Impact of Coronavirus Related Lockdowns on Air Quality in England and Scotland

Dr Murilo Da Silva Baptista

This work is presented in partial fulfilment of course PX5008, Introduction to Data Science

19th of November 2021

Binti Mat Zahid, Farisa Dania

Gambedotti, Federico

Gaze, Shaun

Gould, Nicole Kirsten

Walls, Victoria

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 Abstract

In December 2019, the first cases of the Coronavirus (COVID-19) Pandemic occurred, since then there have been several studies addressing the environmental impact of lockdowns worldwide. Some studies have indicated that air quality has improved because of the restrictions placed on travel. It was found that COVID-19 caused Particulate Matter 2.5 (PM2.5) and Daily Air Quality Index (DAQI) to decrease by about 7 µg/m3 and 5-points in China, respectively. Similarly, in the Kingdom of Saudi Arabia, a 91.12% lower concentration of Particulate Matter 10 (PM10) in the air across nine cities was noted. These studies concluded that their results were due to the reduction of industrial and travel pollution because of the lockdown. In the United Kingdom, national lockdown was enforced on March 26 2020, as such this study has anticipated a similar trend in air quality from this point. It was decided that this would be assessed by comparing air quality data – in particular, PM2.5 and PM10 for 2019 and 2020 across England and Scotland, the results found that air quality levels for these variables did show considerable improvements for individual sites because of lockdowns. Based on our analyses, missing data lead to an overall statistically insignificant result.

2 Introduction

Air pollution is currently the focal point of the COP26 conference, it had been found by the UK Natural Environment Research Council that this is the link between cause and effect of climate change (European Commission, 2015). It was stated by the UN in 1972 that breathing clean air is a human right (United Nations, 1972). WHO clearly states that good air quality requires on average PM < 10 μg/m^3 but 74% of all European sites record reading above this threshold, causing 477,000 premature deaths (EEA, 2020 Report). For these reasons, institutions should work to tackle pollutants through innovative technologies (Redondo-Bermúdez *et al.,* 2021) and legislations (Calderon and Keirstead, 2012, respectively).

PM variables are a significant factor in air quality due to their effects on respiratory health. PM is a collection of microscopic factors of varying shape and composition. Due to their size, PM can be inhaled and could potentially pose harm to the human body (California Air Resources Board, 2021) with PM2.5 being the more harmful of the two (United States Environmental Protection Agency, 2021). As a result, we aimed to test the hypothesis whether lockdowns reduce the levels of PM in the Scotland and England.

3 Data Model, Source and Explanation

The key source of the data used in this analysis comes from the DAQI and UK Air Information Resource Page, both of which can be found through the Department for Environment, Food & Rural Affairs (DEFRA, 2021b; DEFRA, 2021a, respectively) website. Since this data comes from multiple ministerial and non-ministerial departments of the UK government as well as a significant number[2] of public bodies, any occurrence of individual biases in the formation of this resource have been counteracted and the resulting data will thus be credible.

3.1 Daily Air Quality Index (DAQI)

To obtain the data for the DAQI across England and Scotland for Jan 2019 – Dec 2020, the selection menus on the DEFRA (2021b) DAQI regional data page were utilised. There were two submenus along with a date selector for regions and agglomeration regions, where “All” was selected for both. DAQI data was split yearly to allow for straightforward comparison between the two time periods. This resulted in CSV files for each year, which were downloaded and stored in a directory “DataDir” ready to be imported into.

When opening Mathematica, the directory did not automatically set to the required one. This was corrected by using an initialization cell and by inputting the command **SetDirectory[FileNameJoin[{NotebookDirectory[],"DataDir"}]]** for all notebooks.

The format of the DAQI datasets consisted of a series of daily numerical values. A set of 32 keys (date column and 31 stations measuring air quality) can be found on the first row. The first 7 entries of the 2019 DAQI dataset can be seen in **TableForm[]** as an example of the format in Figure 1.

Figure 1: Table Showing DAQI 2019 Format Example

Date	Central Scotland	East Midlands	Eastern	Greater London	Highland	North East	North East Scotland	North Wales
01/01/2019	3	3	3	2	3	3	3	3
02/01/2019	3	3	3	2	3	3	3	3
03/01/2019	3	3	3	2	3	3	2	3
04/01/2019	3	5	3	3	3	2	2	3
05/01/2019	3	3	2	3	3	2	2	3
06/01/2019	3	2	2	2	3	2	2	3

3.2 Particulate Matter (PM10 and PM2.5)

Obtaining the PM data was more challenging than that of the DAQI. Early on it was found that individual datasets for the home nations had to be downloaded, cleaned, and joined instead of using one single dataset. Specifically, the cleaning included sorting monitoring sites by availability of data on PM and using these to form a fresh dataset.

To do so, an Intersection was done on the 2019 - 2020 datasets to obtain the necessary columns in the PM dataset that contained viable data. This was done using **Intersection[]**. Figure 2 displays the result of the initial query.

Figure 2: Mathematica Query Result after Running the Intersection Command

```
intersectingScotSites = Intersection[scotSites19[[All, 3]], scotSites19[[All, 4]]]
{, Aberdeen Market Street 2, Aberdeen Union Street Roadside, East Dunbartonshire Kirkintilloch, Falkirk West Bridge Street,
 Fife Cupar, North Ayrshire Irvine High St, South Lanarkshire Lanark, South Lanarkshire Rutherglen, West Dunbartonshire Clydebank, West Lothian Linlithgow High Street 2}
```

The dataset for both years consisted of a series of numerical values with two rows per site, where the first row displayed values for PM10, and the second row displayed values for PM2.5 on an hourly basis for the entire year.

In Figure 3, one monitoring site and the first 3 hours of readings for Scotland can be seen from 2019.

Figure 3: Example of Scotland Monitoring Site Readings for PM2.5 and PM10 in Mathematica

```
In[-]:= scot1[[1 ;; 6, 1 ;; 5]] // Grid
                                            Date                       Tue 1 Jan 2019 00:00:00 GMT   Tue 1 Jan 2019 00:00:00 GMT   Tue 1 Jan 2019 00:00:00 GMT
                                            Time                       Tue 30 Nov -2 01:00:00 GMT    Tue 30 Nov -2 02:00:00 GMT    Tue 30 Nov -2 03:00:00 GMT
Out[-]= Aberdeen Market Street 2  PM10 particulate matter (Hourly measured)   4.1                 3.6                 7.7
                                            Status                     V ugm-3 (FIDAS)      V ugm-3 (FIDAS)      V ugm-3 (FIDAS)
                                  PM2.5 particulate matter (Hourly measured)  8.2                 6.2                 5.2
                                            Status                     V ugm-3 (Ref.eq)     V ugm-3 (Ref.eq)     V ugm-3 (Ref.eq)

In[-]:= (*Request all values from the second row, as these should be a time value*)
```

Performing Transposes, the resulting dimensions for Scotland in 2019 and 2020 respectively were {42,8762} and {42,8786}. With England having {122,8762} and {122,8786} respectively.

