OULUN AMMATTIKORKEAKOULU

Liubov Zelenaia

AIR POLLUTION FORECAST IN EUROPE

AIR POLLUTION FORECAST IN EUROPE

Liubov Zelenaia
Thesis
Spring 2023
Master's degree programme in Data Analytics and Project Management
Oulu University of Applied Sciences

ABSTRACT

Oulu University of Applied Sciences
Master’s degree programme in Data Analytics and Project Management

---

Author(s): Liubov Zelenaia
Title of the thesis: Air pollution forecast in Europe
Thesis examiner(s): Ilpo Virtanen
Term and year of thesis completion: Spring 2023 Pages: 94 + 2 appendices

---

According to the European Environment Agency air pollution is the greatest environmental health risk in Europe. Wherever the person lives, it is difficult to avoid. Air pollution may severely impact the environment and health. Despite recent improvements in the region's air quality the levels of air pollutants in Europe continue to exceed both the strictest World Health Organization recommendations and European Union requirements.

Several questions were considered during the thesis research such as how much excess pollutants there are in the air, what is the cause of such pollution, and how it will evolve in the future. The study focused on the European region. The main work was done by gathering needed data and predicting factors and pollutant levels in the next ten years based on retrospective data.

Thousands of stations around Europe monitor air quality following the concentration of various pollutants. The stations report data to several databases of different organizations yearly. The research was based on open-source data from 1969 to 2022, which allowed accurate and efficient detection of air quality anomalies. The data were pre-processed for better performance and used for the forecast with Python language tools such as LinearRegression and MLPRegressor of the scikit-learn package. The models were evaluated using the R^2 score for testing and training data. The thesis also provided a visual representation of the gathered and predicted data in graphs and tables.

The predicted most probable levels of six air pollutants were compared with the Year Average Common Air Quality Index (YACAQI) of these pollutants. The models of the predictions were accurate by more than 80% and stable in different scenarios which makes them reliable as a foundation for further research.

The thesis work found that most air pollutant levels are predicted to decrease in the next ten years, with only ozone levels going to increase. The main air pollution factors are GDP per capita and fossil fuel consumption such as oil, gas, and coal. Oil and coal consumption are predicted to decrease, while gas consumption will increase to the level of oil consumption. Despite that, the levels of pollutants will be still below the year average which will not affect citizens' health.

Keywords: air pollution, Europe, forecast, Python, LinearRegression, MLPRegressor, YACAQI

CONTENTS

1 INTRODUCTION

Air pollution is still a problem in many countries due to human activity, which leads to a worsening environment, health, and life expectancy (European Environment Agency 2022). Only 13 out of 131 countries and regions did not exceed the pollution level standards of the World Health Organization in 2022. Among these countries are Estonia, Finland, and Iceland. Every year, millions of people die from diseases related to air pollution. (IQAir 2023.) Understanding air pollution and its projection in the near future is essential to ensure safety.

The thesis focuses on air pollution in the European region based on the gathered data and forecasts its development for the next 10 years.

The thesis work covers questions such as:

- Why and how does air pollution take place?
- What are the main factors and main pollutants?
- How do pollutant levels affect Europe citizens' health? And what are the limits?
- What are the main causes of pollution in each European country?
- What will be the most probable level of main pollutants level in Europe and what will it be in different scenarios the level in these scenarios?
- Do the results exceed the Year Average Common Air Quality Index (YACAQI)?

The goals of the thesis are to conduct research on air pollution in Europe, explain the reasons behind it, and forecast the pollution levels using different scenarios. Air quality is an important factor in terms of attracting investment and the overall attractiveness of the country. It can be useful to know the pollution hazard level while planning a vacation to the European region or to get acquainted with the air pollution topic.

The main strength of the thesis is its ability to give an overview of the air pollution situation in Europe as it focuses on several pollutants and helps to introduce the fundamentals of this topic. The thesis work ends to pollution forecast which can be useful for further air pollution-related scientific studies.

1.1 The state of the atmosphere

Air pollution discussion begins with the structure of the atmosphere. The focus is on how parts of the troposphere or stratosphere are out of the norms. The sources of the pollution are often associated with smoke from industrial pipes, smog spread, dead trees followed by acid rain, and natural disasters. As a result, the polluted mass is released into the lowest layers of the atmosphere, the troposphere and stratosphere, where people and animals live. (Vallero 2014, 3.)

Four gases make up the atmosphere: argon (Ar), molecular oxygen, nitrogen (N_2 and O_2), and carbon dioxide (CO_2). Air pollution occurs in the atmosphere within 50 kilometers of the Earth's surface - troposphere and stratosphere. The temperature of the Earth's atmosphere changes with the altitude, as well as the density of its elements. The temperature profile of the Earth's atmosphere is presented in Figure 1. The Earth warms up because of the solar energy being absorbed by its surface and then re-emitted at longer infrared wavelengths. The air becomes less dense with altitude moving from the troposphere through the stratosphere, chemosphere, and ionosphere. (Vallero 2014, 3.)

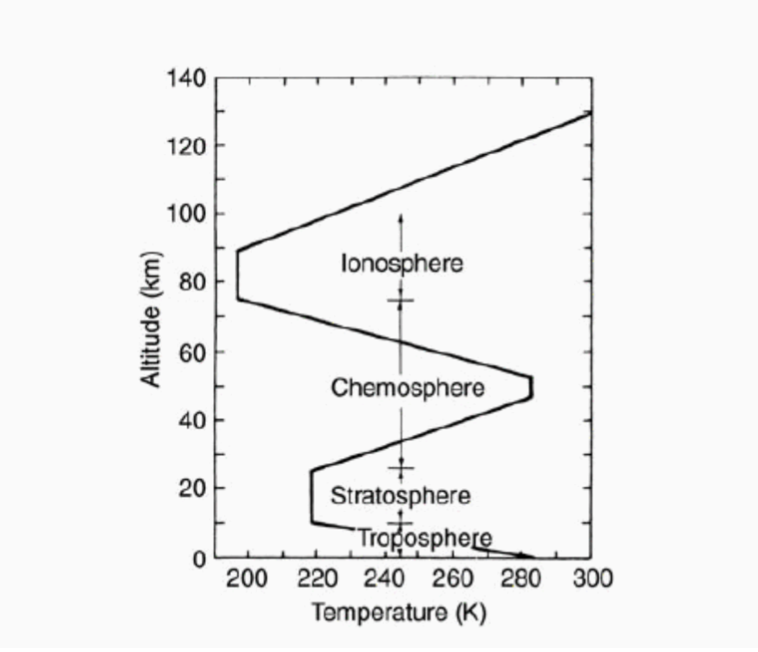


FIGURE 1. The temperature profile of the Earth's atmosphere.

Scientists studying air pollution are interested in the ionosphere and chemosphere because of how they absorb and release solar energy. It affects the quantity and spectral distribution of cosmic rays as well as solar energy that enters the stratosphere and troposphere. The stratosphere is of interest to scientists because of the global pollutants' transportation from the surface nuclear tests, volcanic eruptions, and solar energy dissipation. The troposphere is the area where people reside and is the focus of air pollutant science. The composition of gases in the air without pollution is illustrated in Table 1. The represented gaseous concentrations are mean values and do not exist in the atmosphere because of human activity but also because air is more complex than a mixture of gases. The air includes water vapor as well as other vapors, liquid, and solid matter. (Vallero 2014, 4.)

TABLE 1. Gaseous Composition of the Atmosphere (Dry, Unpolluted air).

	ppm (vol.)	$\mu g\ m^{-3}$
Nitrogen	780,000	8.95×10^{8}
Oxygen	209,400	2.74×10^{8}
Water	-	-
Argon	9300	1.52×10^{7}
Carbon dioxide	315	5.67×10^{5}
Neon	18	1.49×10^{4}
Helium	5.2	8.50×10^{2}
Methane	1.0-1.2	$6.56\text{-}7.87 \times 10^{2}$
Krypton	1.0	3.43×10^{3}
Nitrous oxide	0.5	9.00×10^{2}
Hydrogen	0.5	4.13×10^{1}
Xenon	0.08	4.29×10^{2}
Organic vapors	c. 0.02	-

1.2 The science of air pollution

Despite its size, the atmosphere has been and can be polluted. Science distinguishes between an unpolluted and a polluted atmosphere in various ways. The most useful thing to do is to compare the state of the atmosphere today with some baseline, such as the state of the atmosphere before industrialization in the nineteenth century. Another approach is to look at the principles by which pollutants enter, move, transform in the atmosphere, and are removed from the atmosphere. Based on this assessment, the degree of atmospheric pollution can be determined. (Vallero 2014, 46-47.)

The definition of air pollution depends on the context and the impact of a set of circumstances. The same chemical compounds come from a natural source such as volcanoes and cause the same harmful effects as anthropogenic sources. Therefore, the “unpolluted” air is just a benchmark that shows the degree of air pollution. Governments have established and continuously evaluate the impact of gases and particular matters in the atmosphere. (Vallero 2014, 47.)

Pollution changes the air, water, or soil in a way that can make it harmful to people or nature. Pollutants come in a variety of forms, including chemicals, dust, noise, and radiation. These pollutants have many different sources. Some of those sources are diffuse, such as transport or agriculture, whereas others are linked to a specific place, such as a factory or power plant. (European Environment Agency 2020.)

The six "criteria air pollutants" under National Ambient Air Quality Standards (NAAQS) established in 1970 include (Vallero 2014, 47):

1. particulate matter (PM).
2. ozone (O_3).
3. carbon monoxide (CO).
4. sulphur dioxide (SO_2).
5. nitrogen dioxide (NO_2).
6. lead (Pb).

Gas or vapor exists as an individual molecule in a random motion, and particles on the other hand are aggregates of many molecules. PM is a general classification of air particles such as dust, dirt, soot, smoke, and liquid droplets. Unlike other pollutants (O_3, CO, SO_2, NO_2, and Pb), PM are a mixture of particles from various sources and different sizes, composition, and properties. The chemical composition of PM is very important and very variable which can tell a lot about its source. Particle size ranges from ultrafine to large and is determined by how it was formed. Small particles remain floating in the air for longer periods. Larger particles of an aerodynamic diameter of 10 microns are found in smoke or soot, while very small particles < 2.5 microns can only be seen indirectly. The particles can be emitted directly into the air from stationary sources such as factories, power plants, open burning, and moving vehicles. Local or non-point sources include construction and agricultural activities such as mining and forest fires. (Vallero 2014, 47-48.)

1.3 Air pollution sources

To avoid air pollution, it is crucial to understand its sources. A prominent source of air pollution is the burning of fuel. Many businesses, including transportation, electricity generation, industry, and residential heating, include the combustion of fuel. There are two types of air pollutants: primary and secondary. (European Environment Agency 2022.)

Primary pollutants are immediately released into the atmosphere. Among primary pollutants are particulate matter (PM), black carbon (BC), sulphur oxides (SO_2), nitrogen oxides (NO_X), ammonia (NH_3), carbon monoxide (CO), methane (CH_4), non-methane volatile organic compounds (NMVOCs), benzene, metals, and polycyclic aromatic hydrocarbons (PAH), such as benzo[a]pyrene (BaP). (European Environment Agency 2022.)

Secondary pollutants come from precursor gases through a combination of microphysical and chemical mechanisms. Among secondary pollutants are PM, ozone (O_3), NO_2, and several oxidized volatile organic compounds (VOCs). Precursor gases are sulphur dioxide (SO_2), NO_X, NH_3, and VOCs. (European Environment Agency 2022.)

Pollutants and their precursor gases may have both natural and anthropogenic origins including the burning of fossil fuels for transportation, industry, and housing, the use of solvents in industrial operations, agriculture, garbage disposal, natural sources, such as plant emissions of volatile organic compounds, wind-borne dust, sea-salt spray, and volcanic eruptions. (European Environment Agency 2022.)

Ground-level ozone is created by chemical reactions that occur when some NO_X, carbon monoxide (CO), NMVOCs, or methane (CH_4) precursor gases are released into the atmosphere. (European Environment Agency 2022.)

Nitrogen oxides come from the road transportation sector and industrial establishments during fuel burning. NO_X is a group of gases that produce particulate matter and ozone and includes nitrogen monoxide (NO) and nitrogen dioxide (NO_2). (European Environment Agency 2022.)

Particulate matter is created from solid and liquid aerosol particles of various sizes and chemical compositions. PM is either formed in the atmosphere because of precursor pollutant emissions, such as SO_2, NO_X, NH_3, and NMVOCs, or directly released into the atmosphere as particles. PM is released through a variety of human-made sources, including combustion and non-combustion sources. Sea salt and Saharan dust are two examples of natural sources of PM emissions. (European Environment Agency 2022.)

Sulphur dioxide is created and released when fossil fuels, primarily coal, and oil, are burned to create power. Large industrial operations tend to have higher SO_2 concentrations. Environmental concerns arise from SO_2 emissions since they are a significant precursor of $PM_{2.5}$ concentrations in the environment. (European Environment Agency 2022.)

A polycyclic aromatic hydrocarbon (PAH) called benzo(a)pyrene (BaP) can be discovered in tiny particulate matter (PM). The primary sources of BaP in Europe are coke and steel production, trash burning, domestic home heating, particularly the burning of wood and coal, and road traffic. (European Environment Agency 2022.)

