hypothesis (H_0):** There is no significant relationship between economic activity and air pollution levels.

• **Alternative Hypothesis (H_a):** There is a significant relationship between economic activity and air pollution levels, indicating that changes in economic activities correspond to changes in pollution levels.

For the Clim680 project, I decided to focus on ***the USA*** because the country's economic structure and environmental policies provide a compelling context for such an analysis. The USA has a high GDP per capita, which enables significant investments in pollution control technologies, regulatory enforcement, and public health initiatives.
By investigating this relationship, I aim to assess whether the USA's economic prosperity translates into lower air pollution levels, as suggested by the Environmental Kuznets Curve, which indicates that pollution levels initially increase with economic growth but decline as income reaches higher levels.

Methodology

Mathematical Expression

The regression model used to analyze the relationship between economic activity and air quality indicators (e.g., dust PM2.5, dust PM10, total PM2.5, total PM10, and dust AOD) can be expressed as:

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t \quad (1)$$

Where:
y_t: Represents the air quality indicators at time t.
x_t: Represents the first difference of GDP per capita (GDP_diff), serving as a measure of economic activity at time t.
β_0: The intercept, indicating the baseline value of the air quality indicator when there are no changes in GDP per capita.
β_1: The slope coefficient, quantifying the change in the air quality indicators associated with a one-unit change in the first difference of GDP per capita.
ε_t: The error term, capturing unexplained variability in y_t.

This modeling approach draws inspiration from Vandenbroucke and Zhu (2017), who investigated the relationship between economic activity and pollution in a cross-country context. Their analysis highlighted the importance of assessing pollution relative to the economic activity that generates it. While their study utilized log-transformed GDP per capita and total PM2.5 data to account for non-linear relationships and cross-country variability, this study focuses on U.S. data and examines direct relationships between GDP differences and air quality indicators. The choice of variables and the methodology reflect the need to adapt their approach to a single-country context, where the variability in economic and pollution measures is narrower.

Data

1. Air Quality Data

Original Air Quality Data Sources

The primary dataset used in this study is the World Meteorological Organization (WMO) Sand and Dust Storm Warning Advisory and Assessment System (SDS-WAS), Global

Dust Reanalysis Ensemble. This dataset integrates multiple dust modeling systems to provide a high-resolution and comprehensive reanalysis of global dust distributions.
The dataset combines outputs from four global dust reanalysis models:

- CAMS-RA: Spatial resolution of 0.75° × 0.75° with 3-hourly temporal resolution, starting in 2003.
- MERRA-2: Spatial resolution of 0.5° × 0.625° with hourly temporal resolution, starting in 1980.
- NAAPS-RA: Spatial resolution of 1° × 1° with daily temporal resolution, starting in 1980.
- SILAM: Spatial resolution of 0.25° × 0.25° with daily temporal resolution, starting in 2003.

Processed Ensemble Data

For this study, the ensemble mean of these models is used to create a single dataset, combining the strengths of the individual models. The processed dataset spans from 2003 to 2022, with the following characteristics:

- Spatial Resolution: High resolution of 0.1° × 0.1°.
- Temporal Resolution: Daily values.

Variables of Interest

The dataset includes multiple parameters relevant to dust pollution analysis, such as:

- Dust PM2.5
- Dust PM10
- Dust aerosol optical depth (AOD)
- Total PM2.5
- Total PM10

2. GDP per Capita Data

GDP per capita data is retrieved from FRED, Federal Reserve Bank of St. Louis.

- **Frequency**: Quarterly
- **Units**: Chained 2017 Dollars, Seasonally Adjusted Annual Rate

By using chained dollars (2017 as the base year), the data accounts for the changing value of money over time, making it suitable for comparing economic performance across different years. It reflects real growth, focusing on actual economic output while discounting the effects of price changes.

3. Chicago Fed National Activity Index

The index data is retrieved from FRED, Federal Reserve Bank of St. Louis.

- **Frequency**: Monthly converted to quarterly.

The index data is useful for composite analysis due to its nature as a standardized index with both positive and negative values, indicating economic activity relative to historical trends. This characteristic allows for clear classification into distinct categories—high (positive values), low (negative values), and moderate (values near zero)—making it good for assessing how different levels of economic activity correlate with environmental or atmospheric variables.

4. Access and Data Path

- Dust Dataset Path: /groups/ESS3/sara/data/GlobDust/ensembled/daily
  - Organized into subdirectories based on statistical measures, including mean and median.
- GDP per Capita Dataset Path: /home/salmadan/GDPperCapita.csv
- Chicago Fed National Activity Index Dataset Path: /home/salmadan/CFNAI.csv

Table 1 provides an overview of the summary statistics for the key variables analyzed in this study, including their means, standard deviations, and ranges, offering a snapshot of the underlying data distribution.

Table 1: Summary Statistics

Statistics	**Count**	**Mean**	**Std**	**Min**	**Max**
Air Pollution Variables					
Dust PM2.5	518400000.0	2.678963	10.291	0.013	590.0775
Dust PM10	518400000	9.605349	39.094	0.042	2858.38
Dust AOD	518400000	0.01854374	0.047	0.0002	0.9432814
Total PM2.5	518400000	9.026587	10.245	1.426	582.037
Total PM10	518400000	25.04368	30.237	3.538	2751.336
Economic Activities Variables					
GDP Per Capita	80	57375.8	50462	54386	4095
National Activity Index	80	-0.110792	0.580	-2.39	1.43

Figure 1 illustrates the time series of GDP per capita (left axis) and the National Activity Index (CFNAI, right axis) from 2003 to 2022, highlighting key economic trends and fluctuations. Both variables exhibit non-stationary behavior, with GDP showing a consistent upward trend over the years, punctuated by sharp declines during significant economic downturns, such as the 2008 global financial crisis and the COVID-19 pandemic in 2020. The CFNAI, while centered around zero, demonstrates substantial variability, reflecting fluctuations in economic activity. The apparent non-stationarity in both series, characterized by the lack of a constant mean or variance over time, underscores the importance of verifying stationarity before conducting further analyses.

Figure 1: Time Series of GDP per Capita and the National Activity Index

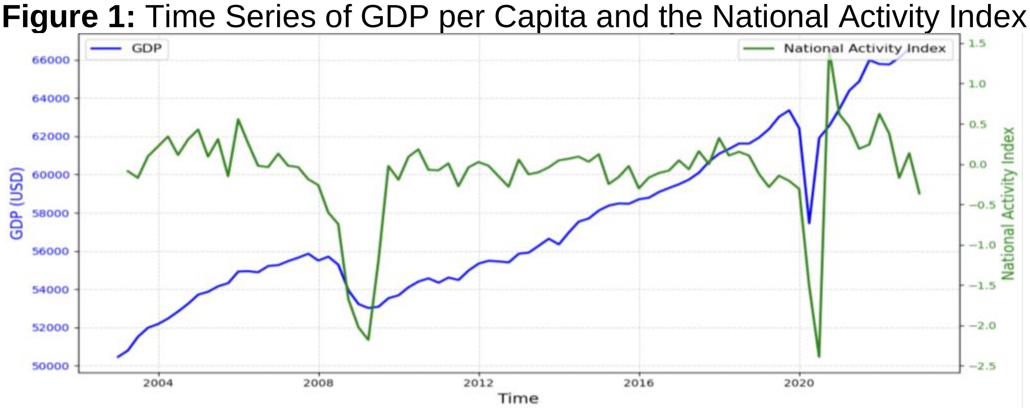
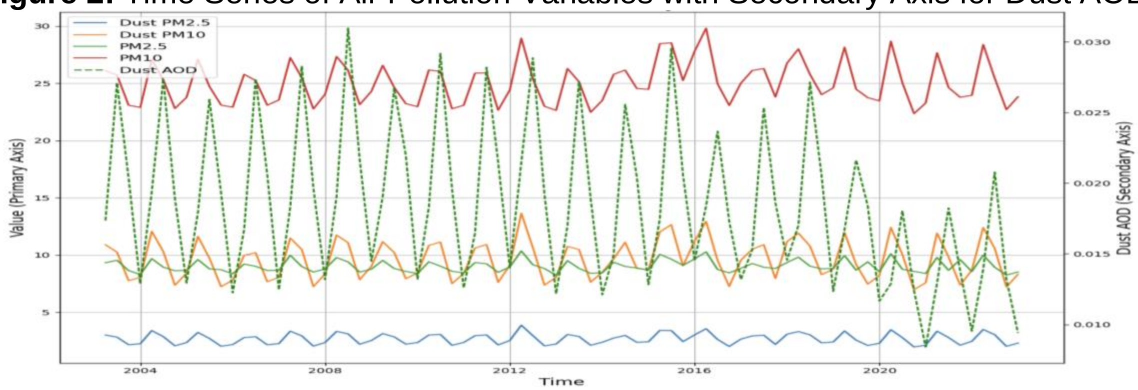


Figure 2 presents the time series of Dust PM2.5, Dust PM10, total PM2.5, and total PM10 concentrations (primary axis) alongside Dust AOD (secondary axis) from 2003 to 2022. The variables exhibit clear seasonal cycles, likely driven by seasonal variations in dust activity. Dust AOD follows a similar seasonal trend, though at a smaller scale and captured on the secondary axis due to its lower magnitude.

