Seasonal and Geographic Air Pollution Trends: A Remote Sensing Perspective

Alison Holderbaum

Department of Geography and Environmental Sustainability, University of Oklahoma

GIS 4133: Fundamentals of Remote Sensing

Dr. Yanhua Xie

December 3, 2024

Abstract

This paper will look at trends in air pollution in the United States by season and state. Air pollutants such as sulfur dioxide, nitrogen dioxide, and carbon monoxide can lead to health and environmental concerns. Understanding what areas and times of year have higher levels of these pollutants can help guide climate policy. Sentinel-5P data is used to find the average level of air pollution in each state for each season. Data for each pollutant is combined and normalized to give a complete picture of how much pollution is in each state. Google Earth Engine is used for most of the analysis. There was not a strong connection found between air pollution and either of the variables. More research could be done in this area with different data or methods to ensure that there actually is not a trend.

Keywords: Air pollution, particulate matter concentration, remote sensing, seasonal differences, regional differences.

Introduction

Despite being invisible in most areas, air pollution has a tremendous impact on human and environmental health. Short term exposure to air pollution can cause respiratory illness such as asthma, wheezing, and shortness of breath (Manisalidis et al., 2020). Long term exposure can lead to an increased chance of severe respiratory and cardiovascular disease as well as mental health issues (Manisalidis et al., 2020). Although air pollution specifically refers to particulate matter held in the air and atmosphere, an increase in air pollution also leads to more pollution in water and vegetation. This has an impact on all parts of the environment and human health due to the prevalence of particulate matter beyond what is just in the air. It is important for society to understand what areas tend to have higher concentrations of atmospheric pollutants in order to properly attempt to reduce pollution and increase health.

While there has been a large amount of research done to look at the connection between air pollution and either location or season, there is significantly less research looking at the connection between the two variables. A study of pollution in the 88 largest metropolitan areas in the United States found that levels of pollution may correspond with different regions (Dominici et al., 2003). Some of this variation might be from the energy used in each region. For example, there is an increase in burning wood for energy in the Northeast during the colder part of the year that increases air pollution compared to warmer parts of the year (Peng et al., 2005). Similar variations occur for other regions and seasons. While variations in particulate matter concentrations across different regions have been found, these differences may not be significant due to overlapping error bars (Peng et al., 2005). It has been found that particulate matter concentration may be higher in “California and across the Midwest and Southeast” (Dominici et al., 2003). However, this conclusion does not match mortality trends in those areas (Dominici et al., 2003). This can be discounted due to the unreliability of mortality as a climate predictor. While it is commonly used, it naturally tends to increase in colder seasons and decrease in warmer seasons (Peng et al., 2005). This follows the natural trends of infectious diseases such as the flu which lead to an increase in mortality.

Despite the inclusivity of previous research, there is still some evidence that there are differences in air pollution tied to both location and season. However, the causes, specific trends and significance of these variations may vary depending on how the research is performed.

Additional research shows that there is not a trend in pollution during the winter with random outliers (Bell et al., 2007). It also finds a significant increase in the Northeast during summer, contrary to research that suggests that the pollution concentration in the Northeast should increase during the winter due to the burning of firewood. More research needs to be done on this topic, especially due to the large variation in the conclusions from previous research. There is especially a need for research that looks at both the geographical and seasonal impacts on air pollution.

Research Questions

Research Question 1: Is there a significant difference in air pollution between different seasons of the year?

Research Question 2: Is there a state or region of the United States that has higher air pollution? Does this difference in air pollution vary seasonally?

Research Question 3: Are there any apparent large-scale trends to air pollution or is it seemingly random?

Methodology

To analyze the connection between geography and air pollution, we need to find the average concentration of atmospheric particulate matter for each state. In order to get a representation of air quality in each study area, a combination of air pollutants are looked at. The Sentinel-5P satellite, operated by the European Union's Copernicus mission, provides data on multiple atmospheric pollutants, allowing us to look at similarly structured data for different aspects of air pollution. The data used for this project will include sulfur dioxide, nitrogen dioxide, and carbon monoxide. Sulfur dioxide is primarily released into the atmosphere by anthropogenic processes such as burning fossil fuels to create energy. It can cause acid rain, difficulty breathing, and harm the growth of vegetation (US EPA, 2016a). Similar to other air pollutants, nitrogen dioxide can cause respiratory illnesses. It also has environmental impacts including acid rain and haze that reduces visibility. Nitrogen dioxide primarily comes from vehicle emissions and power plants (US EPA, 2016b). Carbon monoxide also is released from burning fossil fuels including vehicle emissions. While carbon monoxide is known for causing dizziness or unconsciousness, this is unlikely to happen while outside. There is still a health risk for atmospheric carbon monoxide especially if you have heart problems (US EPA, 2016c). All three of these compounds cause health and environmental risks. They also are largely created by activities that could be reduced or modified to ensure that there are less pollutants released into the atmosphere. Air pollution impacts everyone no matter where in the United States they live, however there may be higher health and environmental risk in certain areas or seasons.