Air pollution sources in Europe are presented in Figure 2 (European Environment Agency Signals 2013).

European Environment Agency

Sources of air pollution in Europe

Air pollution is not the same everywhere. Different pollutants are released into the atmosphere from a wide range of sources, including industry, transport, agriculture, waste management and households. Certain air pollutants are also released from natural sources.

1 / Around 90 % of ammonia emissions and 80 % of methane emissions come from **agricultural activities.**

2 / Some 60 % of sulphur oxides come from **energy production and distribution.**

3 / Many **natural phenomena,** including volcanic eruptions and sand storms, release air pollutants into the atmosphere.

4 / **Waste (landfills), coal mining and long-distance gas transmission** are sources of methane.

5 / More than 40 % of emissions of nitrogen oxides come from **road transport.**

Almost 40 % of primary $PM_{2.5}$ emissions come from transport.

6 / **Fuel combustion** is a key contributor to air pollution – from road transport, households to energy use and production.

Businesses, public buildings and households contribute to around half of the $PM_{2.5}$ and carbon monoxide emissions.

Source: EEA

FIGURE 2. Sources of air pollution.

1.4 Air pollution history and the industrial revolution

The programs to reduce air pollution are relatively new, yet the effort to do so dates to ancient times. Tribes moved for a variety of reasons, for example to escape the smell of human, animal, and vegetable waste. Tribes used fire in a way that incomplete combustion products flooded the air inside homes. The invention of the chimney made it easier to get cooking aromas and combustion products out of houses, but an open chimney in a fireplace produced smoke both inside and outside. Seneca, a Roman philosopher, characterized the air at Rome in A.D. 61 as heavy with the smell of its smoking pipes and all the lethal fumes and soot they contained. (Vallero 2014, 74.)

Eleanor of Aquitaine, wife of King Henry II of England, found the air pollution from burning wood at Tutbury Castle in Nottingham intolerable and was forced to leave in 1157. Consequently, it was prohibited to burn coal in London, and in 1306, Edward I issued a royal proclamation mandating the use of sea coal in furnaces. Elizabeth I had to ban coal burning in London again because it persisted despite royal edicts prohibiting it. By 1661, King Charles II and Parliament received a brochure from John Evelyn regarding the fume, unpleasantness of the air, and smoke in London since the city's pollution had become so serious. Students studying air pollution are still advised to read it. (Vallero 2014, 74-75.)

Prior to the Industrial Revolution, the primary industries producing air pollutants were metallurgy, ceramics, and the canning of animal products. The use of steam for energy generation led to the Industrial Revolution, and steam turbines required boilers that used fossil or plant-based fuels. Coal served as the primary fuel for most of the nineteenth century, while oil was employed to generate steam at the end of the century. Smoke and ash from burning coal or oil in boilers were the main sources of air pollution. (Vallero 2014, 75-76.)

The beginning of the 20th century saw significant technological advances. The electric motor took the place of the steam engine as one technological advancement. It transferred smoke and ash emissions from the plant's boiler system to the power generation boiler system. Oil replaced coal in these systems, reducing ash emissions. The most notable shift was the quick rise in vehicle production from zero to millions by 1925. (Vallero 2014, 77.)

From the 1950s to the 1970s attention was paid to reducing nitrogen oxide emissions from combustion processes and sulphur oxide emissions by eliminating sulphur oxides from flue gases and fuel desulfurization. By 1980, mathematical models for air pollution were being actively researched. At the same time, the chemistry of air pollution developed and has remained a significant research challenge ever since. Since then, air quality monitoring devices have been in use all over the world. (Vallero 2014, 79.)

Unpredictable global climate change and stratospheric ozone depletion were two worldwide environmental disasters that the media began to highlight in the 1990s. Concerns about climate change have been expressed regarding both cooling trends brought on by PM and sulphate emissions into the same atmosphere as well as warming trends brought on by the accumulation of greenhouse gases in the atmosphere. (Vallero 2014, 79.)

1.5 Health aspects of air pollutant hazards

Numerous harmful health impacts are brought on by exposure to excessive air pollution levels. Heart problems, respiratory infections, and lung cancer are all result of bad air quality. Health impacts are linked to both short- and long-term exposure to air pollution. People who are already ill have more severe effects. More vulnerable groups include children, the elderly, and poor people. (WHO 2019.)

By entering receptors through the air, chemicals, and microorganisms can spread disease and harm public health. Physical hazards can raise the risk of melanoma by increasing ultraviolet radiation exposures caused by air pollutants that travel to the stratosphere and interact with ozone. Hazardous air pollutants harm the environment and cause cancer and chronic health problems such as reproductive problems or birth deformities. These compounds are distinct from the typical air pollutants, which are pervasive and have a negative impact on public health. Toxic air pollutants are typically spatially bound to "hot spots" such as industrial and urban regions, but pollutant criteria are frequently employed to establish general ambient air quality based on common norms. (Vallero 2014, 204.)

The cause of the ozone air quality to human health is described in Table 2. The index is used for daily calculations of Ozone, $PM_{2.5}$, NO, PM_{10}, and CO_2 pollutants. (Vallero 2014, 209, 211.)

TABLE 2. Air Quality Guide for Ozone.

Air quality index	Protect your health
Good (0-50)	No health impacts are expected when air quality is in this range
Moderate (51-100)	Unusually sensitive people should consider limiting prolonged outdoor exertion
Unhealthy for sensitive Groups (101-150)	The following groups should limit prolonged outdoor exertion: • People with lung disease, such as asthma • Children and older adults • People who are active outdoors
Unhealthy (151-200)	The following groups should avoid prolonged outdoor exertion: • People with lung disease, such as asthma • Children and older adults • People who are active outdoors Everyone else should limit prolonged outdoor exertion
Very unhealthy (201-300)	The following groups should avoid all outdoor exertion: • People with lung disease, such as asthma • Children and older adults • People who are active outdoors Everyone else should limit outdoor exertion

1.5.1 Respiratory effects of air pollutants

Breathing is the most important way in which humans and other breathing animals encounter air pollutants. Through the extensive interaction of the respiratory system with the surrounding atmosphere, trace gases, and air pollutants are delivered to the respiratory system. Respiratory effects are the most immediate health effects associated with air pollutants. (Vallero 2014, 247.)

Many respiratory illnesses, both acute and chronic, are brought on by air pollution. Depending on the exposure intensity, acute diseases can include mild irritation, inflammation, allergic reactions, poor lung function, and entire respiratory failure. Lung cancer, asthma, cardiovascular disease, and chronic obstructive pulmonary disease (COPD) are examples of chronic diseases. Oxidative stress brought on by air pollution is a major factor in the emergence of such chronic diseases. (Vallero 2014, 255.)

The principal air pollutants linked to respiratory system effects include ozone, sulphur oxides, carbon monoxide, nitrogen oxides, and particulate matter. There are many chemical substances, including polycyclic aromatic hydrocarbons, that are linked to respiratory system cancer. (Vallero 2014, 255.)

The organic portion of the aerosol, including PM, is linked to cancer. Mesothelioma, lung cancer, and asbestosis have all been linked to asbestos and other fibres long-term consequences. Black lung disease, also known as pneumoconiosis, has been linked to coal dust. Silicosis has been linked to dust created from the crushing of rocks that contain silica. Byssinosis, or so-called brown lung illness, has been linked to exposure to textile fibres and can be brought on by bacteria found in cotton. (Vallero 2014, 255.)

Children and people who have asthma are particularly vulnerable to the negative effects of exposure to high tropospheric (ground level) ozone concentrations. Individuals doing intense activity or exercise are additionally in danger due to their increased doses of breathing and higher respiratory rate. (Vallero 2014, 255.)

1.5.2 Cardiovascular effects of air pollutants

The term "cardiovascular disease" is used to refer to illnesses that have an impact on the heart or blood vessels. Atherosclerosis, or the build-up of plaque on artery walls are examples of several health issues linked to heart disease. For patients with heart problems, this build-up can result in blood clots that obstruct blood flow and cause a heart attack or stroke. (EPA 2022.)

Upon ingestion, molecular oxygen (O_2) needs to be transported by the circulatory system throughout the body for energy consumption and cell growth. The cardiovascular system and heart are harmed by reduced throughput of O_2, which results in an increase in carboxyhaemoglobin and a decrease in haemoglobin. The cardiovascular system is further harmed by the toxicity of hazardous air pollutants on the respiratory system. (Vallero 2014, 257.)

The relationship between exposure to air pollutants and cardiovascular consequences has several mechanisms and stages. For instance, three hypotheses have been set up relating gaseous and airborne air pollution to arrhythmia. These mechanisms are presented in Figure 3. The pollutant may translocate, which causes a spike in blood pressure and endothelial dysfunction that results in arrhythmias. The pollutant can result in oxidative stress, inflammation, and coagulation, all of which raise the risk of myocardial ischemia. Arrhythmia is the result of oxidative stress, myocardial ischemia, and coagulation. The pollutant may also change the autonomic nervous system, which changes how the heart is regulated by the autonomic nervous system. As a result, heart rate variability (HRV), repolarization issues, and rhythm alterations cause arrhythmias. (Vallero 2014, 264 -265.)

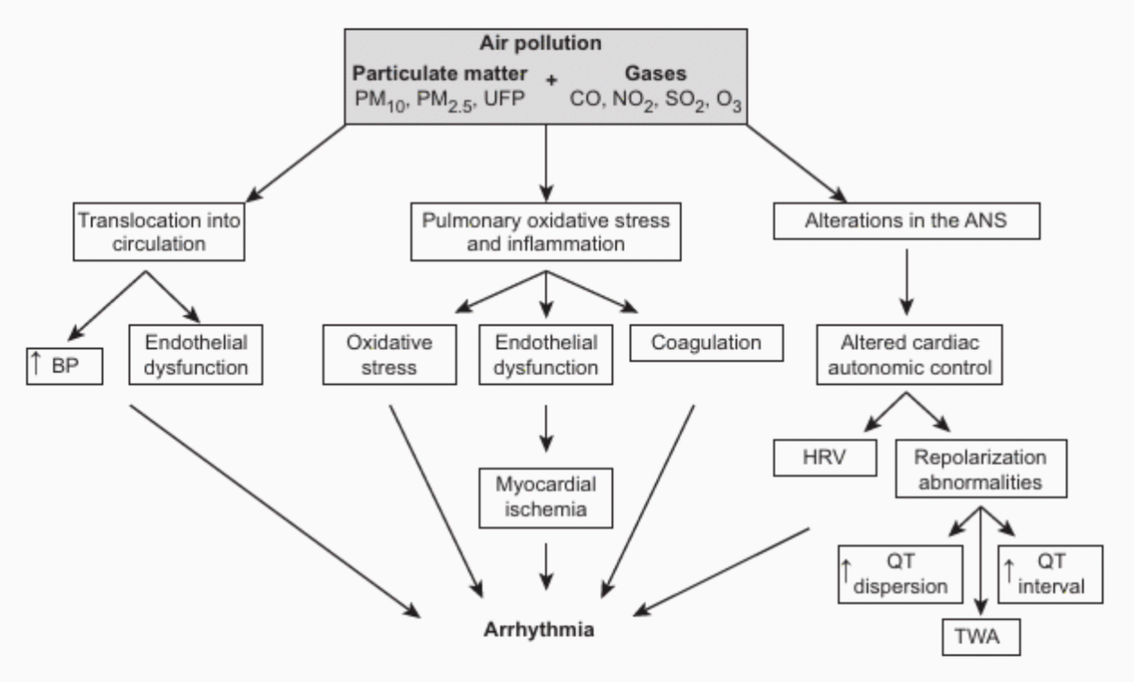


FIGURE 3. Possible biological mechanisms linking air pollution with cardiac arrhythmia. ANS, autonomic nervous system; BP, blood pressure; CO, carbon monoxide; HRV, heart rate variability; O_3, ozone; NO_2, nitrogen dioxide; PM, particulate matter; SO_2, sulphur dioxide; TWA, T wave alternans; UFP, ultrafine particles.

Particulate matter is a significant air pollutant that has been linked to both acute and chronic cardiovascular diseases. It includes increased heart rate and blood pressure, fibrinogen and blood clotting factors, arterial vasoconstriction, inflammatory mediators, endothelial damage or dysfunction, and decreased heart rate variability. This then leads to acute coronary syndromes, myocardial ischemia, malignant ventricular arrhythmias, increased plaque vulnerability, or increased potential for acute thrombosis. (Vallero 2014, 265.)

Studies on second-hand smoke (SHS), also known as environmental tobacco smoke which is the main source of indoor PM while a smoker is present, provided additional evidence of these changes. Acute myocardial infarctions can be brought on by high environmental particle concentrations, but stress brought on by exposure is linked to chronic cardiovascular disease. (Vallero 2014, 265.)

A good illustration of the process underlying the long-term cardiovascular consequences of air pollution is exposure to carbon monoxide (CO). Tachycardia, tachypnoea, rhabdomyolysis, palpitations, cardiac arrhythmias, hypertension, myocardial ischemia, and cardiac arrest are some cardiovascular consequences of CO exposure. PM_{10}, $PM_{2.5}$, NO_2, O_3, and SO_2 are typical factors that correlate with CO concentrations as CO is not the only measurable air contaminant in the environment. Moreover, cardiovascular air pollutants such as PM, NO_2, benzene, and 1,3-butadiene are concentrated alongside high-traffic areas. It is common to find O_3, PM, a variety of airborne poisons, nitrogen oxides, and sulphur oxides in metropolitan areas. (Vallero 2014, 266.)

