Czech University of Life Sciences
Faculty of Environmental Sciences

Department of Water Resources and Environmental Modelling

Study Program: Environmental Data Science

Czech University of Life Sciences Prague

Development of a web application for air quality monitoring in Czech Republic

BACHELOR THESIS

Author: Denis Maxheimer

Supervisor: doc. Ing. Mgr. Ioannis Markonis, Ph.D.

Year: 2022

Czech University of Life Sciences Prague

Faculty of Environmental Sciences

BACHELOR THESIS TOPIC

Author of thesis:	Denis Maxheimer
Study programme:	Environmental Data Science
Track:	Informatics
Thesis supervisor:	doc. Mgr. Ing. Ioannis Markonis, Ph.D.
Supervising department:	Department of Water Resources and Environmental Modeling
Language of a thesis:	English

Thesis title:

Development of a web application for air quality monitoring in Czech Republic

Objectives of thesis:

This thesis aims in developing a web application, which will provide a variety of air quality aspects by statistical visualization. It will cover the whole region of Czech Republic, highlighting regions that are prone to air pollution.

Methodology:

The thesis methodology will follow this, or a similar structure:

1. Create an algorithm for accessing and collecting data from API
a. Store data on PostgreSQL server. Tables design.
b. Decide on which variables are necessary.
c. Provide some statistics for received data.
2. Theoretical part, where should be decided on additional data, which can provide new variables and connections between them.
a. Except for air quality variables, additional candidates are weather, industrial zones, health issues statistics.
3. Provide algorithms for accessing additional data and create dependencies from air quality datasets
4. Aggregate statistics for areas
a. Anthropogenic/natural activity pollution
b. Most probable consequences from air pollution according to the area
5. Do visualization and mapping (to be decided on the best software for this)
6. Create web application design
a. Create backend (server) part
b. Create frontend (client) part
c. Combine with actual datasets
7. Final tests and setup on server

The proposed extent of the thesis: 50 pages

Keywords: air quality, web application, air pollution monitoring, statistical vizualization

Recommended information sources:

1. Gurjar, B. R., Molina, L. T., & Ojha, C. S. P. (Eds.). (2010). Air pollution: health and environmental impacts. CRC press.
2. Vallero, Daniel. Fundamentals of air pollution. Academic press, 2014.

Expected date of thesis defence: 2021/22 SS - FES

Advisor of thesis: Pavla Dagsson Waldhauserová

Electronically approved: 29. 3. 2022
prof. Ing. Martin Hanel, Ph.D.
Head of department

Electronically approved: 29. 3. 2022
prof. RNDr. Vladimír Bejček, CSc.
Dean

Author's Statement

I hereby declare that I have independently elaborated the diploma/final thesis with the topic of: Development of a web application for air quality monitoring in Czech Republic and that I have cited all the information sources that I used in the thesis and that are also listed at the end of the thesis in the list of used information sources. I am aware that my diploma/final thesis is subject to Act No. 121/2000 Coll., on copyright, on rights related to copyright and on amendment of some acts, as amended by later regulations, particularly the provisions of Section 35(3) of the act on the use of the thesis. I am aware that by submitting the diploma/final thesis I agree with its publication under Act No. 111/1998 Coll., on universities and on the change and amendments of some acts, as amended, regardless of the result of its defence. With my own signature, I also declare that the electronic version is identical to the printed version and the data stated in the thesis has been processed in relation to the GDPR.

In Prague on 29.03.2022

Denis Maxheimer

Acknowledgment

Thank you doc. Ing. Mgr. Ioannis Markonis for leading this work and for his inspiring suggestions. And thank you Mgr. Pavla Dagsson Waldhauserová, Ph.D. for bringing new amazing ideas to the project.

Denis Maxheimer

Thesis Title:

Development of a web application for air quality monitoring in Czech Republic

Author: Denis Maxheimer

Study Field: Environmental Data Science

Study Program: Informatics

Type of thesis: Bachelor Thesis

Supervisor: doc. Ing. Mgr. Ioannis Markonis, Ph.D.

Department of Water Resources and Environmental Modelling

Consultant: Mgr. Pavla Dagsson Waldhauserová, Ph.D.

Department of Water Resources and Environmental Modelling

Abstract(EN): This thesis aims in developing a web application, which will provide a variety of air quality aspects by statistical visualization. It will cover the whole region of Czech Republic, highlighting regions that are prone to air pollution.

Abstract(CZ): Tato práce si klade za cíl vyvinout webovou aplikaci, která bude demonstrovat různé aspekty kvality ovzduší pomocí statistické vizualizace. Prace bude pokrývat celou zemi České republiky s důrazem na regiony náchylné ke znečištění ovzduší.

Key words: air pollution, Czech Republic, live monitoring, online application

Contents

List of abbreviations used

EEA European Environmental Agency

API Application Programming Interface

CHMI Czech Hydrometeorological Institute

CO Carbon monoxide

NO2 Nitrogen dioxide

SO2 Sulfur dioxide

O3 Ozone

URL Uniform Resource Locator

CSS Cascade Style Sheets

BEM Block Element Modifier

GIS Geographic Infromation System

WHO World Health Organization

csv Comma Separated Values

List of Figures

Chapter 1

Introduction

Nowadays a lot of attention is focused on saving the environment and decreasing the negative human impact on it. Comparing to the others types of pollution, air pollution is happening in a shorter time scale and in most cases can be noticed by anyone, since air quality directly affects our health and lifestyle.

Applying modern data collection, processing and visualization tools, the analysis of current conditions of air pollution can be brought to social masses. The possible connection of air quality with regular weather conditions, industrial activity, vehicles usage and common health issues is estimated on air pollution data, known pollutant behavior and information about administrative regions.

This thesis presents the development of an on-line air quality monitoring application for the major cities and towns of the Czech Republic. Based on the results of previous approaches on finding the connection between ambient air pollution and local life quality, this work includes a common range of pollutants and higher temporal coverage. The data are collected with the use of python programming language, and subsequently they are processed and saved in the PostgreSQL database. For their visualization the python matplotlib package is used.

Chapter 2

Objectives

This work is focused on creating a web application for sharing the state of air pollutant concentration in the Czech Republic cities. The ideal aim of a resulting application is to make a clear and simple summary for the any end-user. To this end the objectives of the thesis are:

- Analyse the historical levels and trends of air pollution in Czech Republic by regions.
- Find problematic places and compare live data results with the historical ones.
- Develop an application for collecting and processing the data.
- Classify processed live data and provide a brief summary for the end-users.
- Create a web application for displaying results of the summaries.
- Raise the social awareness and involvement on the air pollution issues in local areas.

Chapter 3

Literature Review

The ambient air pollution major threat is present since 1950s, when the coal-combustion power plants were put into the operation emitting large amounts of SO2, NOx and other pollutants into the air [1]. Ambient air pollution monitoring was initialized and maintained by the CHMI[1] since 1964, expansion of a monitoring network and the smog warning systems was performed during the next years covering Ostrava region in 1969 and Brno region in 1970 [1]. Other organisations were involved in the ambient air quality measurement and monitoring, such as: Public Health Service, Organization for the Rationalization of Power Plants, Czech Geological Survey, Forestry and Game Management Research Institute, Water Management Research Institute and the Research Institute of Plant Production [1]. The first measurements in the industrial areas were followed by the analyzing of elevated SO2 concentrations impact on health. Children organisms responded to the higher SO2 levels same way as if they were in high mountains with the lower O2 pressure. An increase in the total surface of erythrocytes ensured the transfer of the O2 to tissues easier [2]. In 1987 an estimated 2-3% of the total mortality in the Czech Republic could be connected with the air pollution there [3]. Back in few decades, an air pollution problem was considered as the local scale problem. However due to

[1]Czech Hydrometeorological Institute

the long-range transport it has developed into the regional-scale issue [4].

A positive decrease trend of the air pollution can be referenced back to the 1980s [5]. However, first noticeable effect of an ambient air quality improvement was after passing through the communist past, which was triggered in November 1989 by the Velvet Revolution [1]. Newly adopted strict legislation measures, investments into desulphurisation remedies installed at large emission sources and a transfer from burning lignite to gas heating, brought the reduction of 90% of the SO2 emissions [1]. An important motivation in reducing pollution for the Czech Republic also was the preparation for the EU accession [6].

From 1989 till 1998 an incredible improvement in the air pollutant emissions reduction can be seen [7, 8, 9]. This is due to the transition from the coal burning to gas in many places and the newly strict emission limits to be met in 1998 (Figures 3.1 and 3.2, taken from article [1]).

Nowadays the number of air measuring stations had been reduced or modernized. There are 198 measuring stations in the Czech Republic, 127 of them are maintained by the CHMI and 80 out of 127 stations are automated [11]. The air quality had also improved since then, however some places still have problems with the suspended particles, ground-level ozone and benzo[a]pyrene, similarly to other areas in Europe [12]. Current directions in an air pollution monitoring are CO, SO2, NOx, PM2.5, PM10, benzene, benzo[a]pyrene and toxic metals in aerosol [1].

Self-cleaning process of the atmosphere- atmospheric deposition, which cleans an ambient air from pollutants, however it acts as a contamination source for the hydrosphere, pedosphere and biosphere [13]. An improvement of air quality and atmospheric deposition decreasing in the same time, have also been demonstrated by long-term biomonitoring, including mosses and tree bark analysis [14, 15, 16, 17, 18, 19].

A suspended particles air pollution is considered to be a major problem in the Czech Republic and other countries [12]. The Moravian-Silesian region is considered as both Czech and European hotspot of the PM and BaP [11, 12]. The particles are

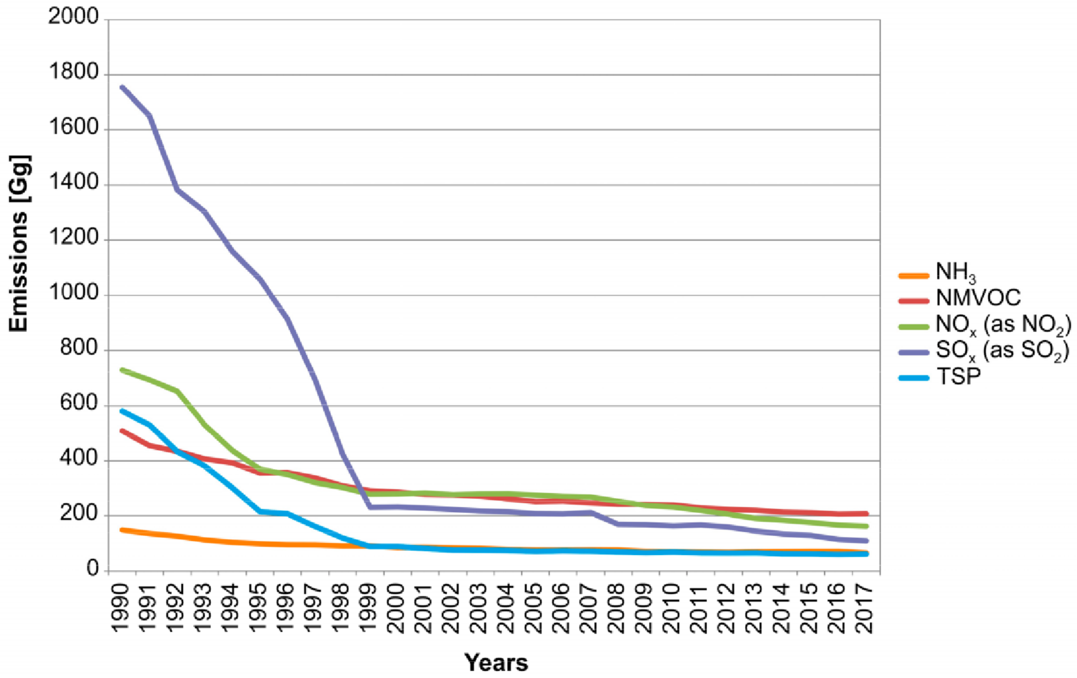

Figure 3.1: Overall emissions of the SO2, NOx, NH3, NM VOC and TSP in the Czech Republic, according to the European Monitoring and Evaluation Programme (EMEP) data [10]

emitted from the industries, such as: steel and coke plants and also from a solid fuel (either coal or wood) house heating. It is estimated that about 20% of the households in Czech Republic are heated individually using solid fuels [20]. Different kinds of transport also contribute to the particulate matter pollution. It can be serious in the populated cities such as Prague, where a lot of roads cross the city center and residential neighbourhoods [1, 21, 22].