To properly compare air pollution in different areas, we first need to come up with a way to calculate the relative quality of the air. Measurements for the different air pollutants may not be directly comparable. For example, a high value of sulfur dioxide might be represented by a much lower value than what is an acceptable value of carbon monoxide. In order to reduce the impact from these differences, the air pollution values need to be normalized. This can be done by converting the data for each pollutant to z-scores. This creates data that is all on the same scale with a mean of zero and a standard deviation of one. The scaled data makes it easier to compare values from different pollutants against each other. It also allows us to understand extreme high and low values without needing to understand typical pollution ranges. Once the data is prepared in a way that makes it easier to compare, we still need to continue to process it to be able to create data that is easy to understand. To have a complete view of how polluted the air is, we need to look at all three pollutants together. Since the data has been converted to z-scores, the data for each pollutant can be combined with minimal additional adjustments.

There are natural variations in air pollution over the course of the year due to climate and other variations. In order to reduce the impact of these variations, seasonal combined air pollution data is created. This allows us to look at data for each season separately to compare differences between the seasons. Limiting the data to each season also reduces the amount of data that needs to be processed. For this project, the seasons are defined as winter being from December 1, 2020, to February 28, 2021, spring being from March 1, 2021, to May 31, 2021, summer being from June 1 to August 31, 2021, and fall being from September 1, 2021, to November 30, 2021. This gives us a full year of data split into seasons that are easier to manage. We will be able to see if there is a season that has a significant difference in air pollution compared to the others. It also lets us look to see if there are regions of the United States that have higher pollution during different parts of the year.

In order to have a more accurate view of air pollution levels in the United States, we also should look at the data on a state-by-state level. This allows us to see regional trends as well as small variations due to different climate laws and regulations. TIGER data for US Census States from 2018 is used to create data on the state level. TIGER data is used because it is easily available in Google Earth Engine, and it provides an accurate outline of state boundaries. Even though it is from a different year than our climate data, there is not a large enough change in state boundaries from year to year for this to be an issue. There are a few states and census areas that are removed to make data processing easier for this project. Alaska, Hawaii, and territories such as Puerto Rico are removed to focus on the continental United States. Once seasonal air pollution data is processed, it is summarized for each state. This creates data that gives us the mean amount of air pollution in each state in the specific season. The data can then be summarized and turned into maps to show the results of the analysis.

Google Earth Engine is used for most of the analysis. This is because it is easy to import and process data in the cloud. All of the data used in this project is built into Google Earth Engine, making it easy to adapt code to work with the data as well as find documentation for the data. The code for this project is contained in appendix A. Simplified, the code imports air pollution data then processes it and exports it as data for each state and season. A large amount of the code is repeated through loops as well as similar code modified for each pollutant. The code starts by importing state outlines and creating arrays that hold the start and end dates for the season. Then for each season, looped code is used to import, normalize, and filter air pollution data. This allows the same code to be used with only the season dates being changed through iteration. After the air pollution data is created, data for each season is combined and averaged. Once this data is completely processed, it is exported to a table. This is done to reduce the amount of processing time as it is not efficient to create map layers with data of this scale. The data held in the table can then be rejoined with the same state outline data in ArcGIS Pro to make maps showing the results. This allows for more flexibility in what maps can be created to display the data. It also makes it easier to create cartographically detailed maps for each season that can be viewed simultaneously.

Results

The analysis found that there is not a significant difference in pollution when looked at either by state or by season. While there are differences in air pollution levels between states, they do not create coherent regional trends. There are also differences between individual state data from multiple seasons. There is not a noticeable trend to any of these differences. It was expected that there would be regional trends in air pollution. States that are close to each other have similar climates and environments. They also are likely to use similar energy sources and have many of the same anthropogenic sources of air pollution. A difference between the seasons would also be expected. Changes in the burning of fuel for energy related to the change in seasons would change the amount of air pollution. It is also possible that changes in climate over the course of the year would lead to differing levels of air pollution for each season.

Additionally, it was expected that a singular season or region would be identified as having the highest or lowest amount of air pollution. While each state and season has its own reasons to be expected to have a change in air pollution, evidence of these reasons do not show up in the data. Maps displaying the data can be found in appendix B. They allow a more compete view of the data beyond what can be summarized succinctly. Maps have been created for all four seasons as well as for combined year round data to provide the most complete visual representation of the data created by the analysis. They were visualized using natural breaks to help reduce the differences between the seasonal data.

It is found that over the course of the entire year, many of the states in the Northeast have lower levels of pollution than the rest of the United States. This contrasts with what was found in both Peng et al.'s 2005 and Bell et al.'s 2007. Many states in the South and Midwest have average or high levels of pollution. Two areas of interest are California and Montana. Both of these states have much lower levels of air pollution compared to surrounding states. This could be to variables such as the large amount of open land in Montana or strict environmental regulations in California. The seasonal views of the data show a much different picture. In the summer, more states have average levels of pollution and there are fewer strong outliers. Many states that had extremely high or low levels or air pollution now have slightly less extreme levels of pollution. The same is true for winter, but with more states that still have high levels of pollution. Each season also has different states that have higher levels of pollution rather than there being one state that is consistently the most polluted.