Figure 2: Time Series of Air Pollution Variables with Secondary Axis for Dust AOD



5. Stationarity Testing

Formal statistical tests, such as the augmented Dickey–Fuller (ADF) test or autocorrelation function (ACF) analysis, should be applied to avoid spurious regressions. Therefore, the Augmented Dickey-Fuller (ADF) test was conducted to evaluate the stationarity of GDP per capita, the National Activity Index, and the pollution variables (Dust PM2.5, Dust PM10, Dust AOD, total PM2.5, and total PM10).

The formula for the ADF test is as follows:

$$\Delta y_t = \alpha + \gamma y_{t-1} + \sum_{i=1}^{p} \beta_i \, \Delta y_{t-1} + \varepsilon_t \tag{2}$$

Where:
α: Constant term
y_{t-1}: The first difference of the time series y_t.
γy_{t-1}: Lagged level of the series, where γ is the coefficient being tested for stationarity.
$\sum_{i=1}^{p} \beta_i \, \Delta y_{t-1}$: Lagged differences of y_t to account for higher-order autocorrelation.
p: The number of lagged terms included in the model, often selected using criteria such as AIC or BIC.
ε_t: White noise error term.
$H_0 : \gamma = 0$ vs. $H_1 : \gamma < 0$ With H_0 indicating nonstationarity.

The ADF test for GDP per capita yielded a high p-value, indicating that the null hypothesis of non-stationarity cannot be rejected, confirming that the series is non-stationary. For the National Activity Index, the test showed an ADF statistic of -4.78 and a p-value below 0.05, leading to the rejection of the null hypothesis and confirming that the index series is stationary. Similarly, Dust PM2.5, Dust PM10, and total PM2.5 were also found to be stationary, with their null hypotheses rejected at the 5% significance level. However, Dust

AOD and total PM10 remain non-stationary, as their p-values exceed 0.05, indicating the null hypothesis cannot be rejected. These results highlight that while stationary variables can be used directly in time series modeling, non-stationary variables require transformations like differencing or detrending to ensure reliable statistical analysis.

Results

The results of this study provide an explanation of the relationships between economic activity and air quality indicators. The findings are organized into six key sub-sections:

1. **Regression Analysis of GDP per Capita and Air Pollution Variables**: Examines the statistical relationships between economic activity and air quality variables.
2. **Concentration Maps of Air Pollution variables for 2022**: Highlights spatial and seasonal variations in the variables for the last year of the dataset.
3. **Anomaly Analysis**: Investigates quarterly deviations from long-term averages to identify patterns in air pollution variables influenced by natural and anthropogenic factors.
4. **Description of Composite Anomalies for High, Low, and Moderate Phases**: Explores the spatial distribution of air quality anomalies during different phases of the National Activity Index (CFNAI), emphasizing regional patterns and their connection to economic activity.
5. **Differences in Composite Anomalies of Air Pollution Variables Across National Activity Index (CFNAI) Phases**: Analyzes the shifts in air pollution anomalies between CFNAI phases to capture the dynamic impacts of economic fluctuations.
6. **Regression Map: Total PM2.5 Anomalies vs. National Activity Index**: Visualizes the spatial relationships between PM2.5 anomalies and the CFNAI, highlighting regions with statistically significant correlations and the underlying interplay between economic and environmental factors.

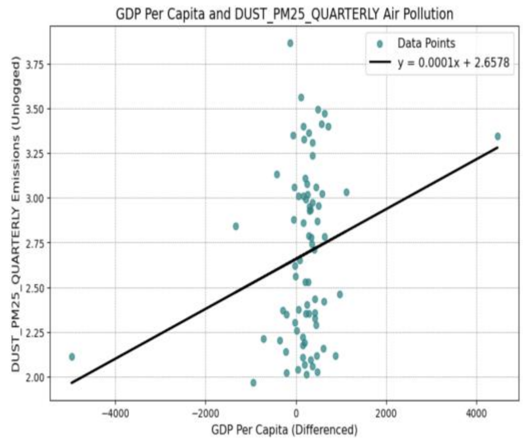
1. Regression Analysis of GDP per Capita and Air Pollution Variables

Regression analyses reveal modest but statistically significant relationships between changes in GDP per capita and air quality indicators, such as total PM2.5 and total PM10 concentrations, with the exception of dust AOD, which showed no significant association. Detailed regression outputs for the logged variables are provided in Annex A.

Dust PM2.5

The regression analysis for quarterly dust PM2.5 concentrations against GDP per capita shows a *statistically significant relationship*, with a p-value of 0.030 for the coefficient of GDP_diff. This indicates that GDP per capita differences have a positive but modest effect on dust PM2.5 levels. However, the R^2 suggests that GDP_diff explains only *6.2%* of the variance in dust PM2.5 concentrations, highlighting that other factors not included in the model likely contribute significantly to PM2.5 levels. *The constant term is also highly significant* (p-value < 0.001), with a value of 2.6578, while the Durbin-Watson statistic of 1.861 suggests minimal autocorrelation in the residuals, supporting the reliability of the model's assumptions.

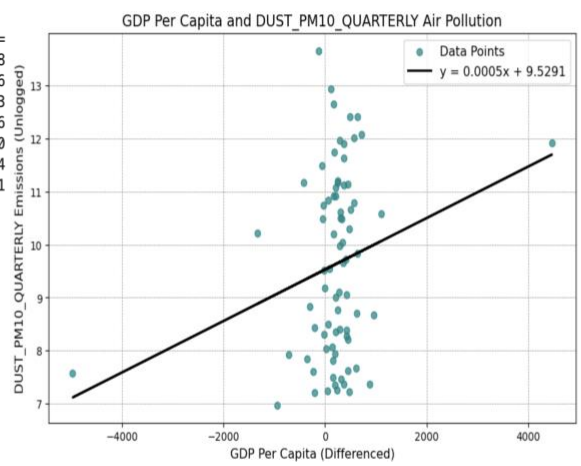
```
Regression Summary for dust_pm25_quarterly:
                            OLS Regression Results
==============================================================================
Dep. Variable:     dust_pm25_quarterly   R-squared:                       0.062
Model:                             OLS   Adj. R-squared:                  0.049
Method:                  Least Squares   F-statistic:                     4.894
Date:                 Mon, 25 Nov 2024   Prob (F-statistic):             0.0300
Time:                         20:48:49   Log-Likelihood:                -49.509
No. Observations:                   76   AIC:                             103.0
Df Residuals:                       74   BIC:                             107.7
Df Model:                            1
Covariance Type:             nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          2.6578      0.055     47.999      0.000       2.547       2.768
GDP_diff       0.0001   6.31e-05      2.212      0.030    1.39e-05       0.000
==============================================================================
Omnibus:                        6.351   Durbin-Watson:                   1.861
Prob(Omnibus):                  0.042   Jarque-Bera (JB):                3.427
Skew:                           0.300   Prob(JB):                        0.180
Kurtosis:                       2.150   Cond. No.                         900.
==============================================================================
```



Dust PM10

The regression analysis for dust PM10 concentrations against GDP per capita indicates a *statistically significant relationship*, with a p-value of 0.036 for the coefficient of GDP_diff. This suggests that GDP per capita differences have a small but statistically significant positive effect on dust PM10 levels. However, the R^2 implies that GDP_diff explains only *5.8%* of the variance in dust PM10. *The constant term is highly significant* (p-value < 0.001), with a value of 9.5291. The Durbin-Watson statistic of 1.827 indicates minimal autocorrelation in the residuals, supporting the reliability of the regression model's assumptions.

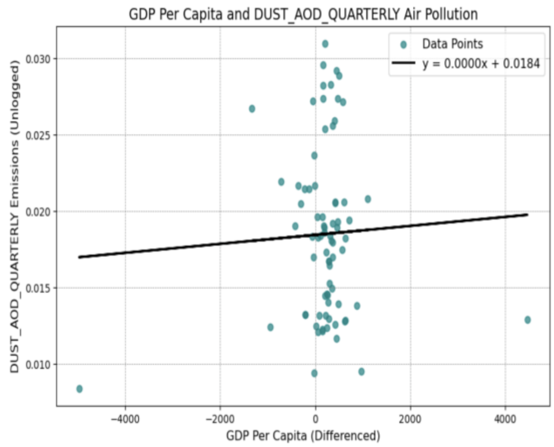
```
Regression Summary for dust_pm10_quarterly:
                            OLS Regression Results
==============================================================================
Dep. Variable:     dust_pm10_quarterly   R-squared:                       0.058
Model:                             OLS   Adj. R-squared:                  0.046
Method:                  Least Squares   F-statistic:                     4.583
Date:                 Mon, 25 Nov 2024   Prob (F-statistic):             0.0356
Time:                         20:48:49   Log-Likelihood:                -146.70
No. Observations:                   76   AIC:                             297.4
Df Residuals:                       74   BIC:                             302.1
Df Model:                            1
Covariance Type:             nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          9.5291      0.199     47.904      0.000       9.133       9.925
GDP_diff       0.0005      0.000      2.141      0.036    3.36e-05       0.001
==============================================================================
Omnibus:                        7.631   Durbin-Watson:                   1.827
Prob(Omnibus):                  0.022   Jarque-Bera (JB):                3.422
Skew:                           0.245   Prob(JB):                        0.181
Kurtosis:                       2.083   Cond. No.                         900.
==============================================================================
```



Dust Aerosol Optical Depth (AOD)

The regression analysis for dust AOD against GDP per capita shows no *statistically significant relationship*, as the p-value for the coefficient of GDP_diff is 0.695. The R^2 indicates that GDP_diff explains just *0.2%* of the variance in dust AOD, and the adjusted R^2 is negative, suggesting that the model does not provide a meaningful explanation for the variation in the dependent variable. The constant term is significant (p-value < 0.001), with a value of 0.0184. The Durbin-Watson statistic of 1.639 suggests low autocorrelation in the residuals, but the overall model fails to identify any substantial relationship between GDP_diff and dust AOD.