The ground-level ozone has been measured in Czech Republic since 1993, however it is considered as an air pollutant only recently. There was no decrease of O3[2] in correspondence to the decrease of O3 precursors gases [1]. The correlation of the ozone levels with the meteorology and air pollution in a form of precursors was shown on five Czech rural sites, where O3 levels differed accordingly to the local

[2]Ozone

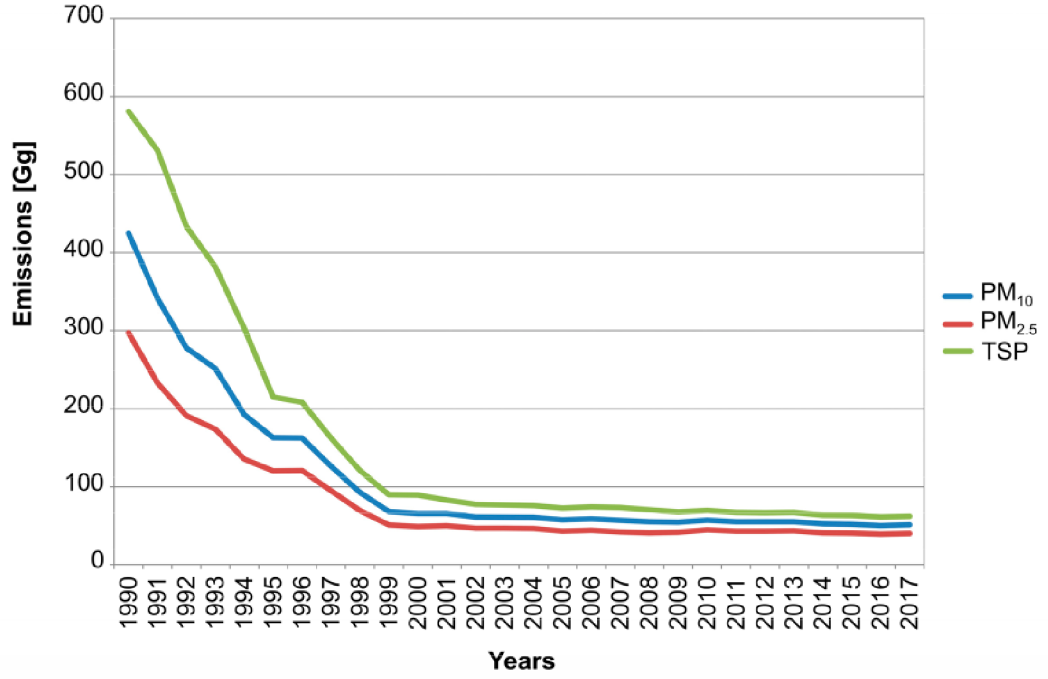

Figure 3.2: Overall emissions of the TSP, PM10 and PM2.5 in the Czech Republic, according to the EMEP data [10]

air temperature, solar radiation, relative humidity and ambient NOx concentrations [23]. Spatial pattern of the ozone emissions correlates with the global solar radiation map across the Czech Republic [24]. The highest ozone levels are found in the northern and southern parts of Czech Republic (3.3). In the northern part, the pollution correlates with the presence of NOx pollution, whereas in the southern part any other air pollution can not be concluded [1].

The air pollution and air quality topics developed over time since 1950s. The measuring systems are improved according to the actual possible pollution problems. The scale of ambient air pollution issues had increased to the regional-scale and its impacts on human health and environment are collated [1].

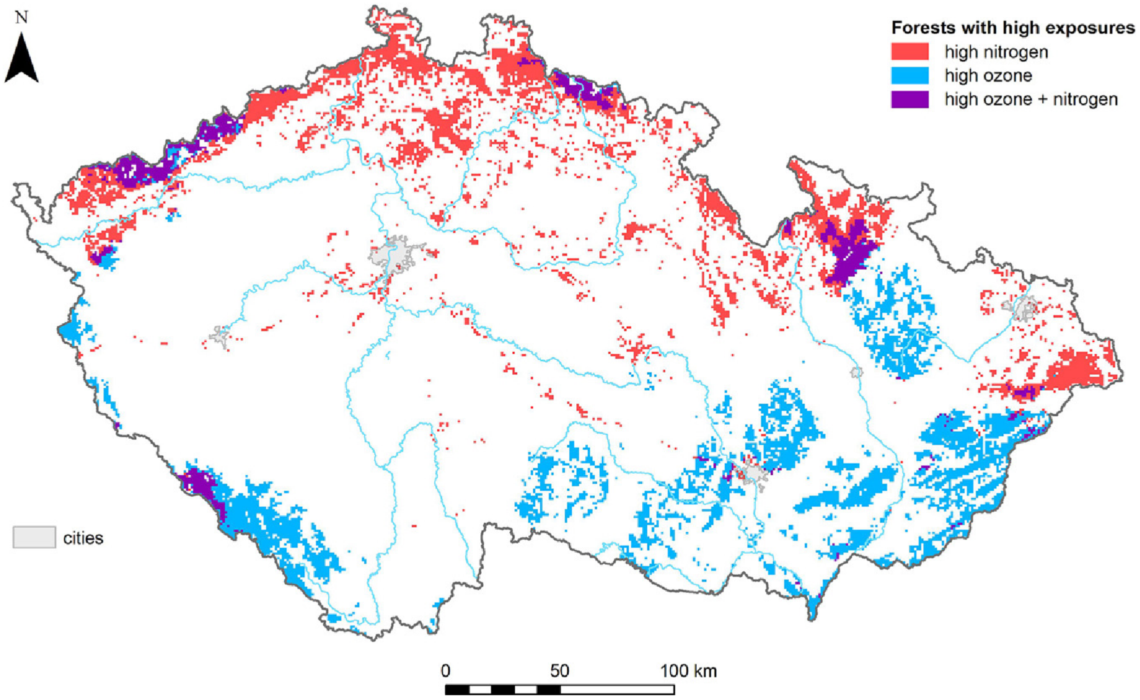

Figure 3.3: Spatial distribution of the forest areas under the highest O3 exposures and N deposition [23]

Chapter 4

Data & Methods

4.1 Data

Data for this work is collected from the European Environmental Agency API[1] using a python code. These sources are easy to process using the python programming language and contain enough of the accessible data, however it is not directly collected by the organizations. For the data on cities and measuring stations, additional tables containing names and spatial information in the database were created. Then it can be referenced in the other tables.

For running the data collecting scripts in a particular time, the python schedule package had been used. Then everything is being run on the local server in a three separate parts with a daily period:
- 23:30 the data for last 2 days is collected,processed and saved to the daily table
- 00:00 the summarizing script is run to update the current pollution situation, which is displayed on the website
- 00:00 the check for if today is first day of the month, if so, the monthly aggregating script is run for previous month

Every run of the code is provided with the safety checks for duplicate values before

[1]Application Programming Interface

doing any aggregations.

4.1.1 GIS data

One of the aims of this project is to cover the most area of the Czech Republic. Therefore it depends on the CHMI sensors coverage. To estimate the values for the whole country, following approach had been used. First, the polygon data had been collected from the Google Earth Engine data set in resolution of the second level administrative areas [25]. The result is an area with the largest city in it. The data is saved in the 'regions' data table with the PostGIS extension to keep a geometry type data correctly (4.1).

However, these areas yet can not be used to directly reference records from the EEA[2] data sets. Further step is to get location of the CHMI stations, which are present in the live data set. Each record in the downloaded data has an info on the station name and sampling point with the x and y coordinates. For storing this information, another table named 'stations' had been created. It stores the data of the station name, id, location and referencing to the region polygon containing the corresponding station (4.1).

The intersection of the station location and region is checked with the python shapely package. Region polygons are extracted from the data table using GeoJSON format. Station sampling point coordinates are projected to the ESPG:4326 coordinate system and then using the '.contains()' method from a shapely package the matching area for each station is found. Then the each station info is saved in the data table and referenced with the region id (4.1).

Such approach makes this data collecting algorithm works independently from the maintaining programming code. The error will not be thrown, when the new stations are met, since this lets the script automatically update the database along with the CHMI adding more stations to the data set.

[2]European Environmental Agency

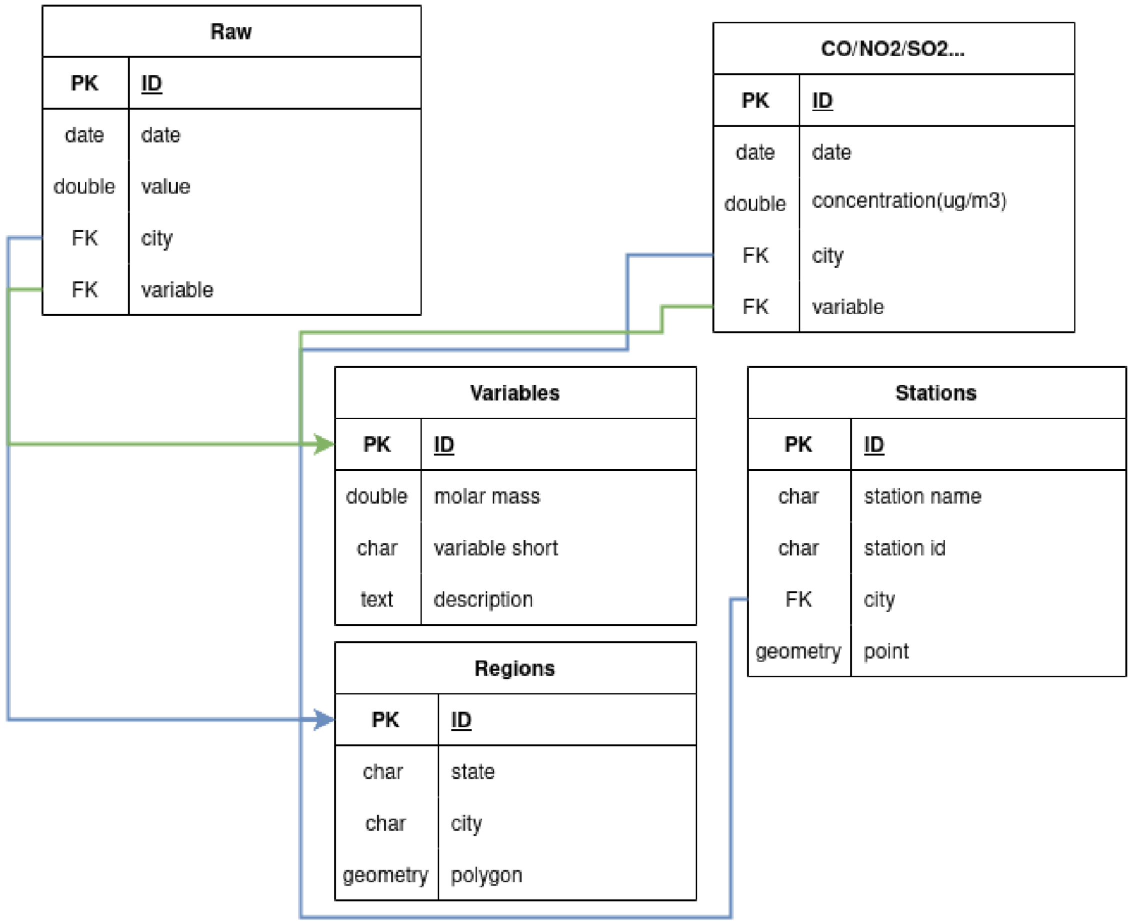

Figure 4.1: Main database diagram

4.1.2 Variables

There is a very wide range of the pollutants information for the historical data. However, for the live data only general variables are available, which are: Carbon Monoxide(CO), Nitrogen Monoxide(NO), Nitrogen Dioxide(NO2), Ozone(O3), Particulate Matter in size of 2.5 nm and 10 nm(PM 2.5 and PM 10), Sulfur Dioxide(SO2). This is a descriptive list of variables, analyzing which can give a clearer understanding on the sources of pollution in different areas. In the data table each variable is stored with a brief description on the sources and possible damage caused by the high concentrations of the pollutant.

4.1.3 Historical data

For collecting historical data, the EEA downloading service had been used [26]. It provides the annual data sets for the combinations of city and variable. Since the project is aimed at a near live measurements, only variables, which are accessible with the live API had been used.

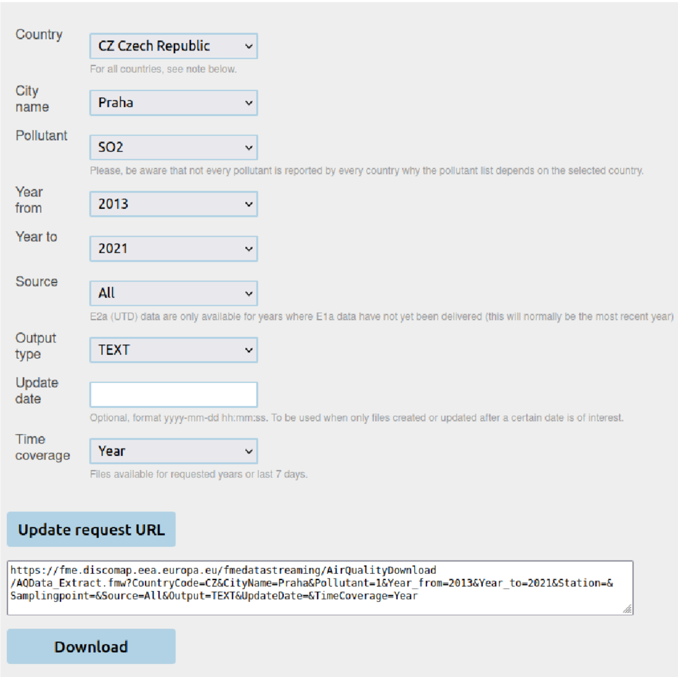

Figure 4.2: EEA download service for air data

Using the python code a data parser had been created. The parser sends request to a corresponding URL[3] with the country, city name, pollutant name and measuring period as the URL parameters. Callback response is a list of another URLs for downloading the filtered data. Downloaded data had been stored in the folders named as 'City-name Variable-name Year-range'.

Data table 'cities' had been created to store the available cities names for the historical data. Locations of the cities had been collected using an API tool, which find an approximate city coordinates by the city name. Then a reference to the

[3] Uniform Resource Locator

regions table was made using the python shapely library with the .contains method.

Further use for the historical data is to have some baseline values for the classifying live measurements as 'High', 'Normal' or 'Low'. 20 and 60 percent quantiles are found for the historical average monthly values and compared with the live data.

After the python parser finished downloading, another algorithm is made to slice and put the required data to the database for further use in the analysis. In the database the data is aggregated to the tables by pollutant, they store timestamp, concentration value and location reference.

4.1.4 Live data

For collecting live data, another EEA service had been used. Using the python code, request is made every day at 23:30, collecting the pollution data for the all available pollutant - city pairs for the last two days. The data itself is taken from the CHMI database, that measures an air pollutant concentration with the sensors in the corresponding cities. Firstly, the python script downloads the data sets in a csv[4] format and saves them locally. However, to save the storage space on the server, every new download replaces previously downloaded file. Then only valid and verified records are taken to the general data table 'raw'. In 'raw' each record is referenced by the variable id and region id in which the station is located.

In the next step the daily script is running. First all the duplicates, if exist, are removed, to prevent the incorrect values after aggregation. The script takes average daily value from the 'raw' table by combination of the city, variable, date. The corresponding row is inserted to the daily table.