The data for the Spring and Fall show interesting regional variations in air pollution. The Fall result map has states with extremely high and low levels of air pollution near each other. There is a distinct lack of a regional trend or a singular region that has higher or lower levels of pollution. The Spring seasonal map shows small groups of close by states that have similar pollution levels. For example, Alabama, Mississippi, and Georgia all have lower than expected amounts of air pollution compared to nearby states. This comes as a surprise because these states all have much higher levels of air pollination in other seasons. There is also a closer regional trend between states that have high of average levels of air pollution.

However these slight trends to the data do not lead to any helpful conclusions about the connection between air pollution and potential influences such as region, land area, or population. The seasonal maps show the same lack of strongly supported trends as the year round data. For each season, there are variations between states with extremely low and extremely high pollution. However, which states are on the far ends of the data, varies from season to season without one state being an overall outlier. States that have very high levels of pollution in one season may have very low levels of pollution in others with a very small region trends. The results of the analysis help find flaws in how it was performed as well as ideas for how to improve the analysis or do additional research.

Discussion

This analysis did not return the expected results based on prior research on the topic. However, that does not mean that this research did not lead to helpful information about the connection between air pollution, geography, and seasonal changes. Understanding what was found in this study helps guide what could be done to further research this topic. More research could be done using different pollution data or information on a different group of air pollutants. Data representing the particulate matter, either PM2.5 or PM10 could provide a more accurate view of air pollution, especially if health impacts are a key part of the research. It also may be worth looking at each air pollutant separately. There may be some trends that get averaged out when data for all three pollutants are combined. For example, there may be a region and season that has significantly higher amounts of one pollutant and significantly lower amounts of another pollutant. These would even out and cause the combined data to appear as if the state has a median level of pollution.

The method of normalizing the data could also be improved. The data needed to be normalized in order to properly combine the different air pollutants to minimize the differences between how pollutants are measured. However, later normalization, especially the use of z-scores individually calculated for each season may lead to a flaw in the data. This method helps to compare data from each state to other states during the same season, but it makes season to season comparisons less accurate. If data was normalized and calculated differently, changes between seasons could be more apparent. The current way of analyzing the data creates individual z-scores for each season. This means that the individual maps for each season provide an accurate compassion of pollution levels for each season, but cannot be accurately used to compare between seasons. In order to have a more complete and accurate view into the geographical and seasonal impacts on air pollution, more analysis should be conducted.

Conclusion

Overall, this analysis looked at the connection between air pollution and seasonal and geographical variables. While it may not have had the expected results, ideas were generated for further research that could build off of this analysis. This research used Sentinel-5P data for sulfur dioxide, nitrogen dioxide, and carbon monoxide as well as TIGER Census data state outlines to acquire the data for the research. Google Earth Engine and ArcGIS Pro were used to perform the research due to ease of use. Even though there was not a strong trend to the results, they show that it is possible that there are other variables at play that determine the amount of air pollution. It is also possible that different ways of processing the data and running the analysis could return results that are closer to what was predicted.

Works Cited

Bell, M. L., Dominici, F., Ebisu, K., Zeger, S. L., & Samet, J. M. (2007). Spatial and Temporal Variation in PM2.5 Chemical Composition in the United States for Health Effects Studies. *Environmental Health Perspectives*, *115*(7), 989–995. https://doi.org/10.1289/ehp.9621

Dominici, F., McDermott, A., Zeger, S. L., & Samet, J. M. (2003). National maps of the effects of particulate matter on mortality: Exploring geographical variation. *Environmental Health Perspectives*, *111*(1), 39–44. https://doi.org/10.1289/ehp.5181

Manisalidis, I., Stavropoulou, E., Stavropoulos, A., & Bezirtzoglou, E. (2020). Environmental and Health Impacts of Air Pollution: A Review. *Frontiers in Public Health*, *8*. https://doi.org/10.3389/fpubh.2020.00014

Peng, R. D., Dominici, F., Pastor-Barriuso, R., Zeger, S. L., & Samet, J. M. (2005). Seasonal Analyses of Air Pollution and Mortality in 100 US Cities. *American Journal of Epidemiology*, *161*(6), 585–594. https://doi.org/10.1093/aje/kwi075

US EPA, O. (2016a, June 2). *Sulfur Dioxide Basics* [Overviews and Factsheets]. https://www.epa.gov/so2-pollution/sulfur-dioxide-basics

US EPA, O. (2016b, July 6). *Basic Information about NO2* [Overviews and Factsheets]. https://www.epa.gov/no2-pollution/basic-information-about-no2

