Investigation of NO_X, $PM_{2.5}$, and O_3 (ppm) trends in Ontario from 2003 to 2022

Mohamad Damaj - 10109, Moham An, Jo Zhu

2025-03-18

1. Description of Datasets

The datasets analyzed in this report are three, one for each of NO_X, $PM_{2.5}$, and O_3, all containing the same variables. These datasets contain records of the pollutant levels (in parts per million, ppm) from 2003 (considered the earliest of all three) to 2022. Each dataset contains detailed information of pollutant levels in the form of the following variables:

```
##  [1] "Year"            "Station Number"  "City"            "Location"
##  [5] "Type"            "Valid Hour"      "10th Percentile" "30th Percentile"
##  [9] "50th Percentile" "70th Percentile" "90th Percentile" "99 Percentile"
## [13] "Mean"            "1-Hour Maximum"  "24-Hour  Maximum"
```

1. Year: The calendar year when the pollutant data was recorded.
2. Station Number: A unique numerical identifier for the monitoring station.
3. City: The name of the city where the pollutant was measured.
4. Location: A more specific description of the monitoring location (street, area).
5. Type: The type of monitoring station or measurement protocol used.
6. Valid Hour: The number of valid hourly measurements recorded for that pollutant in the given year.
7. Percentiles: Percentile statistics showing the distribution of pollutant levels (in ppm) throughout the year.
8. Mean: The annual average concentration of the pollutant (ppm).
9. 1-Hour Maximum: The highest recorded concentration within a one-hour period during the year.
10. 24-Hour Maximum: The highest average concentration recorded over a 24-hour period in that year.

2. Background of the Data

The data used in this analysis was collected across various cities in Ontario, throughout the period from 2003 to 2022. These measurements were compiled and made available by the Ontario Ministry of the Environment, Conservation and Parks.

The information is publically avaliable and is helpful in assiting researchers, policymakers, and public health officials to investigate air quality trends, evaluate the impact of environmental regulations, and inform policy decisions. It provides a long-term view of pollutant levels in various cities in Ontario, allowing for improving environmental and public health outcomes in Ontario.

3. Overall Research Question

The aim of this paper is to investigate the spatial and temporal trends of air pollutants (NO_X, O_3, and $PM_{2.5}$) across various Ontario cities from 2003 to 2022.

1. How have annual mean concentrations (in ppm) for each pollutant changed from 2003 to 2022, and which years show the most significant shifts?

2. Which cities consistently rank among the highest (or lowest) in terms of average pollutant levels, and do these rankings shift over time?

3. How do the four pollutants correlate with each other across different cities and years, and what might this indicate about broader air quality patterns?

4. How do the four pollutants project into future years based on current trends?

4. Tables

4.1 The Top 10 Cities With the Highest Pollutant Concentration (in ppm)

NO_X Table

City	Mean
Toronto West	31.32
Toronto East	21.82
Toronto Downtown	20.41
Toronto North	19.67
Hamilton Downtown	19.36
Windsor Downtown	18.00
Burlington	17.74
Hamilton West	17.63
Windsor West	16.89
Brampton	15.95

O_3 Table

City	Mean
Port Stanley	32.84
Tiverton	31.93
Grand Bend	31.28
Parry Sound	30.56
Chatham	29.88
Belleville	29.63
Kingston	29.61
Newmarket	29.13
Hamilton Mountain	28.61
Peterborough	28.55

$PM_{2.5}$ Table

City	Mean
Sarnia	9.57
Windsor West	9.03
Hamilton Downtown	8.91
Etobicoke West	8.44
Windsor Downtown	8.36
Hamilton West	8.06
Hamilton Mountain	7.91
Toronto West	7.86
Toronto North	7.59
Kitchener	7.52

Table 1: Top 10 Cities with Highest Concentration of Each Pollutant (in ppm)

The table above is three separate tables each one for a respective pollutant, they are sorted in descending order and only the top 10 are being shown based on the overall Mean of the City, over all years from 2003 to 2022; as such, the cities with the high concentration (in ppm) will be listed first. Therefore, we can draw a few key observations about pollutant concentrations across these Ontario cities:

- The highest mean NO_x readings (31.32 ppm) appear at Toronto West, followed closely by other Toronto stations (Toronto East, Toronto Downtown) and industrial/urban areas like Hamilton and Windsor.