1.5.3 Cancer and air pollutants

The main cause of lung and other cancers is exposure to airborne pollution. Depending on the PM concentration, the chance of developing lung cancer after exposure to particulate matter rises. There are many different chemical components in PM, many of which are known carcinogens. Aerosol characteristics also matter because PM from various sources is often associated with cancer. (Vallero 2014, 271.)

Many cancers, including the larynx, lung, and nasal cavity, have been connected to nickel (Ni) and its products. Steel, copper, brass, permanent magnets, storage batteries, glazes, dental fillings, and dental amalgam all include nickel. Exposure can happen by ingesting, inhalation, or skin contact. Lung and nasal cancer have been connected to occupational nickel dust exposure. The primary nickel compound in dust, nickel subsulfide (Ni_3S_2), is a combination of other nickel compounds. (Vallero 2014, 299.)

Studies of beryllium workers have shown that beryllium (Be) compounds can cause lung cancer. The aerospace and defence sectors employ beryllium compounds as metals for electrical parts, medical equipment, rocket fuel, ceramic manufacture, dental uses, sporting goods, and as additions to plastics and glass. Beryllium emissions into the atmosphere mostly result from the burning of coal and fuel oil. Inhalation of Be concentrations in atmospheric air, tobacco smoke, and food are the main ways the population is exposed. (Vallero 2014, 300-301.)

Lung cancer risk is increased by the influence of cadmium (Cd) and its products. Working near zinc and lead mining and processing facilities, producing cadmium powder, welding cadmium-plated steel, and using cadmium-containing solders all result in occupational exposure to Cd. The main application of Cd in the metal coating is to stop corrosion. The items are produced using industrial procedures that release cadmium into the air, surface water, groundwater, and topsoil. Cadmium can then be absorbed by the flora and transmitted through meat consumption. The greatest non-occupational human exposure to cadmium comes from contaminated soil, which tobacco plants can absorb and use because cadmium is released when plant material is burned during smoking. For non-smokers, food is usually the main source of Cd. Gene mutations, DNA strand breaks, chromosomal damage, cell transformation, and reduced DNA repair are all genetic harm caused by cadmium compounds. (Vallero 2014, 301.)

Compounds with hexavalent chromium (Cr^{6+}) have been linked to lung cancer. The metal was employed for tanning and dyeing in the leather and textile industries. Chromium is widely diffused in air, water, soil, and food. The industries that produce stainless steel, weld, chrome plate, and tan leather are those that expose workers to the most chromium. (Vallero 2014, 302.)

Lead (Pb) is a metal, and two of its derivatives, lead acetate, and lead phosphate, are carcinogenic to humans. Lead acetate is a chemical compound that is used in a variety of products, including explosives, paints, varnishes, pigment inks, coatings for metals, dyes for permanent hair colours, and washes for the treatment of poison ivy. For some polymers and glass, lead phosphate acts as a stabilizer. The substances can be ingested, inhaled, or injected through the skin. (Vallero 2014, 303.)

Many types of skin, lung, bladder, kidney, and liver cancer have been connected to arsenic metalloid compounds. Lung cancer risk has repeatedly been linked to occupational exposure to inhaled arsenic, especially in mining and copper smelting. Other uses for arsenic include glass, pesticides, herbicides, and wood preservatives. The major ways that exposure happens are through air inhalation, water, and food. (Vallero 2014, 303.)

1.5.4 Reproductive and hormonal effects of air pollutants

Many air contaminants can disrupt the female reproductive system. Certain air contaminants can potentially have an impact on a growing fetus. These substances are referred to as developmental and reproductive poisons. Teratogens are compounds that result in birth abnormalities, and mutagens or carcinogens are among these compounds. The Agency for Toxic Substances Disease Registry lists reproductive toxicants. For most of them, inhalation is the predominant method of exposure. There are numerous chemical classes, some of which are almost non-volatile, such as atrazine, and others are highly volatile, such as chlorinated solvents. (Vallero 2014, 313.)

Toxins modify the chemical signals sent between glands and receptors in various body cells, which has an impact on every stage of a person's life. Several substances behave differently depending on the environment. Glands, which are specialized cell clusters that generate, store, and release hormones into the bloodstream, make up the endocrine system. Adrenal glands, chemoreceptor organs, gonads, hypothalamus, pancreatic islets, parathyroid, pineal, pituitary, and thyroid glands all appear to be affected by air pollution in terms of glandular activity. Steroids, growth hormones, and thyroxine are among the hormones that are impacted. (Vallero 2014, 313.)

Endocrine-disrupting chemicals (EDCs) can mimic hormones, interfere with healthy hormone production, change the pattern of natural hormone synthesis and metabolism, and modify hormone receptor levels. Pesticide residues (such as DDT, endosulfan, and methoxychlor), polychlorinated biphenyls (PCBs), dioxins, alkylphenols (such as nonylphenol), plastic additives (such as bisphenol A, diethyl phthalate), polycyclic aromatic hydrocarbons (PAHs), and pharmaceutical hormones (e.g. 17b-estradiol, ethinyl estradiol) are among the anthropogenic EDCs. (Vallero 2014, 313-314.)

Most of the endocrine disruptors represented in Table 4 are semi-volatile and persistent organic contaminants (POPs). Airborne POPs are very harmful due to their frequent hormonal activity and detrimental effects on new-borns and pregnant children. This is evidenced by the frequent detection of these substances in breast milk and umbilical cord blood of infants. Most POPs' effects take place away from their sources and venues of use. POPs can travel over great distances because of their tenacity. Long-range atmospheric transport can transmit contaminants thousands of kilometres to locations such as the Arctic that would otherwise be thought of as being largely pristine. (Vallero 2014, 316-317.)

Certain EDCs act as antagonists by blocking hormone receptors. For instance, an estrogenic substance may have the same physiological effect—feminization—as an anti-androgenic substance that inhibits the estrogen receptor. The first is an androgen antagonist, whereas the second is an estrogen agonist. Others directly affect any number of proteins that regulate a hormone's distribution to its intended target cell. (Vallero 2014, 324.)

1.5.5 Neurological effects of air pollutants

Neurotoxicants in the air can affect normal cycles of cell generation and cell death in neurons. Neuronopathy is a collective and general term associated with damage to nerve cells, primarily in the neuronal body. For example, neurotoxicity and neuronal apoptosis in the cerebral cortex are caused by pollutants. Toxicity can also cause ischemic and seizure-induced brain damage. (Vallero 2014, 328.)

Environmental neurotoxins are widely spread and range from well-known poisonous metals such as lead and mercury to chemical substances created specifically to damage the nervous system, including pesticides containing organochlorine and organophosphate. Polychlorinated biphenyls (PCBs) and polycyclic aromatic hydrocarbons (PAHs) are additional substances that turn out to be neurotoxic. Several routes are used to transmit these and other neurotoxins in the environment, but the air is the main path that exposes people to them. (Vallero 2014, 327.)

Myelin functions in the nervous system as the equivalent of an electrical insulator, maintaining ionic currents and preventing their loss. The condition known as myelinopathy, which results from xenobiotics interfering with peripheral nervous system (PNS) functions, causes numbness, muscle weakness, incoordination, and paralysis. Toxic leukoencephalopathy is a result of brain demyelination and can cause illnesses from headaches to cognitive dysfunction, paralysis, and even death. (Vallero 2014, 327.)

Even if there is no damage to myelinating cells, contaminants can damage nerve structure. This group includes hexachlorophene, organotin compounds (Sn), and other contaminants that lead to reversible swelling between myelin layers. On the other hand, prolonged exposure to carbon monoxide (CO) and cyanide compounds (CNA) directly harms myelin-producing cell bodies in the PNS and the central nervous system (CNS). Myelin-producing cells are also harmed by inorganic lead (Pb). (Vallero 2014, 327-328.)

Since ancient times, metals have been known to be neurotoxic. The parent metal, salt, cation, or one of its organometallic compounds is absorbed and disseminated in the body after exposure. Some compounds are eliminated from the body, while some of them or products of their transformation may accumulate in lipids and other substrates. For neurotoxicity to take place, the metal must travel to the neuronal target at a concentration high enough to mechanically change how the nervous system normally functions. Proteins, amino acids, and nucleic acids are examples of nucleophilic macromolecules that tend to react with metals. (Vallero 2014, 330-331.)

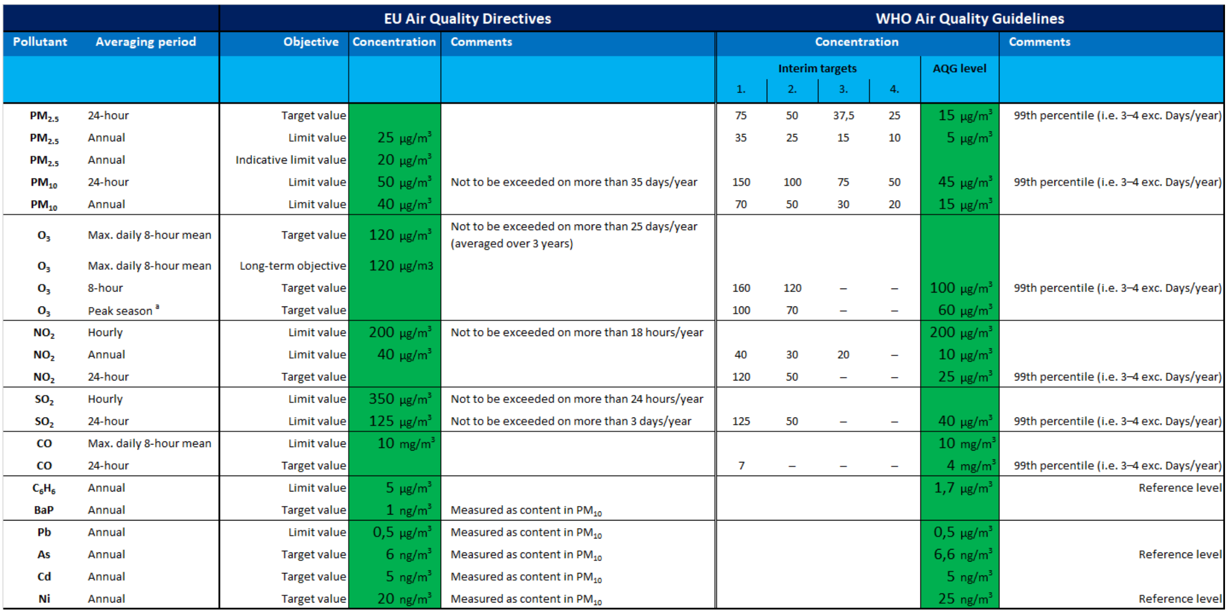
Common air contaminants and hazardous compounds have the potential to be neurotoxic. The causes of PM toxicity, including neurotoxicity, are still being studied. Moreover, if the PM components include a material that is already harmful, such as Pb, Hg, benzene, PAHs, PCBs, or dioxins, then the particles merely act as a delivery system. Organochlorine, organophosphate, and other pesticides frequently have similar mechanisms of action, and AChE is a crucial neurotoxic mechanism. (Vallero 2014, 338.)

1.6 Air quality standards

Air quality standards are concentrations that have been measured over time and are considered acceptable in the context of what is known scientifically about the impact of each pollutant on human health and the environment. They may also be used as a reference point to assess if air pollution becomes better or worse. The duration of surpassing the norm for concentration is known as the exceedance. (UK Air 2021.)

Limit values are mandated by law parameters that cannot be exceeded. Limit values are established for specific pollutants and include the concentration value, the measurement timeframe, and the permitted number of exceedances annually, if any. Some pollutants have multiple limit values, each covering different endpoints or averaging periods. The same procedures are used to set target values as limit values. In order to accomplish them, all necessary steps must be taken. (UK Air 2021.)

European Union standards and recommendations from the World Health Organization (WHO) are presented in Figure 4 (European Environment Agency 2021). These guidelines are applied throughout various time periods since the observed health impacts linked to various contaminants occur during various exposure intervals.