US EPA, O. (2016c, July 13). *Basic Information about Carbon Monoxide (CO) Outdoor Air Pollution* [Overviews and Factsheets]. https://www.epa.gov/co-pollution/basic-information-about-carbon-monoxide-co-outdoor-air-pollution

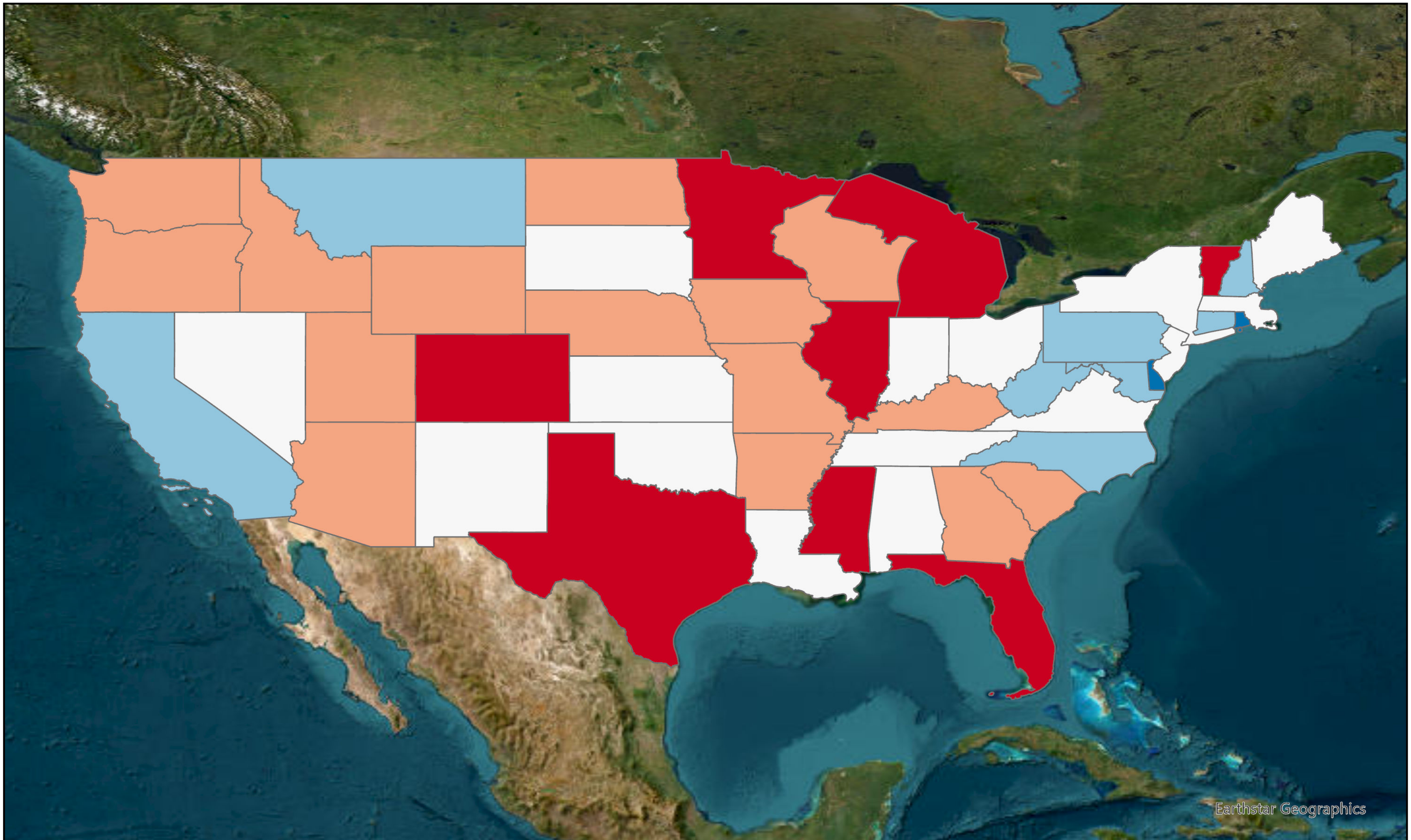
Appendix A – Google Earth Engine Code

Link to code: https://code.earthengine.google.com/6a855207e30c69ce5008ce6dccd90722

```
//Imports the state outlines
var states = ee.FeatureCollection('TIGER/2018/States');
//Creates arrays holding the start and end dates used in the processing
var startDates = ['2020-12-01', '2021-03-01', '2021-06-01', '2021-09-01'];
var endDates = ['2021-02-28', '2021-05-31', '2021-08-31', '2021-11-30'];
//Creates empty array to hold the processed data
var seasonalPollution = [];

//Loop that runs once for each season
//Creates normalized and combined pollution data for each season
var count = 4;
for (var i = 0; i < count; i++) {

  // Imports sulfur data
  //Finds the mean and standard deviation
  //Calculates the z-score
  var sulfur = ee.ImageCollection('COPERNICUS/S5P/OFFL/L3_SO2')
    .filterDate(startDates[i], endDates[i])
    .map(function(image) {
      return image.select('SO2_column_number_density').rename('pollution');
    });
  var meanSulfur = sulfur.mean();
  var stdDevSulfur = sulfur.reduce(ee.Reducer.stdDev());
  var normalizedSulfur = sulfur.map(function(image) {
    return image.subtract(meanSulfur).divide(stdDevSulfur);
  });

  // Imports nitrogen data
  //Runs the same processing as on the sulfur data
  var nitrogen = ee.ImageCollection('COPERNICUS/S5P/OFFL/L3_NO2')
    .filterDate(startDates[i], endDates[i])
    .map(function(image) {
      return image.select('tropospheric_NO2_column_number_density').rename('pollution');
    });
  var meanNitrogen = nitrogen.mean();