4 Methodology

4.1 Data Pre-Processing: Daily Air Quality Index (DAQI)

<u>Step 1: Import the Dataset and Remove Redundant Rows</u>

The DAQI dataset was imported by using **Import["AQ_Regions.xlsx,{"CSV", "Data"}].** The first and the last five rows were displayed to check for missing values. This was done using **data[[1;;5]]** and **data[[-5;;-1]]** represented in **//TableForm**.

The first and last two rows were redundant and thus had to be removed, using **data[[2;;-3]]** taking the original dataset and selecting every row from the second to the third last.

Figure 4: Example of Imported DAQI Dataset with Removed Rows

```
data2019[[1 ;; 5]] // TableForm
Form=
AQ Daily index (https://uk-air.defra.gov.uk//air-pollution/daqi?view=more-info) data from UK-AIR website
Date                                                                        Central Scotland   East Midlands   Eastern
01/01/2019                                                                  3                  3               3
02/01/2019                                                                  3                  3               3
03/01/2019                                                                  3                  3               3

data2019[[-5 ;; -1]] // TableForm
Form=
29/12/2019  2 2 2 3 3 2 1 2 2 3 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
30/12/2019  3 2 2 2 3 2 2 2 2 3 2 2 2 3 2 2 2 2 3 2 3 2 2 2 2 2 2 2 2 2
31/12/2019  2 2 3 3 3 3 2 2 2 4 3 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3

End of file
```

<u>Step 2: Remove Missing Data</u>

Inspecting input forms using **InputForm /@ data[[2]],**a blank string was identified at the end of each row. To remove any blank elements, **Table[DeleteCases[data2019cleaningV1[[k]],""],{k,1, Length[data2019cleaningV1]}]** was used to loop across each list deleting any blank string, thus reducing the amount of redundant data.

<u>Step 3: Convert the Dates into European Date Format</u>

It was found that Mathematica may misinterpret the date format as MM/DD/YYYY. To match the required format the date was converted into the form DD/MM/YYYY matching the date format from DEFRA website. This was done using **(DateObject[{#, {"Day", "/", "Month", "/", "Year"}}, "Day"] & /@ dateswitch2019v1)**, the result of which was set equal to the first column in the original dataset.

Step 4: Create Association for the Regions

To define a series of associations, the 32 keys from the DAQI dataset are taken and threaded to each value in the list of lists to form a dataset using **Dataset@(AssociationThread[key2019, #] & /@ data2019cleaningV2[[2 ;;]])**

Step 5: Further Removal of Redundant Data

The final stage for processing was to remove any data relating to Northern Ireland and Wales, due to a vast amount of the data being missing. The data was removed by using **KeyDrop[]** which drops all data relating to this site. The resultant format is in Figure 5.

Figure 5: Visualisation of DAQI Dataset after Cleaning

Date	Central Scotland	East Midlands	Eastern	Greater London	Highland	North East	North East Scotland	North West & Merseyside	Scottish Borders
Wed 1 Jan 2020	2	3	3	4	3	2	1	3	2
Thu 2 Jan 2020	3	3	2	2	3	2	1	2	3
Fri 3 Jan 2020	3	3	3	3	3	3	2	3	3
Sat 4 Jan 2020	3	3	2	2	3	3	2	3	3
Sun 5 Jan 2020	3	3	3	2	3	3	2	3	3
Mon 6 Jan 2020	3	3	3	2	3	3	2	3	3
Tue 7 Jan 2020	3	3	2	2	3	3	1	3	3
Wed 8 Jan 2020	3	3	2	3	3	3	2	3	3
Thu 9 Jan 2020	3	2	3	2	3	3	2	2	3
Fri 10 Jan 2020	3	3	3	2	3	2	2	2	3
Sat 11 Jan 2020	3	3	3	3	3	3	2	3	3
Sun 12 Jan 2020	3	3	3	3	3	3	1	3	3
Mon 13 Jan 2020	3	3	3	3	3	3	1	3	3

Step 6: Export Files

The clean data was then exported into file “AQ_Cleaned_2019.csv” using **Export["AQ_Cleaned_2019.csv", data2019Clean, "CSV"]** as a CSV file. This step was also repeated for the 2020 data and both files were stored in “DataDir”.

4.2 Data Pre-Processing: Particulate Matter (PM10 and PM2.5)

Step 1: Import the Dataset

To prepare the PM datasets, data was imported using **Import["England Sites 2019.xlsx", {"Data", 1, 4 ;; -5}]**. This made the data accessible from the 4th until the 5th to final row in the first sheet.

Step 2: Remove the “Status” Rows

The next step was to remove all the “Status” rows. This was achieved using the **Position[england19trans, "PM10 particulate matter (Hourly measured)"]** and **Position[england19trans, "PM2.5 particulate matter (Hourly measured)"]** to retrieve the positions of only PM2.5 and PM10. The result, which can be seen in Figure 6, was in the form of lists, to combine all the data **PMAlleng19 = Join[eng19PM10trans[[1]], eng19PM25Trans[[1]]] // Sort** was used.

Figure 6: Example Output of PM Dataset after Sorting

```
{3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47
 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87,
 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121}
```

A new table containing the date, time, site names and PM type was created using **Join[england20trans[[1 ;; 2]], england20trans[[PMAlleng20]]]**. Redundant data was removed, and the output is displayed in Figure 7.