		EU Air Quality Directives			WHO Air Quality Guidelines					
Pollutant	Averaging period	Objective	Concentration	Comments	Concentration					Comments
					Interim targets				AQG level	
					1.	2.	3.	4.		
$PM_{2.5}$	24-hour	Target value			75	50	37,5	25	15 $\mu g/m^3$	99th percentile (i.e. 3–4 exc. Days/year)
$PM_{2.5}$	Annual	Limit value	25 $\mu g/m^3$		35	25	15	10	5 $\mu g/m^3$	
$PM_{2.5}$	Annual	Indicative limit value	20 $\mu g/m^3$							
PM_{10}	24-hour	Limit value	50 $\mu g/m^3$	Not to be exceeded on more than 35 days/year	150	100	75	50	45 $\mu g/m^3$	99th percentile (i.e. 3–4 exc. Days/year)
PM_{10}	Annual	Limit value	40 $\mu g/m^3$		70	50	30	20	15 $\mu g/m^3$	
O_3	Max. daily 8-hour mean	Target value	120 $\mu g/m^3$	Not to be exceeded on more than 25 days/year (averaged over 3 years)						
O_3	Max. daily 8-hour mean	Long-term objective	120 µg/m3							
O_3	8-hour	Target value			160	120	–	–	100 $\mu g/m^3$	99th percentile (i.e. 3–4 exc. Days/year)
O_3	Peak season [a]	Target value			100	70	–	–	60 $\mu g/m^3$	
NO_2	Hourly	Limit value	200 $\mu g/m^3$	Not to be exceeded on more than 18 hours/year					200 $\mu g/m^3$	
NO_2	Annual	Limit value	40 $\mu g/m^3$		40	30	20	–	10 $\mu g/m^3$	
NO_2	24-hour	Target value			120	50	–	–	25 $\mu g/m^3$	99th percentile (i.e. 3–4 exc. Days/year)
SO_2	Hourly	Limit value	350 $\mu g/m^3$	Not to be exceeded on more than 24 hours/year						
SO_2	24-hour	Limit value	125 $\mu g/m^3$	Not to be exceeded on more than 3 days/year	125	50	–	–	40 $\mu g/m^3$	99th percentile (i.e. 3–4 exc. Days/year)
CO	Max. daily 8-hour mean	Limit value	10 mg/m^3						10 mg/m^3	
CO	24-hour	Target value			7	–	–	–	4 mg/m^3	99th percentile (i.e. 3–4 exc. Days/year)
C_6H_6	Annual	Limit value	5 $\mu g/m^3$						1,7 $\mu g/m^3$	Reference level
BaP	Annual	Target value	1 ng/m^3	Measured as content in PM_{10}						
Pb	Annual	Limit value	0,5 $\mu g/m^3$	Measured as content in PM_{10}					0,5 $\mu g/m^3$	
As	Annual	Target value	6 ng/m^3	Measured as content in PM_{10}					6,6 ng/m^3	Reference level
Cd	Annual	Target value	5 ng/m^3	Measured as content in PM_{10}					5 ng/m^3	
Ni	Annual	Target value	20 ng/m^3	Measured as content in PM_{10}					25 ng/m^3	Reference level

FIGURE 4. Air quality directives and guidelines.

According to European regulations, the Year Average Common Air Quality Index (YACAQI) provides a general assessment of air quality level each year. The ratio of pollutant actual levels divided by European target values is used to generate the annual index. Table 3 represents annual restrictions that can be used. To calculate the ozone index, the number of days where the urban background limit of 120 μgm^{-3} as an 8-hourly average is exceeded is normalized by 25, the maximum number of exceedances allowed by the EU. (McHugh et al. 2014.)

TABLE 3. Annual Statistics and Values for Each Pollutant.

Pollutant	City annual value (background, traffic, industrial index)	Value for normalisation
NO_2	Annual average concentration	*40 μgm^{-3}
CO	Annual average concentration	5,000 μgm^{-3}
SO_2	Annual average concentration	**20 μgm^{-3}
PM_{10}	Annual average concentration	*40 μgm^{-3}
$PM_{2.5}$	Annual average concentration	*25 μgm^{-3}
Benzene	Annual average concentration	*5 μgm^{-3}
Pollutant	**City annual value (ozone index)**	**Value for normalisation**
Ozone	Number of exceedences of 120 μgm^{-3} by the maximum daily 8-hour average	*25 (exceedences)

Notes: *EU limit value for health; ** EU limit value for vegetation.

2 THEORETICAL FOUNDATION

2.1 Data sources

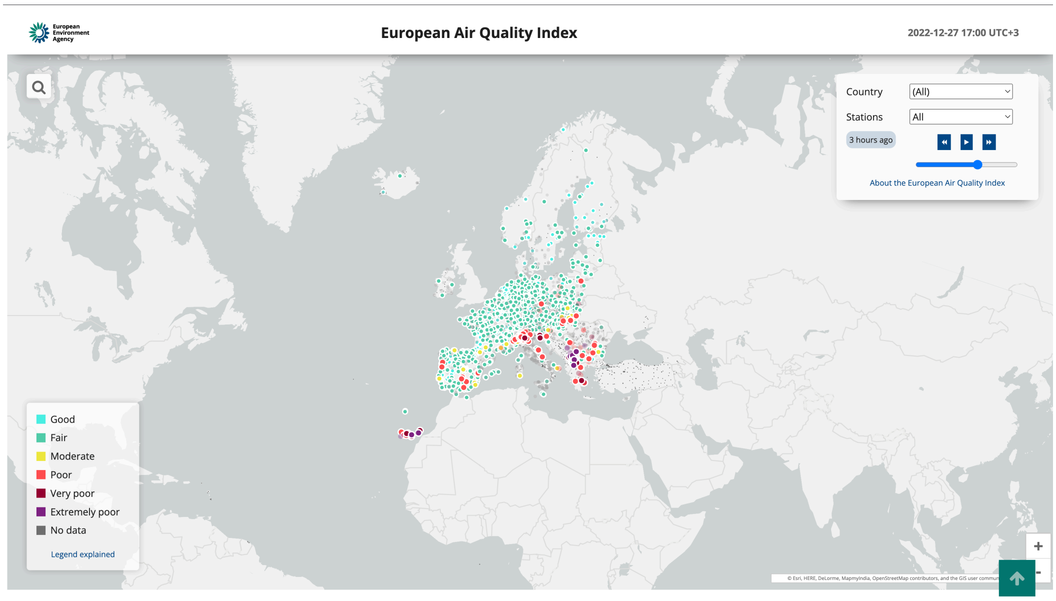
In order to effectively predict air pollution, it is necessary to have reliable data from valid sources. This is especially important in the context of data forecasting, where several open sources provide data of varying sizes and quality on different aspects of air pollution. This section provides an overview of the most reliable data sources and definitions of their databases that form the basis of the thesis project.

2.1.1 European Environment Agency

In order to support Europe's environmental and climate goals, the European Union operates the European Environment Agency (EEA) which is the source of expertise and data.

1. Interactive data views and services representing air quality data.

   Air quality live: European Air Quality Index (EAQI) is presented in Figure 5 and combines information on five different air pollutants (Particles < 2.5 µm, Particles < 10 µm, Nitrogen dioxide, Ozone, Sulphur dioxide) to show the current state of air quality in Europe.

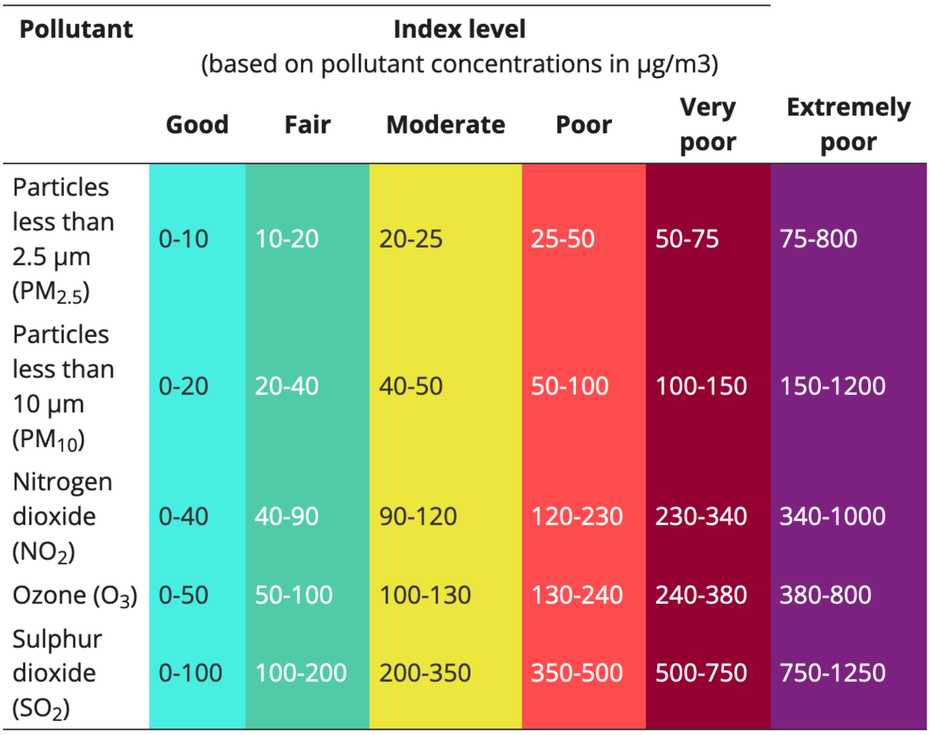


Country (All)

Stations All

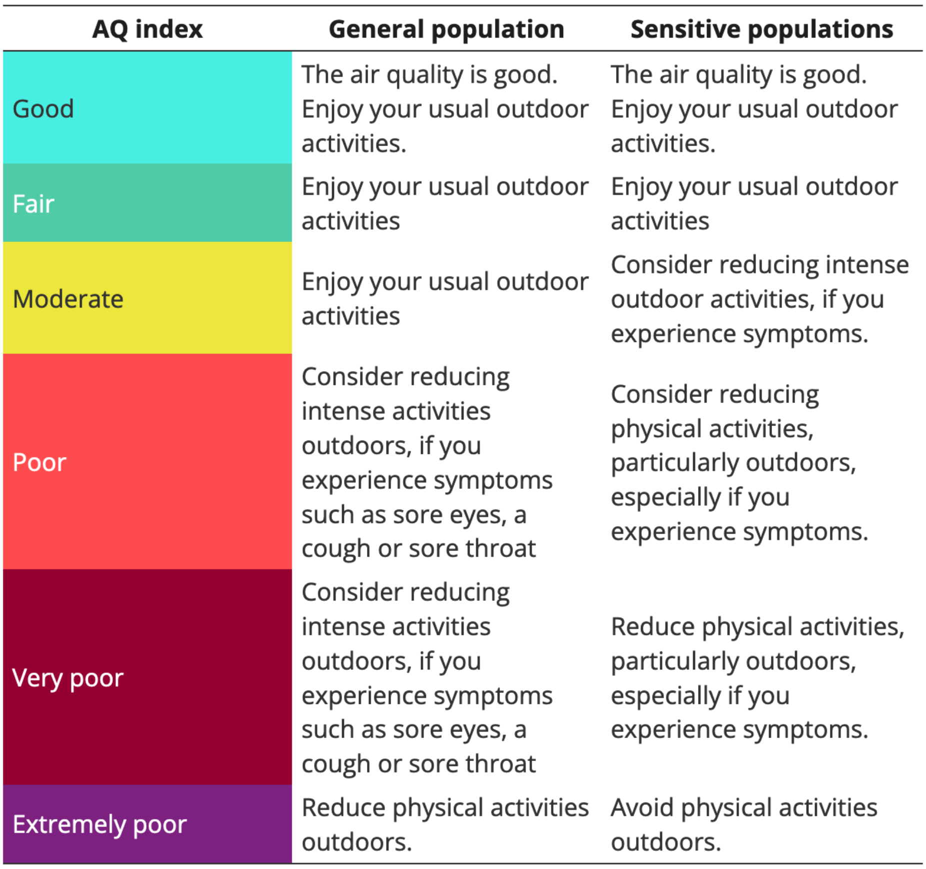
FIGURE 5. European Air Quality Index.

The World Health Organization's definition of the relative risks associated with short-term exposure to $PM_{2.5}$, O_3, and NO_2, is included in the report on the project's Health Hazards of Air Pollution in Europe and is used to define the limits. For NO_2, O_3, and SO_2, hourly concentrations are fed into the calculation of the index. For PM_{10} and $PM_{2.5}$, the 24-hour running means for the past 24 hours are fed into the calculation of the index. The limits are presented in Figure 6. Health-related suggestions that advise both the public and vulnerable individuals complement the index limits and are presented in Figure 7. (European Environmental Agency 2022.)



Pollutant	**Index level** (based on pollutant concentrations in µg/m3)					
	Good	**Fair**	**Moderate**	**Poor**	**Very poor**	**Extremely poor**
Particles less than 2.5 µm ($PM_{2.5}$)	0-10	10-20	20-25	25-50	50-75	75-800
Particles less than 10 µm (PM_{10})	0-20	20-40	40-50	50-100	100-150	150-1200
Nitrogen dioxide (NO_2)	0-40	40-90	90-120	120-230	230-340	340-1000
Ozone (O_3)	0-50	50-100	100-130	130-240	240-380	380-800
Sulphur dioxide (SO_2)	0-100	100-200	200-350	350-500	500-750	750-1250

FIGURE 6. Air quality limits.



AQ index	General population	Sensitive populations
Good	The air quality is good. Enjoy your usual outdoor activities.	The air quality is good. Enjoy your usual outdoor activities.
Fair	Enjoy your usual outdoor activities	Enjoy your usual outdoor activities
Moderate	Enjoy your usual outdoor activities	Consider reducing intense outdoor activities, if you experience symptoms.
Poor	Consider reducing intense activities outdoors, if you experience symptoms such as sore eyes, a cough or sore throat	Consider reducing physical activities, particularly outdoors, especially if you experience symptoms.
Very poor	Consider reducing intense activities outdoors, if you experience symptoms such as sore eyes, a cough or sore throat	Reduce physical activities, particularly outdoors, especially if you experience symptoms.
Extremely poor	Reduce physical activities outdoors.	Avoid physical activities outdoors.