After the daily values had been collected, script for the monthly data checks completeness of the daily data for the last month. If there are values with a date earlier than the first day of the month, and there are values with a date later than the last day of the month, in such case the monthly data is taken as completed. In a

[4]Comma Separated Values

further processing, the average mean monthly value is found as well as the monthly: minimum, maximum, standard deviation, particles per million (ppm), particles per billion (ppb). The values are stored in the 'monthly' table along with the year, month, city and variable information.

Such structure approach of the script works so, that the longer script is running, the more precise classification of the monthly values is made, since it is based on the historical data. If a new sensor is placed in a city, which had not been measured yet, in an annual time scale it will be classified next year.

4.1.5 Methods

The thesis methodology follows the following structure:

- 1. Create an algorithm for accessing and collecting data from APIs
- a. Store data on PostgreSQL server. Tables design
- b. Decide on which variables are necessary
- c. Provide some statistics for received data.
- 2. Theoretical part, where should be decided on additional data, which can provide new variables and connections between them.
- 3. Provide algorithms for accessing additional data and create dependencies from the air quality datasets
- 4. Historical data analysis
- a. Anthropogenic/natural activity pollution
- b. Most probable consequences from the air pollution according to the area
- 5. Data visualization
- 6. Create web application design

- a. Create backend (server) part
- b. Create frontend (client) part
- c. Combine with actual data sets and algorithms
- 7. Final tests and deployment on server

4.1.6 Web application

The web application had been created using the python Django framework. The structure goal was to have a user-friendly web site without unnecessary content (4.3 and 4.4). There is a single button which leads to the list of the available cities. For each city there is a list of the available pollutants along with summary and short description of the sources and negative effects from the high pollutant concentration (4.5).

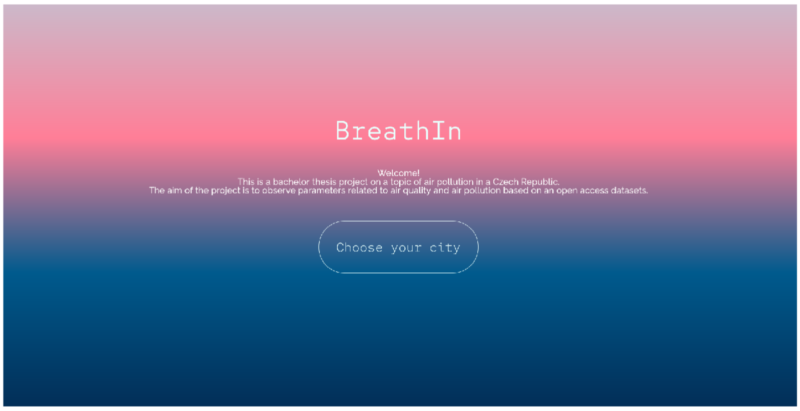

Figure 4.3: Web site title page

4.1.7 Django framework

Django is a python framework for creating a web application. It's been selected due to the high feature variety for further development and personal familiarity with this framework. Web application consists of a single page and interconnected with

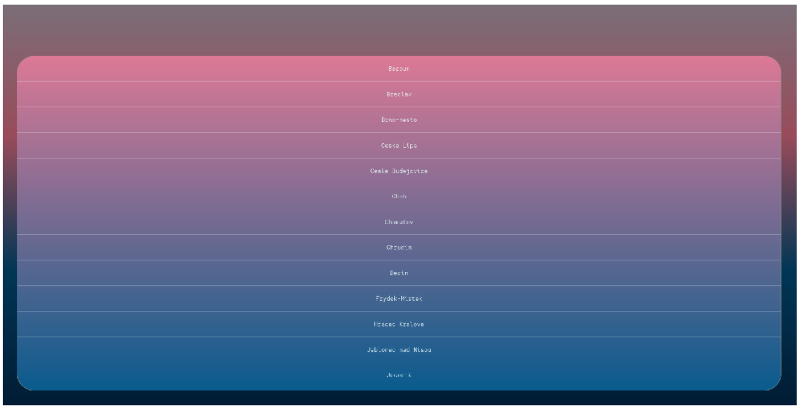

Figure 4.4: Web site city selection menu

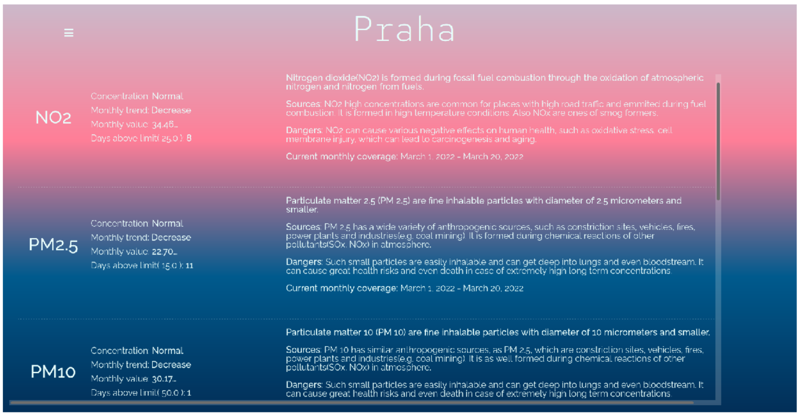

Figure 4.5: Web site city summary page

the regions and summary data tables to provide the information by location. Styling was made using the CSS[5] language and BEM[6] methodology.

4.1.8 Python scripts

Using the python programming language the data processing had been made. General packages used in processing are pandas- for the data manipulation, psycopg2- for the PostgreSQL database interconnection, shapely and geojson- for the GIS[7] data manipulation. Basic structure of the processing algorithm consists of a two

[5]Cascade Style Sheets
[6]Block Element Modifier
[7]Geographic Infromation System

files for each kind of operation: class object file and run file using corresponding object. Example workflow for a near-live time raw data collecting:

- main.py - contains EEA object with necessary methods for data collection (B)
- run.py - contains a function for running methods from main object and implementing them (C)
- db.py - class object with methods on database manipulation, used every time information from the database is requested or recorded there (A)

Overall code structure for the project is following:

Historical data

historical.py -> run historical.py - collecting historical EEA data
historical/main.py -> historical/run.py - analysing collected historical data

Live data

main.py -> run.py - collecting near-live data from EEA API
daily.py -> run daily.py - processing collected data on a daily time scale
monthly.py -> run monthly.py - aggregating daily data to monthly scale
core/main.py -> core/run.py - processing daily data to provide a summary for the web site

4.1.9 Data summary

The summary for each pair pollutant-city is provided in three aspects:

- Daily value classification is interpreted as Low, Normal or High. Daily concentration is classified by comparing the current value with the historical values for the same month. 20 percent quantile is used for classifying Low concentration, 60 percent quantile is used for the High values classification. Such quantiles had been selected in a regard to the actual resulting values, since

>80% quantile is not expected to be exceeded and 20-60% range corresponds more to the common values range.

- Last 30 days concentration trend is calculated using the linear model fitting. Based on the resulted coefficients trend is classified as the Stable, Increase/Decrease or Intense increase/decrease.
- Number of days in month exceeding limit or guideline value are taken for the daily concentrations from the EEA and WHO[8] recommendations. If limit value is not stated, the guideline value was taken. Then the daily concentration is compared with the limit value and if exceeded, a counter increased by 1. During the calendar month total of the exceeding can be observed.

[8] World Health Organization

Chapter 5

Analysis

It is always said that the air pollution is a visible pollution with bare eye. However, currently there is an ecological trend in Europe on controlling air pollutant emissions. The purpose of the work is to provide a tool, which could briefly summarize current monthly values on pollutants concentrations and show correlation of this values to the historical ones.

For a live data the summary script had been created to provide an estimated result on the concentration class, monthly concentration trend and how many days the values exceeded a daily European limits on the pollutants concentration. Final result must be understandable for anyone without background knowledge about air pollution. This might raise people awareness on a possible air quality problematic in their local area, which is the next step in creating a stable and clean environment.

Historical data analysis in this work is performed by aggregating pollutant historical concentrations data on an annual and monthly time scale. For each pollutant compound five most polluted cities by average for period 2013-2021 and by maximum annual values have been selected using written python script. Bar plots used for visualizing average annual pollution levels and line plots are used for monthly average levels. Combining these results with the common facts about regions are expected to give a better understanding of a pollution situation in the corresponding region. It is known that the regions of Czech Republic such as Moravia-Silesia are more polluted due to the high concentration of industrial production there and the coal mining. It is expected to find a more descriptive correlations for cities in this region. Most of the researches on air quality aspects are focused on the Ostrava city. According to them concentration there changes seasonally for corresponding gases [27].

Further literature review had been made on each pollutant, to have a facile understanding on a possible damage caused by the high concentrations. For the end user of the web site, results on this review are provided in a form of short descriptions.

5.1 CO

5.1.1 Sources and dangers

Carbon monoxide(CO) is a common pollutant for inhibited areas. It's present in both outdoor and indoor air pollution. CO[1] is produced from the methane and non-methane hydrocarbon oxidation, fuel combustion and vegetation burning [28]. Sources of CO are vehicles that burn fossil fuels in the outdoor and gas stoves or heaters indoor, and a tobacco smoke present in both. The CO atmospheric lifetime ranges from a few week to a few month, related to the season and location [28].

High CO concentration is common for urban and industrial areas. Its values can be higher in winter season due to the intense house heating. The possible major source of CO in summer is the forest fires, due to a more active photochemistry, as the solar radiation results in a high ON levels [28].

CO has a great impact on a human health, since it can cross the placenta to gain access to the fetal circulation and the developing brain [29]. Since CO is colorless, odorless and non-irritating its toxicity often could be underappreciated and misdiagnosed [29].

[1]Carbon monoxide

5.1.2 Historical data

Historical data analysis of CO showed that the highest concentrations for the observing time 2013-2021 (5.1) of this gas are in Ostrava city (5.1). City and region around it are familiar for the coal mining and high industrial level, which polluting an ambient air with the CO and other pollutants intensively.

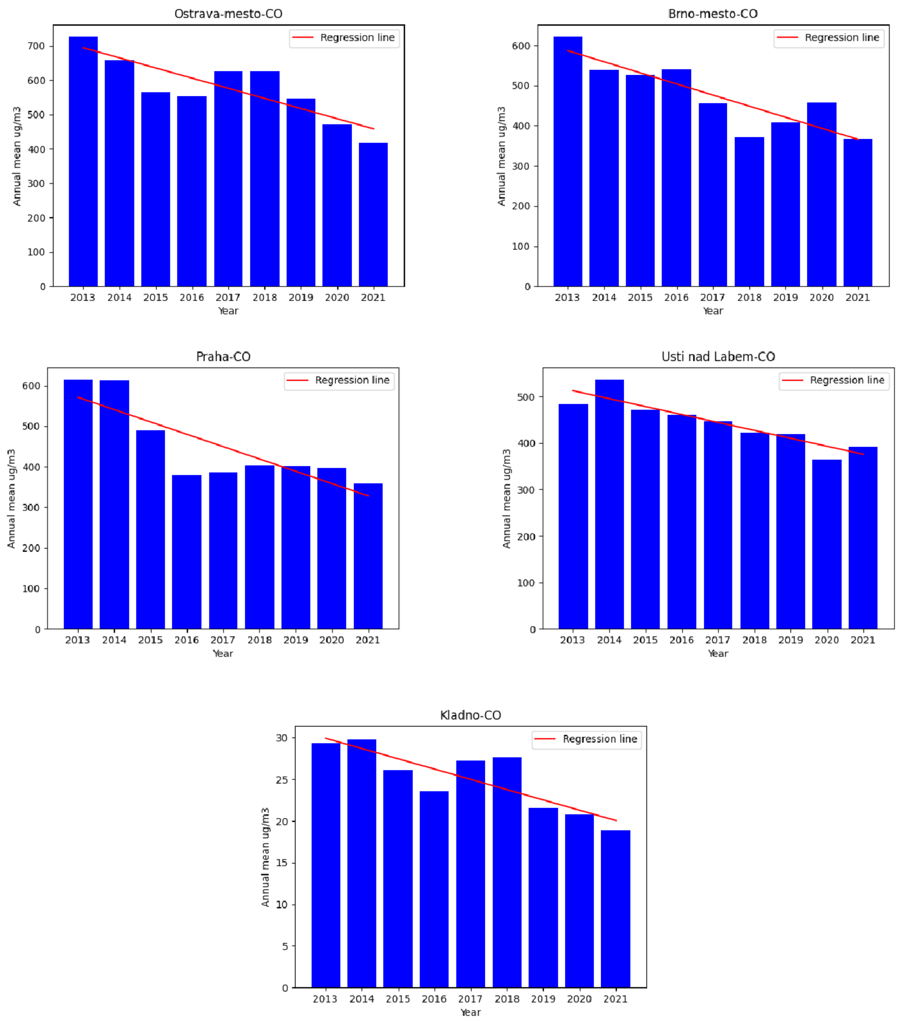

Figure 5.1: CO annual average for period 2013-2021

Rest of the cities in the table (5.2) having a relatively high population except Kladno, which is located close to the capital city- Prague. High population results in a higher traffic pollution and a pollution from heating or cooking. CO concentration

1	Ostrava-mesto	576.00 ug/m3
2	Brno-mesto	475.92 ug/m3
3	Praha	449.54 ug/m3
4	Usti nad Labem	444.39 ug/m3
5	Kladno	387.79 ug/m3

Table 5.1: CO - average for the period 2013-2021

in these regions is normal, however in order to reduce a negative impact of a CO pollution, more controlling measures on the transport emissions and indoor pollution might be taken. The positive decreasing annual trend is seen on the bar chart (5.1) for all cities from the list.

Three out of five positions for maximum annual values of CO are for the Ostrava other two are for the Brno (5.2). Monthly analysis of the corresponding years shows high CO concentration at the beginning of the year, however the monthly mean concentrations have either a stable or decreasing trend (5.2). In years 2013-2014 average monthly values barely exceeded a concentration edge of 1000 ug/m3. In year 2017 CO concentration was lower than in 2013, but still got close to the values of 1000 ug/m3, which is relatively high value for the average monthly concentration (5.2).