  var stdDevNitrogen = nitrogen.reduce(ee.Reducer.stdDev());
  var normalizedNitrogen = nitrogen.map(function(image) {
    return image.subtract(meanNitrogen).divide(stdDevNitrogen);
  });
```

```
  // Imports carbon data
  //Runs the same processing as on the sulfur and nitrogen data
  var carbon = ee.ImageCollection('COPERNICUS/S5P/OFFL/L3_CO')
    .filterDate(startDates[i], endDates[i])
    .map(function(image) {
      return image.select('CO_column_number_density').rename('pollution');
    });
  var meanCarbon = carbon.mean();
  var stdDevCarbon = carbon.reduce(ee.Reducer.stdDev());
  var normalizedCarbon = carbon.map(function(image) {
    return image.subtract(meanCarbon).divide(stdDevCarbon);
  });

  // Combines the normalized pollution values
  var combinedPollution = normalizedSulfur.mean()
    .add(normalizedNitrogen.mean())
    .add(normalizedCarbon.mean());

  // Calculates the mean pollution for each state
  var statePollution = combinedPollution.reduceRegions({
    collection: states,
    reducer: ee.Reducer.mean(),
  });

  // Adds the state pollution data to the empty seasonal pollution array
  seasonalPollution.push(statePollution);
}

var seasons = ['Winter', 'Spring', 'Summer', 'Fall'];

// Exports the final data to a table
for (var i = 0; i < seasonalPollution.length; i++) {
  Export.table.toDrive({
    collection: seasonalPollution[i],
    description: seasons[i] + 'PollutionFINAL',
    folder: 'final project',
    fileNamePrefix: seasons[i] + 'PollutionFINAL',
    fileFormat: 'CSV'
  });
}
```

Appendix B – Result Maps



Z-Score Value

- -5.38 - -2.14
- -2.13 - -0.176
- -0.175 - 0.214
- 0.214 - 0.505
- 0.505 - 1.250

Year Round Pollution

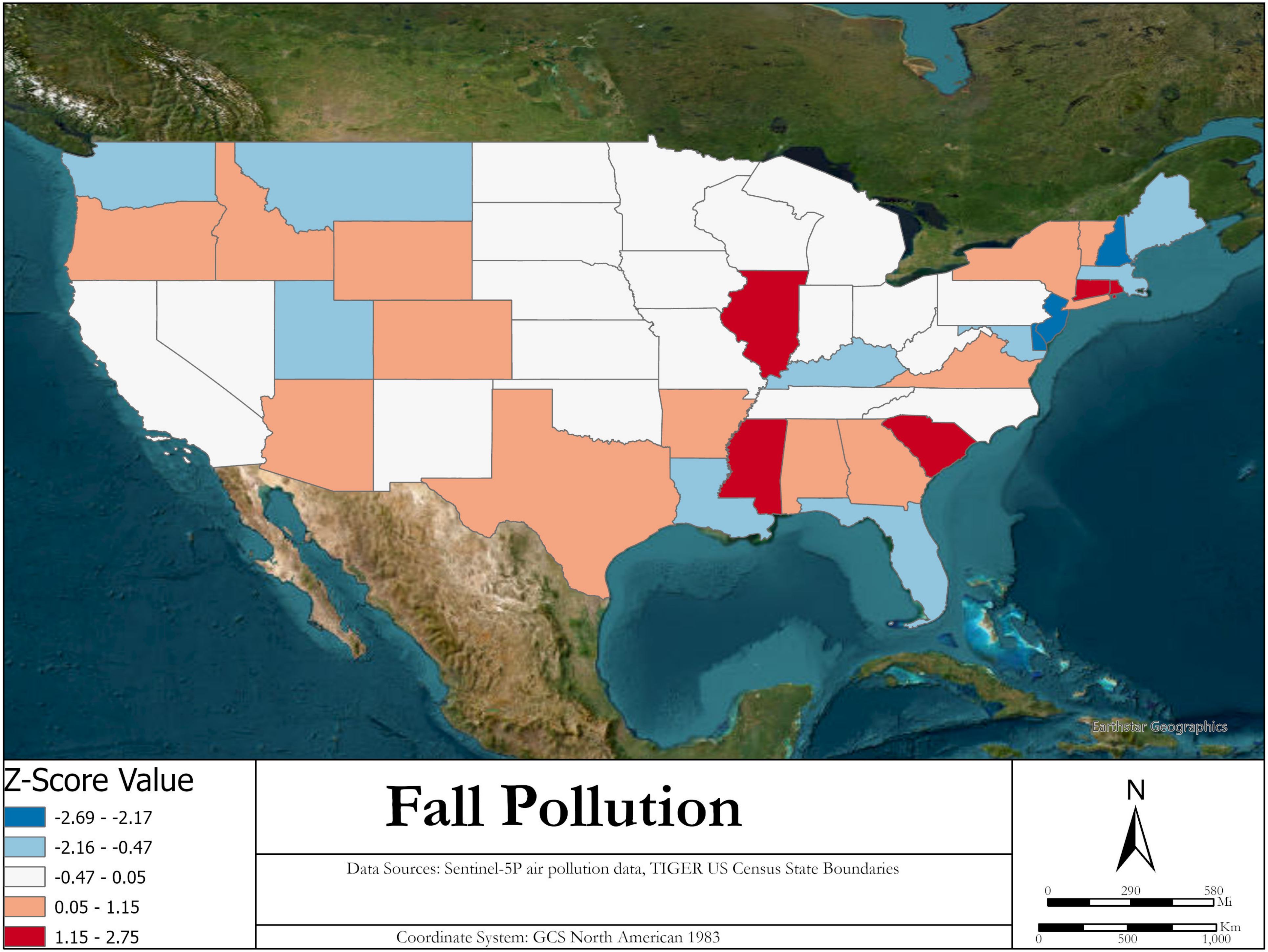
Data Sources: Sentinel-5P air pollution data, TIGER US Census State Boundaries