FIGURE 7. Health-related suggestions according to the air quality index.

2. Air Quality e-Reporting (AQ e-Reporting) database.

The data includes information on air pollutants that has been gathered over several years by EEA member countries, including European Union members, EEA cooperating countries, and other nations. The database includes statistics on a few air pollutants as well as data on annual air quality monitoring. It also contains details about the air quality monitoring networks that are engaged, including their stations, measurements, modelling methodologies, zones, assessment regimes, compliance achievements, plans, and projects for the member states of the EU and the EEC. (European Environmental Agency 2022.)

Air quality annual statistics are computed by the EEA based on observations that countries upload to the Central Data Repository. The data table contains aggregated air quality values that are calculated by the EEA (CDR). The table definition is presented in Table 4. The calculations were made for the years from 1969 to 2022. The covered location includes Albania, Andorra, Austria, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Kosovo under UNSCR 1244/99, Latvia, Lichtenstein, Lithuania, Luxembourg, Malta, Montenegro, Netherlands, North Macedonia, Norway, Poland, Portugal, Romania, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, United Kingdom. It covers more than 100 air pollutants. (European Environmental Agency 2022.)

TABLE 4. Air Quality annual statistics table definition.

Field name	Description	Data type
Country	Country or territory name.	string
Air Quality Network	Inspire identifier (Local Id) of air quality network, given by data provider.	string
Air Quality Network Name	Name of air quality measurement network, given by data provider.	string
Air Quality Station EoI Code	EoI code of air quality measurement station (as in AirBase), given by data provider.	string
Air Quality Station Name	Name of air quality measurement station, given by data provider.	string
Sampling Point Id	Inspire identifier (Local Id) of sampling point, given by data provider.	string
Air Pollutant	Air polluting substance, the level of which is measured and reported to the EEA (see notation in Data Dictionary: http://dd.eionet.europa.eu/vocabulary/aq/pollutant).	string
Air Pollutant Description	Description of air polluting substance, level of which is measured and reported to the EEA (see also in Data Dictionary: http://dd.eionet.europa.eu/vocabulary/aq/pollutant).	string
Data Aggregation Process Id	Id of the process of data aggregation into annual values (see in Data Dictionary: http://dd.eionet.europa.eu/vocabulary/aq/aggregationprocess).	string

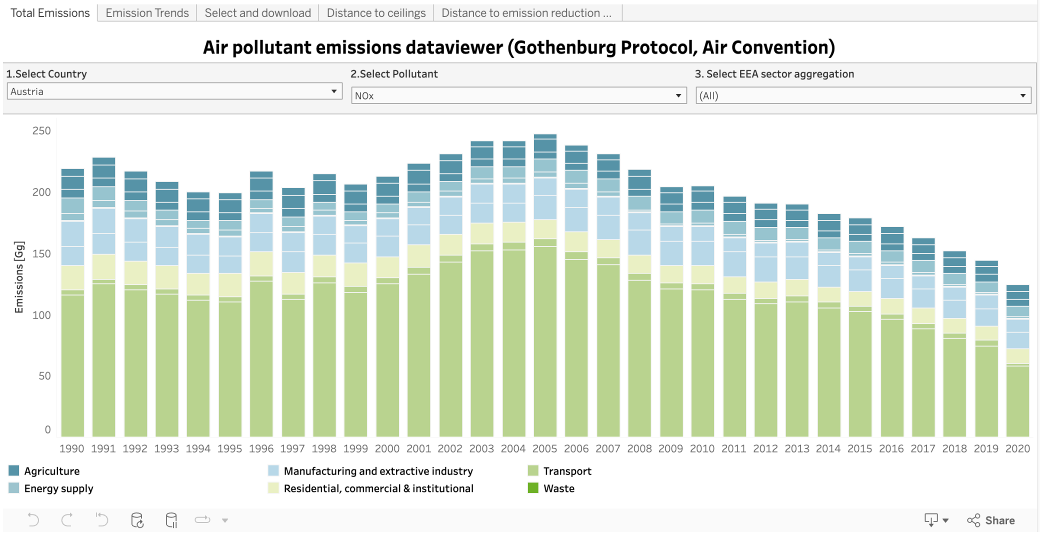
Data Aggregation Process	Description of the process of data aggregation into annual values (see in Data Dictionary: http://dd.eionet.europa.eu/vocabulary/aq/aggregationprocess).	string
Year	Year for which primary data have been reported/ statistics were calculated.	numeric
Air Pollution Level	Concentration or level of air polluting substance, here given as an aggregation of air pollutant concentration values from primary observation time series.	numeric
Unit Of Air Pollution Level	Unit of concentration or level of air-polluting substance (see in Data Dictionary: http://dd.eionet.europa.eu/vocabulary/uom/concentration).	string
Data Coverage	The proportion of valid measurements included in the aggregation process within the averaging period is expressed as a percentage. If Data Coverage is < 75% for the averaging period of a year, annual statistics should not be included in air quality assessments, if Data Coverage is < 85% (in a year), annual statistics should not be included in compliance checks.	numeric
Verification	Information based on verification flags found in reported time series (see in Data Dictionary: http://dd.eionet.europa.eu/vocabulary/aq/observationverification).	numeric
Air Quality Station Type	Type of Air Quality Measurement Station - information whether it is measuring background, industrial, or traffic-related air pollution (see in Data Dictionary: http://dd.eionet.europa.eu/vocabulary/aq/stationclassification).	string

Air Quality Station Area	Area of Air Quality Measurement Station - information on whether it is measuring air pollution in urban, suburban, rural (etc.) environments (see in Data Dictionary: http://dd.eionet.europa.eu/vocabulary/aq/areaclassification).	string
Longitude	Longitude of the location of AQ measurement for which the statistics were calculated.	numeric
Latitude	Latitude of the location of AQ measurement for which the statistics were calculated.	numeric
Altitude	The altitude of the location of the AQ measurement for which the statistics were calculated.	numeric
City	City name.	string
City Code	City code as specified in Eurostat's Urban Audit.	string
City Population	The population of the city.	string
Source Of Data Flow	Specification of data flow which is the source of raw data used for the statistics calculation.	string
Calculation Time	Date and time of the latest statistics calculation.	datetime
Link to raw data (only E1a/validated data from AQ e-Reporting)	Link to raw data (only E1a/validated data from AQ e-Reporting).	string

3. EU National Emission Ceilings (NEC) Directive collects data from national air pollutant emission inventories and projections submitted by EU Member States. The data collection is recommended for anyone wishing to assess how well EU Member States are doing in fulfilling their obligations to reduce emissions under the NEC Directive. It includes data on air pollutant emissions (ammonia (NH_3), non-methane volatile organic compounds (NMVOCs), nitrogen oxides (NO_x), particulate matter 2.5 ($PM_{2.5}$), and sulphur dioxide (SO_2)

reported annually by the Member States to the European Commission following Directive 2016/2284 of the European Parliament and Council to reduce national emissions of certain air pollutants are included in the National Emission Reductions Commitments (NEC) Directive emission inventory. The description was made the for years 1980-2020. (European Environmental Agency 2022.)

4. Air pollutant emissions data viewer (Gothenburg Protocol, LRTAP Convention) is presented in Figure 8. Following the UNECE Convention on Long-range Transboundary Air Pollution (LRTAP), the air pollutant emissions data viewer (LRTAP Convention) gives access to the information in the EU emission inventory report 1990-2020. It has different sections as Total emissions, Emission Trends, Select and download, Distance to ceilings, and Distance to emission reduction. (European Environmental Agency 2022.)



Air pollutant emissions dataviewer (Gothenburg Protocol, Air Convention)

1.Select Country: Austria

2.Select Pollutant: NOx

3. Select EEA sector aggregation: (All)

FIGURE 8. Air pollutant emissions data viewer (Gothenburg Protocol, Air Convention).

5. Air Quality Health Risk Assessments is a dataset that provides estimates of health risks from exposure to the three major pollutants ($PM_{2.5}$, NO_2, and O_3-SOMO35) at the NUTS3 and country levels. In addition, the dataset has population-weighted average and average concentrations for PM_{10}, $PM_{2.5}$, NO_2, and O_3 (SOMO35). The calculations were made for the years 2005-2020. Following the NUTS 2021 classification (Nomenclature of territorial units for statistics), the administrative boundaries of the EEA38 and the UK were divided into the country (NUTS0), NUTS1, NUTS2, and NUTS3 regions, as well as some additional territories such as Andorra, Monaco, San Marino, and Vatican City. Major socioeconomic regions fall under the NUTS1 group, basic regions fall under the NUTS2 group, and small regions fall under the NUTS3 group, which are used for specialized diagnosis. (European Environmental Agency 2022.)

6. Indicators and reports.
    a. Emissions of the main air pollutants in Europe. Graph and table are presented in Figure 9 and Figure 10. This indicator tracks trends since 2005 in anthropogenic emissions of five major air pollutants - NO_X, NH_3, SO_X, NMVOCs, and $PM_{2.5}$. All these pollutants, directly or indirectly, harm human health, vegetation, and ecosystems. Pollutant emissions in the EU decreased significantly between 2005 and 2020: SO_X emissions decreased by 79.5%, NO_X emissions decreased by 48.5%, NMVOC emissions decreased by 31%, and $PM_{2.5}$ emissions decreased by 32%. These declines are largely attributable to lower emissions from the energy, industrial, and transportation sectors, in part because of sector-specific emission limit values established by other EU laws such as the Industrial Emissions Directive, the Large Combustion Plant Directive, and Euro standards for vehicles. NH_3 emissions also decreased, but only by 8.5%, and even slightly increased between 2013 and 2017. Agriculture is the source of more than 90% of NH_3 emissions and has not made much in recent years. (European Environmental Agency 2022.)

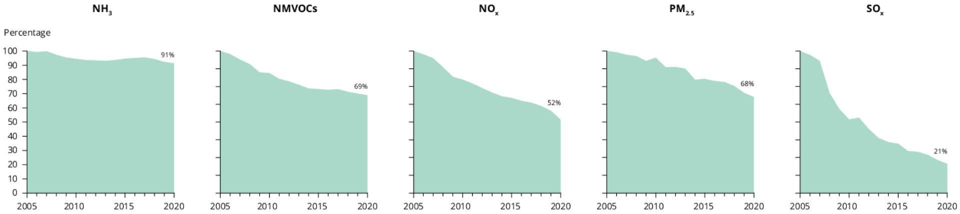
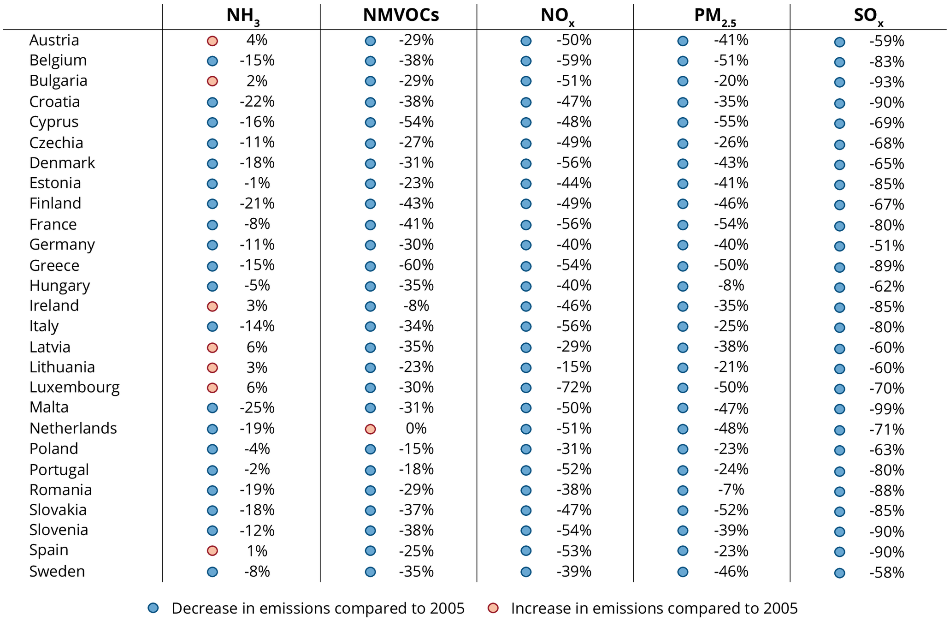


FIGURE 9. Percentage emission reductions of main air pollutants in 2020 compared to 2005 levels.