1	Ostrava-mesto	2013	1088.26 ug/m3
2	Brno-mesto	2013	933.75 ug/m3
3	Ostrava-mesto	2017	924.15 ug/m3
4	Ostrava-mesto	2014	871.82 ug/m3
5	Brno-mesto	2017	829.71 ug/m3

Table 5.2: CO - maximum annual values

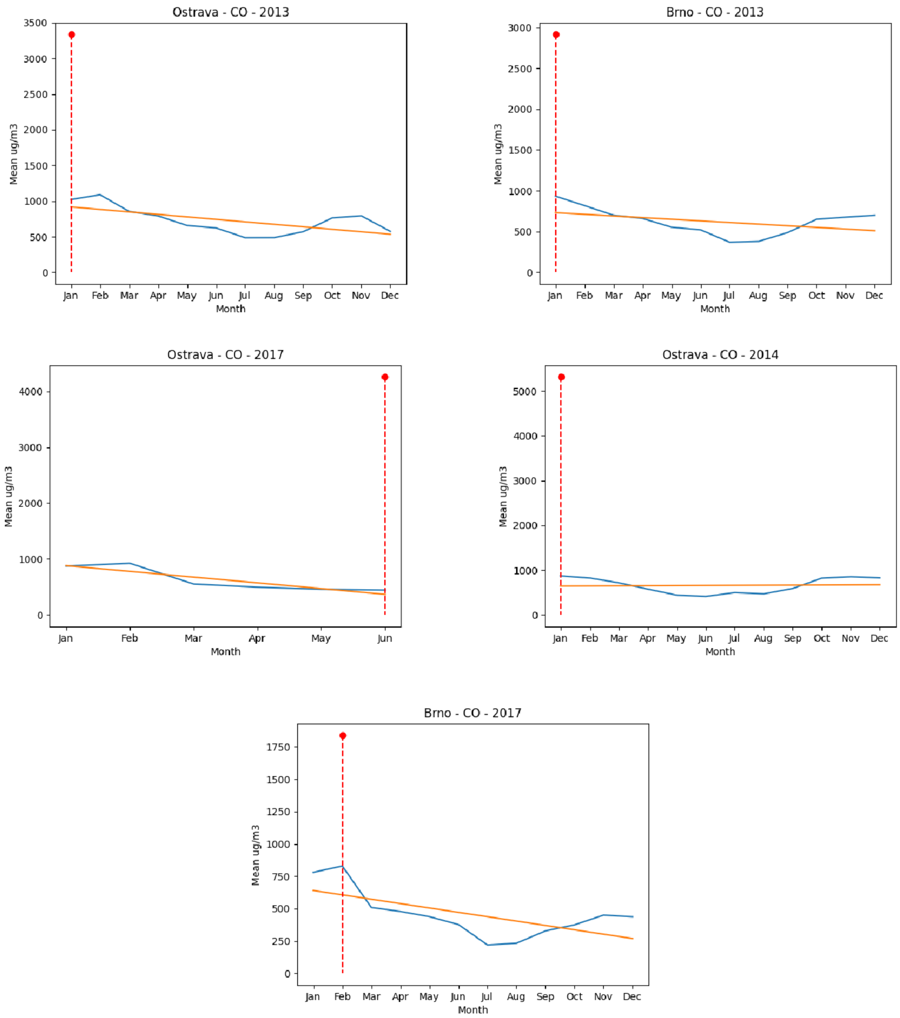

Figure 5.2: CO monthly average

5.1.3 Live data

From the resulting values on the final web site, the CO high pollution presence can not be concluded in any of the cities. This is both due to the complexity of observing CO on a short time scale and to the good control of the CO pollution across Czech Republic (5.3).

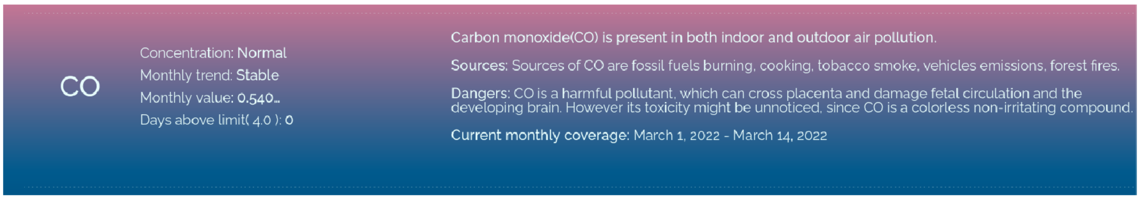

Figure 5.3: Live data summary from the website on the CO for Ostrava in March 2022

5.2 NOx and NO2

5.2.1 Sources and dangers

Nitrogen dioxide(NO2) is a pollutant formed during the combustion of a fossil fuel through the oxidation of atmospheric nitrogen and of nitrogen form fuels such as coal and oil [30]. Due to its short lifetime in the atmosphere NO2[2] is one of the easiest to measure pollutants. The high concentration of NO2 is common for areas with a high road traffic.

NO2 and the produced nitrogen dioxide ion(NO2-) could cause many negative effects on the human organism, such as oxidative stress, injury of the cell membrane and result in various injuries, carcinogenesis and aging [31].

5.2.2 Historical data

NO2 historical data correlates with the population density in Czech Republic. Cities from the summary list are in the 10 most populated cities in Czech republic except Most. Such values are expected for the cities with a high road traffic (5.3). NO2 annual trend for the cities from the list is similarly decreasing (5.4). For the all observed cities for a period of 2013-2021, the annual concentration of NO2 exceeded average annual value of 40 ug/m3 only 4 times (5.3). There was a noticeable decrease in years 2020-2021 connected with the COVID-19 pandemic lockdown (5.4).

However, improvement of a public transport and controlling vehicles emissions would greatly contribute to the overall air pollution in Czech republic. NO2 is yet most noticeable and under-controlled pollution in the country. This is an important pollutant to control, since the possible health negative consequences are serious. According to the recent local news, around 13 electric public buses have been launched in the Prague city. This is a great and important step in reducing vehicles pollution. Such approach might decrease both NO2 and CO emissions.

[2] Nitrogen dioxide

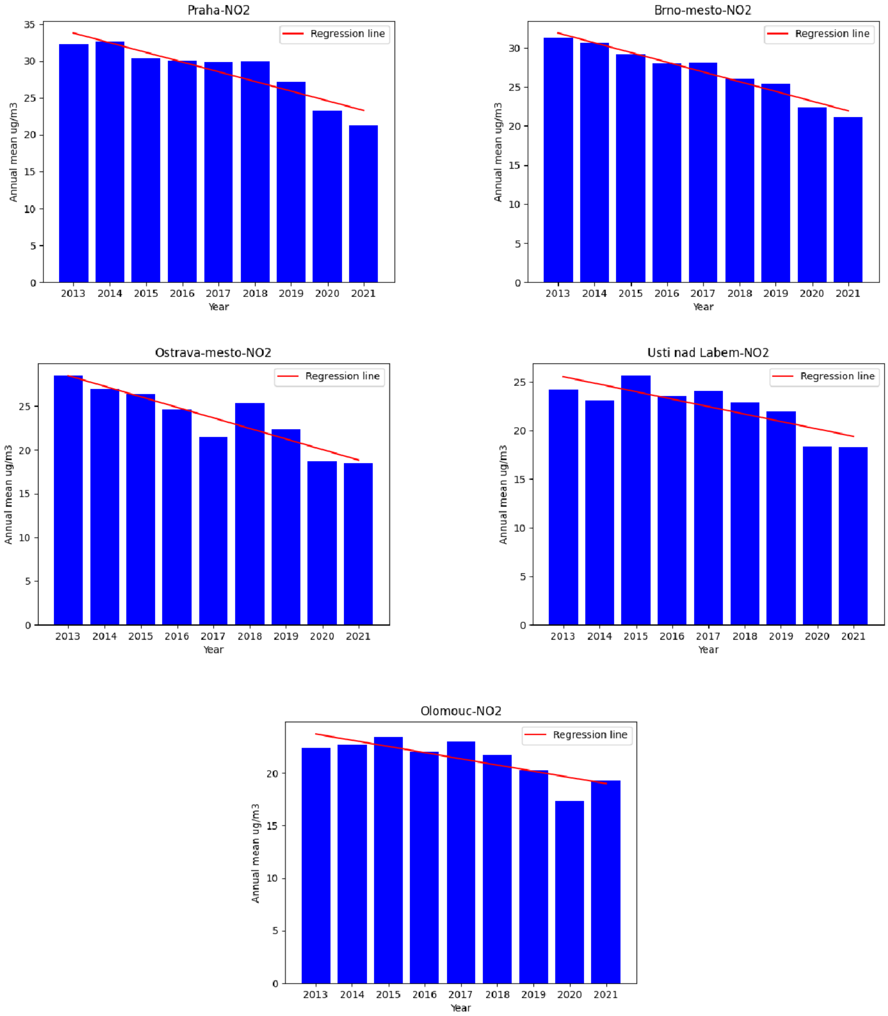

Figure 5.4: NO2 annual average for the period 2013-2021

According to the analysis of the maximum annual values, Prague capital city is struggling most from the NO2 pollution (5.4). It is the most populated city in Czech Republic, which leads to the high traffic in a city. Monthly trend for the cities from the table is mainly stable and not decreasing, however due to the weather conditions, values of NO2 are lower in Spring-Summer seasons and higher in Autumn-Winter (5.5).

1	Praha	28.53 ug/m3
2	Brno-mesto	26.91 ug/m3
3	Ostrava-mesto	23.66 ug/m3
4	Usti nad Labem	22.47 ug/m3
5	Olomouc	21.35 ug/m3

Table 5.3: NO2 - average all observing time

1	Praha	2014	42.98 ug/m3
2	Olomouc	2017	42.37 ug/m3
3	Praha	2017	42.10 ug/m3
4	Brno-mesto	2017	41.43 ug/m3
5	Usti nad Labem	2017	38.82 ug/m3

Table 5.4: NO2 - maximum annual values

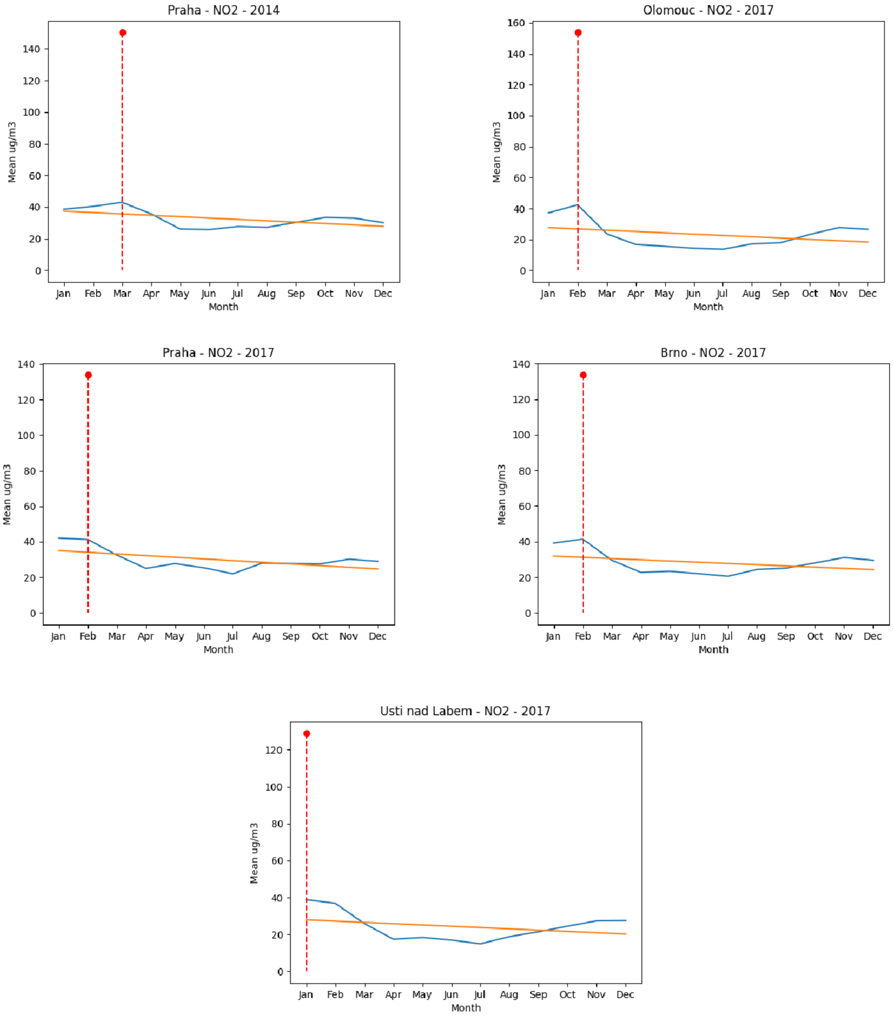

Figure 5.5: NO2 monthly average

5.2.3 Live data

NO2 levels are easily observed and estimated in a near live time. On the web site it is seen, that the NO2 values more frequently exceed the WHO guideline values than other pollutants. Since the largest source of NO2 is vehicles, exceeding concentrations are expected for a larger cities with high road traffic (5.6).

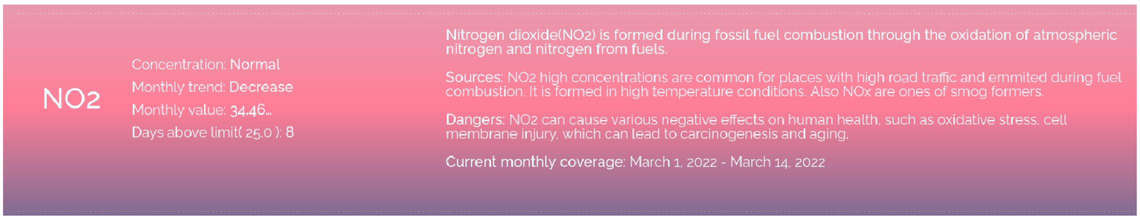

Figure 5.6: Live data summary from the website on the NO2 for Prague in March 2022

5.3 SO2

5.3.1 Sources and dangers

Sources of the SO2[3] are: burning fossil fuels containing sulfur, industries and power plants, production extracting metals from ore. Natural sources of the SO2 such as volcanoes are not common for the Czech Republic, only a side pollution from other European countries is possible. SO2 and its oxides (SOx) can harm human and vegetation health in both low and high concentrations. People with a respiratory deceases and children are more sensitive to the SO2 pollution. SOx in the atmosphere contribute to a particulate matter concentration increasing that pollution as well.