Coordinate System: GCS North American 1983

N

0 290 580 Mi

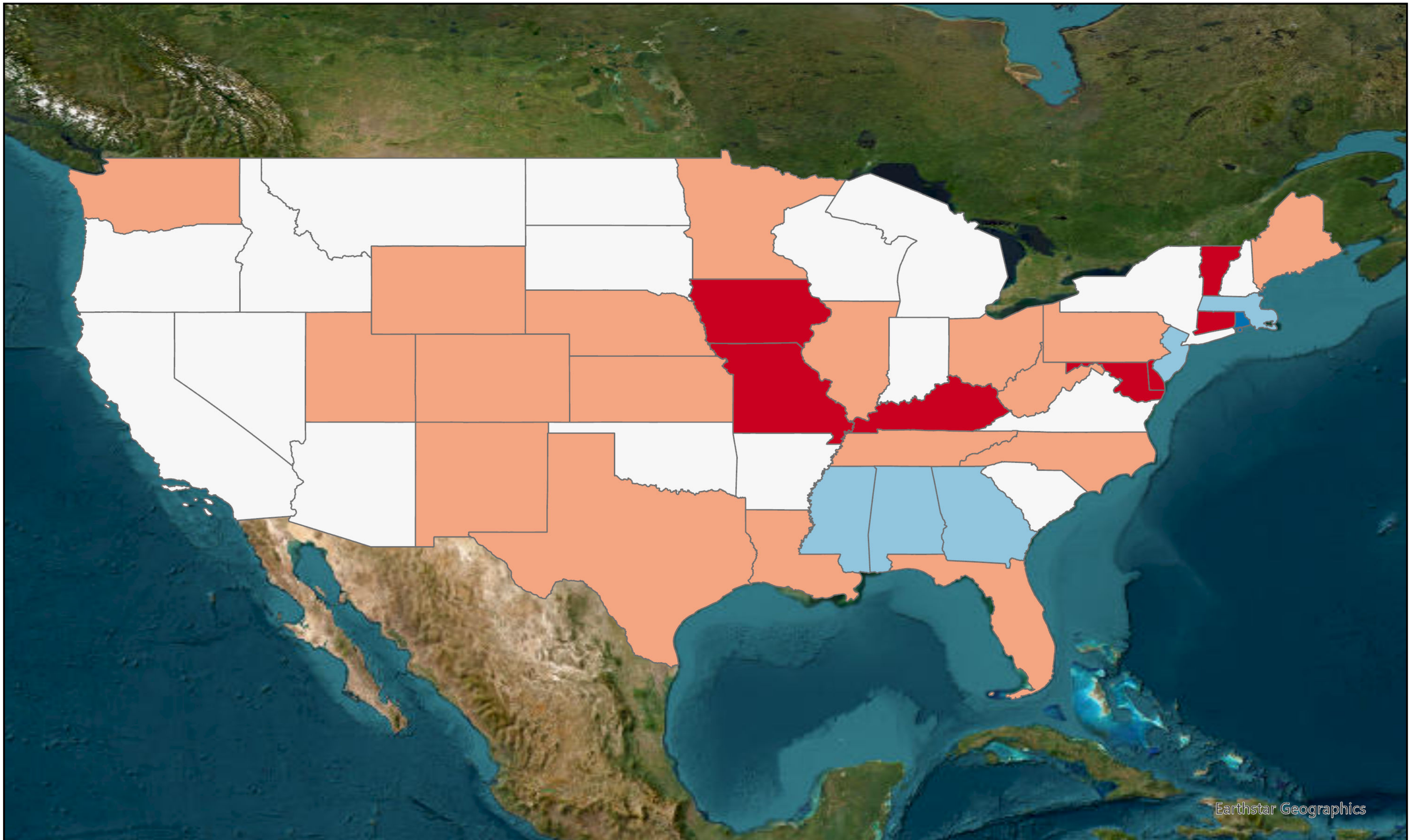
0 500 1,000 Km



Fall Pollution

Data Sources: Sentinel-5P air pollution data, TIGER US Census State Boundaries

Coordinate System: GCS North American 1983