	NH_3	NMVOCs	NO_x	$PM_{2.5}$	SO_x
Austria	○ 4%	● -29%	● -50%	● -41%	● -59%
Belgium	● -15%	● -38%	● -59%	● -51%	● -83%
Bulgaria	○ 2%	● -29%	● -51%	● -20%	● -93%
Croatia	● -22%	● -38%	● -47%	● -35%	● -90%
Cyprus	● -16%	● -54%	● -48%	● -55%	● -69%
Czechia	● -11%	● -27%	● -49%	● -26%	● -68%
Denmark	● -18%	● -31%	● -56%	● -43%	● -65%
Estonia	● -1%	● -23%	● -44%	● -41%	● -85%
Finland	● -21%	● -43%	● -49%	● -46%	● -67%
France	● -8%	● -41%	● -56%	● -54%	● -80%
Germany	● -11%	● -30%	● -40%	● -40%	● -51%
Greece	● -15%	● -60%	● -54%	● -50%	● -89%
Hungary	● -5%	● -35%	● -40%	● -8%	● -62%
Ireland	○ 3%	● -8%	● -46%	● -35%	● -85%
Italy	● -14%	● -34%	● -56%	● -25%	● -80%
Latvia	○ 6%	● -35%	● -29%	● -38%	● -60%
Lithuania	○ 3%	● -23%	● -15%	● -21%	● -60%
Luxembourg	○ 6%	● -30%	● -72%	● -50%	● -70%
Malta	● -25%	● -31%	● -50%	● -47%	● -99%
Netherlands	● -19%	○ 0%	● -51%	● -48%	● -71%
Poland	● -4%	● -15%	● -31%	● -23%	● -63%
Portugal	● -2%	● -18%	● -52%	● -24%	● -80%
Romania	● -19%	● -29%	● -38%	● -7%	● -88%
Slovakia	● -18%	● -37%	● -47%	● -52%	● -85%
Slovenia	● -12%	● -38%	● -54%	● -39%	● -90%
Spain	○ 1%	● -25%	● -53%	● -23%	● -90%
Sweden	● -8%	● -35%	● -39%	● -46%	● -58%

● Decrease in emissions compared to 2005 ○ Increase in emissions compared to 2005

FIGURE 10. Emission reduction of the main air pollutants by Member States from 2005 to 2020.

b. Heavy metal emissions. Graph and table are presented in Figure 11 and Figure 12. This indicator tracks trends in anthropogenic emissions of heavy metals over time since 2005. Heavy metals (such as Cd, Hg, and Pb) are known to be diindirectly toxic to biota. (European Environmental Agency 2022.)

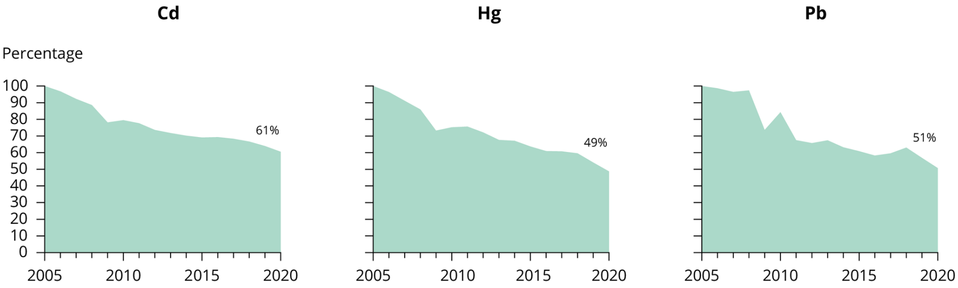
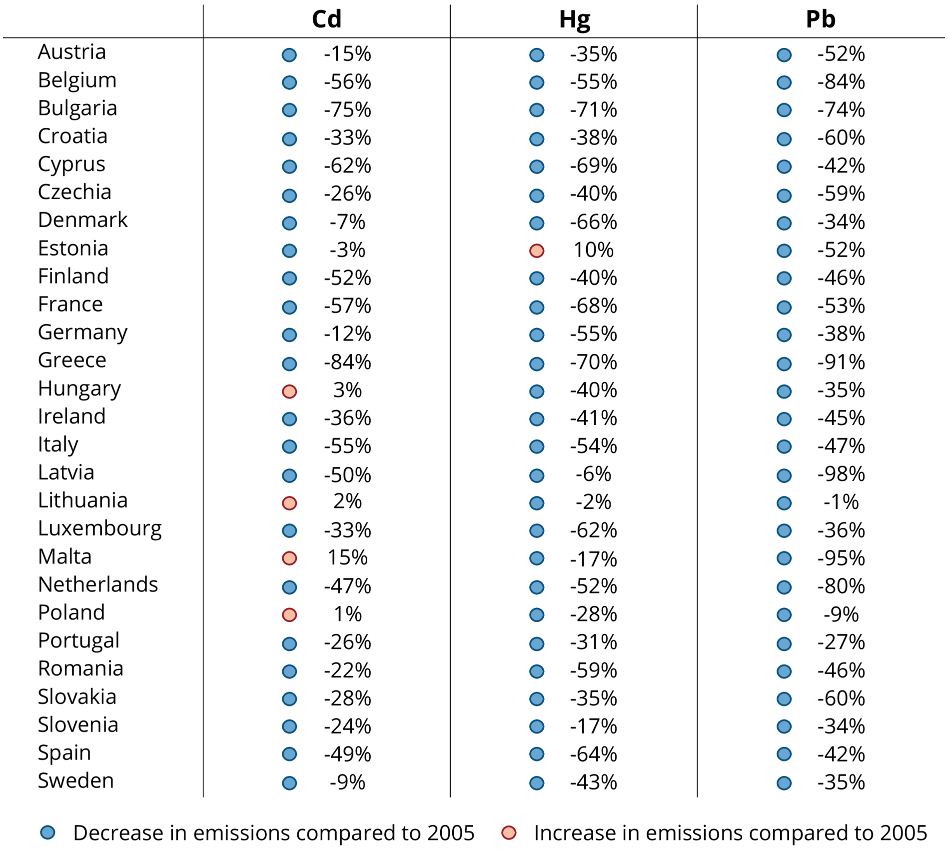


FIGURE 11. Percentage emission reductions in 2020 of primary heavy metals compared with 2005 levels.



	Cd	Hg	Pb
Austria	🔵 -15%	🔵 -35%	🔵 -52%
Belgium	🔵 -56%	🔵 -55%	🔵 -84%
Bulgaria	🔵 -75%	🔵 -71%	🔵 -74%
Croatia	🔵 -33%	🔵 -38%	🔵 -60%
Cyprus	🔵 -62%	🔵 -69%	🔵 -42%
Czechia	🔵 -26%	🔵 -40%	🔵 -59%
Denmark	🔵 -7%	🔵 -66%	🔵 -34%
Estonia	🔵 -3%	🔴 10%	🔵 -52%
Finland	🔵 -52%	🔵 -40%	🔵 -46%
France	🔵 -57%	🔵 -68%	🔵 -53%
Germany	🔵 -12%	🔵 -55%	🔵 -38%
Greece	🔵 -84%	🔵 -70%	🔵 -91%
Hungary	🔴 3%	🔵 -40%	🔵 -35%
Ireland	🔵 -36%	🔵 -41%	🔵 -45%
Italy	🔵 -55%	🔵 -54%	🔵 -47%
Latvia	🔵 -50%	🔵 -6%	🔵 -98%
Lithuania	🔴 2%	🔵 -2%	🔵 -1%
Luxembourg	🔵 -33%	🔵 -62%	🔵 -36%
Malta	🔴 15%	🔵 -17%	🔵 -95%
Netherlands	🔵 -47%	🔵 -52%	🔵 -80%
Poland	🔴 1%	🔵 -28%	🔵 -9%
Portugal	🔵 -26%	🔵 -31%	🔵 -27%
Romania	🔵 -22%	🔵 -59%	🔵 -46%
Slovakia	🔵 -28%	🔵 -35%	🔵 -60%
Slovenia	🔵 -24%	🔵 -17%	🔵 -34%
Spain	🔵 -49%	🔵 -64%	🔵 -42%
Sweden	🔵 -9%	🔵 -43%	🔵 -35%

🔵 Decrease in emissions compared to 2005 🔴 Increase in emissions compared to 2005

FIGURE 12. Percentage emission reductions of primary heavy metals of EU Member States in 2020 compared with 2005 levels.

c. Emissions of air pollutants from transport. The chart is presented in Figure 13. This indicator is based on the assessment of emissions trends of CO, NOx, NMVOCs, SOx, and primary particulates. The calculations were made for the years 1990-2017. (European Environmental Agency 2021.)

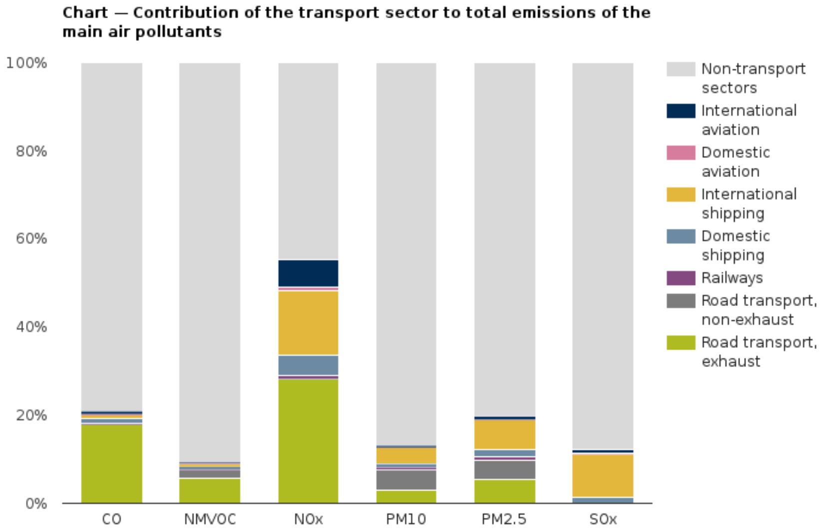


FIGURE 13. Contribution of the transport sector to the total emission of the main air pollutants.

d. Industrial pollutants releases into the air in Europe. Charts are presented in Figure 14 and Figure 15. This indicator tracks trends of industrial emissions of selected air pollutants. The indicator includes releases of carbon dioxide (CO_2), the most significant greenhouse gas, acidifying pollutants (sulphur oxides (SOx), nitrogen oxides (NOx)), and other pollutants that damage human health and the environment, such as particulate matter (PM_{10}), non-methane volatile organic compounds (NMVOCs) and heavy metals (Cadmium (Cd), Lead (Pb), and mercury (Hg). These trends are presented together with the trend of gross value added (GVA) by industry, as an indicator of the economic contribution of the sector. The calculations were made for the years 2010-2020. (European Environmental Agency 2022.)

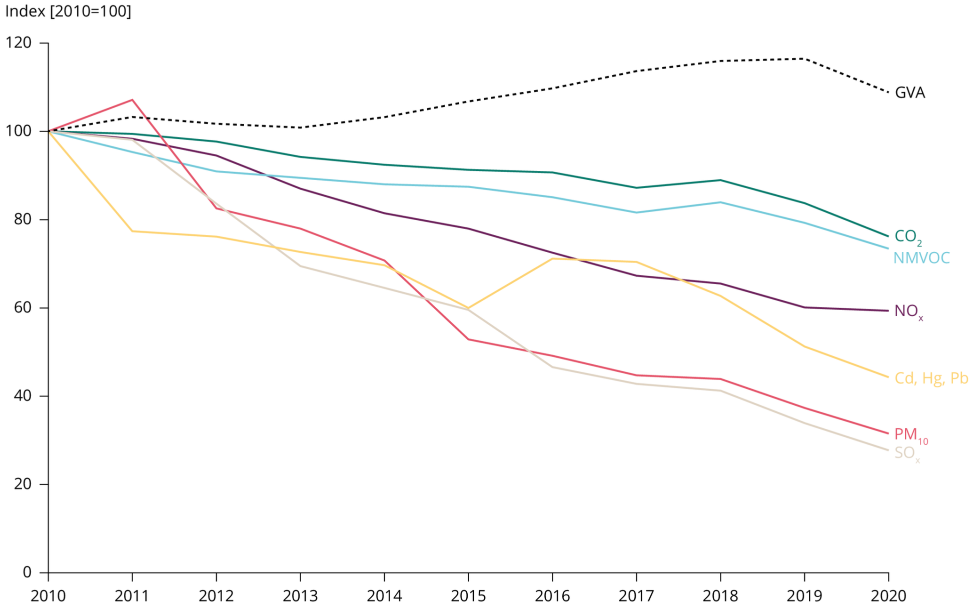
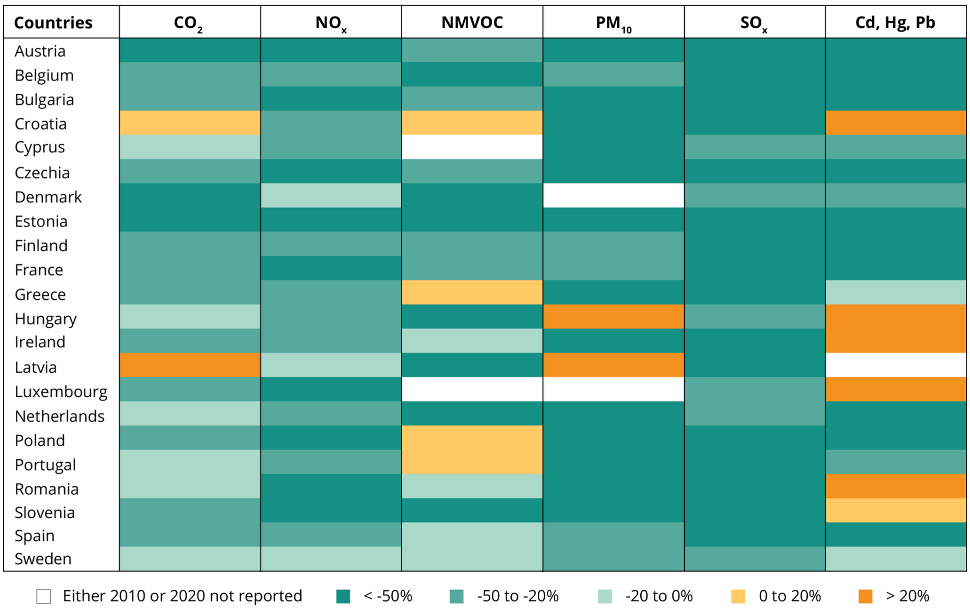


FIGURE 14. Industrial releases of pollutants to air and economic activity in the EU-27.