Figure 7: PM Dataset after Removal of Redundant Rows

	Date	Tue 1 Jan 2019 00:00:00 GMT	Tue 1 Jan 2019 00:00:00 GMT	Tue 1 Jan 2019 00:00:00 GMT
	Time	Tue 30 Nov −2 01:00:00 GMT	Tue 30 Nov −2 02:00:00 GMT	Tue 30 Nov −2 03:00:00 GMT
Bexley - Belvedere	PM10 particulate matter (Hourly measured)	28.2	18.4	14.4
	PM2.5 particulate matter (Hourly measured)	16.	12.7	6.1
Bexley - Belvedere West	PM10 particulate matter (Hourly measured)	31.6	20.8	15.8
	PM2.5 particulate matter (Hourly measured)	19.5	13.7	7.2

Step 3: Turn Hourly Measurements into Daily Values by Averaging Results

After re-selection, the data needed to be partitioned in sets of 24, converting hourly data to daily. Once partitioned, the average value was calculated using **N@(Mean /@ (Partition[eng19dat, 24])).**

Some values contained a "No Data" string, potentially causing a non-numerical value. To overcome this, "No Data" was discounted using a "ReplaceAll" command **eng19part /. ("No data" :> 0)**. The concentrated dataset now has dimensions of {60,365} containing one value for each day of the year.

The averages for the daily values of PM10 and PM2.5 was calculated using the same method – partitioning both individually, then mapping the mean across the list of lists by using **N@Mean/@**.

Step 4: Visualise the Daily Mean Result in a Table

A table was created by forming headings (Site, Particular Matter, Dates) and joining it with the average daily values with the respective sites and PM type. To do that, it was first required to join the latter. Therefore, the command

Partition[Flatten@Table[{england19CleanTEMP[[row, 1;;2]], eng19DailyMean[[row]]}, {row, 1, Length[eng19DailyMean]}], 367] was used. The results were joined by using **Prepend[eng19joinedTab, eng19heading]**. The daily mean in one site is displayed in Figure 8.

Figure 8: PM Dataset after Converting to Daily Mean Values

```
eng19dailyMeanTable[[1 ;; 2, 1 ;; 10]] // Grid
```

Site	Particulate Matter	Tue 1 Jan 2019 00:00:00 GMT	Wed 2 Jan 2019 00:00:00 GMT	Thu 3 Jan 2019 00:00:00 GMT	Fri 4 Jan 2019 00:00:00 GMT	Sat 5 Jan 2019 00:00:00 GMT	Sun 6 Jan 2019 00:00:00 GMT	Mon 7 Jan 2019 00:00:00 GMT	Tue 8 Jan 2019 00:00:00 GMT
Bexley – Belvedere	PM10 particulate matter (Hourly measured)	13.4833	17.8833	21.975	31.3125	24.4083	25.1708	19.2708	20.2958

Step 5: Export the File

The file was exported using **Export["England 19 Particulate Matter (Both) - Daily Mean Concentration.xlsx", eng19dailyMeanTable]** and stored in DataDir.

Step 6: Visualise the Co-ordinates, Sites, Zone and Environment Type in a Table

- Find Co-ordinates of the Monitoring Sites

To calculate the co-ordinates of the monitoring sites, **FindGeoLocation[]** was mapped across all sites. During this process, some elements contained missing values. To progress, these values had their respective co-ordinates re-mapped with the correct co-ordinates. Then, the command **LatitudeLongitude** was used to create a set of coordinates.

- Find Zones and Environment Type of the Monitoring Sites

To find where the monitoring sites fell in the England region, the DAQI dataset was re-imported. This dataset contains all the regions of the respective monitoring sites for both countries. Only the rows with England regions were initially selected from it. The same steps were repeated to find the environment type.

Once obtained, tabulation of the data was possible using **eng19holdValue = Grid[Transpose@Partition[Flatten@Table[{eng19latitude, eng19longitude, eng19sites, CompleteAQZones, CompleteAQET}, 1], 31], Frame -> All]**, the output of which can be viewed in Figure 9.

Figure 9: Tabulated PM Station Data

Latitude	Longitude	Site	Zone	Environment Type
51.4895°	0.147359°	Bexley - Belvedere	Greater London Urban Area	Urban
51.4964°	0.133155°	Bexley - Belvedere West	Greater London Urban Area	Urban
52.1999°	0.128022°	Cambridge Gonville Place	Eastern	Urban
51.524°	-0.143934°	Camden - Euston Road	Greater London Urban Area	Urban
51.7632°	-0.573073°	Dacorum Northchurch High Street	Eastern	Urban
54.968°	-1.60607°	Gateshead Tyne Bridge	Tyneside	Urban
52.5798°	1.73482°	Great Yarmouth South Denes	Eastern	Urban
51.4949°	0.0121162°	Greenwich - John Harrison Way	Greater London Urban Area	Urban

Similar methods were carried out on the remaining datasets to bring them to a pre-processed form.

4.3 Data Pre-Processing: Zones and Environment Types

To determine the closest zone to sites based on their coordinates, this study utilized an excel spreadsheet which contained all sites from the AirQuality regions and used an advanced VLOOKUP. This formula was used to assign the closest zone and then use a regular VLOOKUP to pull the environment type.

The advanced VLOOKUP listed below compares the latitude and longitude of the sites being used in the PM datasets with the ones listed from the AirQuality sites. Once this comparison was made, the information required was extracted and assigned to the individual sites.

=LOOKUP(1,1/FREQUENCY(0,SIN((RADIANS(I$2:I$153-$D3))/2)^2+SIN((RADIANS(J$2:J$153-$E3))/2)^2COS(RADIANS(I$2:I$153))COS(RADIANS($D3))),$H$2:$H$153)

4.4 Techniques

The direction which the analysis would take was not pre-defined. As a result, multiple methods were trialled throughout this report. When observing the pre-processed data for both PM values, it was predicted that average PM would be reduced by COVID-19 because of restricted vehicular travel and most working from home. The daily data was analysed to see to what extent this reduction in PM was visible. Outcomes were visualized in tables for values, as well as several different graphics including Ridgeline graphs and List plots.

4.5 Descriptive Statistics

Several variables were assigned to the data for the Site name and location, environment type and region, this made the data easily identifiable upon extraction for tables and visualizations.

It was possible to obtain descriptive statistics inclusive of Minimum, Maximum, Mean, and Standard deviation. This information was useful in visualizing ridgelines and checking for normality. A significant proportion of the data used was in a temporal format to determine whether the hypothesis – if the lockdowns instigated in the UK because of the COVID-19 pandemic had a measurable impact in the quality of the air – could be accepted or rejected.