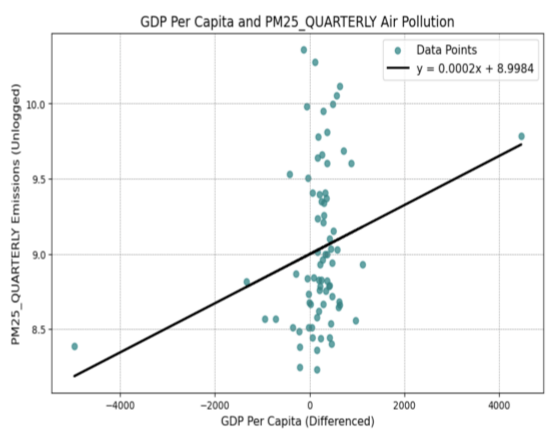
```
Regression Summary for dust_aod_quarterly:
                            OLS Regression Results
==============================================================================
Dep. Variable:     dust_aod_quarterly   R-squared:                       0.002
Model:                            OLS   Adj. R-squared:                 -0.011
Method:                 Least Squares   F-statistic:                    0.1552
Date:                Mon, 25 Nov 2024   Prob (F-statistic):              0.695
Time:                        20:48:49   Log-Likelihood:                 287.80
No. Observations:                  76   AIC:                            -571.6
Df Residuals:                      74   BIC:                            -566.9
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0184      0.001     28.182      0.000       0.017       0.020
GDP_diff    2.939e-07   7.46e-07      0.394      0.695   -1.19e-06    1.78e-06
==============================================================================
Omnibus:                        5.122   Durbin-Watson:                   1.639
Prob(Omnibus):                  0.077   Jarque-Bera (JB):                4.300
Skew:                           0.483   Prob(JB):                        0.116
Kurtosis:                       2.349   Cond. No.                         900.
==============================================================================
```



Total PM2.5

The regression analysis for PM2.5 against GDP per capita indicates a *statistically significant positive relationship*, with a p-value of 0.021 for the coefficient of GDP_diff. The R^2 suggests that GDP_diff explains 7.0% of the variance in PM2.5 concentrations, indicating a modest explanatory power for the model. The constant term is highly significant (p-value < 0.001), with a value of 8.9984. The Durbin-Watson statistic of 2.112 suggests minimal autocorrelation in the residuals, further supporting the reliability of the model. Overall, while GDP_diff exhibits a significant association with PM2.5 levels, the relatively low R^2 value suggests that additional variables likely contribute to PM2.5 variability.

```
Regression Summary for pm25_quarterly:
                            OLS Regression Results
==============================================================================
Dep. Variable:         pm25_quarterly   R-squared:                       0.070
Model:                            OLS   Adj. R-squared:                  0.058
Method:                 Least Squares   F-statistic:                     5.586
Date:                Mon, 25 Nov 2024   Prob (F-statistic):             0.0207
Time:                        20:48:49   Log-Likelihood:                -56.458
No. Observations:                  76   AIC:                             116.9
Df Residuals:                      74   BIC:                             121.6
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          8.9984      0.061    148.308      0.000       8.877       9.119
GDP_diff       0.0002   6.92e-05      2.363      0.021    2.57e-05       0.000
==============================================================================
Omnibus:                        7.112   Durbin-Watson:                   2.112
Prob(Omnibus):                  0.029   Jarque-Bera (JB):                7.410
Skew:                           0.757   Prob(JB):                       0.0246
Kurtosis:                       2.788   Cond. No.                         900.
==============================================================================
```

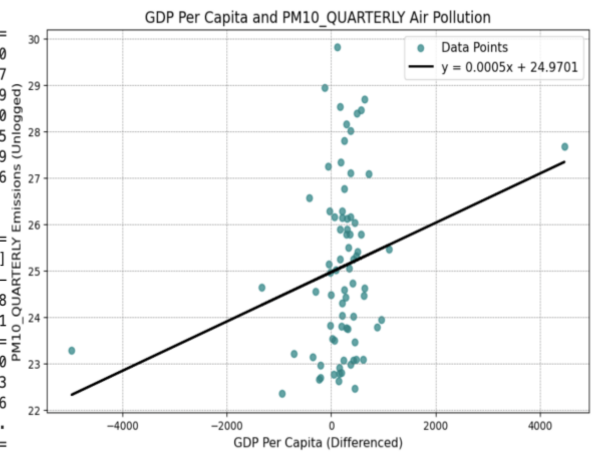


Total PM10

The regression analysis for PM10 concentrations against GDP per capita shows a *statistically significant positive relationship*, with a p-value of 0.033 for the coefficient of GDP_diff. This suggests that GDP per capita have a small but significant effect on PM10 concentrations. The R^2 indicates that GDP_diff explains 6.0% of the variance in PM10. The constant term is highly significant (p-value < 0.001), with a value of 24.9701. The

Durbin-Watson statistic of 1.750 suggests low autocorrelation in the residuals, supporting the validity of the regression model's assumptions.

```
Regression Summary for pm10_quarterly:
                            OLS Regression Results
==============================================================================
Dep. Variable:         pm10_quarterly   R-squared:                       0.060
Model:                            OLS   Adj. R-squared:                  0.047
Method:                 Least Squares   F-statistic:                     4.719
Date:                Mon, 25 Nov 2024   Prob (F-statistic):             0.0330
Time:                        20:48:49   Log-Likelihood:                -152.45
No. Observations:                  76   AIC:                             308.9
Df Residuals:                      74   BIC:                             313.6
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         24.9701      0.215    116.376      0.000      24.543      25.398
GDP_diff       0.0005      0.000      2.172      0.033     4.4e-05       0.001
==============================================================================
Omnibus:                        4.751   Durbin-Watson:                   1.750
Prob(Omnibus):                  0.093   Jarque-Bera (JB):                4.653
Skew:                           0.561   Prob(JB):                       0.0976
Kurtosis:                       2.541   Cond. No.                         900.
==============================================================================
```



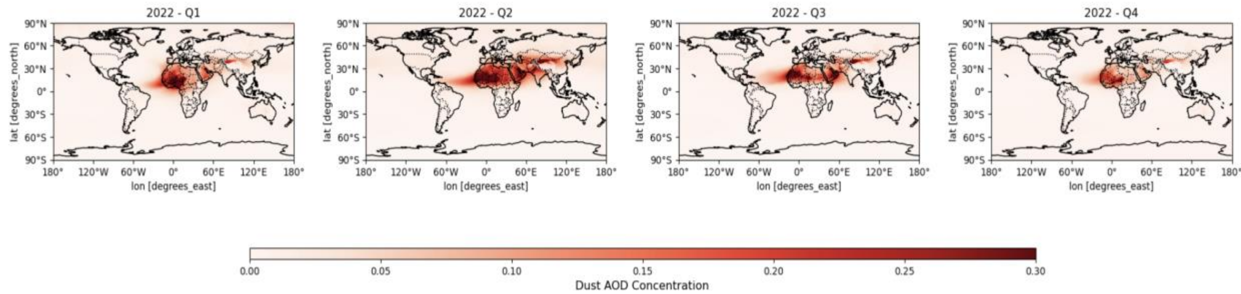
2. Concentration Maps of Air Pollution Variables for 2022

The concentration maps present the spatial distribution of air pollution variables across the four quarters of 2022, the final year in the dataset. Focusing on this year allows for a detailed examination of *recent* air pollution trends. The maps highlight variations in dust PM10, dust AOD, total PM2.5, and total PM10, with a specific emphasis on the United States as the study area.