Z-Score Value

- -5.59
- -5.58 - -0.65
- -0.65 - 0.12
- 0.12 - 0.62
- 0.62 - 1.43

Spring Pollution

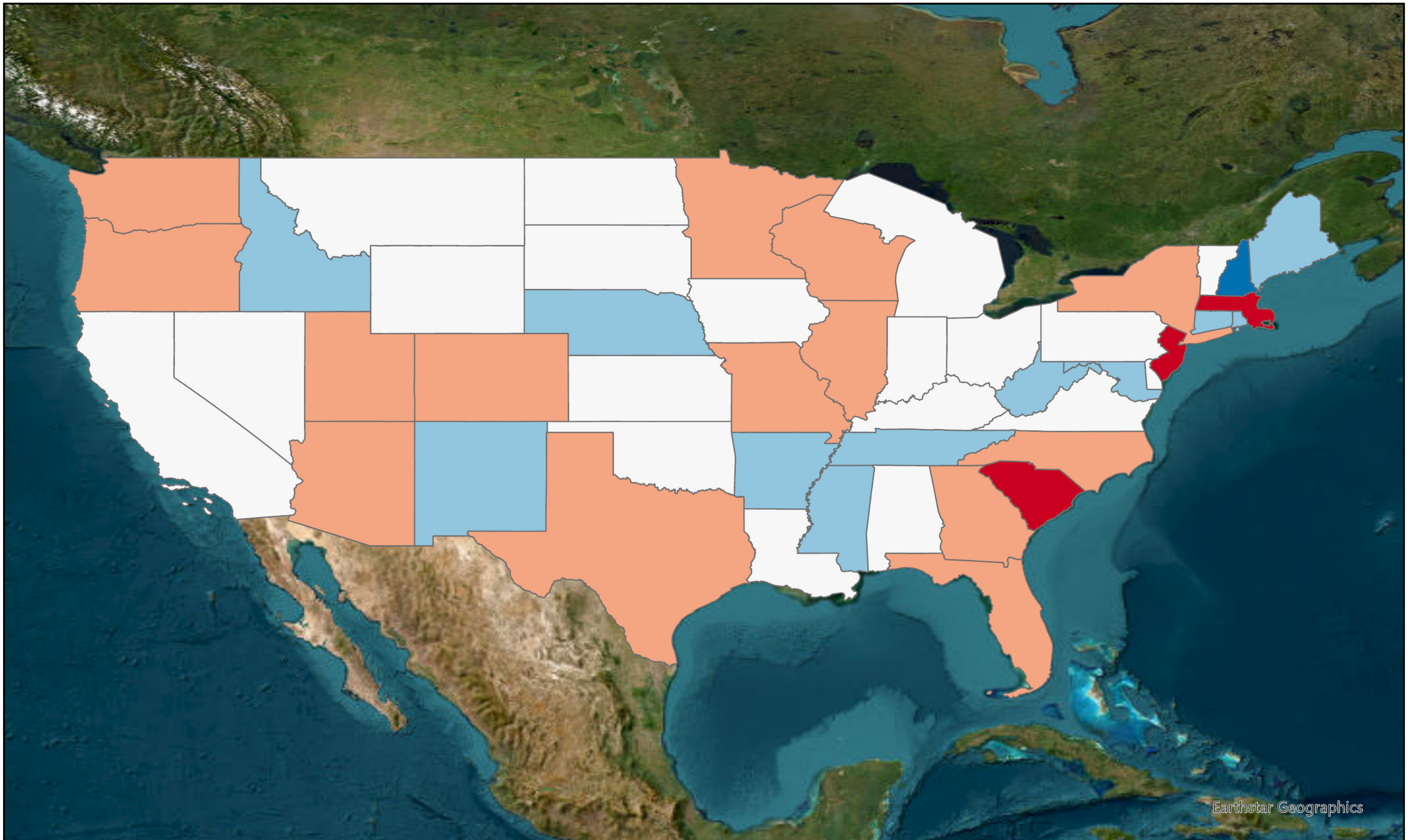
Data Sources: Sentinel-5P air pollution data, TIGER US Census State Boundaries