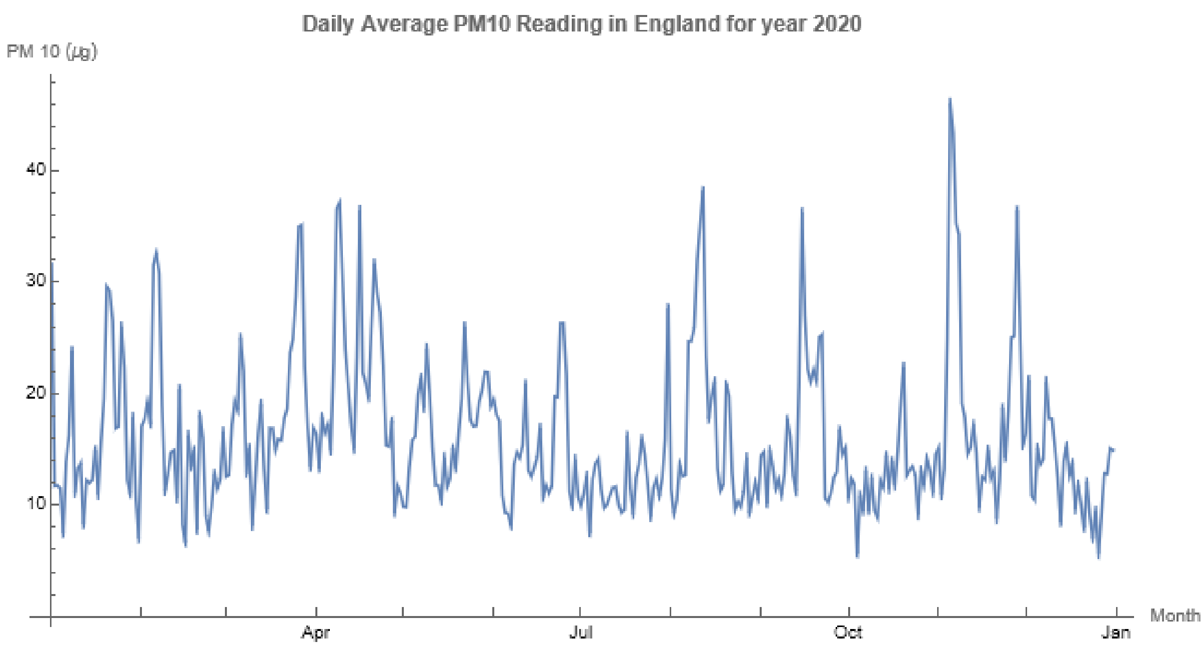
4.6 Modelling Approaches

To visualize the pattern for the average daily PM10 readings across the entirety of England, a series of codes were needed to address tasks such as taking the daily average across all sites and creating a visual of a Temporal object. In calculating the averages, the daily mean for PM10 readings for each site was inspected. Transposing the data allowed the mean command to be mapped across the dataset. This was achieved using **Mean /@**

(Transpose@eng20PM10DailyMean) and gave the daily average over the whole of England.

It was then possible to create pairs of values in the form {Date, Average Reading} by transposing the data , the command **TemporalData[]** could then be utilized to create a temporal object from the pairs. This was then used to visualize the PM10 average readings which can be seen in Figure 10.

Figure 10: Daily Average PM10 Reading in England for the year 2020



4.7 Correlation Between Average Particulate Matters and Average Air Quality Score

One aim was to test whether average PM and average Air Quality scoring were correlated. Using the central Scotland as a direct test of whether this method was feasible, the outcome was determined to have a weak positive correlation (r=0.3). As a result, this method was deemed infeasible. Focus shifted to investigating averages and making use of different graphical methods to investigate average change.

5 Results & Discussion

5.1 Results

According to World Health Organisation (WHO, 2021), there are 4 levels of PM10 and PM2.5 that can affect our health and are categorised as:

1. Good level (air pollution is PM < 10 μg/m^3)
2. Moderate level (air pollution is between 10 and 15 μg/m^3)
3. Poor level (air pollution is between 15 and 25 μg/m^3)
4. Very poor (air pollution is more than 25 μg/m^3)

Tables comprised of the Minimum, Maximum, Mean and Standard Deviation for PM values for the 2019 and 2020 dataset for both countries were created. For each table, the best and worst averages were highlighted. For England, the monitoring site in Slough had the best average of 8.26 μg/m^3 and Maidstone had the worst with 26.33 μg/m^3. As hypothesized, England had a moderate level of Air Pollution with only one site categorized as "Very Poor."

Figure 11: Tables Highlighting the Best and Worst Averages for Sites in England

	Minimum	Mean	Maximum	Standard Deviation
{Bexley - Belvedere}	4.47083	18.6522	70.8167	11.3193
{Bexley - Belvedere West}	4.775	17.4711	69.6625	9.97769
{Cambridge Gonville Place}	0.975	18.5222	63.0875	8.7078
{Camden - Euston Road}	8.2375	21.4346	68.9	9.82563
{Dacorum Northchurch High Street}	3.1375	18.1582	399.079	29.7801
{Gateshead Tyne Bridge}	2.7375	14.1382	71.9083	9.32891
{Great Yarmouth South Denes}	4.29167	19.1321	74.4583	13.2854
{Greenwich - John Harrison Way}	2.54583	14.0537	62.8292	10.1093
{Greenwich - Woolwich Flyover}	8.42917	22.6247	81.3417	10.654
{Hackney - Old Street}	5.25833	19.6312	67.0167	10.5411
{Havering - Rainham}	2.52083	16.7153	64.7625	9.75557
{Heathrow Green Gates}	2.5125	12.7818	55.5417	8.58185
{Heathrow LHR2}	2.79583	13.3706	57.2	9.21743
{Hitchin Stevenage Road Particulates}	1.29583	16.9524	64.3208	10.3073
{Hounslow Brentford}	5.59167	18.7169	71.2208	10.6819
{Hounslow Chiswick}	3.60417	18.3859	57.3458	9.18535
{King's Lynn Stoke Ferry Buckenham Drive}	0.0416667	9.30137	32.5	5.46158
{King's Lynn Stoke Ferry Wretton Road}	3.	10.5638	29.2083	4.71339
{Luton Dunstable Road East}	1.35833	15.676	63.0375	10.5869
{Maidstone Upper Stone Street}	7.11779	26.3292	71.4555	10.769
{North Tyneside Coast Road}	3.49583	17.2554	68.275	10.2015
{Norwich Castle Meadow}	1.95833	18.4315	71.9167	9.71114
{Salford M60}	2.90417	20.7436	72.8292	11.3033
{Scunthorpe East Common Lane Osiris}	0.833333	15.4037	106.	11.5756
{Slough Lakeside 2 Osiris}	2.4375	8.25992	50.1667	9.01064
{South Cambs Girton Rd}	1.63333	16.1904	56.9125	7.49114
{Tower Hamlets - Blackwall}	3.9125	14.6081	74.6083	12.7505
{Tower Hamlets - Victoria Park}	4.70417	17.4519	65.3417	10.0796
{Waltham Forest Dawlish Rd}	2.8625	16.0745	57.0417	8.60731
{Watford Town Hall}	3.70833	14.3986	53.6458	8.99925