5.3.2 Historical data

Processing SO2 data resulted in the following table (5.5). SO2 pollution is common for the industrial areas especially with metal production. Ostrava and Karvina are cities form an industrial zone near the borders of the Czech Republic and Poland. These cities commonly emitting a lot of SO2 into the air which might affect people's physical and mental health, since SO2 is one of the heaviest air pollutants.

The industrial area is continued on the Poland side, from where also high pollutant concentrations may occur. The pollution might get worse, due to the relief elevation on the South-West of Ostrava. It might stop the gas from being blown away by wind and SO2 will affect the area inside for a longer time.

Average annual values for the Karvina city has a decreasing trend (5.7), however they are relatively high, the mean value is close to 10 ug/m3 (5.5). Usti nad Labem and Kladno cities have a stable non-decreasing SO2 annual trend, which can be a signal to create a better SO2 pollution control there. There is a peak of SO2 level in the Ostrava city in year 2018 and stabilising after it with around twice lower concentrations (5.7). Plzen, Kladno and Usti nad Labem cities had their SO2 annual

[3]Sulfur dioxide

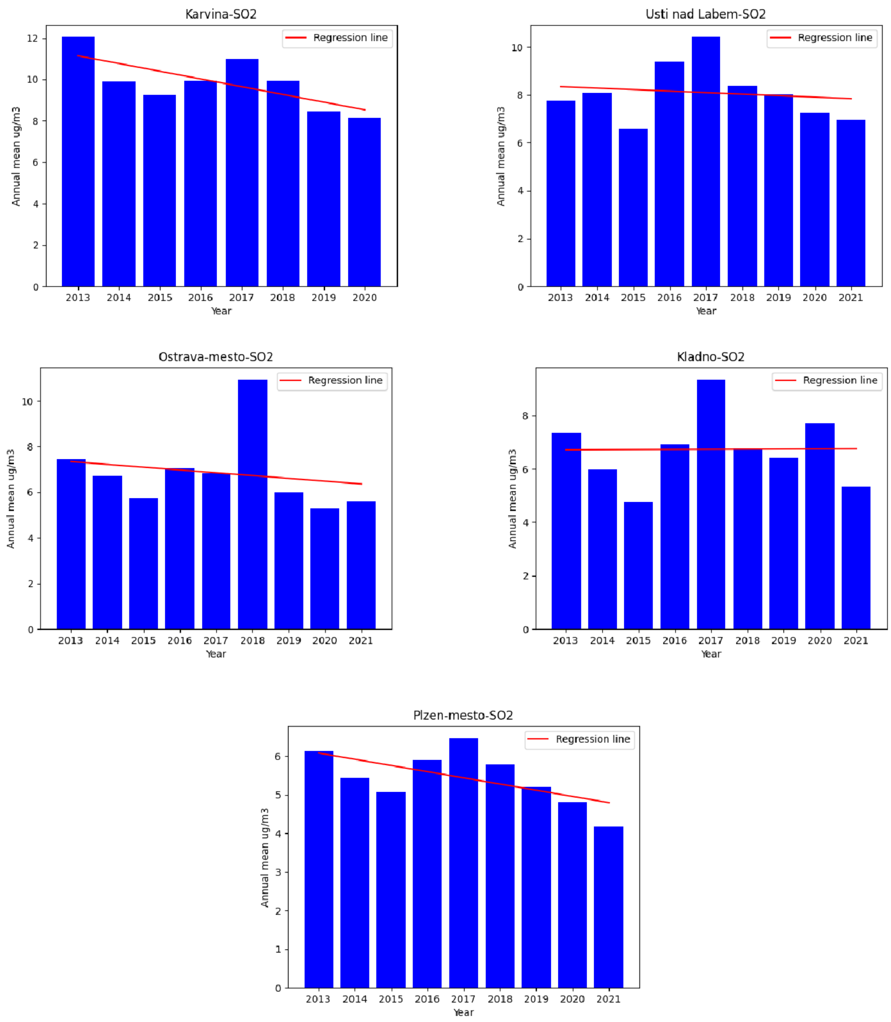

Figure 5.7: SO2 annual average for the period 2013-2021

1	Karvina	9.83 ug/m3
2	Usti nad Labem	8.09 ug/m3
3	Ostrava-mesto	6.85 ug/m3
4	Kladno	6.73 ug/m3
5	Plzen-mesto	5.44 ug/m3

Table 5.5: SO2 - average all observing time

1	Karvina	2017	28.12 ug/m3
2	Karvina	2013	27.23 ug/m3
3	Usti nad Labem	2017	20.62 ug/m3
4	Karvina	2014	19.99 ug/m3
5	Ostrava-mesto	2018	18.65 ug/m3

Table 5.6: SO2 - maximum annual values

level peak in year 2017 (5.7).

Highest maximum annual values for the five pairs city-year resulted in the following table (5.6). Karvina appears in a table three times with the highest values in years 2017, 2013 and 2014 (5.8). From the monthly average SO2 levels plot it is seen, that the highest values are common for the cold months, mostly in January and February (5.8). This is a very important indicator on an ambient air pollution, since the SO2 in Czech Republic has mostly anthropogenic origin, then it can be limited and controlled to prevent serious negative consequences.

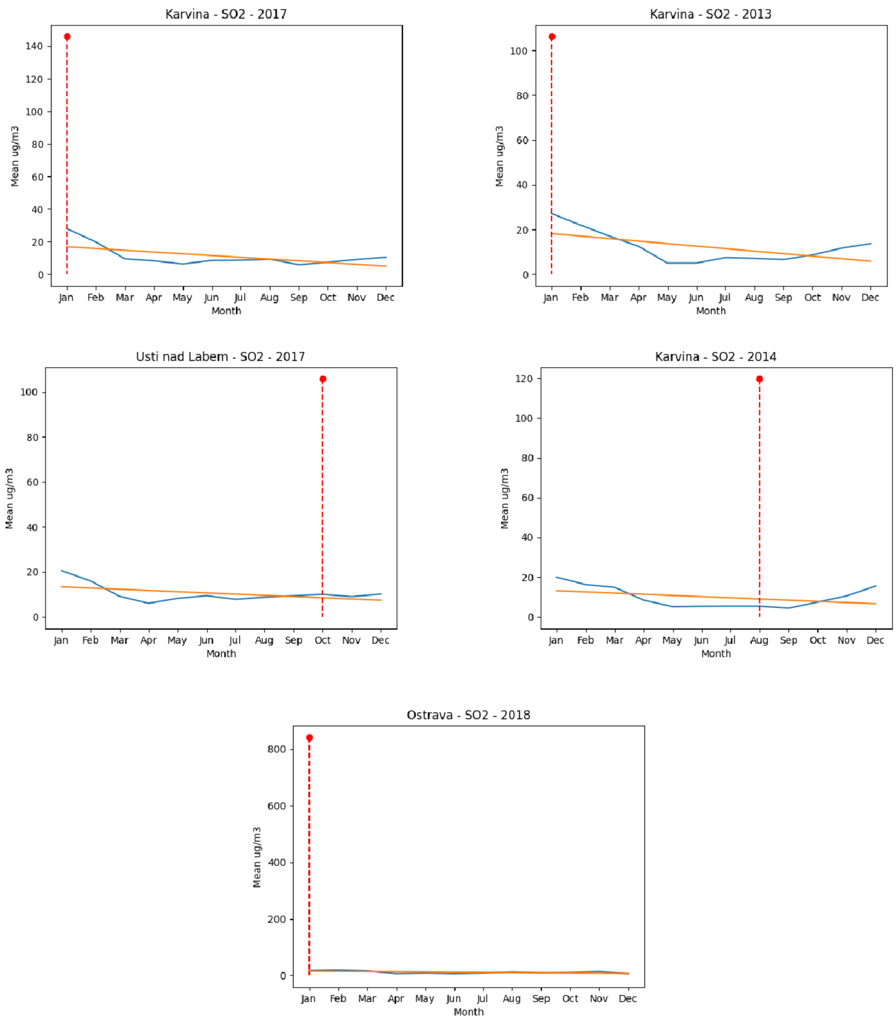

Figure 5.8: SO2 monthly average

5.3.3 Live data

SO2 is well controlled in the Czech Republic and its pollution is hardly observed in a short time scale. However, it still must be monitored in the industrial areas to raise awareness of people leaving nearby in case of a sudden increase due to accidents or weather conditions (5.9).

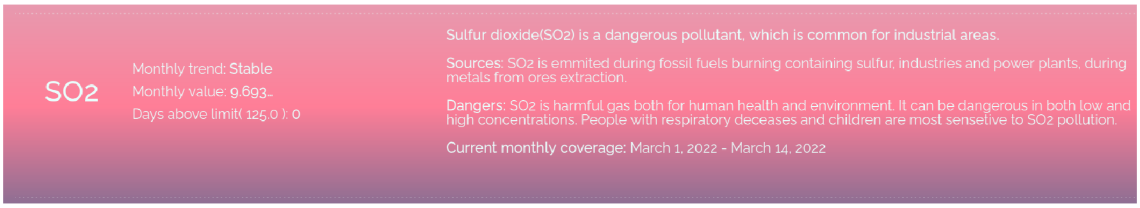

Figure 5.9: Live data summary from the website on the SO2 for Karvina in March 2022

5.4 O3

5.4.1 Sources and dangers

Ozone is a gas formed from the 3 oxygen atoms. It's effect might be both positive and negative. Stratospheric O3, which is far above ground, forms a protective layer, preventing harmful effects from the sun UV radiation. This layer has different thickness depending on the location and sometimes it is reduced due to the anthropogenic impact. Very intense reduction of the Ozone layer thickness is called Ozone hole. However, intrusion of the stratospheric ozone has an impact on the ambient O3, especially during spring [32]. Ground level O3 is quite harmful and dangerous for the environment. It is formed during chemical reactions of the nitrogen oxides and volatile organic compounds in the sunlight conditions. Higher emissions of the other known air pollutants could lead to the high O3 pollution, especially during the sunny weather conditions. That makes it hard to control an ambient O3 pollution levels. O3 is one of the main components of the smog. Its direct negative impact on a human health is yet not proven well, however it leads to the respiratory deceases and breathing issues in a form of smog. Children and people spending more time outdoor in conditions of the high O3 concentrations are affected the most [32].

5.4.2 Historical data

Analysis of O3 historical data brought the following results (5.7). For the average values in years 2013-2021, Hradec Kralove city resulted in the highest average concentration of 58.12 ug/m3 (5.7). Possible O3 pollution problematic can be present in cities from the table, since the same cities have both the highest average annual values for observing time (5.7) and the highest maximum annual values (5.8). Common thing for these cities is a northern geographic location and relief elevations near them. Probable accumulation of the different air pollutants due to the winds can lead to the higher O3 values in an ambient air.

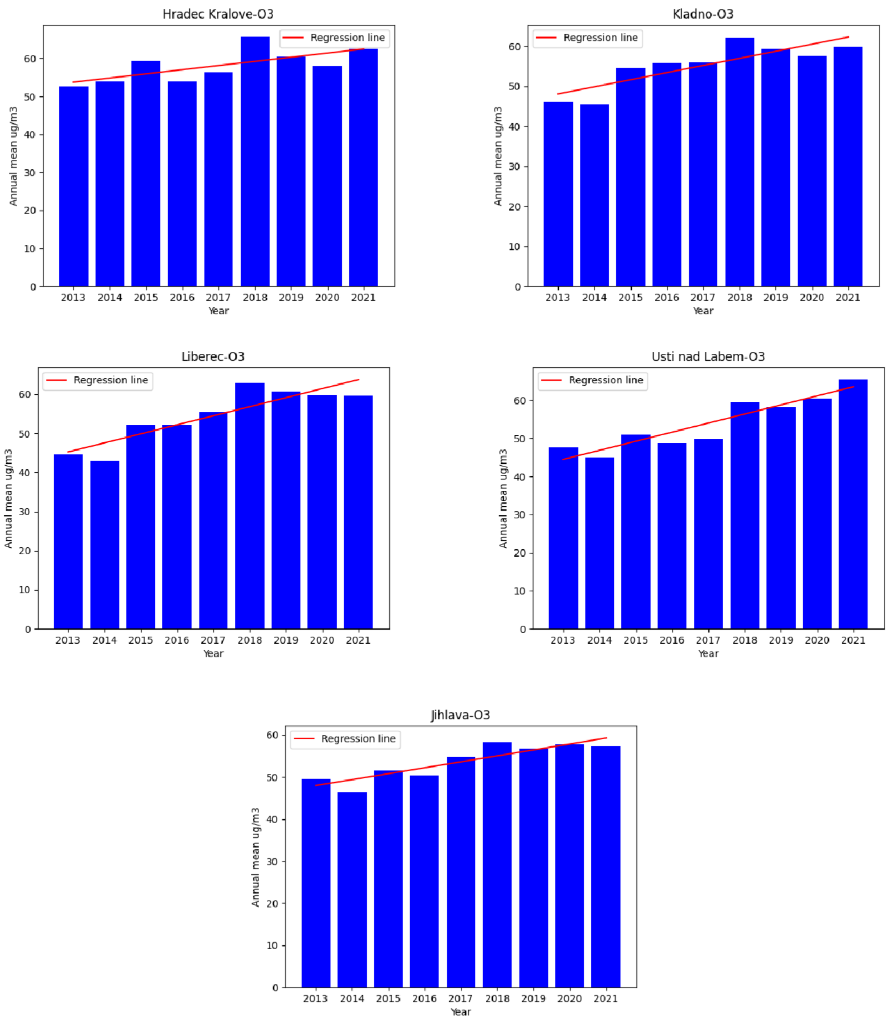

Figure 5.10: O3 annual average for the period 2013-2021

1	Hradec Kralove	58.12 ug/m3
2	Kladno	55.22 ug/m3
3	Liberec	54.50 ug/m3
4	Usti nad Labem	54.02 ug/m3
5	Jihlava	53.61 ug/m3

Table 5.7: O3 - average all observing time

1	Usti nad Labem	2019	100.70 ug/m3
2	Kladno	2018	95.12 ug/m3
3	Liberec	2019	92.37 ug/m3
4	Usti nad Labem	2018	92.27 ug/m3
5	Hradec Kralove	2015	92.18 ug/m3

Table 5.8: O3 - maximum annual values

From the bar chart is seen that for all five cities from the list the annual trend is increasing (5.10). However, resulted values are not close to the harmful concentrations and expected to be reduced along with reducing other air pollutants presence.

