Evaluation of air quality within different states in the US

This manuscript (permalink) was automatically generated from uiceds/project-team492@c870a0a on October 27, 2024.

Published: September 26, 2024

Authors

- **Siyoung Park** ✉
  · siyoung3
  Department of Civil and Environmental Engineering, University of Illinois

- **Tausif E Elahi** ✉
  · tausifeelahi
  -Department of Civil and Environmental Engineering, University of Illinois

- **Tabassum Nanzeeba** ✉
  · nanxee492
  Department of Civil and Environmental Engineering, University of Illinois

- **Rauf Momina** ✉
  · MominaRauf
  Department of Civil and Environmental Engineering, University of Illinois

✉ — Correspondence possible via GitHub Issues or email to Siyoung Park <siyoung3@illinois.edu>, Tausif E Elahi <tausife2@illinois.edu>, Tabassum Nanzeeba <nt23@illinois.edu>, Rauf Momina <momina2@illinois.edu>.

Project Description

Description of data set:

The dataset consists of

1. Air pollutants (O_3, CO, SO_2, NO_2, PM10) in different states (Illinois, California, Florida, North Dakota). of the US

- CSV file 1 Rows: 463,218 // Cols: 7

A data set shows air pollutant concentrations for 5 criteria i.e. O_3, CO, SO_2, NO_2, and PM10. The variable "pred_weight" shows concentrations of pollutants in μg/m^3. The title lat and lon represent the latitude and longitude of the specific place where the data were measured and noted. O_3 and CO are measured as parts per million (ppm) whereas NO_2 and SO_2 are measured as parts per billion (ppb).

Threshold for pollutants:

O_3, CO, SO_2, NO_2 and PM10 concentration thresholds are based on standards set by WHO (World Health Organization) and EPA (Environmental Protection Agency).

- O_3: 0.070 ppm exposure for 8 hours.
- CO: 9 ppm (8 hours) and 35 ppm (1 hour). > Not to be exceeded more than once per year
- SO_2: 75 ppb (1 hour, 3 years average) and 0.5 ppm (3 hours, year)
- NO_2: 100 ppb (1 hour, 3 years average) and 53 ppb (year, annual mean)
- PM10: 150 μg/m^3 (24 hours, 3 years average)

Dataset: https://www.caces.us/data [Accessed: 09/19/2024] In a CSV file

2. Temperature variation of the above states over years (using NOAA, National Centers for Environmental Information dataset from 1991 to 2020).

- CSV file 1 (Bismarck 2.4 NNW, ND, US) Rows: 13 // Cols: 313
- CSV file 2 (Springfield Capital AP, IL, US) Rows: 13 // Cols: 413
- CSV file 3 (Tallahassee AP, FL, US) Rows: 13 // Cols: 413
- CSV file 4 (Sacramento, CA, US) Rows: 13 // Cols: 385

A data set of 4 cities i.e. capital cities of four selected states of the US. The information of capital cities of selected states is given below

Regions	**State**	**City**
1	North Dakota (ND)	Bismarck
2	Mid Illinois (IL)	Springfield
3	South Florida (FL)	Tallahassee
4	California (CA)	Sacramento

The dataset consists of temperature variations in the abovementioned cities from 1991 to 2020. CSV file includes daily max, daily min, mean, standard deviation, cooling degree days, heating degree days, and mean number of days. The data are given in every month from 1991 to 2020.

Dataset: Palecki, Michael; Durre, Imke; Applequist, Scott; Arguez, Anothony; Lawrimore, Jay. 2021: U.S. Climate Normals 2020: U.S. Hourly Climate Normals (1991-2020). [indicate subset used]. NOAA National Centers for Environmental Information. https/doi.org/. Accessed [09/19/2024] In a CSV file

Proposal:

1. In this research, our team will track the air pollutants in different states of the US. O_3, CO, SO_2, NO_2, and PM10 are the subjects of the investigation. The five states of the US, North Dakota (north), Illinois (mid), Florida (south), and California (west) are the regions for measuring air pollutants.

2. The historical trend (temperature change) of the air pollutants will also be investigated along with the US State data. From 1991 to 2020, the variation of climate (such as mean temperature by months) and air pollutants will be compared to evaluate their correlation.