	Minimum	Mean	Maximum	Standard Deviation
{Bexley - Belvedere}	4.47083	18.6522	70.8167	11.3193
{Bexley - Belvedere West}	4.775	17.4711	69.6625	9.97769
{Cambridge Gonville Place}	0.975	18.5222	63.0875	8.7078
{Camden - Euston Road}	8.2375	21.4346	68.9	9.82563
{Dacorum Northchurch High Street}	3.1375	18.1582	399.079	29.7801
{Gateshead Tyne Bridge}	2.7375	14.1382	71.9083	9.32891
{Great Yarmouth South Denes}	4.29167	19.1321	74.4583	13.2854
{Greenwich - John Harrison Way}	2.54583	14.0537	62.8292	10.1093
{Greenwich - Woolwich Flyover}	8.42917	22.6247	81.3417	10.654
{Hackney - Old Street}	5.25833	19.6312	67.0167	10.5411
{Havering - Rainham}	2.52083	16.7153	64.7625	9.75557
{Heathrow Green Gates}	2.5125	12.7818	55.5417	8.58185
{Heathrow LHR2}	2.79583	13.3706	57.2	9.21743
{Hitchin Stevenage Road Particulates}	1.29583	16.9524	64.3208	10.3073
{Hounslow Brentford}	5.59167	18.7169	71.2208	10.6819
{Hounslow Chiswick}	3.60417	18.3859	57.3458	9.18535
{King's Lynn Stoke Ferry Buckenham Drive}	0.0416667	9.30137	32.5	5.46158
{King's Lynn Stoke Ferry Wretton Road}	3.	10.5638	29.2083	4.71339
{Luton Dunstable Road East}	1.35833	15.676	63.0375	10.5869
{Maidstone Upper Stone Street}	7.11779	26.3292	71.4555	10.769
{North Tyneside Coast Road}	3.49583	17.2554	68.275	10.2015
{Norwich Castle Meadow}	1.95833	18.4315	71.9167	9.71114
{Salford M60}	2.90417	20.7436	72.8292	11.3033
{Scunthorpe East Common Lane Osiris}	0.833333	15.4037	106.	11.5756
{Slough Lakeside 2 Osiris}	2.4375	8.25992	50.1667	9.01064
{South Cambs Girton Rd}	1.63333	16.1904	56.9125	7.49114
{Tower Hamlets - Blackwall}	3.9125	14.6081	74.6083	12.7505
{Tower Hamlets - Victoria Park}	4.70417	17.4519	65.3417	10.0796
{Waltham Forest Dawlish Rd}	2.8625	16.0745	57.0417	8.60731
{Watford Town Hall}	3.70833	14.3986	53.6458	8.99925

Table 1: Number of Sites in England that fall within the 4 Levels of Air Pollution

	μg/m^3 <10	10< μg/m^3 <15	15< μg/m^3 <25	μg/m^3 >25

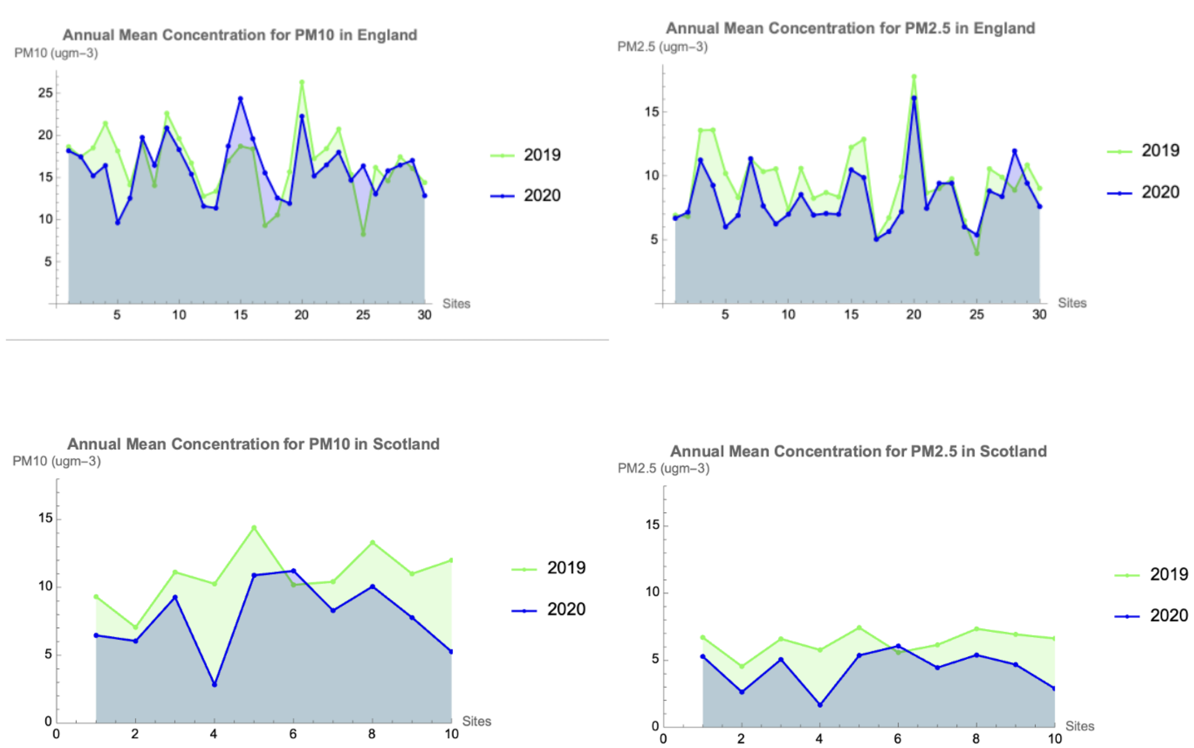
Number of Sites	2	7	20	1
Percentage Represented	6.7%	23.3%	66.7%	3.3%

5.1.1 Annual Mean Concentration Comparison with Both Years (Line Graphs)

The figure and table above suggest a moderate decrease in PM levels from 2019 to 2020, which is better visualised using Line Graphs. Upon overlaying graphs of 2019 and 2020 for all pollutant types, it is visible that a large majority of sites experienced an average reduction in pollutants, with PM10 consisting of 9 sites in Scotland and 20 sites in England, while PM2.5 consisting of 9 sites in Scotland and 25 sites in England.