Analysis of the maximum annual O3 values by year and city resulted in the following table (5.8). The cities are the same as in the previous table. It is seen that the peak of O3 is happening continuously and reaches the maximum in Spring-Summer season, when there is more sunlight and the sunny day lasts longer (5.11).

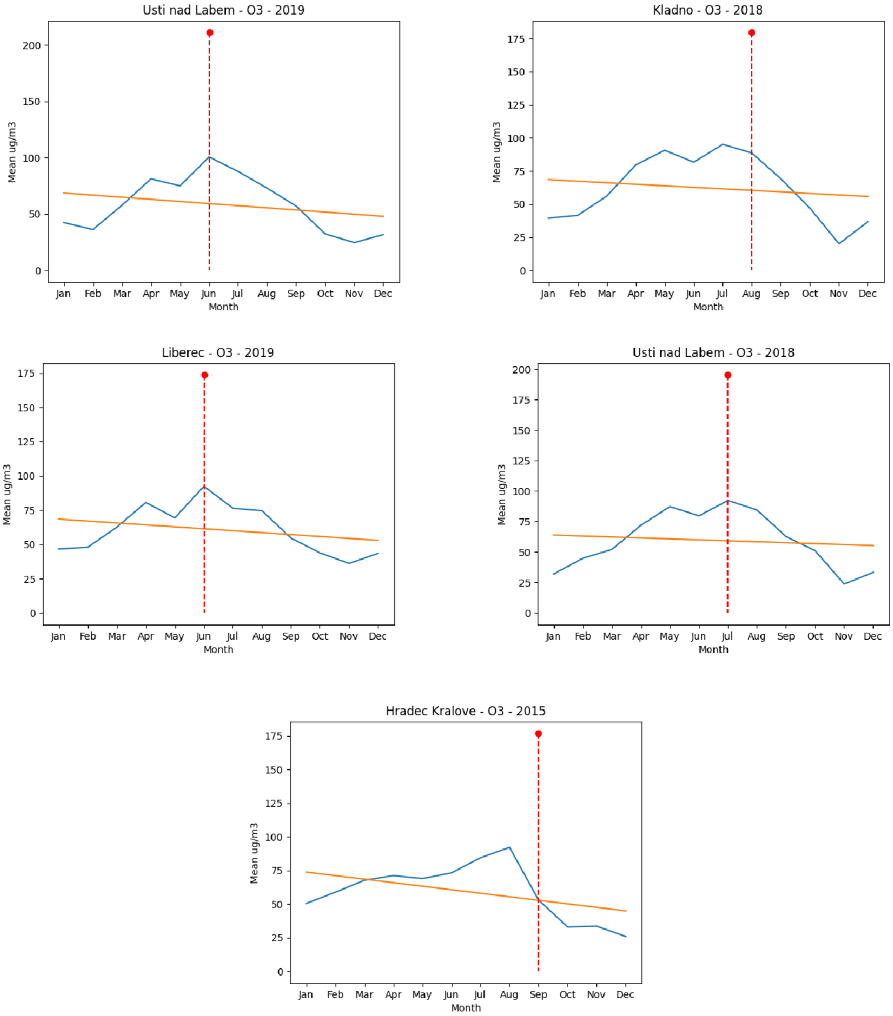

Figure 5.11: O3 monthly average

5.4.3 Live data

O3 levels in a near live time data does not exceed its limits. But must be monitored, since it may be an indicator of the increasing SO2 and NO2 emission, which might not be measured in a corresponding area (5.12).

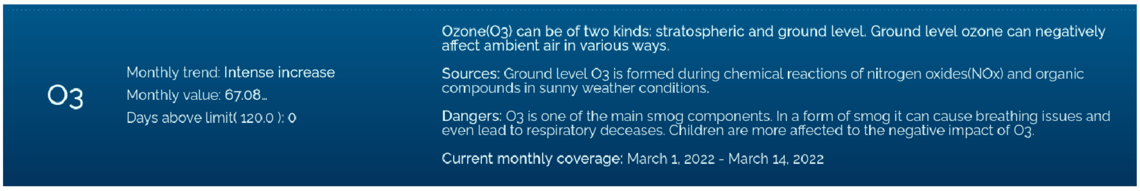

Figure 5.12: Live data summary from the website on the O3 for Hradec Kralove in March 2022

5.5 PM 2.5 and PM 10

5.5.1 Sources and dangers

Particulate matter 2.5 are fine inhalable particles with diameter of a 2.5 micrometers and smaller, PM 10 are particles with diameter of a 10 micrometers and smaller. PM's are of different shapes and made up of different chemicals. There are various anthropegenic sources of the PM 2.5 and PM 10 such as: construction sites, vehicles, fires, power plants and industries(e.g. coal mining). Most of them forms in the atmosphere as a result of the other pollutant chemicals reaction such as SO2 or NOx. Such small particles are easily inhalable and can get deep into lungs and even bloodstream. It can cause the great health risks and even death in case of the extremely high long term concentrations.

5.5.2 Historical data

It is expected to find the highest PM 2.5 (as well as PM 10) values for the most industrialized areas - Ostrava city and the neighbour regions. Averaging annual values proves such assumption (5.9 and 5.10), Karvina, Ostrava, Olomouc and Zlin are located in the same Eastern part of the Czech Republic. However, there is a Most city in the table, which is located in the North-West part of the Czech Republic and absolutely not impacted by the Ostrava industrial zone. Most is an another industrial city, which has the biggest brown coal mining and largest transportation node of that coal in the county, and also chemical production and electrometallurgy. Attention must be paid to the reducing particulate matter pollution levels in the industrial zones, since it strongly affects people living nearby.

Average annual values of PM 2.5 for the Most city show, that since year 2014 values decreased and after a leap of concentration in 2018 continued to decrease (5.13). Same decrease in an annual scale keeps for the Ostrava with a decrease after year 2018 (5.13). PM 10 concentrations correlate with the PM 2.5 values, averaging

1	Karvina	25.03 ug/m3
2	Ostrava-mesto	24.27 ug/m3
3	Olomouc	21.43 ug/m3
4	Most	18.56 ug/m3
5	Zlin	18.55 ug/m3

Table 5.9: PM 2.5 - average all observing time

1	Karvina	34.95 ug/m3
2	Ostrava-mesto	31.54 ug/m3
3	Olomouc	28.35 ug/m3
4	Most	26.93 ug/m3
5	Kladno	24.99 ug/m3

Table 5.10: PM 10 - average all observing time

values for the observing time 2013-2021, shows that almost each city with the high average PM 2.5 concentration also has proportionally high PM 10 concentration (5.14). Unexpectedly Kladno city appeared in the list of the high average PM 10 levels for observing time on a 5th position. In 21th century there were not any heavy industrial productions or coal mining. Such high concentration in a city closely located to the Prague require a further researching to find a cause of that pollution.

According to the analysis of the maximum annual values for corresponding years: Karvina, Ostrava and Olomouc are the most problematic places in Czech Republic in case of the PM air pollution (5.11 and 5.12). High PM2.5 and PM10 values appear in these cities in an Autumn-Winter seasons and decreasing in a Spring-Summer (5.15 and 5.16). The monthly trend for the most polluted pairs city-year is mainly stable. However the annual trend is decreasing, which is a positive sign on reducing the PM ambient air pollution.