Countries	CO_2	NO_x	NMVOC	PM_{10}	SO_x	Cd, Hg, Pb
Austria	■	■	▨	■	■	■
Belgium	▨	▨	■	▨	■	■
Bulgaria	▨	■	▨	■	■	■
Croatia	🟨	▨	🟨	■	■	🟧
Cyprus	░	▨	☐	■	▨	▨
Czechia	▨	■	▨	■	■	■
Denmark	■	░	■	☐	▨	▨
Estonia	■	■	■	■	■	■
Finland	▨	▨	▨	▨	■	■
France	▨	■	▨	▨	■	■
Greece	▨	▨	🟨	■	■	░
Hungary	░	▨	■	🟧	▨	🟧
Ireland	▨	▨	░	■	■	🟧
Latvia	🟧	░	■	🟧	■	☐
Luxembourg	▨	■	☐	☐	▨	🟧
Netherlands	░	▨	■	■	▨	■
Poland	▨	■	🟨	■	■	■
Portugal	░	▨	🟨	■	■	▨
Romania	░	■	░	■	■	🟧
Slovenia	▨	■	■	■	■	🟨
Spain	▨	▨	░	▨	■	■
Sweden	░	░	░	▨	▨	░

☐ Either 2010 or 2020 not reported ■ < -50% ▨ -50 to -20% ░ -20 to 0% 🟨 0 to 20% 🟧 > 20%

FIGURE 15. Change in pollutant releases to air in the EU-27 countries, 2010-2020.

7. Air pollution country fact sheets 2022. The Air Polluting Countries Fact Sheets summarize basic air pollution data for each of the 32 SES member countries and 6 cooperating countries. Past and future emission trends are presented, as well as a brief description of the national air quality situation in each country. (European Environmental Agency 2022.)
8. Global and European land average near-surface temperatures. Graphs are presented in Figure 16. The Copernicus Climate Change Service (C3S) used these datasets to calculate the global yearly averages of near-surface temperatures of land and oceans: ERA5

(C3S/ECMWF), JRA-55 (JMA), GISTEMPv4 (NASA), HadCRUT5 (Met Office Hadley Centre), NOAAGlobalTempv5 (NOAA), and Berkeley Earth and the following datasets to calculate the annual average near-surface temperature in Europe: ERA5 (C3S/ECMWF), JRA-55 (JMA), GISTEMPv4 (NASA), HadCRUT5 (Met Office Hadley Centre), NOAAGlobalTempv5 (NOAA), and Berkeley Earth. The data table includes average near-surface temperature in Europe from different sources and its definition is presented in Table 5. The calculations were made for the years 1850-2021. (European Environmental Agency 2022.)

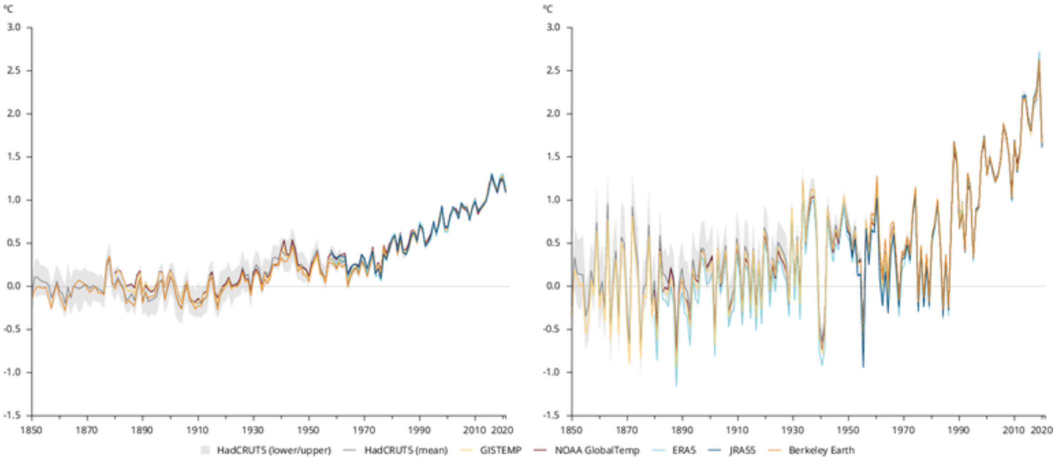


FIGURE 16. Global (left) and European land (right) average near-surface temperatures.

Left chart: The data are averaged over a 12-month calendar year. Data only covers land in Europe and the Arctic. Right chart: The data are averaged over a 12-month calendar year. Europe is the land-only portion of the latitude/longitude box limited by 25°W - 40°E, 34°N - 72°N. (European Environmental Agency 2022.)

TABLE 5. Average near-surface temperatures data table definition.

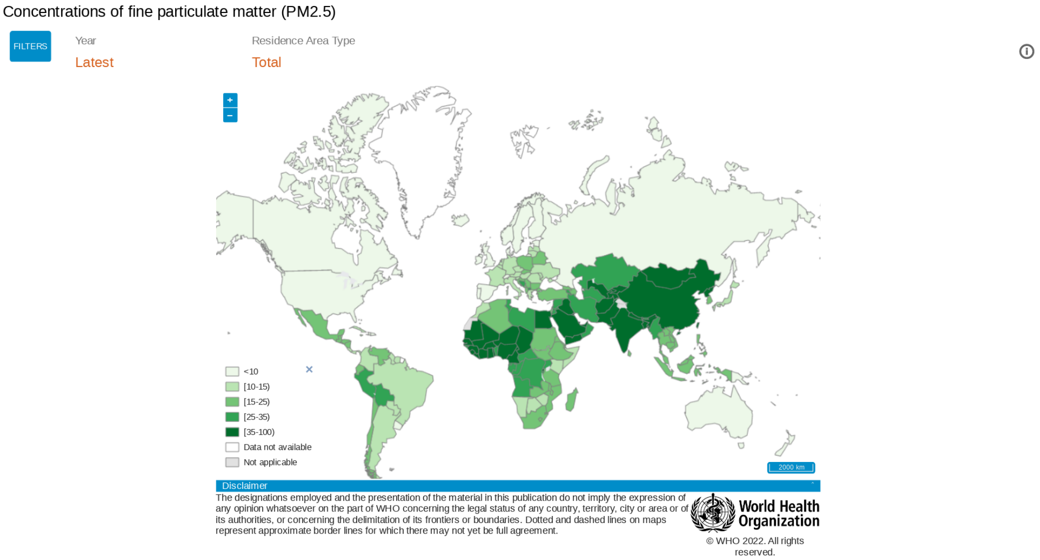
Field name	Description	Data type
Year	Year for which data have been reported (1850-2021).	numeric
HadCRUT5	Annual average near-surface temperature from HadCRUT5 database.	numeric
HadCRUT5max	Annual maximum near-surface temperature from HadCRUT5 database.	numeric
HadCRUT5min	Annual minimum near-surface temperature from HadCRUT5 database.	numeric
NOAAGlobalTemp	Annual average near-surface temperature from NOAAGlobalTemp database.	numeric
GISTEMP	Annual average near-surface temperature from GISTEMP database.	numeric
ERA5	Annual average near-surface temperature from ERA5 database.	numeric
JRA-55	Annual average near-surface temperature from JRA-55 database.	numeric
BerkeleyEarth	Annual average near-surface temperature from BerkeleyEarth database.	numeric

Additional data from European Environment Agency can be found in Appendix 1.

2.1.2 World Health Organization

The World Health Organisation is a United Nations institution that unites nations, allies, and individuals to promote health, maintain global security, and assist the most vulnerable people so that everyone can enjoy the best possible level of health. It also provides data about different aspects of health issues, including air pollution.

1. The WHO Air Quality Database gathers information from ground-based measurements of annual mean concentrations of nitrogen dioxide (NO_2), particulate matter with a diameter of equal to or less than 10 microns (PM_{10}), and particulate matter with a diameter of equal to or less than 2.5 microns ($PM_{2.5}$), to represent an average for the city or town rather than for individual stations. Both categories of pollutants are mostly the result of human activities connected to the burning of fossil fuels. The database presently contains information on the air quality for more than 6000 human settlements in more than 100 nations. Since 2011, the database has been consistently updated every two to three years. The calculations were made from 2000-2021. (World Health Organization 2022.)
2. Concentrations of fine particulate matter ($PM_{2.5}$) data contain an annual mean concentration of particulate matter of fewer than 2.5 microns of diameter ($PM_{2.5}$) [$\mu g/m^3$] in urban areas. The data view is presented in Figure 17. The calculations were made in the years 2010-2019. (World Health Organization 2022.)



Concentrations of fine particulate matter (PM2.5)

Year: Latest

Residence Area Type: Total

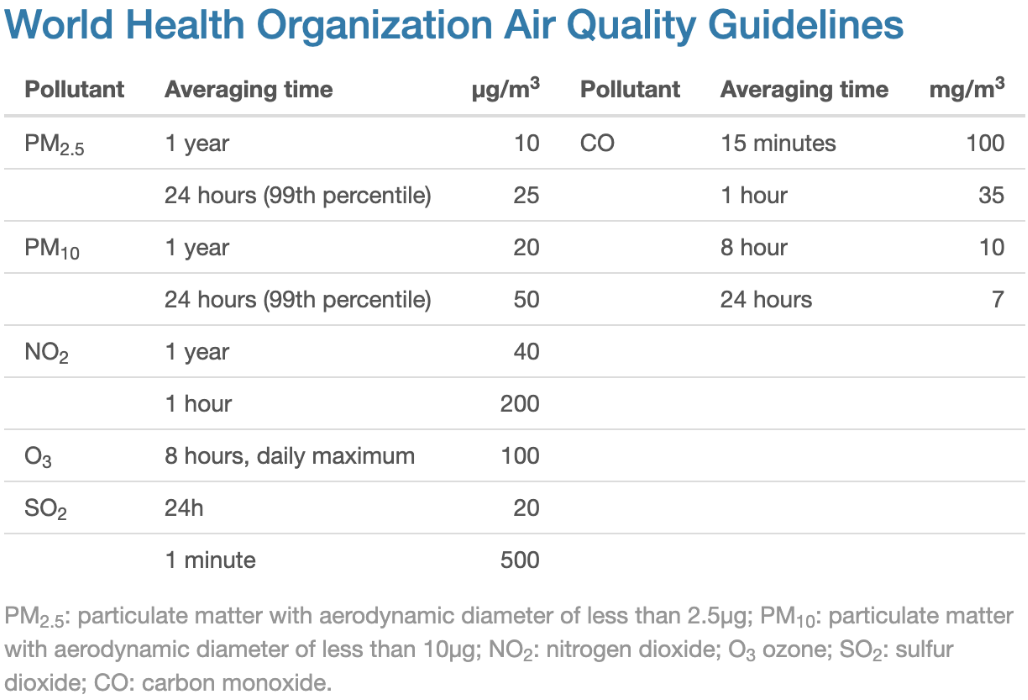
Disclaimer

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of WHO concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

© WHO 2022. All rights reserved.

FIGURE 17. Indicator 11.6.2: Annual mean levels of fine particulate matter (e.g. $PM_{2.5}$ and PM_{10}) in cities (population weighted) on a map.

3. Current national air quality criteria for the five traditional pollutants—particulate matter, nitrogen dioxide, ozone, carbon monoxide, and sulphur dioxide are displayed interactively on a map called Air Quality Standards. The WHO Air Quality Guidelines values are presented as references in Figure 18. (World Health Organization 2022.)



World Health Organization Air Quality Guidelines

Pollutant	Averaging time	µg/m³	Pollutant	Averaging time	mg/m³
$PM_{2.5}$	1 year	10	CO	15 minutes	100
	24 hours (99th percentile)	25		1 hour	35
PM_{10}	1 year	20		8 hour	10
	24 hours (99th percentile)	50		24 hours	7
NO_2	1 year	40			
	1 hour	200			
O_3	8 hours, daily maximum	100			
SO_2	24h	20			
	1 minute	500			

$PM_{2.5}$: particulate matter with aerodynamic diameter of less than 2.5µg; PM_{10}: particulate matter with aerodynamic diameter of less than 10µg; NO_2: nitrogen dioxide; O_3 ozone; SO_2: sulfur dioxide; CO: carbon monoxide.

FIGURE 18. World Health Organization Air Quality Guidelines.

2.1.3 IQAir data

IQAir is a company from Switzerland specializing in monitoring air quality in real-time. The website has an interactive map that shows data about air quality in a specified region: air pollution index, air quality index, main pollutant, the concentration, health recommendations, air quality index forecast, historic air quality graph, real-time air pollution map, data contributor information, weather, ranking, and a webcam. It also has air quality analysis and statistics for almost every country in the world which answers questions such as how bad air pollution is, what is the main cause of the pollution, where is the clearest air, what is the government doing about air pollution, and who monitors air quality.