The visualizations below suggest a reduction in annual mean PM10 and PM2.5 in both England and Scotland.

Figure 12: Comparing the Annual Mean Concentrations from 2019 to 2020 for PM2.5 and PM10 for England and Scotland



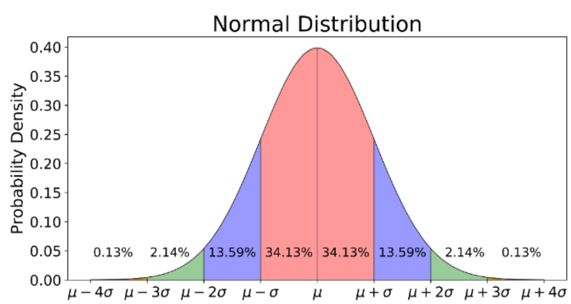
5.1.2 Probability Distribution (Ridgeline Graphs)

It is then possible to use the calculated values of mean (μ) and Standard Deviation (σ) as parameters for a Normal Distribution, with the PDF being:

$$\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

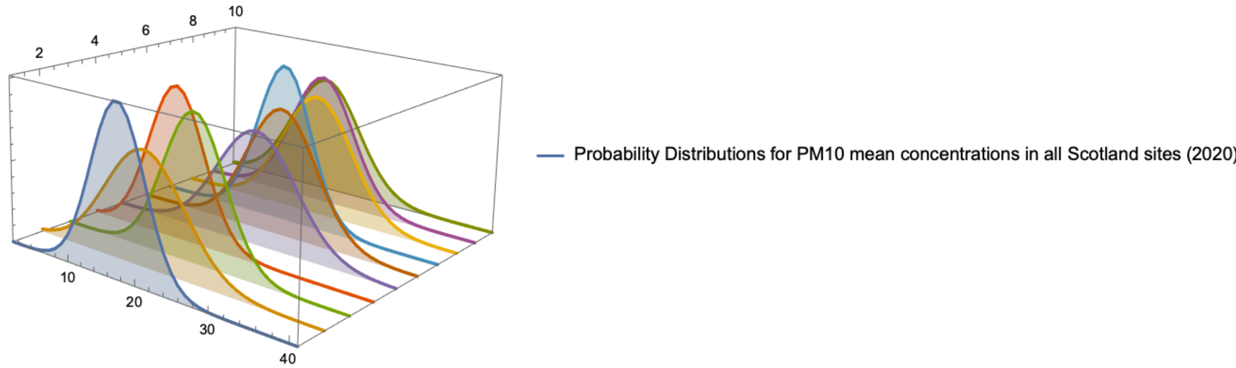
A Normal Distribution PDF has a symmetrical bell-shaped curve with μ being central and represents the average PM value of a site on a particular day, whilst σ constitutes the dispersion of these values in a particular site.

Figure 13: Normal Distribution PDF



Consider the Ridgeline Graphs for Scotland shown in Figure X, it is possible to identify sites with a higher PM by looking at the position of the highest peak along the x-axis, with the lower PM levels being the bell curves with peaks closer to zero. Dispersion can be observed by how wide each of the figures are.

Figure 14: Probability Distribution for PM10 Mean Concentration of All Scotland Sites in 2020



5.1.3 Quantifying the difference in PM10 levels in the sites of Slough and Maidstone in 2019 and 2020

To better understand the changes of PM10 values over the years, English monitoring sites with the highest and lowest average were selected. The Normal Distributions of the PM10 values for 2019 and 2020 have different shapes, moreover the intersection between these areas is minimal which leads us to believe the 2020 values were significantly affected by other external factors.

The probability distribution was constructed to compare with another similar distribution. The test needed to check if both normal models created for the sites Slough and Maidstone were the same by using a Two Sample Z-Tests for means.

The Hypotheses tested are: H0: μ1= μ2H0: $\mu 1= \mu 2$, with H1: μ1≠ μ2H1: $\mu 1\neq \mu 2$.

Table 2: Results of the Two Sample Z-Test

	2019		2020		2019-2020	2019-2020
Sites	Mean μg/m^3	St Dev	Mean μg/m^3	ST Dev	Z-value	P-value
Slough Lakeside 2 Osiris	8.26	9.01	16.37	11.94	0.246	>0.10

Maidstone Upper Stone Street	26.33	10.77	22.27	8.12	-0.301	>0.10

The Paired Z-Test between two normal variables is dependent on the Mean and Standard Deviation:

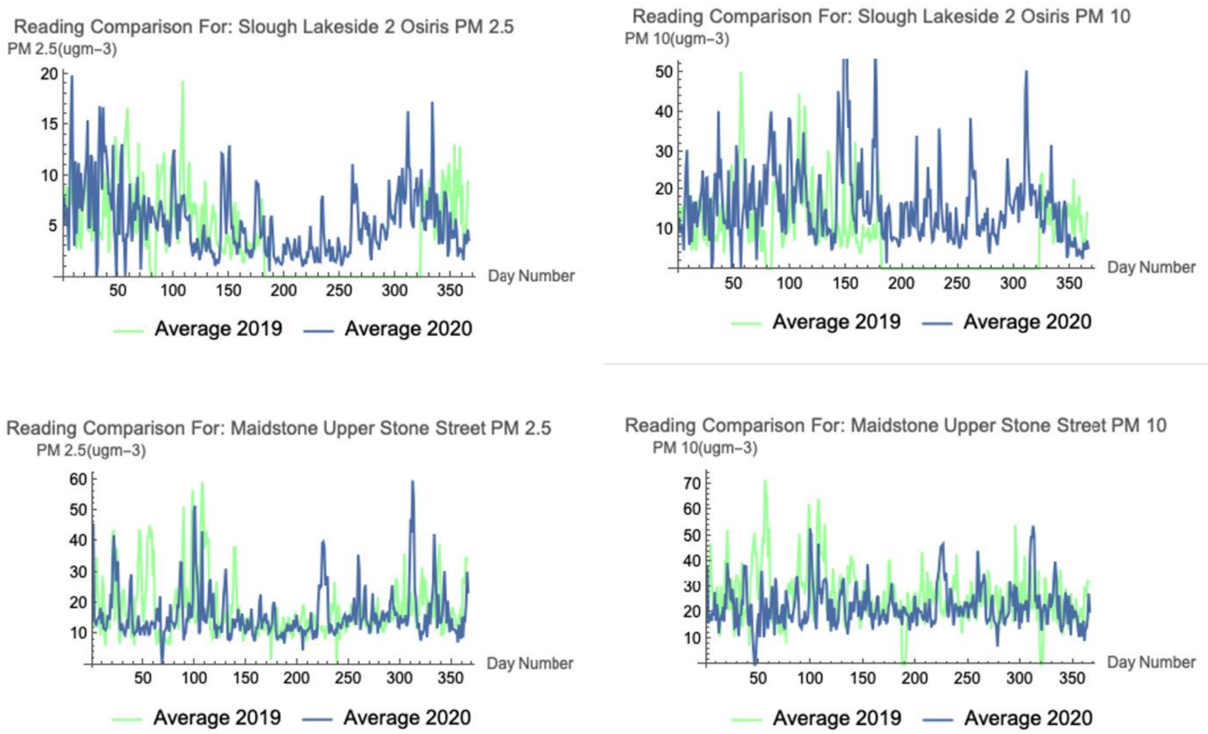
$$Z = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\sigma_{X_1}^2 + \sigma_{X_2}^2}}$$

The resultant Z-Values are 0.246 and –0.301 respectively with p-values above 0.10 which conclude that these datasets do not present a significant difference between the 2019 and 2020 datasets.

5.1.4 Understanding the seasonality of two sites in England

Despite the outcome of the test, it is still possible to understand the air pollution change throughout the years. An overlay graph for the sites of Slough and Maidstone for PM10 and PM2.5 is reported.

Figure 15: Daily Mean Comparison of PM10 and PM2.5 for Slough and Maidstone Sites from 2019 and 2020



5.2 Discussion

5.2.1 Air Quality of England and Scotland

England and Scotland had a moderate to poor level of air quality in 2019, according to the WHO. However, both countries recorded air quality levels for PM2.5 and PM10 better than most countries in Europe (European Environment Agency, 2021):

Table 3: Recorded Air Quality Levels for England, Scotland and Most Countries in Europe for 2019 and 2020

	2019	*2019*	*2020*	*2020*
Pollutant	PM10 mg/m3	PM2.5 mg/m3	PM10 mg/m3	PM2.5 mg/m3
England	16.71	9.55	16.14	8.23
Scotland	10.91	6.37	7.81	4.35

Europe commonest average (2018)	≈18.5	≈13.5	/	/

In 2020, the average PM levels improved. Scotland was able to stay within the “good air quality” threshold of 10mg/m3 setup by the WHO.

In 2019, the worst air pollution means recorded were Fife Cupar (7.43 μg/m^3 PM2.5, 14.40 μg/m^3 PM10) and Maidstone Upper Stone Street (17.78 μg/m^3 PM2.5, 26.33 μg/m^3 PM10). These sites improved in 2020 with both pollutants (Fife 5.37 μg/m^3 PM2.5, 10.89 μg/m^3 PM10 and Maidstone with 16.10 μg/m^3 PM2.5, 22.28 μg/m^3 PM10) representing an overall improvement.

5.2.2 Air Quality in Towns versus Countryside

In the initial stages of the analysis, a comparison between rural and urban sites were considered. However, there was an inadequate number of qualified rural sites. Despite our hypothesis, the only two rural sites (King's Lynn Stoke Ferry Buckenham Drive and King's Lynn Stoke Ferry Wretton Road) recorded higher pollution in 2020 than 2019, but this sample is too small to be representative.

5.2.3 COVID-19 Pandemic Effect on the Air Quality Level

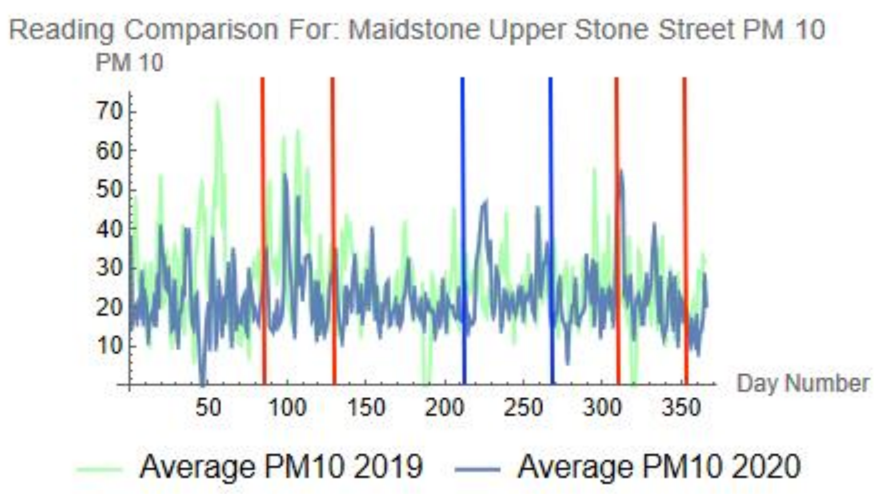
The most significant result of the research is represented by the overlay graphs of England and Scotland for PM2.5 and PM10. 90% of sites recorded lower PM10 values from 2019 to 2020, including extreme decrease in sites such as Falkirk (PM10 from 10.27 μg/m^3 to 5.77 μg/m^3; 44% decrease) and Camden (PM2.5 from 13.60 μg/m^3 to 9.25 μg/m^3; 32% decrease). This suggests that there must have been a significant change in air pollution level in 2020 that might be linked to COVID-19.

Due to the limited time and computational power to develop the research, a Paired Z-Test was only evaluated between the data samples for PM10 in 2019 and 2020 for the sites of Slough Lakeside 2 Osiris and Maidstone Upper Stone Street. The results of z-value

0.246 (p>0.10) and z-value –0.301 (p>0.10) respectively, Thus, having no real statistically significant difference.