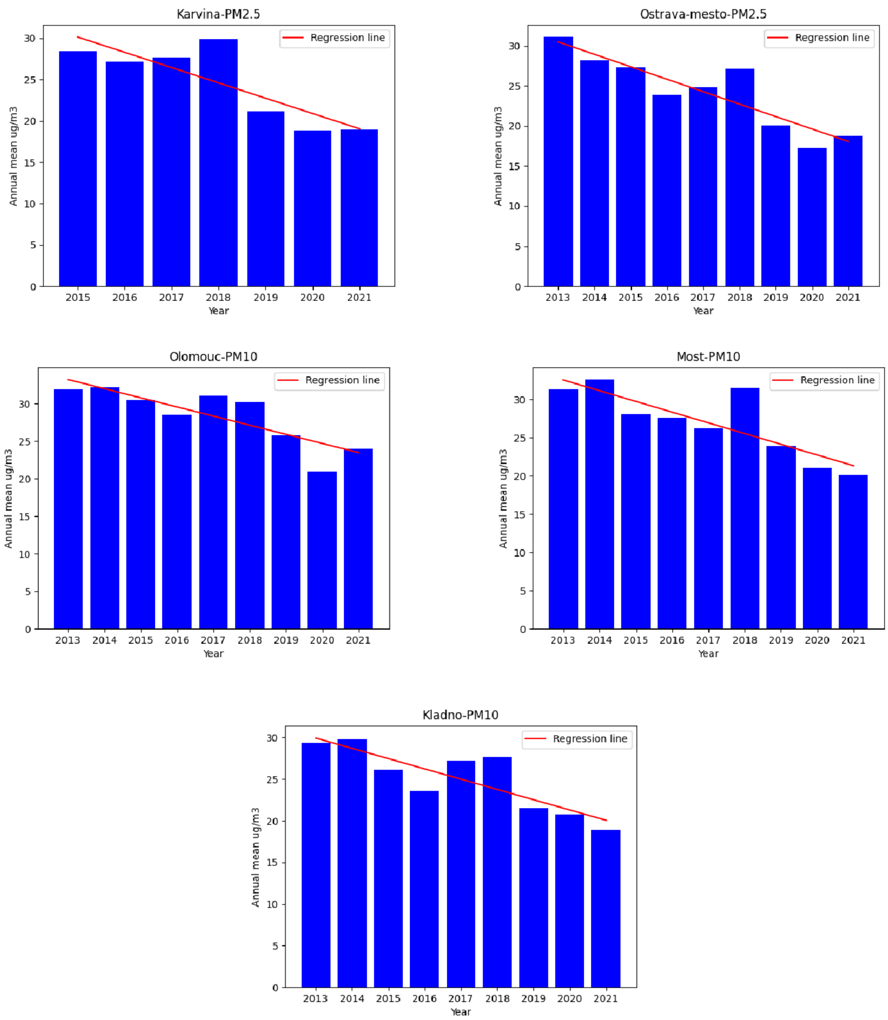

Figure 5.13: PM 2.5 annual average for the period 2013-2021

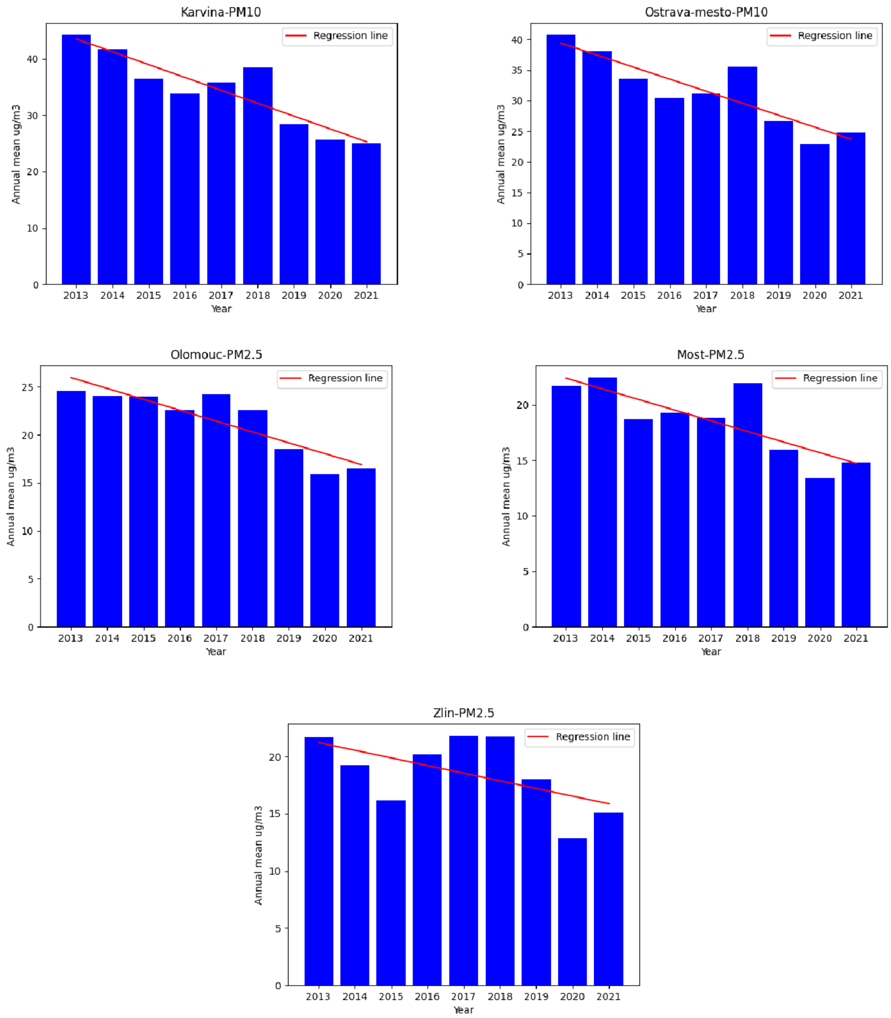

Figure 5.14: PM 10 annual average for the period 2013-2021

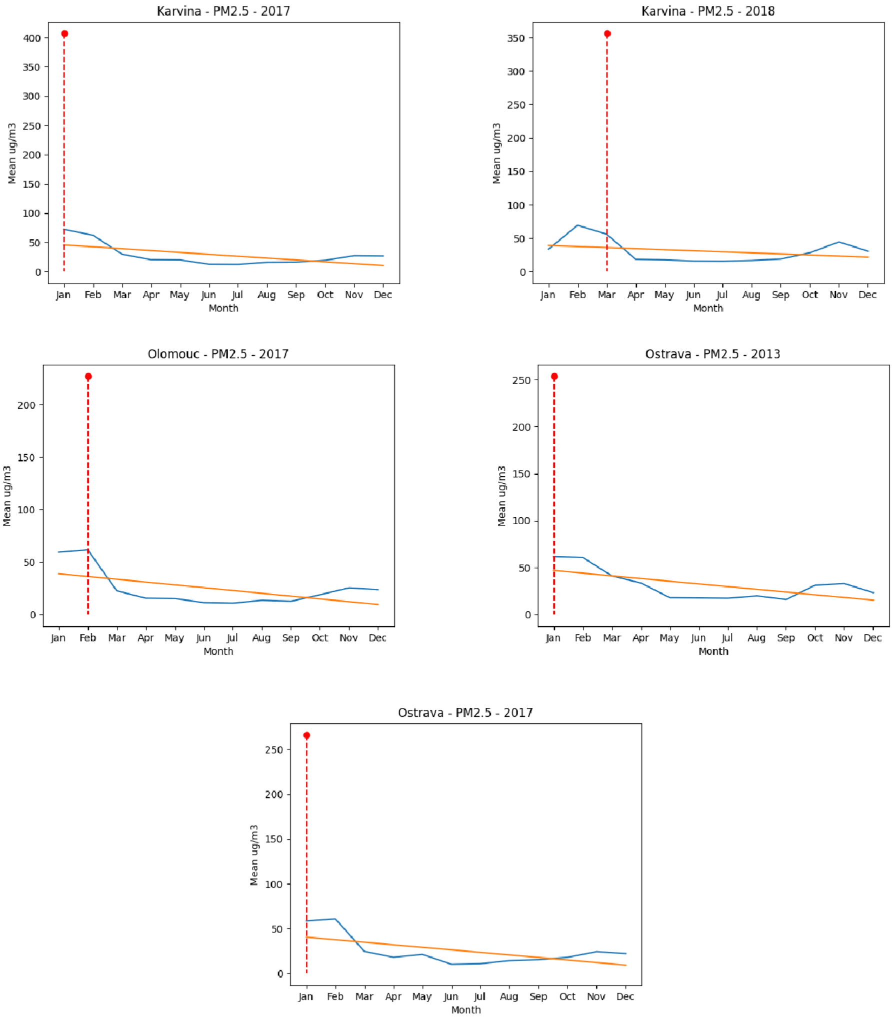

Figure 5.15: PM 2.5 monthly average

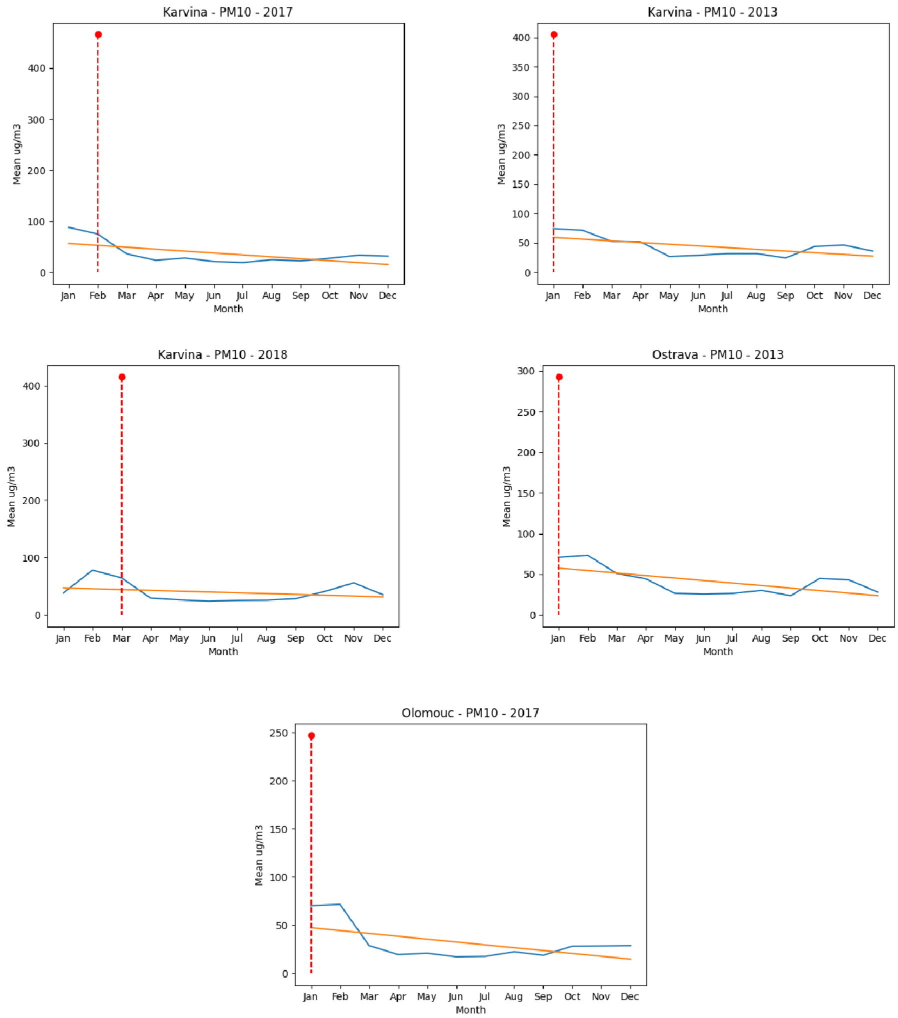

Figure 5.16: PM 10 monthly average

1	Karvina	2017	72.25 ug/m3
2	Karvina	2018	69.30 ug/m3
3	Olomouc	2017	61.71 ug/m3
4	Ostrava-mesto	2013	61.57 ug/m3
5	Ostrava-mesto	2017	60.93 ug/m3

Table 5.11: PM 2.5 - maximum annual values

1	Karvina	2017	88.21 ug/m3
2	Karvina	2013	81.18 ug/m3
3	Karvina	2018	78.05 ug/m3
4	Ostrava-mesto	2013	73.54 ug/m3
5	Olomouc	2017	71.69 ug/m3

Table 5.12: PM 10 - maximum annual values

5.5.3 Live data

PM 2.5 and PM 10 values in a near live time are very descriptive. It is seen that there is still some problematic areas even in a short time scale (5.17). Except cities from the table (5.9 and 5.10), the daily PM2.5 and PM10 limits are exceeded for the other cities as well.

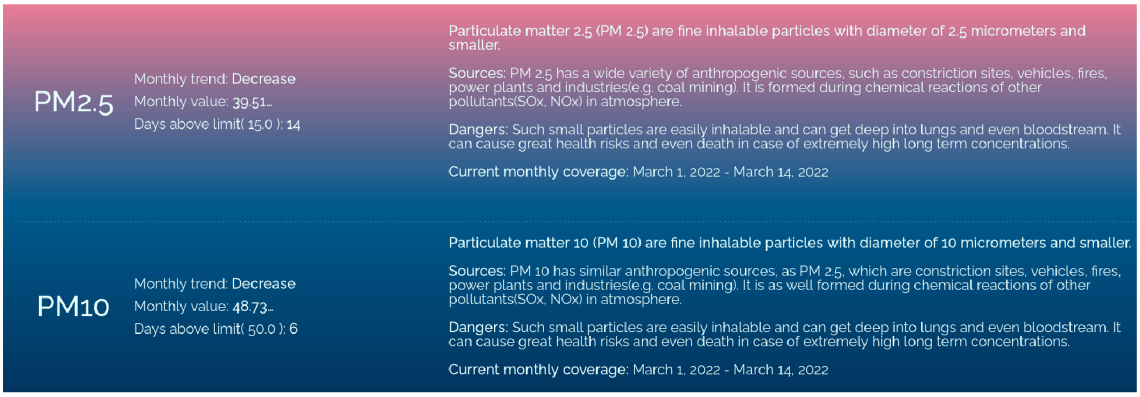

Figure 5.17: Live data summary from the website on the PM2.5 and PM10 for Karvina in March 2022

Chapter 6

Discussion

Based on the research there are more problematic regions except Moravia-Silesia. However the trend of the pollution in such regions is mostly decreasing and not exceeding dangerous values. Steps for the preventing high pollutant concentrations in an ambient air might be different. The approach of this work is to raise the awareness on this life aspect and make people more familiar with the air condition in their administrative region.

The air pollution comes from a lot of different sources, both natural and anthropogenic. In addition, some of the pollutants are interconnected (e.g. SO2 and NO2 with O3), which means that reducing concentration of one pollutant, the dependants compound concentrations will be decreased as well. My assumption is that the air pollution should be limited in a relation to any source and since the source is mostly anthropogenic, bringing the live region pollution problematic to the masses might have an effect on people paying more attention to the emission control.

The industrial production is a very important part of the country development. Emissions from the industrial zones had been reduced in the last decades in most parts of Europe. Nowadays these emissions are still decreasing, however it is definitely hard and long term task to reach the non-harmful level emissions in the industrial areas. Attracting local citizens attention to the air pollution topic might help reducing the emission levels with an individual impact, such as reducing vehicles usage, control of fossil fuel burning, smoking.

Further monitoring tools for the other pollution types might be developed. Then such tools might become as common as a weather forecasts, making people more aware and involved in the pollution topics. That way it can be easier to come up with the pollution control measures and agree on them.

Chapter 7

Conclusion

The task on creating an application for a near live time air pollution monitoring is completed. The data source had been selected according to its validity and spatial coverage of the Czech Republic. CHMI was the choice for the data sets, since it can provide valid and verified data with a great spatial coverage of the Czech Republic and most importantly in a near live time with an easy API access through the EEA service. For the pollutant analysis and possible problematic detection, historical data, measured by the CHMI and collected through the EEA service, had been processed and saved to the database. Resulting database and code organisation make the application mainly autonomic. Analysis of the historical data has been made on each pollutant for the five most polluted cities. Results on the NO2, PM2.5 and PM10 were most representative of an air pollution problems in Czech Republic. Big cities from the Moravia-Silesia region such as Ostrava, Karvina, Olomouc appeared more frequently in the list of the five most polluted cities in a variety of pollutants. The most populated places as Prague, Brno, Usti nad Labem and also Ostrava resulted in a higher historical pollution levels from the CO and NO2 pollutants. There was not observed any presence of a strong SO2 and O3 pollution from the historical data. However, the ozone appeared to be the only pollutant with an increasing annual trend. Further live data collection and processing had been made. The web application was made using the Django framework for the python programming language. The processed summarized data is saved to the database and then displayed on the web site. In a near live time the NO2, PM2.5 and PM10 pollution can be well observed in a variety of cities.

Bibliography

1. HUNOVÁ, Iva. Ambient Air Quality in the Czech Republic: Past and Present. *MDPI-Atmosphere*. 2020, roč. 11, č. 214. ISSN 2073-4433. Available from DOI: http://dx.doi.org/10.3390/atmos11020214.
2. SYMON, K. *Some aspects of air pollution - a public health hazard in Czechoslovakia: inter-regional symposium on criteria for air quality and methods of measurement, Geneva, 6-12 August 1963*. World Health Organization, 1963. Technical documents.
3. BOBAK, Martin; MARMOT, Michael. East-West mortality divide and its potential explanations: proposed research agenda. *BMJ*. 1996, roč. 312, č. 7028, pp. 421–425. ISSN 0959-8138. Available from DOI: 10.1136/bmj.312.7028.421.
4. RAO, S. Trivikrama; GALMARINI, Stefano; PUCKETT, Keith. Air Quality Model Evaluation International Initiative (AQMEII): Advancing the State of the Science in Regional Photochemical Modeling and Its Applications. *Bulletin of the American Meteorological Society*. 2011, roč. 92, č. 1, pp. 23–30. ISSN 00030007, ISSN 15200477. Available from DOI: 10.2307/26226796.
5. ŠAUER, Petr; MÁDR, Marek. ECONOMIC GROWTH AND AIR POLLUTION IN THE CZECH REPUBLIC: DECOUPLING CURVES. *International Journal of Economics and Finance Studies*. 2012, roč. 4, pp. 205–213.
6. PAVLÍNEK, Petr; PICKLES, John. Environmental Pasts/Environmental Futures in Post-Socialist Europe. *Environmental Politics*. 2004, roč. 13, č. 1, pp. 237–265. Available from DOI: 10.1080/09644010410001685227.

7. FIALA, Jaroslav; RIEDER, Mark; DVOŘÁKOVÁ, Miriam; LIVOROVÁ, Hana. Trends in the atmospheric and hydrological cycle of sulphur at catchments in the Czech Republic. *Atmospheric Environment*. 2001, roč. 35, S55–S62. ISSN 1352-2310. Available from DOI: `https://doi.org/10.1016/S1352-2310(00)00516-1`. Selected Papers Presented at the Venice Conference.

8. HŮNOVÁ, I.; ŠANTROCH, J.; OSTATNICKÁ, J. Ambient air quality and deposition trends at rural stations in the Czech Republic during 1993–2001. *Atmospheric Environment*. 2004, roč. 38, č. 6, pp. 887–898. ISSN 1352-2310. Available from DOI: `https://doi.org/10.1016/j.atmosenv.2003.10.032`.