Coordinate System: GCS North American 1983

N

0 290 580 Mi

0 500 1,000 Km



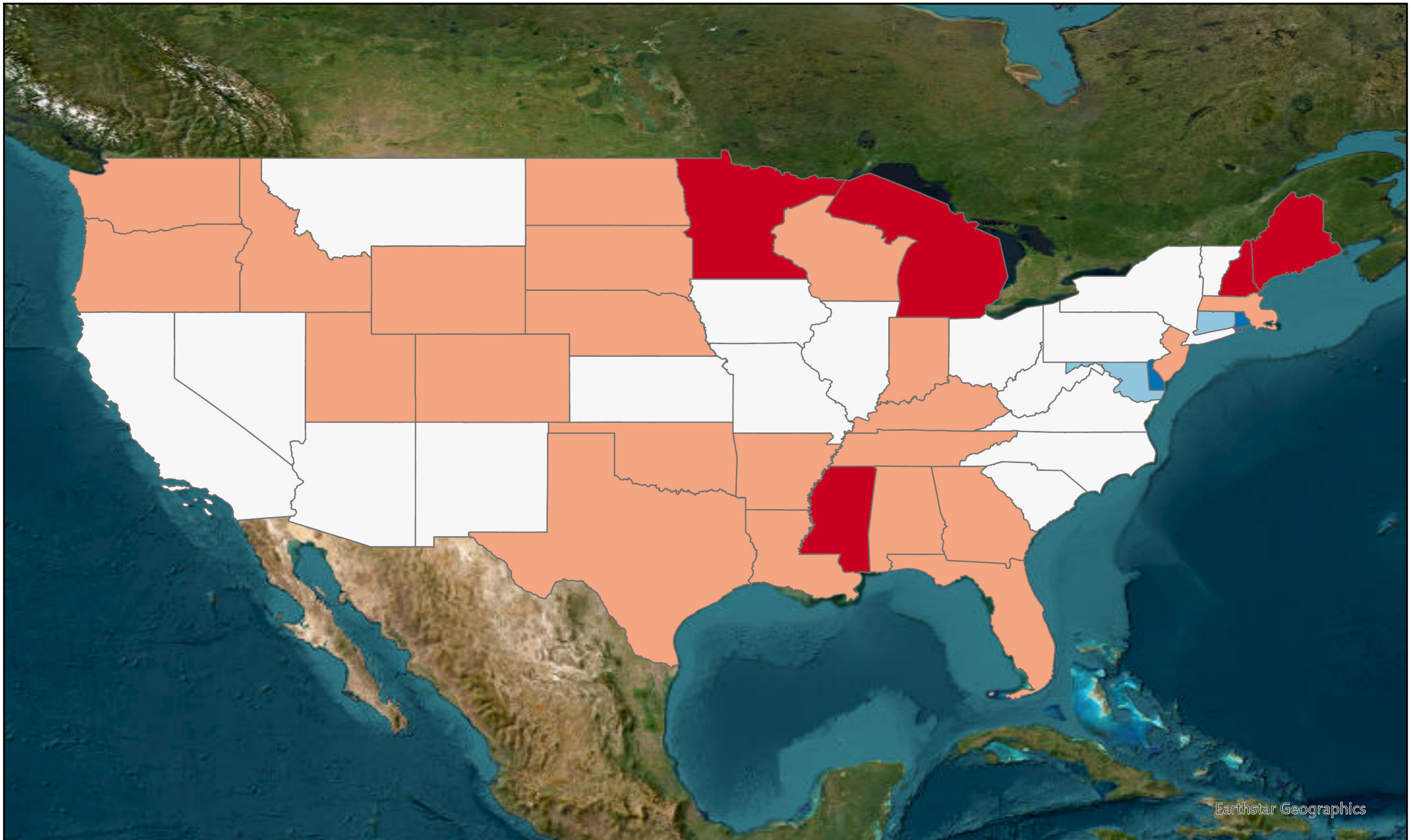
Z-Score Value

- -3.46 - -2.05
- -2.04 - -0.49
- -0.49 - 0.28
- 0.28 - 1.25
- 1.25 - 2.49

Summer Pollution

Data Sources: Sentinel-5P air pollution data, TIGER US Census State Boundaries

Coordinate System: GCS North American 1983



Z-Score Value

- -4.19 - -3.27
- -3.26 - -0.95
- -0.95 - 0.06
- 0.06 - 0.73
- 0.73 - 1.49

Winter Pollution

Data Sources: Sentinel-5P air pollution data, TIGER US Census State Boundaries

Coordinate System: GCS North American 1983

N

0 290 580 Mi

0 500 1,000 Km