An argument that supports the lockdowns impacts on the air pollution levels can be made through the graphs that compare the yearly air pollution levels of a specific site. Making use of the UK lockdown timeline (Institute for Government, 2021), two lockdowns (between red lines) and one "return to normality" phase (purple) were identified (Institute for Government, 2021). The lockdown phases strongly contributed to the PM10 levels reduction during the first lockdown (days 85 to 132), while during the ease of restrictions (days 217 to 260) the level of pollution rose to a similar level to 2019.

Figure 16: Comparison of PM10 Values from 2019 to 2020 for Maidstone Site with COVID-19 Lockdown Indicators



Although statistically non-significant, a decreasing trend of the level of PM10 and PM2.5 can be observed.

6 Conclusion

This study analyses the change in air pollution levels between the years 2019 and 2020 in England and Scotland, and its linkage to the Covid Lockdown. The main result shows that there has been a visible reduction of air pollution, with 32 sites recording improvements in PM, but the difference in PM levels between years was not statistically significant, rejecting the initial hypothesis. It is therefore possible to conclude that the Covid Pandemic affected air quality, but its contribution was non-significant.

Covid Lockdowns have affected the production of emissions produced by means of transport and heavy industries, improving air quality levels across the world (Ming *et al*., 2020). Studies like this one and conducted in the US, the UK, China, Brazil and Saudi Arabia (Zangari *et al*.,2020; Jephcote, 2021; Nakada and Urban, 2020; Aljahdali *et al*., 2021, respectively), found no-significant impact of Lockdown on air quality in the short-term, but it is thought that an extended lockdown would purify the air we breathe.

The decision was made not to make use of MongoDB or MySQL due to the resultant dataset not being vast. If this study were to be scaled up in size and duration, MySQL would be useful for specific search term querying. Future analyses have potential to include plotting predictors which make use of splines in a nonlinear model style fit so that we can predict future values based upon many things such as how the values are clustered (or rather, in knots).

Another issue of open-source data is its completeness and correctness. Missing information and storage issues could be avoided by installing more technologically advanced air monitoring devices. In this case the sample size would be larger, and the analysis would yield more precise results.
Moreover, Particulate Matter is only one of the many air pollutants that impacts human health and climate change. There is a broad range of poisonous gases that contribute to air pollution including carbon monoxide, nitrogen dioxide, ozone, and sulphur dioxide. If this study were extended, it would be possible to collect these datasets and evaluate the impact of Covid Lockdowns on them. Finally, it would be possible to extend the research to create an air quality forecasting tool with heat maps like the ones in weather apps.

7 Bibliography

Aljahdali, M.O., Alhassan, A.B., and Albeladi, M.N., 2021. Impact of novel coronavirus disease (COVID-19) lockdown on ambient air quality of Saudi Arabia. *Saudi Journal of Biological Sciences*, **28(2)**, pp.1356–1364.

Calderon, C. and Keirstead, J., 2012. Modelling frameworks for delivering low-carbon cities: advocating a normalized practice. *Building Research and Information,* **40(4)**, pp.504-517

California Air Resources Board, 2021. Inhalable Particulate Matter and Health (PM2.5 and PM10). https://ww2.arb.ca.gov/resources/inhalable-particulate-matter-and-health#:~:text=5%3F-, last accessed 19/11/2021

Department for Environment, Food & Rural Affairs, 2021a. ENV02 - Air quality statistics. https://www.gov.uk/government/statistical-data-sets/env02-air-quality-statistics, last accessed 19/11/2021

Department for Environment, Food & Rural Affairs, 2021b. DAQI Regional Data. https://uk-air.defra.gov.uk/data/DAQI-regional-data, last accessed 19/11/2021

European Commission, 2015. Horizon 2020, air pollution and climate change – a vicious circle. https:// ec.europa.eu/programmes/horizon2020/en/news/air-pollution-and-climate-change-–-vicious-circle, last accessed 11/11/21

European Environment Agency, 2021. Air Quality in Europe - 2020 Report. https://www.eea.europa.eu/publications/air-quality-in-europe-2020-report, last accessed 19/11/2021

Institute for Government, 2021. Timeline of UK government coronavirus lockdowns. https://www.instituteforgovernment.org.uk/charts/uk-government-coronavirus-lockdowns, last accessed 19/11/2021

Jephcote, C., Hansell, A.L., Adams, K., and Gulliver, J., 2021. Changes in air quality during COVID-19 “lockdown” in the United Kingdom. *Environmental Pollution*, **272(116011)**, p.116011.

Li, L., Zhou, Z., Kalo, M., and Piltner, R., 2016. Spatiotemporal Interpolation Methods for the Application of Estimating Population Exposure to Fine Particulate Matter in the Contiguous U.S. and a Real-Time Web Application. *International Journal of Environmental Research and Public Health,* **13(8)**, 749

Ming, W., Zhou, Z., Ai, H., Bi, H., and Zhong, Y., 2020. COVID-19 and Air Quality: Evidence from China. *Emerging Markets Finance and Trade*, **56(10)**, pp.2422–2442.

Nakada, L.Y.K. and Urban, R.C., 2020. COVID-19 pandemic: Impacts on the air quality during the partial lockdown in São Paulo state, Brazil. *Science of The Total Environment*, **730**, p.139087.

Redondo-Bermúdez, MdC., Gulenc, I.T., Cameron, R.W. and Inkson, B.J., 2021. ‘Green barriers’ for air pollutant capture: leaf micromorphology as a mechanism to explain plants capacity to capture particulate matter. E*nvironmental Pollution, 2***88,** 117809

Ropkins, K., and Tate, J.E., 2021. Early observations on the impact of the COVID-19 lockdown on air quality trends across the UK. *Science of The Total Environment*, **754**, 142374.

United Nations, 1972. 1972 Declaration of the UN conference on the Human environment, Stockholm Declaration. https://www.un.org/en/conferences/environment/stockholm1972, last accessed 11/11/21

United States Environmental Protection Agency, 2021. Particulate Matter (PM) Basics. https://www.epa.gov/pm-pollution/particulate-matter-pm-basics, last accessed 19/11/2021

Word Health Organisation, 2021. Monitoring health for the SDGs, world health statistics. https:// apps.who.int/iris/bitstream/handle/10665/342703/9789240027053-eng.pdf, last accessed 19/11/2021

Zangari, S., Hill, D.T., Charette, A.T., and Mirowsky, J.E., 2020. Air quality changes in New York City during the COVID-19 pandemic. *Science of The Total Environment*, **742**, 140496