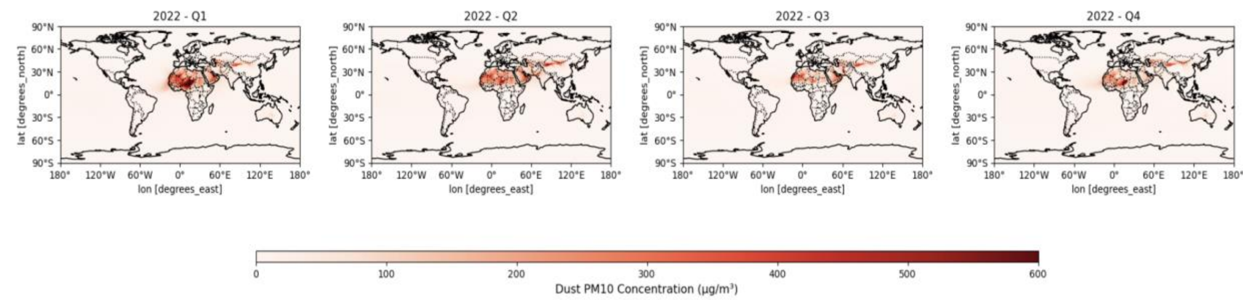
Within the USA, dust-related pollutants such as dust PM10 and dust AOD show relatively *low concentrations* compared to global hotspots like northern Africa and the Middle East. The highest dust levels in the USA are observed in the southwestern states, attributed to localized natural dust sources. These concentrations remain fairly consistent across quarters, reflecting the limited impact of seasonal variations on dust emissions in the USA.

In contrast, anthropogenic pollutants such as total PM2.5 and total PM10 exhibit more spatial variability, with elevated levels in urban and industrialized regions, particularly in the Midwest and along the East Coast. Seasonal trends suggest higher total PM2.5 concentrations during the colder months, likely driven by residential heating, while total PM10 levels peak slightly during the spring and summer, possibly influenced by increased dust activity and other atmospheric conditions.

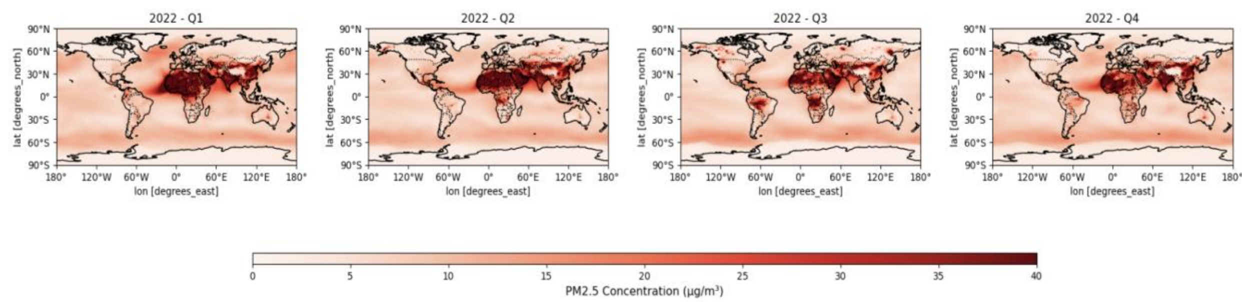
Dust AOD Concentration Maps for 2022



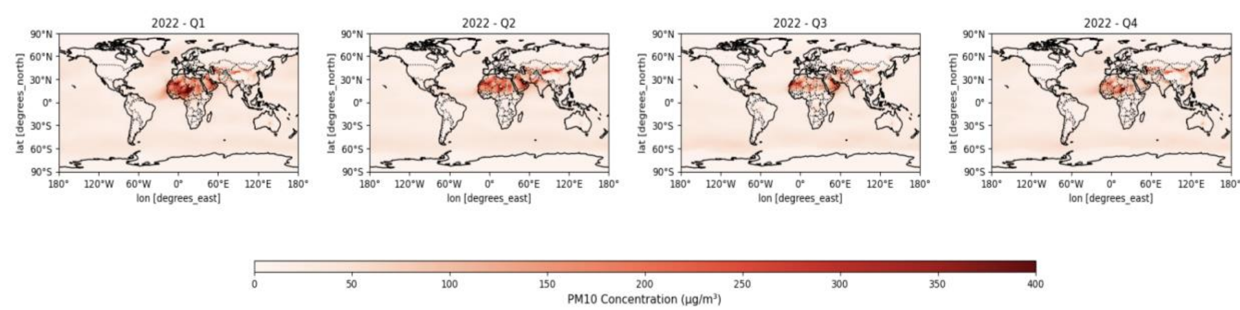
Dust PM10 Concentration Maps for 2022



PM2.5 Concentration Maps for 2022



PM10 Concentration Maps for 2022

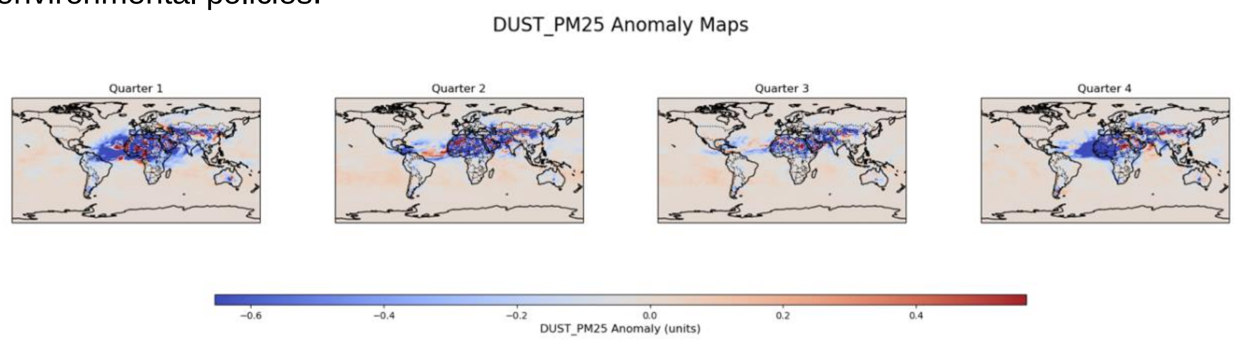


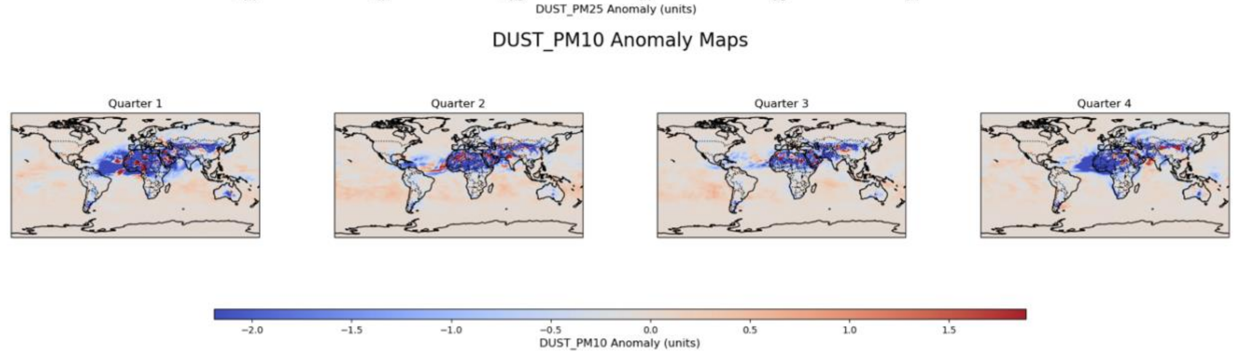
3. Anomaly Analysis

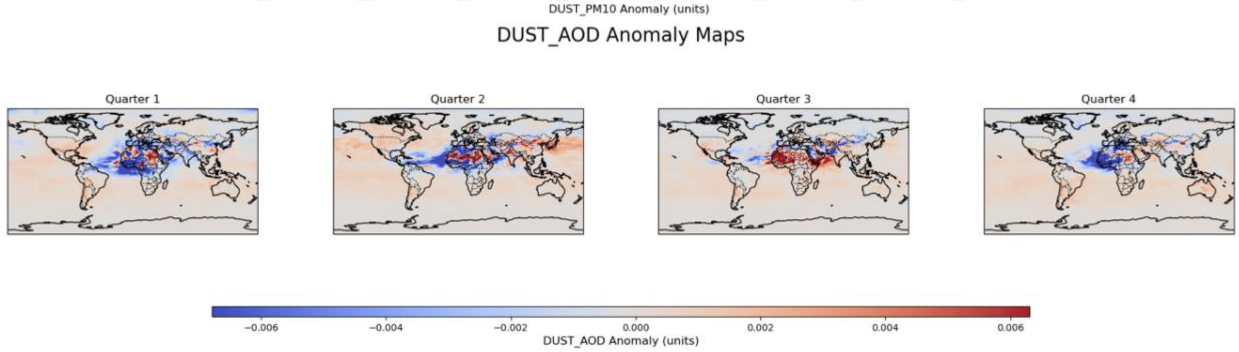
The anomaly maps for dust PM2.5, dust PM10, dust AOD, total PM2.5, and total PM10 illustrate quarterly deviations from long-term averages, showing significant spatial and temporal variations in air pollution. In the USA, positive anomalies in dust PM2.5 and dust PM10 are observed in the southwestern and western regions during dry seasons (Quarters 2 and 3), while negative anomalies dominate the eastern USA during wetter seasons (Quarters 1 and 4). Globally, notable positive anomalies align with major natural dust sources, such as the Sahara Desert and the Middle East.