9. KOPÁČEK, Jiří; VESELÝ, Josef. Sulfur and nitrogen emissions in the Czech Republic and Slovakia from 1850 till 2000. *Atmospheric Environment*. 2005, roč. 39, č. 12, pp. 2179–2188. ISSN 1352-2310. Available from DOI: `https://doi.org/10.1016/j.atmosenv.2005.01.002`.

10. MONITORING, European; PROGRAMME, Evaluation. *The Emissions Database*. 2019. Available also from: `https://www.ceip.at/webdab-emission-database`.

11. INSTITUTE, Czech Hydrometeorological. *Air Pollution in the Czech Republic in 2018; Graphical Yearbook (Czech/English)*. Czech Hydrometeorological Institute: Prague, Czech Republic, 2019, 2019. Report.

12. AGENCY, European Environmental. *Air Quality in Europe—2019 Report*. EEA: Copenhagen, Denmark, 2019, 2019. EEA Report No. 10/2019. ISSN 1977-8449. Available also from: `https://www.eea.europa.eu/publications/air-quality-in-europe-2019/air-quality-in-europe-2019/viewfile#pdfjs.action=download`.

13. SEINFELD, J.H.; PANDIS, S.N. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change. 2nd Edition*. New York: John Wiley & Sons, 2006.

14. SUCHARA, I.; SUCHAROVÁ, J. Distribution of 36 element deposition rates in a historic mining and smelting area as determined through fine scale biomonitoring techniques, Part II: Relative long-term accumulated atmospheric deposition levels. *Water Air Soil Pollution*. 2004, č. 153, pp. 229–252. Available from DOI: https://doi.org/10.1023/B:WATE.0000019915.70739.bf.

15. SUCHAROVA J.; Suchara, I. Distribution of 36 element deposition rates in a historic mining and smelting area as documented through fine-scale biomonitoring techniques. Part I: Relative and absolute current atmospheric deposition levels detected by moss analyses. *Water Air Soil Pollution*. 2004, č. 153, pp. 205–228. Available from DOI: https://doi.org/10.1023/B:WATE.0000019944.33209.83.

16. CONKOVA, M.; KUBIZNAKOVA, J. Lead isotope ratios in tree bark pockets: An indicator of past air pollution in the Czech Republic. *Science of The Total Environment*. 2008, roč. 404, č. 2, pp. 440–445. ISSN 0048-9697. Available from DOI: https://doi.org/10.1016/j.scitotenv.2008.04.025. BIOGEOCHEMISTRY OF FORESTED ECOSYSTEM - Selected papers from BIOGEOMON, the 5th International Symposium on Ecosystem Behaviour, held at the University of California, Santa Cruz, on June 25–30, 2006.

17. SUCHARA, I. Temporal and Spatial Changes in Spruce Bark Acidity at the Scale of the Czech Republic in the Last Two Decades, and the Current Abundance of Epiphytic Lichen Hypogymnia physodes. *Water Air Soil Pollution*. 2012, č. 223, pp. 1685–1697. Available from DOI: https://doi.org/10.1007/s11270-011-0975-6.

18. SUCHARA, Ivan; SUCHAROVÁ, Julie; HOLÁ, Marie. The influence of contrasting ambient SO2 concentrations in the Czech Republic in 1995 and in 2010 on the characteristics of spruce bark, used as an air quality indicator. *Ecological Indicators*. 2014, roč. 39, pp. 144–152. ISSN 1470-160X. Available from DOI: https://doi.org/10.1016/j.ecolind.2013.12.009.

19. SUCHARA, Ivan; SUCHAROVÁ, Julie; HOLÁ, Marie. Spatiotemporal Changes in Atmospheric Deposition Rates Across The Czech Republic Estimated in The Selected Biomonitoring Campaigns. Examples of Results Available For Landscape Ecology and Land Use Planning. *Journal of Landscape Ecology*. 2015, roč. 8, č. 2, pp. 10–28. Available from DOI: doi:10.1515/jlecol-2015-0002.

20. T., Horák J. & Hopan F.; Šyc M.& Machálek P.& Krpec K.& Ocelka T.& Tomšej. Estimation of selected pollutant emissions from solid-fuel combustion in small heating appliances. *Chem. Listy*. 2011, roč. 105, pp. 851–855. Available also from: http://www.chemicke-listy.cz/ojs3/index.php/chemicke-listy/article/view/1028.

21. BRANIŠ, M. Air quality of Prague: traffic as a main pollution source. *Environ Monit Assess*. 2009, roč. 156, pp. 377–390. Available from DOI: https://doi.org/10.1007/s10661-008-0491-3.

22. ONDRÁČEK, J.; SCHWARZ, J.; ŽDÍMAL, V.; ANDĚLOVÁ, L.; VODIČKA, P.; BÍZEK, V.; TSAI, C.-J.; CHEN, S.-C.; SMOLÍK, J. Contribution of the road traffic to air pollution in the Prague city (busy speedway and suburban crossroads). *Atmospheric Environment*. 2011, roč. 45, č. 29, pp. 5090–5100. ISSN 1352-2310. Available from DOI: https://doi.org/10.1016/j.atmosenv.2011.06.036.

23. HŮNOVÁ, Iva; KURFÜRST, Pavel; BALÁKOVÁ, Lea. Areas under high ozone and nitrogen loads are spatially disjunct in Czech forests. *Science of The Total Environment*. 2019, roč. 656, pp. 567–575. ISSN 0048-9697. Available from DOI: https://doi.org/10.1016/j.scitotenv.2018.11.371.

24. V., Tolasz R.& Miková T.& Valeriánová A.& Voženílek. *Climate Atlas of Czechia*. CHMI: Praha, Czech Republic, 2007.

25. *Global Administrative Unit Layers 2015, Second-Level Administrative Units*. FAO GAUL, 2015. Available also from: https://developers.google.com/earth-engine/datasets/catalog/FAO_GAUL_2015_level2?hl=en#description.

26. *Download service for E1a and E2a data*. European Environmental Agency, 2019. Available also from: https://discomap.eea.europa.eu/map/fme/AirQualityExport.htm.

27. BUZEK, F.; CEJKOVA, B.; HELLEBRANDOVA, L.; JACKOVA, I.; LOLLEK, V.; LNENICKOVA, Z.; MATOLAKOVA, R.; VESELOVSKY, F. Isotope composition of NH 3 , NO x and SO 2 air pollution in the Moravia-Silesian region, Czech Republic. *Atmospheric Pollution Research*. 2016, roč. 1, č. 1, pp. 1–12. ISSN 1309-1042. Available from DOI: http://dx.doi.org/10.1016/j.apr.2016.08.011.

28. CLERBAUX, Cathy; EDWARDS, David P.; DEETER, Merritt; EMMONS, Louisa; LAMARQUE, Jean-François; TIE, Xue Xi; MASSIE, Steve T.; GILLE, John. Carbon monoxide pollution from cities and urban areas observed by the Terra/MOPITT mission. *Geophysical Research Letters*. 2008, roč. 35, č. 3. ISSN 0094-8276. Available from DOI: https://doi.org/10.1029/2007GL032300.

29. LEVY, Richard J. Carbon monoxide pollution and neurodevelopment: A public health concern. *Neurotoxicology and Teratology*. 2015, roč. 49, pp. 31–40. ISSN 0892-0362. Available from DOI: https://doi.org/10.1016/j.ntt.2015.03.001.

30. GILBERT, Nicolas L; WOODHOUSE, Sandy; STIEB, David M; BROOK, Jeffrey R. Ambient nitrogen dioxide and distance from a major highway. *Science of The Total Environment*. 2003, roč. 312, pp. 43–46. ISSN 0048-9697. Available from DOI: https://doi.org/10.1016/S0048-9697(03)00228-6.

31. GRAZULEVICIENE, Regina; MAROZIENE, Ligita; DULSKIENE, Virginija; MALINAUSKIENE, Vilija; AZARAVICIENE, Ada; LAURINAVICIENE, Dovile; JANKAUSKIENE, Konstancija. Exposure to urban nitrogen dioxide pollution and the risk of myocardial infarction. *Scandinavian Journal of Work, Environment and Health*. 2004, roč. 30, č. 4, pp. 293–298. ISSN 0355-3140. Available also from: http://www.jstor.org/stable/40968791.

32. LIPPMANN, Morton. HEALTH EFFECTS OF OZONE A Critical Review. *Air and Waste Management Association*. 1989, roč. 39, č. 5, pp. 672–695. ISSN 0894-0630. Available from DOI: https://doi.org/10.1080/08940630.1989.10466554.

Appendices

A Database usage

```
import psycopg2
from dotenv import dotenv_values

env = dotenv_values()

# Database object
class LocalDB:
def __init__(self):
    # Parameters for PostgreSQL database connection
    self.params = {
        'database': env['DATABASE'],
        'host': env['HOST'],
        'user': env['USER'],
        'password': env['PASSWORD'],
        'port': env['PORT']
    }

# Method for executing a query
```

```
  without fetching data (e.g. INSERT, UPDATE)
def execute(self, query, data_tuple=()):
    # Connect to database
    conn = psycopg2.connect(**self.params)
    if conn:
        # On success create cursor for query
        cur = conn.cursor()
        # Execute query with passed values
        if len(data_tuple) > 0:
            cur.execute(query, data_tuple)
        else:
            cur.execute(query)
        # Close cursor and commit
        cur.close()
        conn.commit()


# Method for fetching data via query
def fetchall(self, query):
    # Connect to database
    conn = psycopg2.connect(**self.params)
    if conn:
        # On success create cursor
        cur = conn.cursor()
        # Execute query which must return data table
        cur.execute(query)
        # Save result to a variable
        res = cur.fetchall()
        # Close cursor and return resulted data
```

```
        cur.close()
        return res


# Method for removing duplicated rows from raw table
def remove_duplicates(self, source):
    conn = psycopg2.connect(**self.params)
    if conn:
    cur = conn.cursor()
    cur.execute(f'DELETE FROM public."raw_{source}"
                USING (SELECT id, row_number() '
                f'OVER (PARTITION BY date, value, city, variable
                ORDER BY id) AS rn FROM public."raw_{source}") x '
                f'WHERE public."raw_{source}".id = x.id
                AND x.rn > 1;')
        cur.close()
        print('Duplicates have been removed.')
```

B Main class object

```
# Main EEA object for collecting API live data
class EEA:
def __init__(self):
        # Near-live data API url
        self.url = "http://discomap.eea.europa.eu/map/fme/latest"
        # Initialize database object
        self.db = LocalDB()
```

```
# Method saving data from API in local storage (temporarily)
def get_data(self, variable):
        # Filename created from current processing variable(e.g. CZ_CO.csv)
        fileName = f"../data/eea_live/CZ_{variable}.csv"
        downloadFile = f'{self.url}/CZ_{variable}.csv'
        print(f'Downloading: {downloadFile}')
        # Downloading data from EEA API
        file = requests.get(downloadFile).content
        # Saving data to the local file
        output = open(fileName, 'wb')
        output.write(file)
        output.close()
        print('Saved locally as: %s ' % fileName)
        # Extracting data from .csv file using pandas and return it
        data = pd.read_csv(fileName)
        return data

# Method for finding measuring station location by coordinates
def find_city(self, row):
# Selecting from database list of regions
(collected from Google Earth Engine dataset covering whole country)
polygons = self.db.fetchall(f'SELECT id,city,st_asgeojson(polygon)
                              FROM public."regions";')

# For each region its id in database,
  name and geometry are extracted to variables
for polygon in polygons:
city_id = polygon[0]
```

```
city_name = polygon[1]
geojson_data = geojson.loads(polygon[2])

try:
p = Polygon([tuple(l) for l in geojson_data['coordinates'][0]])
except KeyError:
p = Polygon([tuple(l) for l in geojson_data['geometries'][1]
                                        ['coordinates'][0]])

# Creating geometry point from x and y coordinates of station
point = Point(row['samplingpoint_x'], row['samplingpoint_y'])
# Projection fro ESPG:4979 to ESPG:4326
project = pyproj.Transformer.from_proj(4979, 4326, always_xy=True)
# Project point
point_transform = transform(project.transform, point)
# Check if polygon contains point(=region contains station)
contains = p.contains(point_transform)
# On success save station properties and reference
on region in database

if contains:
self.db.execute(f'INSERT INTO public."stations"
            (station_name, station_id, city, source, point) '
                    f'VALUES (\'{row["station_name"]}\',
                          \'{row["station_code"]}\', '
                f'{city_id}, 9,
                st_geomfromtext(\'{point_transform}\', 4326))
print(f'Created new record in stations
      for {row["station_name"]}/{city_name}')
```

```
return city_id
```

C Main script run

```
# Script run for collecting source data and saving it
def run():
# Initialize database object
db = LocalDB()
# Initialize main EEA object
process = EEA()

# Select list of variables for EEA near-live data
variables = db.fetchall(f'SELECT id, variable_short
                                FROM public."variables" WHERE source_id=9;')

# For each variable download a dataset using get_data method
for variable in variables:
data = process.get_data(variable[1])
try:
# For each row in resulted dataset check
if there is existing station record in database
for i, row in data.iterrows():
check_station = db.fetchall(f'SELECT station_name,city
                                    FROM public."stations" '
                                  f'WHERE station_name=\'{row["station_name"]}\'
                                 AND station_id=\'{row["station_code"]}\'')

# In station is not in databaset,
```

```
its location found and saved using find_city method
if len(check_station) < 1:
    city = process.find_city(row)
else:
    city = check_station[0][1]

# Select date and concentration columns from dataset
date = row["value_datetime_begin"].split(' ')[0]
value = row["value_numeric"]

# Exclude NA values and insert result to the raw_eea table
if not pd.isna(value):
db.execute(f'INSERT INTO public."raw_eea" (date,value,city,variable)
           f'VALUES (\'{date}\', {value}, {city}, {variable[0]})')

except KeyError:
    continue

# After all data from API is collected daily calculation is launched
print('RUNNING DAILY CALCULATION')
system('python run_daily.py')



# Part for scheduling to run code every day at the same time
schedule.every().day.at("23:30").do(run)

while True:
schedule.run_pending()
print('Waiting till 23:30...')
```