- Port Stanley shows the highest O_3 levels (32.84 ppm), with other high concentrations at Tiverton, Grand Bend, and Parry Sound—generally smaller or semi-rural communities.

- Sarnia tops the $PM_{2.5}$ list (9.57 ppm), followed by Windsor (West and Downtown) and Hamilton stations, reflecting the influence of industrial facilities and cross-border pollution.

4.2 The Top 10 Regions and Years with Highest Concentration of Each Pollutant (in ppm)

In order to view the pollutant concentration changes by region, we grouped the cities by region based on the map of Ontario from the Ministry of Natural Resources and Forestry, and calculated the mean of the cities in each region grouped by year.

NO$_X$ Table

Year	Region	Mean
2003	Central Ontario	31.45
2005	Central Ontario	28.17
2003	Western Ontario	26.98
2004	Central Ontario	26.94
2006	Central Ontario	23.80
2007	Central Ontario	21.68
2004	Western Ontario	21.55
2008	Central Ontario	20.07
2005	Western Ontario	20.00
2004	Northern Ontario	19.04

O$_3$ Table

Year	Region	Mean
2010	Western Ontario	29.54
2022	Western Ontario	29.49
2007	Eastern Ontario	29.46
2010	Eastern Ontario	29.20
2021	Western Ontario	29.10
2018	Eastern Ontario	28.97
2012	Western Ontario	28.82
2016	Western Ontario	28.79
2013	Western Ontario	28.57
2008	Eastern Ontario	28.54

PM$_{2.5}$ Table

| Year	Region	Mean
2005	Western Ontario	9.48
2014	Western Ontario	9.27
2003	Western Ontario	8.93
2015	Western Ontario	8.79
2013	Western Ontario	8.72
2005	Central Ontario	8.56
2004	Western Ontario	8.44
2014	Central Ontario	8.33
2007	Western Ontario	8.30
2003	Central Ontario	8.26

Table 2: Top 10 Regions and Years with Highest Concentration of Each Pollutant (in ppm)

```
## [1] "| Year|  Mean NOX|" "|----:|---------:|" "| 2003| 25.078500|"
## [4] "| 2004| 23.423000|" "| 2005| 22.427143|" "| 2006| 17.871482|"
```