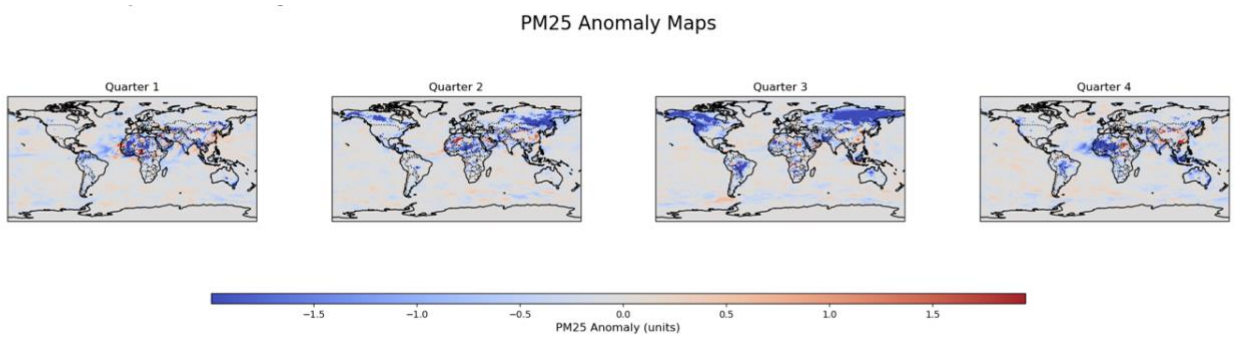
Dust AOD anomalies highlight the contribution of arid regions to atmospheric optical depth, with the USA showing generally neutral anomalies.

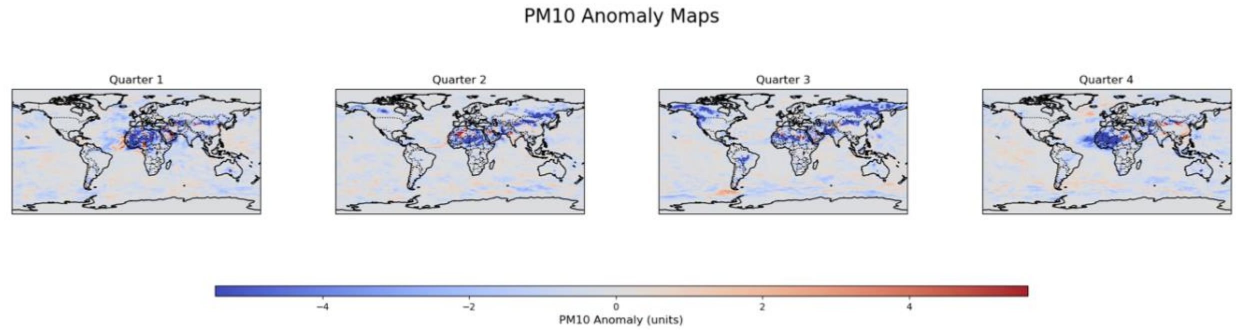
For total PM2.5 and total PM10, the combined influence of anthropogenic and natural factors is evident, with positive anomalies in urban and industrialized regions during winter and negative anomalies in the western USA during summer. These patterns highlight the interplay of emissions, meteorological conditions, and regional characteristics in shaping air quality, providing insights for public health and environmental policies.





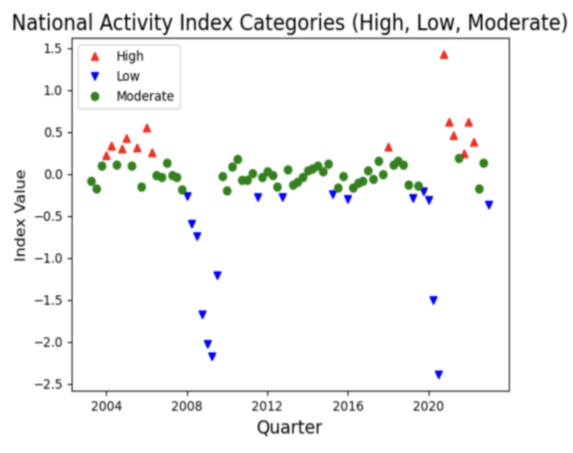


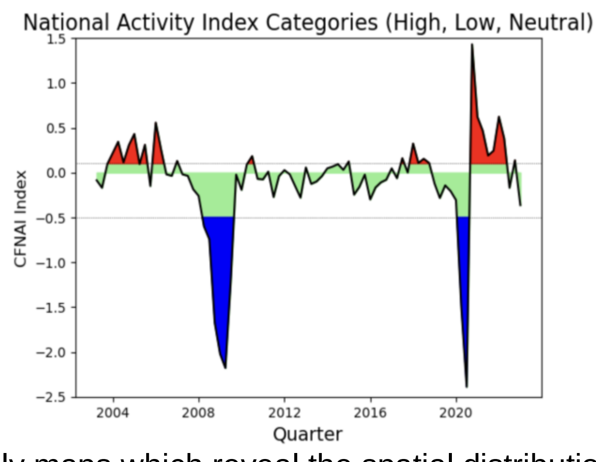




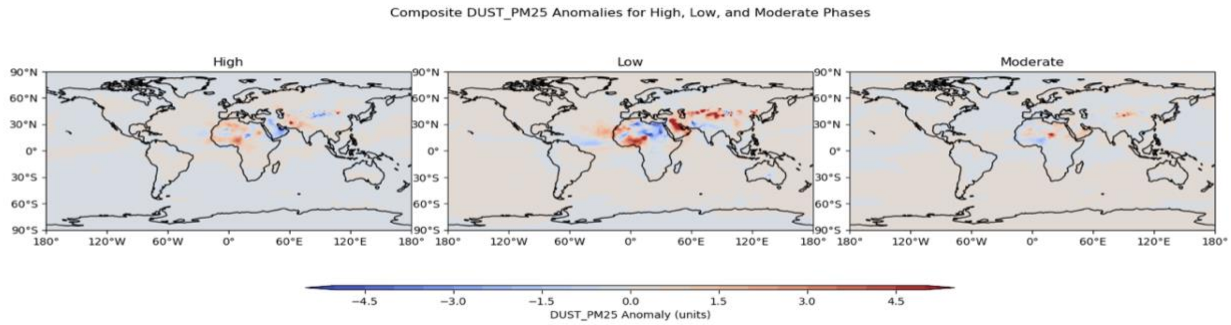
4. Description of Composite Anomalies for High, Low, and Moderate Phases

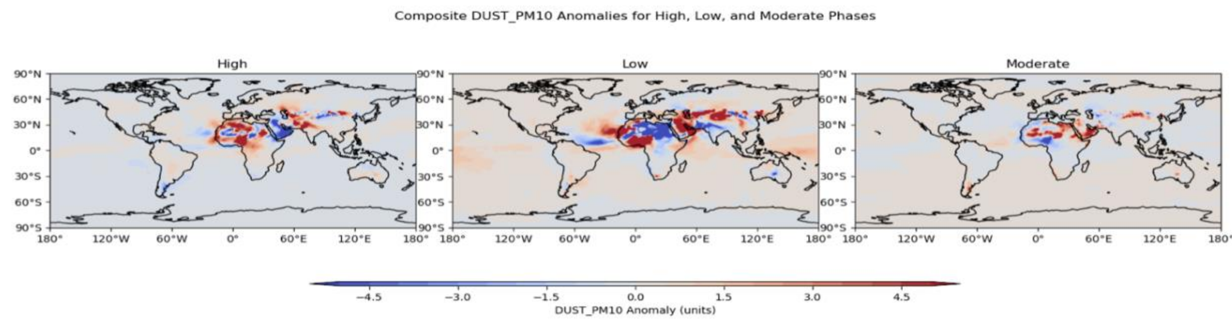
The figures below show the classification of the National Activity Index (CFNAI) into high, low, and neutral (or moderate) categories across the dataset's timeline. These classifications provide a visualization of economic activity fluctuations and their alignment with significant events, such as recessions or economic booms.