3. The ultimate goal of this research will be to alert each investigated States of increasing air pollutant and to come up with ideas for mitigating them. Also, the monthly temperature and air pollutants will be compared to notice when the air pollutants are maximized with an understanding of their distribution.

4. Exploratory Data Analysis

1.1 Description and Characterization of Dataset

The dataset we are going to use is obtained from Bechle et al. [1] which was used for estimating air pollution in terms of NO2 from 2000 to 2010. The dataset contains spatial and temporal concentration of NO2 in ppb at different locations of the different states in US. It also contains Geographic Information System (GIS) data on land-use features such as impervious surfaces, population density, length of different types of roads-residential, major and total etc. These are commonly used as proxies for different pollution sources [2]. Based on the dataset, NO2 concentration varies significantly from state to state depending on different land-use pattern and the value range between 0.31~34.21 ppb for different states. The distribution of the NO2 pollutants across the US based on location are shown in Figure 1. Some of the explanations of the dataset are provided below:

Impervious_100: This represents the percentage of impervious surfaces such as roads, buildings etc. within a 100-meter buffer around the measuring station. Major_1000: It refers to the length of the major roads within a 1-kilometer radius around the measuring location. Resident_500: This indicates length of the roads within 500-meter radius of the monitoring station. Total_100: It represents the length of all the roads including major, minor and residential within a 100-meter buffer zone around the measuring station. Population_100: It denotes the population density within the 100-meter buffer area around the measuring station.

1.2 Preliminary Analysis and Plots

From the given dataset, we did some preliminary analysis to visualize the dataset and the summary of the observations are described briefly: First, we tried to find out if there is any direct relationship between any of the land use characteristics and NO2 concentrations measured at the monitor station. For this preliminary analysis, we considered the effect of this land use pattern within 100m, 5000m and 10000m radius of the station. The reason for selecting these three radii was to cover short, medium and long-distance land use behavior around the station. Figure 1-6 presents the effect of different land-use characteristic on the NO2 concentration.

For impervious surfaces, for all three cases, we can clearly see there is a trend that with the increase of impervious surfaces around the station, the concentration of NO2 increases gradually (Figure 2). As the impervious surface increases, it indicates there is increase in roads, sidewalks, parking lots, buildings, traffic and also there is decrease in vegetation areas and soil surface. Therefore, all these impervious surfaces are kind of indicator of high volume of vehicles, high population density which contributes to high NO2 emission and also the absence of natural filtration effect with the absence of vegetation is another major source of NO2 emission.

In case of Major roads, we clearly see with the increasing length of major roads, there is clear increase in the concentration of NO2 (Figure 3). Moreover, visually, it looks like there is a steep increase in the concentration of NO2 initially with the increase of major roads, but the rate of increase slows down as the length of major roads increases further. The reason of such increase is understandable since the production of NO2 is directly influenced by the volume of traffic and high traffic areas will release more NO2 as the more diesel vehicles will be on the road contributing to high NO2 emissions. Similar trend is observed for the relationship between NO2 concentration and residential roads and total roads (Figure 4-5).

In case of population, it looked like an exponential curve which might describe the pattern very well where initially with the increase in population there is a drastic increase in NO2 concentration which saturates at a certain point (Figure 6).

Figure 7 shows the relationship between NO2 concentration and the distance from the coast. In general, as the distance from the coast will be less, there should be lower concentration of NO2 due to the ventilation from the winds. However, in the figure we can clearly see higher concentration of NO2 in some of the places which are closest to coast. It indicates that although coastal distance have effect on NO2 but it should be analyzed in combination with other land use pattern because even if the place is closer to coast but if there is high population density and roads, it will have higher NO2. Overall, all of these land-use characteristics have their own effect on the NO2 concentration and in some cases, there is strong relationship with NO2.

Sample Image

Figure 1: Caption describing the image.

References

- https://www.epa.gov/criteria-air-pollutants/naaqs-table
- https://www.eea.europa.eu/publications/status-of-air-quality-in-Europe-2022/europes-air-quality-status-2022/world-health-organization-who-air