Same thing w NAs, we prob should na.rm when we use mean O-O

Year	Mean NOX	Mean PM25	Mean O3
2003	25.07850	8.042500	26.06824
2004	23.42300	7.296216	24.31750
2005	22.42714	8.357568	27.16641
2006	17.87148	7.037949	25.83487
2007	15.79867	7.029500	27.83250
2008	15.03903	6.447500	27.40625
2009	13.90548	5.418750	26.05525
2010	12.34758	5.750500	28.41500
2011	12.35829	5.936250	27.09825
2012	11.23257	5.885641	27.84675
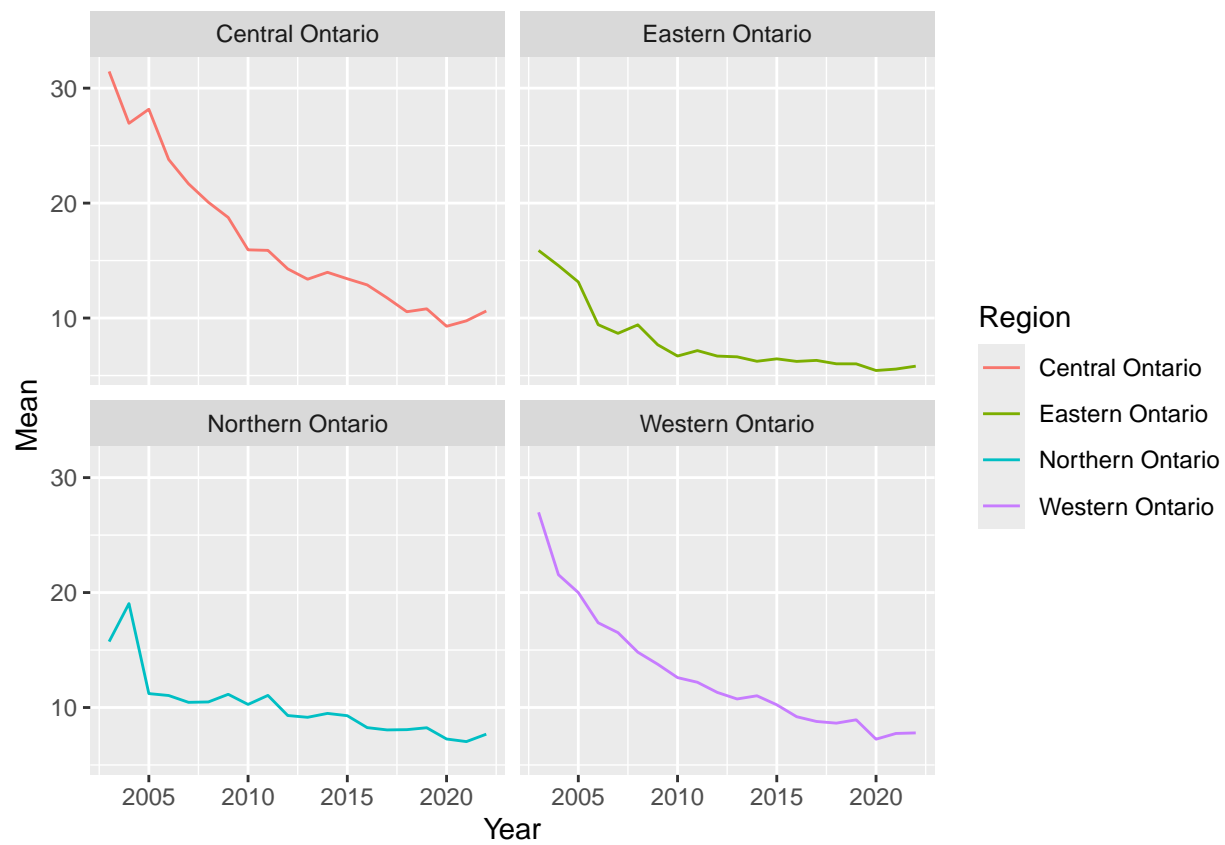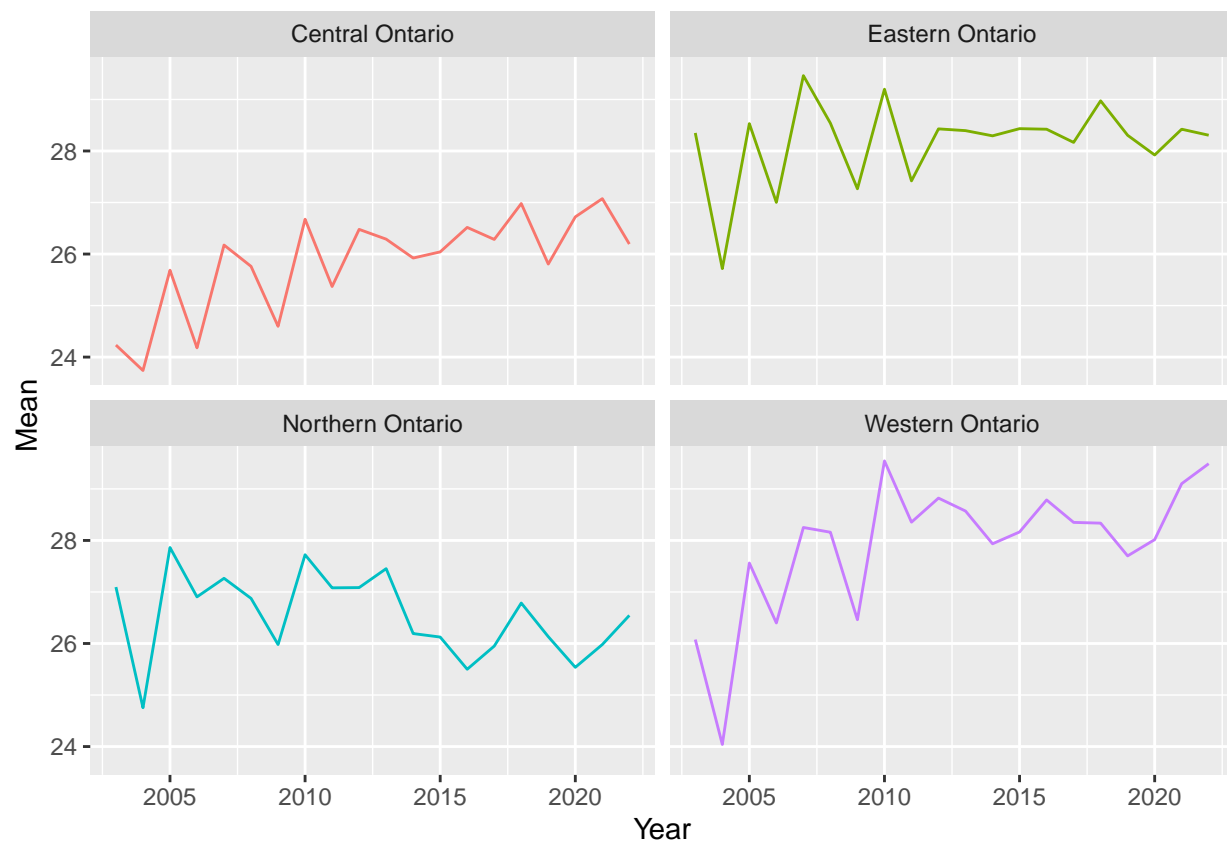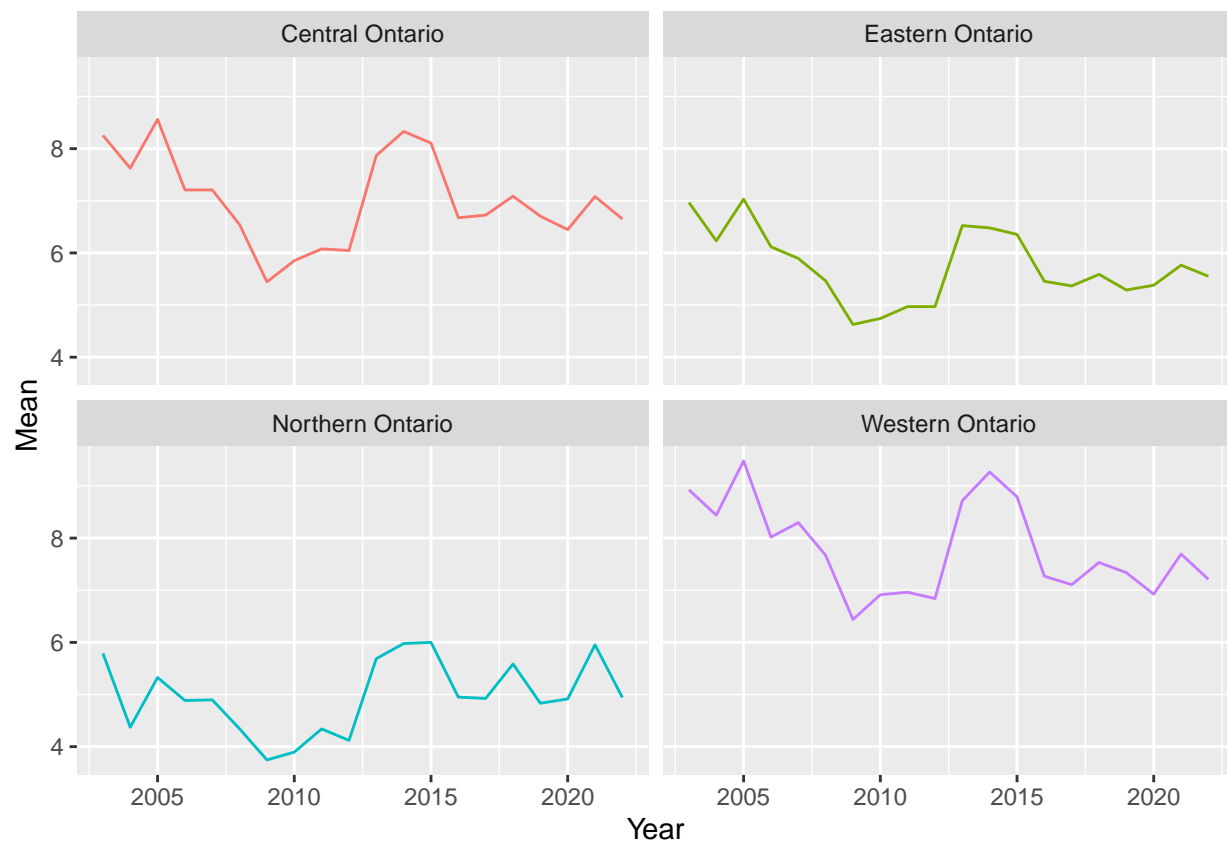
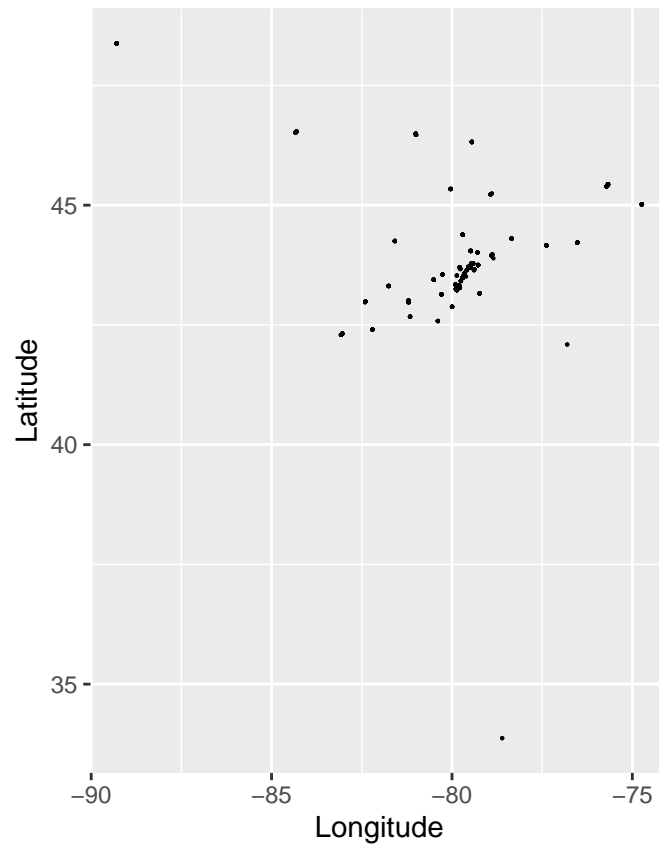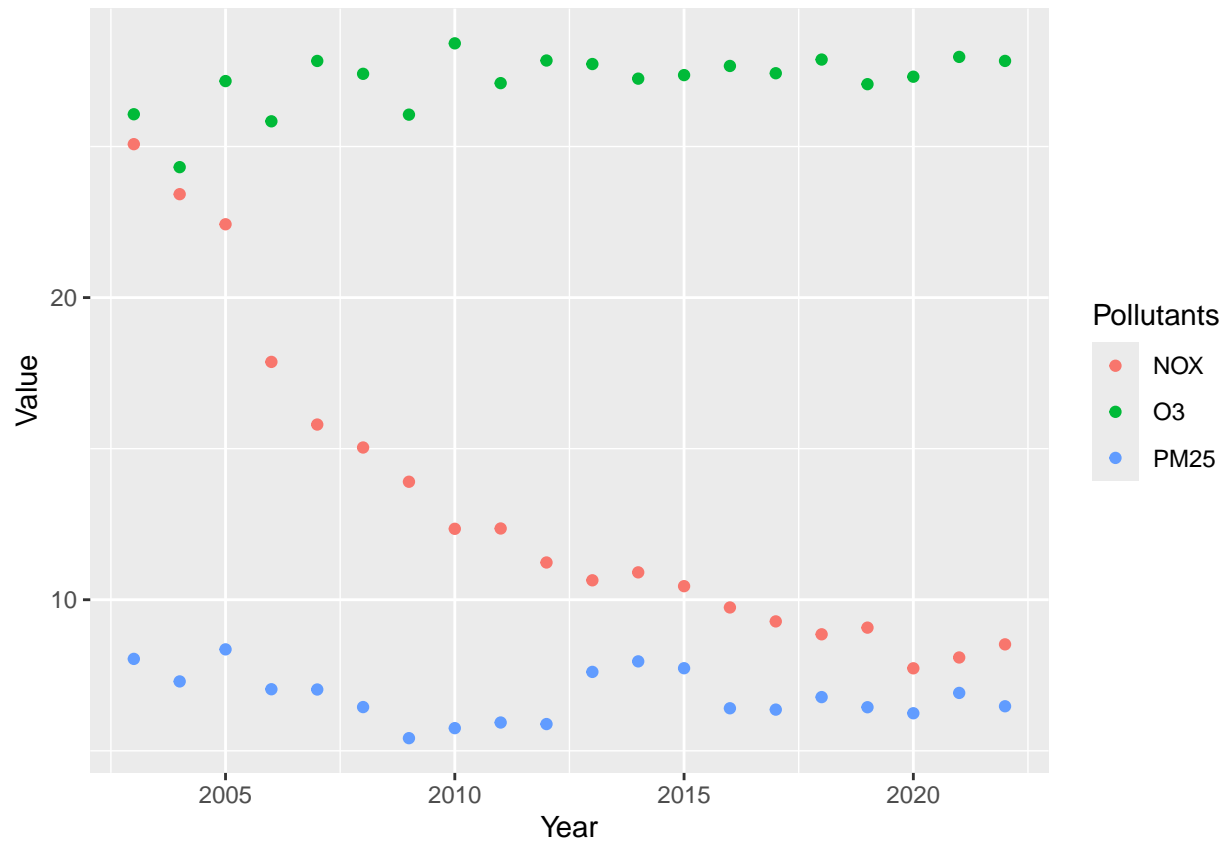## Map cities

```
## New names:
## * `` -> `...1`
```

Paired T-test

compare the value of two years (2003, 2022), x_2003, x_2022 H_0: u_1 - u_2 = 0 H_a: u_1 - u_2 != 0

```
## 
##  One Sample t-test
## 
## data:  values$diff
## t = 8.2745, df = 19, p-value = 1.012e-07
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##   11.31407 18.97593
## sample estimates:
## mean of x
##    15.145


## 
##  One Sample t-test
## 
## data:  values$diff
## t = -3.1488, df = 32, p-value = 0.00354
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##   -2.7273569 -0.5847644
## sample estimates:
```

```
## mean of x
## -1.656061


##
##   One Sample t-test
##
## data:  values$diff
## t = 7.4834, df = 27, p-value = 4.748e-08
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  1.050617 1.844383
## sample estimates:
## mean of x
##    1.4475
```