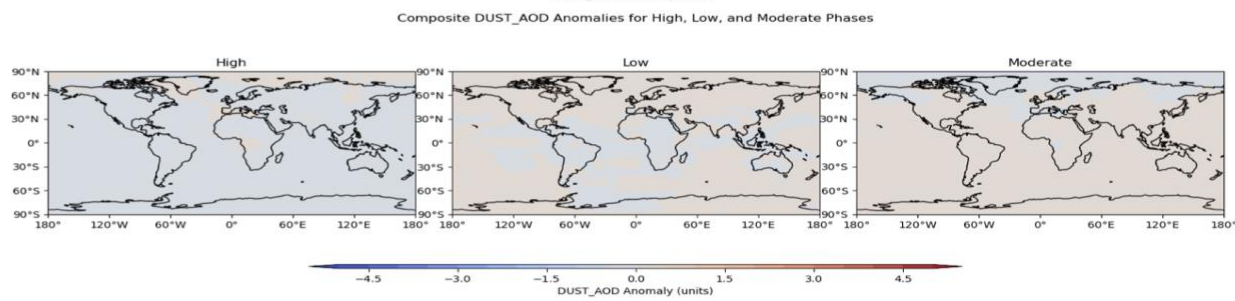


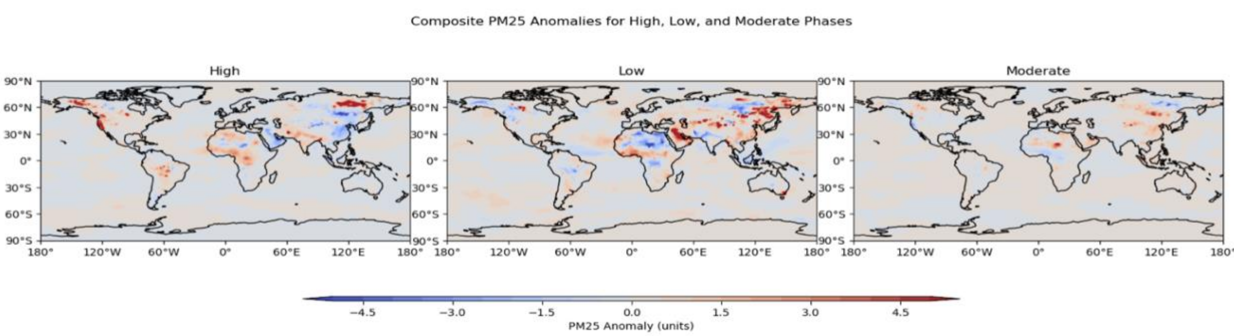


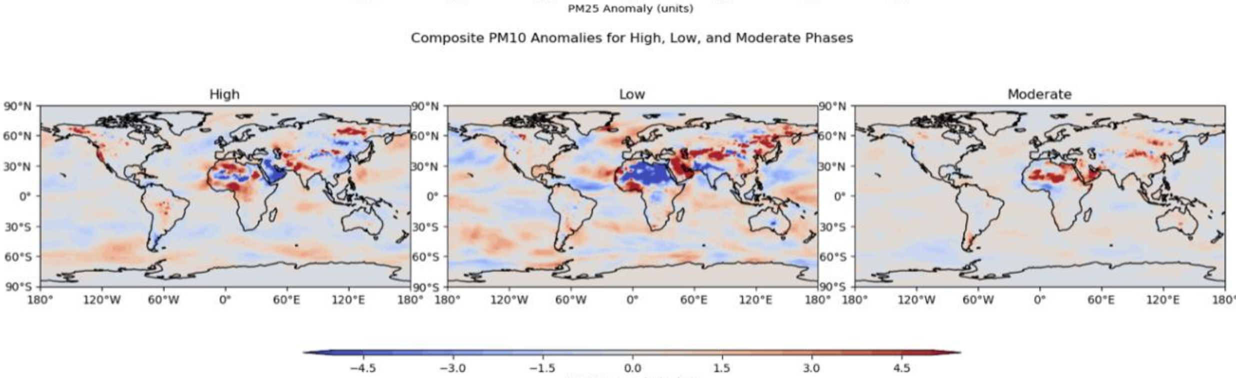
The figures below show composite anomaly maps which reveal the spatial distribution of deviations in air pollution variables across the High, Low, and Moderate phases of the National Activity Index. During the High phase, robust economic activity corresponds with positive anomalies in dust-related concentrations over arid regions like Northern Africa, the Middle East, and South Asia, as well as anthropogenic pollution hotspots in eastern Asia. In contrast, the Low phase shows negative anomalies, reflecting reduced emissions and improved air quality in dust-prone and industrial regions. The Moderate phase displays weaker and more diffuse anomalies, indicating a less impact of intermediate economic activity.


Composite DUST_PM25 Anomalies for High, Low, and Moderate Phases
High
Low
Moderate
DUST_PM25 Anomaly (units)


Composite DUST_PM10 Anomalies for High, Low, and Moderate Phases
High
Low
Moderate
DUST_PM10 Anomaly (units)

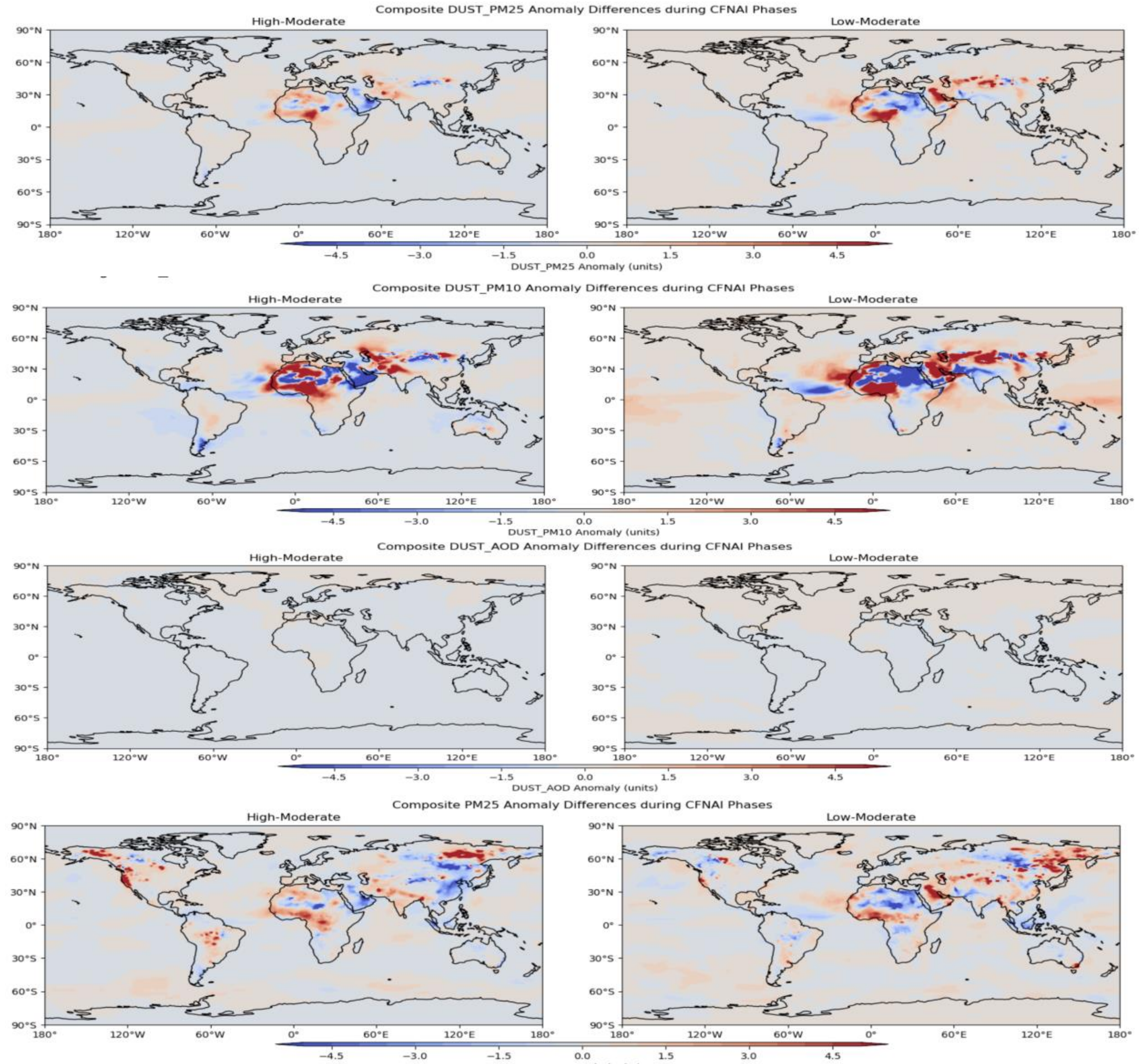

Composite DUST_AOD Anomalies for High, Low, and Moderate Phases
High
Low
Moderate
DUST_AOD Anomaly (units)

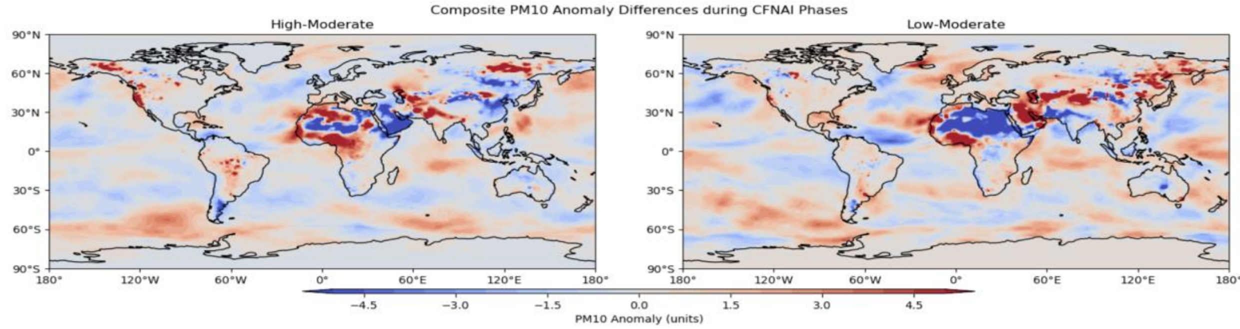

Composite PM25 Anomalies for High, Low, and Moderate Phases
High
Low
Moderate
PM25 Anomaly (units)


Composite PM10 Anomalies for High, Low, and Moderate Phases
High
Low
Moderate
PM10 Anomaly (units)

5. Differences in Composite Anomalies of Air Pollution Variables Across National Activity Index (CFNAI) Phases

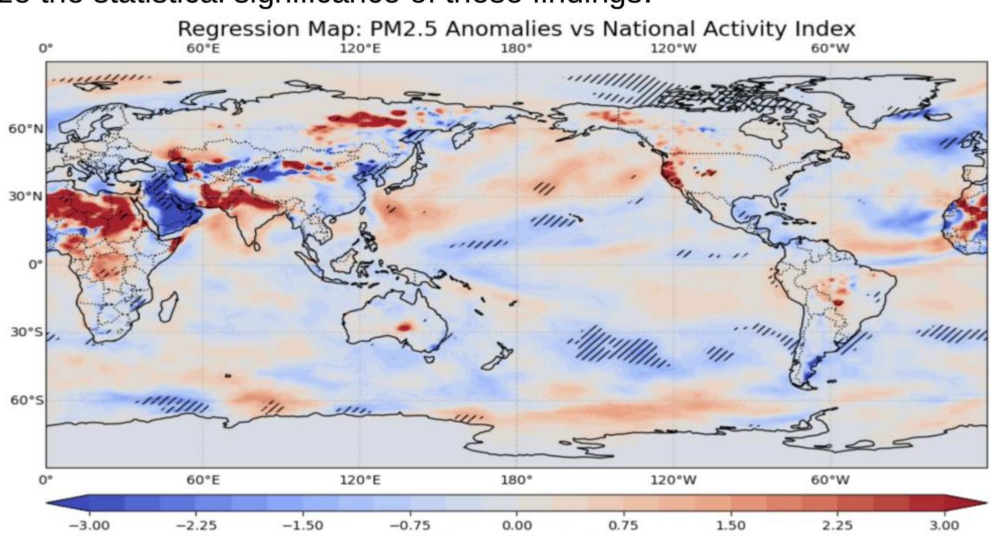
The maps show composite differences in air pollution anomalies between different CFNAI phases (High-Moderate and Low-Moderate) with a particular focus on the USA. Significant positive anomalies are observed in High phases compared to Moderate phases, particularly over regions with high dust activity, such as the Middle East and parts of the USA. Negative anomalies during Low phases relative to Moderate phases are evident over the USA, emphasizing reduced concentrations during periods of lower economic activity. Interestingly, the Dust AOD anomaly maps show minimal variations across CFNAI phases, suggesting that the optical depth of dust aerosols is less sensitive to CFNAI-defined economic fluctuations.





6. Regression Map: Total PM2.5 Anomalies vs. National Activity Index

The regression map illustrates the relationship between PM2.5 anomalies and the National Activity Index, focusing on total PM2.5, which includes both anthropogenic pollutant sources, such as industrial and vehicular emissions, and natural sources, such as dust and wildfire smoke. The map uses color gradients to indicate the regression coefficients, with *red areas* showing a positive relationship (higher PM2.5 levels with increased economic activity) and *blue areas* indicating a negative relationship (lower PM2.5 levels with increased economic activity). Regions with "/" markings represent statistically significant results, highlighting areas where the relationship between PM2.5 and economic activity is robust. Within the USA, regions with positive regression coefficients underscore the impact of industrial and urban emissions, while negative coefficients may reflect successful emission reduction measures or shifts toward cleaner economic activities. Outside the USA, the Middle East presents a surprising inverse relationship (negative coefficients), which can be attributed to its unique economic structure and natural environment. The region's reliance on oil production, typically conducted in sparsely populated areas, and investments in advanced emission controls may mitigate urban PM2.5 levels. Furthermore, the dominance of natural dust storms in the region's PM2.5 composition, driven by meteorological factors rather than human activity, contributes to this negative relationship. The "/" markings in these areas emphasize the statistical significance of these findings.



Discussion and Conclusions

This study examines the relationships between economic activity and air pollution indicators, with a focus on the United States. By analyzing regression models, spatial concentration patterns, anomaly trends, and composite maps across varying economic phases*, the findings reveal statistically significant associations between GDP per capita changes and particulate matter concentrations*. The results highlight the complexity of air quality drivers, where both anthropogenic and natural factors influence pollution levels. For instance, natural dust activity and seasonal meteorological variations contribute significantly to particulate matter concentrations, particularly in regions like the southwestern United States.

Composite anomaly maps further illustrate the influence of economic activity, as measured by the National Activity Index (CFNAI), on air pollution patterns during high, moderate, and low phases. *High CFNAI phases correspond to increased particulate matter in key regions, reflecting heightened economic and industrial activities*, *whereas low CFNAI phases show reduced concentrations, suggesting emission declines during periods of economic slowdown*. The regression map of PM2.5 anomalies against the CFNAI provides additional insights, particularly in the United States, where industrialized and urban regions *exhibit stronger positive relationships with economic activity*.

Despite its contributions, this study has notable *limitations*. A key limitation is the presence of missing daily NetCDF files, with 20 files unavailable out of the expected total. This missing data could impact the accuracy of computed quarterly averages and potentially bias the results, particularly if the missing data corresponds to periods with atypical dust levels. For instance, the absence of data from high-dust events could underestimate the relationship between economic activity and dust-related air quality indicators. Furthermore, the reduced dataset may have weakened the statistical power of the regression analysis, leading to lower R^2 values and less precise p-values. While resampling techniques were employed to mitigate these effects, the potential for unmeasured biases remains a consideration when interpreting the findings.

In conclusion, this study highlights the modest and regionally variable impacts of economic activity on air quality, emphasizing the interplay between anthropogenic emissions and natural factors. While the focus on the United States provides useful insights into the role of economic structure and environmental policy, the study's limitations point to the need for *future research* using more complete datasets.

References

World Health Organization (WHO). (2018). *Air pollution and health*. Retrieved from https://www.who.int

Vandenbroucke, G., & Zhu, H. (2017). *Measures of Pollution. Economic Synopses,* (2017, No. 9). Federal Reserve Bank of St. Louis. Retrieved from https://fraser.stlouisfed.org/title/economic-synopses-6715/measures-pollution-624497

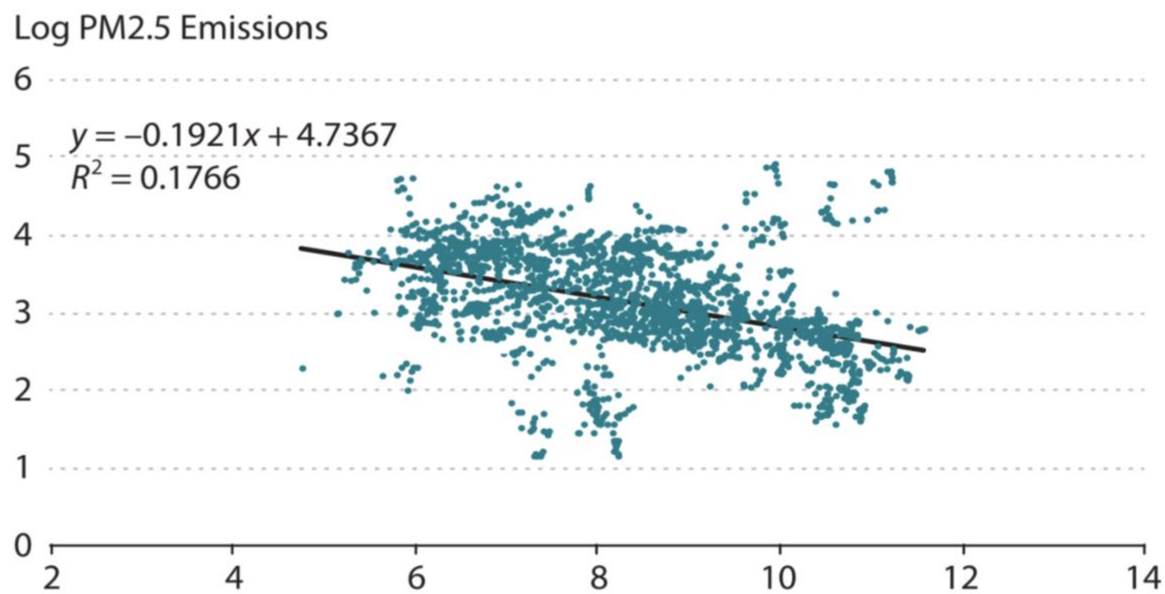
Enders, W. (2014). *Applied Econometric Time Series* (4th ed.). Wiley.

U.S. Bureau of Economic Analysis, Real gross domestic product per capita [A939RX0Q048SBEA], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/A939RX0Q048SBEA,November 18, 2024.

Federal Reserve Bank of Chicago, Chicago Fed National Activity Index [CFNAI], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CFNAI, November 27, 2024.

Appendix A

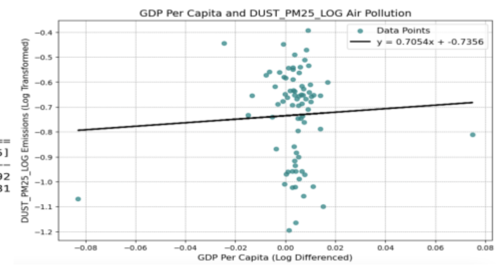
Vandenbroucke and Zhu (2017) employed logarithmic transformations of GDP per capita and Total PM2.5 emissions in their analysis, as illustrated in the figure below. The use of log serves two main purposes. First, it helps linearize potentially nonlinear relationships, making regression models more interpretable and suitable for statistical analysis. Second, it mitigates the influence of outlier observations. Following this methodology, I applied logarithmic transformations to the non-stationary variables, followed by computing their first differences to address stationarity issues and prepare the data for further analysis.



The regression results (see below) for the log-transformed variables indicate *very low R^2* values across all models, suggesting that the independent variable (Log_GDP_diff) explains only a minimal portion of the variability in the dependent variables. Additionally, the p-values for the coefficients of Log_GDP_diff are consistently above the significance level of 0.05. Upon further review of the methodology employed by Vandenbroucke and Zhu (2017), it appeared that their analysis involved cross-country panel data, which inherently introduces greater variability and allows for the exploration of broader patterns across different economies. This justifies their use of logarithmic transformations, as panel data often encompass a wide range of GDP values and total PM2.5 emissions spanning multiple orders of magnitude. By contrast, my analysis focuses on data from the United States, where GDP per capita and pollution variables likely exhibit less variability, reducing the necessity for log transformations.
To test this hypothesis, I re-ran the regressions without log transformations (see Result section).

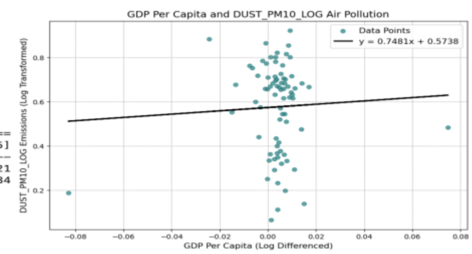
Dust PM2.5:

```
Regression Summary for dust_pm25_log:
                            OLS Regression Results
==============================================================================
Dep. Variable:           dust_pm25_log   R-squared:                       0.003
Model:                             OLS   Adj. R-squared:                 -0.010
Method:                  Least Squares   F-statistic:                    0.2230
Date:                 Mon, 25 Nov 2024   Prob (F-statistic):              0.638
Time:                         20:11:56   Log-Likelihood:                 20.699
No. Observations:                   77   AIC:                            -37.40
Df Residuals:                       75   BIC:                            -32.71
Df Model:                            1
Covariance Type:             nonrobust
================================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
--------------------------------------------------------------------------------
const         -0.7356      0.022    -33.491      0.000      -0.779      -0.692
Log_GDP_diff   0.7054      1.494      0.472      0.638      -2.270       3.681
==============================================================================
Omnibus:                        5.527   Durbin-Watson:                   1.726
Prob(Omnibus):                  0.063   Jarque-Bera (JB):                5.244
Skew:                          -0.579   Prob(JB):                       0.0726
Kurtosis:                       2.457   Cond. No.                         70.0
==============================================================================
```



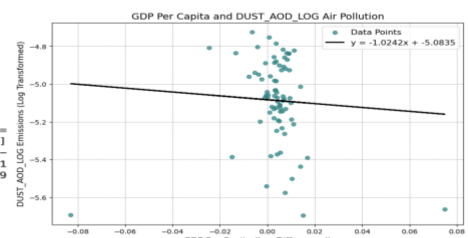
Dust PM10:

```
Regression Summary for dust_pm10_log:
                            OLS Regression Results
==============================================================================
Dep. Variable:           dust_pm10_log   R-squared:                       0.003
Model:                             OLS   Adj. R-squared:                 -0.010
Method:                  Least Squares   F-statistic:                    0.2121
Date:                 Mon, 25 Nov 2024   Prob (F-statistic):              0.646
Time:                         20:11:56   Log-Likelihood:                 14.251
No. Observations:                   77   AIC:                            -24.50
Df Residuals:                       75   BIC:                            -19.82
Df Model:                            1
Covariance Type:             nonrobust
================================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
--------------------------------------------------------------------------------
const          0.5738      0.024     24.026      0.000       0.526       0.621
Log_GDP_diff   0.7481      1.624      0.461      0.646      -2.488       3.984
==============================================================================
Omnibus:                        5.856   Durbin-Watson:                   1.665
Prob(Omnibus):                  0.053   Jarque-Bera (JB):                5.814
Skew:                          -0.630   Prob(JB):                       0.0546
Kurtosis:                       2.528   Cond. No.                         70.0
==============================================================================
```



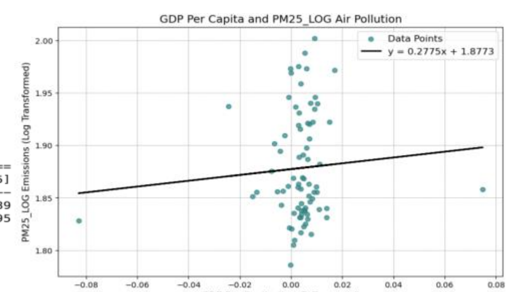
Dust AOD

```
Regression Summary for dust_aod_log:
                            OLS Regression Results
==============================================================================
Dep. Variable:            dust_aod_log   R-squared:                       0.004
Model:                             OLS   Adj. R-squared:                 -0.009
Method:                  Least Squares   F-statistic:                    0.3242
Date:                 Mon, 25 Nov 2024   Prob (F-statistic):              0.571
Time:                         20:11:56   Log-Likelihood:                 6.3956
No. Observations:                   77   AIC:                            -8.791
Df Residuals:                       75   BIC:                            -4.104
Df Model:                            1
Covariance Type:             nonrobust
================================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
--------------------------------------------------------------------------------
const         -5.0835      0.026   -192.214      0.000      -5.136      -5.031
Log_GDP_diff  -1.0242      1.799     -0.569      0.571      -4.607       2.559
==============================================================================
Omnibus:                       12.079   Durbin-Watson:                   0.606
Prob(Omnibus):                  0.002   Jarque-Bera (JB):               12.554
Skew:                          -0.931   Prob(JB):                      0.00188
Kurtosis:                       3.670   Cond. No.                         70.0
==============================================================================
```



Total PM2.5

```
Regression Summary for pm25_log:
                            OLS Regression Results
==============================================================================
Dep. Variable:                pm25_log   R-squared:                       0.006
Model:                             OLS   Adj. R-squared:                 -0.007
Method:                  Least Squares   F-statistic:                    0.4579
Date:                 Mon, 25 Nov 2024   Prob (F-statistic):              0.501
Time:                         20:11:56   Log-Likelihood:                 120.22
No. Observations:                   77   AIC:                            -236.4
Df Residuals:                       75   BIC:                            -231.8
Df Model:                            1
Covariance Type:             nonrobust
================================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
--------------------------------------------------------------------------------
const          1.8773      0.006    311.283      0.000       1.865       1.889
Log_GDP_diff   0.2775      0.410      0.677      0.501      -0.540       1.095
==============================================================================
Omnibus:                        7.075   Durbin-Watson:                   2.094
Prob(Omnibus):                  0.029   Jarque-Bera (JB):                6.115
Skew:                           0.601   Prob(JB):                       0.0470
Kurtosis:                       2.322   Cond. No.                         70.0
==============================================================================
```



Total PM10:

```
Regression Summary for pm10_log:
                            OLS Regression Results
==============================================================================
Dep. Variable:                pm10_log   R-squared:                       0.014
Model:                             OLS   Adj. R-squared:                  0.001
Method:                  Least Squares   F-statistic:                     1.081
Date:                 Mon, 25 Nov 2024   Prob (F-statistic):              0.302
Time:                         20:11:57   Log-Likelihood:                 98.516
No. Observations:                   77   AIC:                            -193.0
Df Residuals:                       75   BIC:                            -188.3
Df Model:                            1
Covariance Type:             nonrobust
================================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
--------------------------------------------------------------------------------
const          2.8684      0.008    358.786      0.000       2.852       2.884
Log_GDP_diff   0.5652      0.544      1.039      0.302      -0.518       1.648
==============================================================================
Omnibus:                        6.701   Durbin-Watson:                   1.883
Prob(Omnibus):                  0.035   Jarque-Bera (JB):                3.320
Skew:                           0.261   Prob(JB):                        0.190
Kurtosis:                       2.127   Cond. No.                         70.0
==============================================================================
```