2.1.4 The World Bank

The World Bank offers developing countries technical assistance, policy recommendations, research, analysis, and support. The analytical work typically acts as the foundation for World Bank financing and contributes to steering the investments of developing nations. It also provides free and open global development data:

1. Population, total – European Union. Total population is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship. The values shown are midyear estimates. The data table definition is presented in Table 6. The calculations were made for the years 1960-2021. (The World Bank 2022.)

TABLE 6. Population, total table definition.

Field name	Description	Data type
Country Name	Country or territory name.	string
Country Code	Country or territory code.	string
Indicator name	Population, total	string
Indicator Code	SP.POP.TOTL	string
1960-2021	The fields include the population number for each year.	numeric

2. GDP per capita (current US$) - European Union. GDP per capita is gross domestic product divided by midyear population. GDP is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for the depreciation of fabricated assets or for the depletion and degradation of natural resources. Data are in current U.S. dollars. The data table definition is presented in Table 7. The calculations were made for the years 1960-2021. (The World Bank 2022.)

TABLE 7. GDP per capita (current US$) table definition.

Field name	Description	Data type
Country Name	Country or territory name.	string
Country Code	Country or territory code.	string
Indicator name	GDP per capita (current US$)	string
Indicator Code	NY.GDP.PCAP.CD	string
1960-2021	The fields include GDP per capita number for each year.	numeric

2.1.5 Our World in Data

The charity Global Change Data Lab, which owns, publishes, and maintains both the website and the data tools, has partnered with scholars at the University of Oxford to create the website Our World in Data. It provides free and open global data:

1. Fossil fuel consumption (TWh) includes the amount of primary energy from fossil fuels that are consumed each year. This is the sum of energy from coal, oil, and gas. The data table definition is presented in Table 8. The calculations were made for the years 1965-2021. (Ritchie, Roser & Rosado 2022).

TABLE 8. Fossil fuel consumption (TWh) table definition.

Field name	Description	Data type
Entity	Country or territory name.	string
Code	Country or territory code.	string
Year	Year for which data have been reported (1965-2021).	numeric
Fossil fuels (TWh)	Number of fossil fuel consumption in terawatt-hours.	numeric

2. Fossil fuel consumption by type – oil, coal, and gas (TWh). It includes the amount of primary energy from oil, coal, and gas that is consumed each year. The data table definition

is presented in Table 9. The calculations were made for the years 1965-2021. (Ritchie, Roser & Rosado 2022).

TABLE 9. Fossil fuel consumption by type table definition.

Field name	Description	Data type
Entity	Country or territory name.	string
Code	Country or territory code.	string
Year	Year for which data have been reported (1965-2021).	numeric
Gas Consumption - TWh	Number of gas consumption in terawatt-hours.	numeric
Coal Consumption - TWh	Number of coal consumption in terawatt-hours.	numeric
Oil Consumption - TWh	Number of oil consumption in terawatt-hours.	numeric

2.2 Factors

In order to make predictions, it is necessary to understand the factors that cause an increase or decrease in air pollution. The thesis work defines factors such as population, GDP per capita, fossil fuel consumption, and temperature as having the greatest impact on air pollution levels. This section provides an overview of these factors used in the forecast.

2.2.1 Population

For most of human existence, population growth has been extremely slow, with diseases, climatic changes, and other socioeconomic issues. It took until 1804 for the first billion people to be reached. Since then, the population has increased significantly due to continual developments in technology, medicine, and nutrition. (Dovers & Butler 2019.)

Consumption of resources such as land, food, water, air, fossil fuels, and minerals as well as waste products such as air and water pollution, dangerous compounds, and greenhouse gases are the two main ways that so many people damage the environment. (Dovers & Butler 2019.)

The environment could face a disaster because of population growth. An analysis of the data shows that environmental health has decreased along with population growth. Because of the consequences that so many people have had on the planet, scientists have named this period the Anthropocene. Unlike prior geological epochs, where diverse geological and climate processes characterized the time periods, the projected Anthropocene age is named for the primary impact humans and their actions are having on the environment. In essence, people in the world are a new geophysical force. (Dovers & Butler 2019.)

2.2.2 Gross domestic product

The gross domestic product (GDP) is a measure of economic activity that captures the market value of all final products and services—those created for direct consumption are produced throughout time in all economic sectors of a given nation for domestic use, export, and accumulation. It is frequently used as a benchmark for the state of the nation's economy, and when the GDP is increasing, both businesses and employees are happier. (Callen 2012.)

The population's income is an important factor. The environment is under pressure from both the lowest and highest income levels because of the unequal income distribution. Many of the world's poorest people use unsustainable amounts of resources, such as burning garbage, tires, or plastics to survive. They might even be forced to destroy natural resources such as forests or animal populations. On the other end of the spectrum, people with the greatest salaries tend to utilize an excessive number of resources because of the cars, houses, and lifestyle choices they make. (Dovers & Butler 2019.)

On a national level, environmental degradation and economic growth are also related. The least developed nations often have lower levels of industrial activity, which results in less environmental damage. Most developed countries have figured out how to advance technology and reduce their harmful environmental effects. The most environmental impact typically takes place in developing nations with high rates of resource consumption. (Dovers & Butler 2019.)

2.2.3 Fossil fuels consumption - gas, coal, oil

Although there is a link between poverty and environmental deterioration, unsustainable patterns of consumption and production, especially in industrialized nations, are much more concerning. All consumption has a big impact on the ecosystem. The mass production of many items — many of which are not necessary for living comfortably — uses a lot of energy. Furthermore, there is a lot of pollution and garbage being produced. (Dovers & Butler 2019.)

Excessive consumption has an impact on the environment that extends beyond the local area or even the country making issues worse. For instance, the amount of CO_2 in the atmosphere and the ensuing effects on the ecosystem are influenced by using fossil fuels for energy. Richer countries can rely on imports that use a lot of resources and produce waste in poorer countries. They can enjoy the products without having to worry about the pollution or short-term impacts of the industries that made them. (Dovers & Butler 2019.)

Figure 19 illustrates the absolute contribution of CO_2 emissions in Europe by source, including cement manufacturing, gas flaring, oil flaring, and coal. Industrial-scale coal-fired electricity was first introduced in North America and Europe in the 1700s. It was not until the late 1800s that the production of oil and gas increased emissions. Most fuels used today are solid and liquid, but gas production also plays an important role. Flaring and cement manufacturing are still fairly few on a global scale. (Ritchie, Roser & Rosado 2020.)

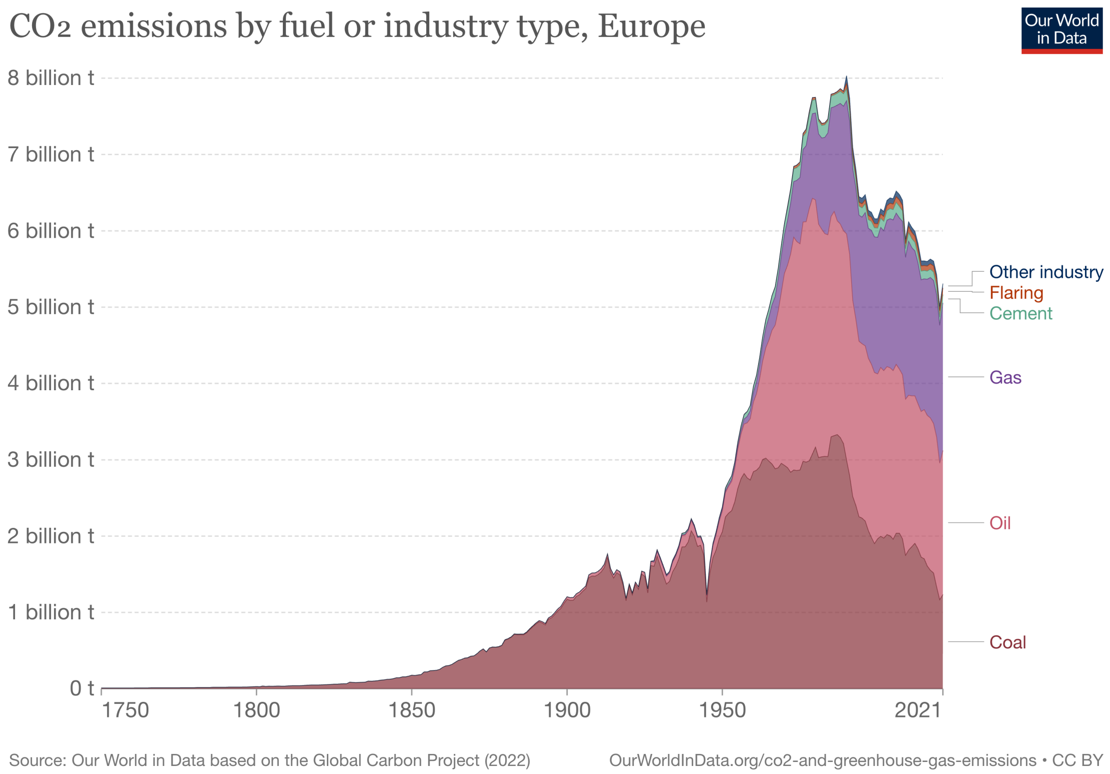


FIGURE 19. CO_2 emissions by fuel or industry type, Europe.

Not all people ~~on Earth~~ are equally responsible for the environment. In some parts of the world, where consumption and resource use patterns are excessive, the fundamental needs of entire communities are not being met. In such areas, population densities tend to be much higher.

2.2.4 Temperature

Air temperature has an impact on airflow and, as a result, on the movement of air pollution. Because the Earth's surface absorbs solar energy, air near the ground is warmer than air higher up in the troposphere. The cooler, heavier air in the higher troposphere sinks while the warmer, lighter air at the surface rises. (Center for Science Education 2021.)

The warm air near the ground often removes pollution from the atmosphere, but in the winter, this layer of warm air acts as a lid to keep cold air at the surface. Thermal inversion occurs in the winter when cool air and pollution are held close to the ground by a layer of warm air above. In hot, sunny weather, certain types of pollution, such as ground-level ozone, are produced more quickly. The chemical interactions that produce the hazardous ozone in our atmosphere require sunlight.

Throughout the summer, ozone frequently reaches dangerous levels in urban areas or close-by rural areas, especially during extreme heat waves. The soils might become extremely dry during a heat wave due to drought conditions and it is the time when forest fires occur more frequently. Fires cause both carbon monoxide and particle pollution. Figure 20 represents how rising, warm air frequently aids in the dispersion of pollution from near the surface. (Center for Science Education 2021.)

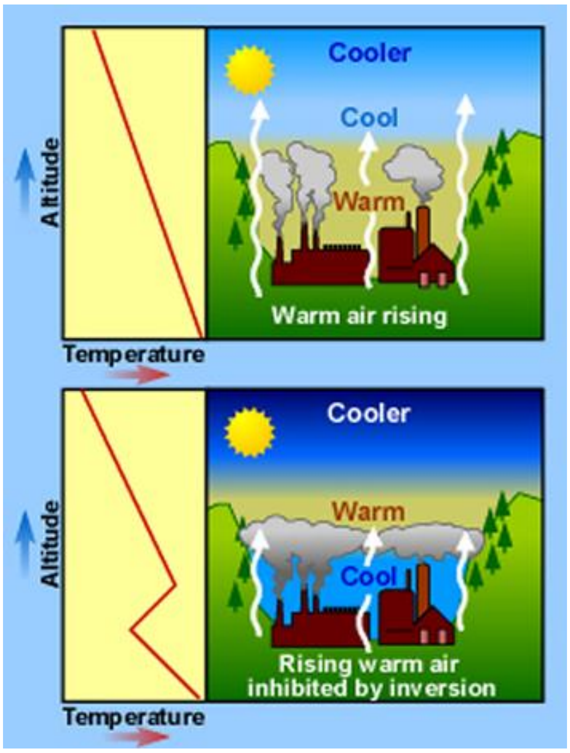


FIGURE 20. Temperature and air pollution.

2.3 Air pollution causes in Europe

The most common cause of air pollution in Europe is emissions from transport, especially from vehicles with diesel engines, which produce more pollutants than petrol such as soot (or black carbon), microscopic particles, gases, dust, sulfur dioxide (SO_2), and nitrogen dioxide (NO_2). Another big cause is power generation from coal power plants, and domestic heating with wood, coal, and organic waste which emit particulate matter ($PM_{2.5}$), nitrogen oxide (NO), sulfur dioxide (SO_2), and carbon dioxide (CO_2). The cause, which is related to the traffic, but does not produce as many pollutants are friction from breaking and tires, regardless of the fuel used, which produce particulate matter and microscopic rubber particles. Almost every country in the European region suffers from such causes.

In many parts of the United Kingdom, transportation is the main cause of air pollution. Both diesel and petrol vehicles emit nitrous oxide (N_2O) and particulate matter (PM) pollutants, with diesel emitting more of both than petrol. Other significant sources of air pollution in the UK include industry, farming, burning fuels in homes for heating or energy, and emissions from power generation, particularly those from fossil fuels such as coal power plants. (IQAir 2021.)

In the Netherlands, diesel engine combustion is the primary source of air pollution. Nitrogen dioxide (NO_2) is created when oxygen and nitrogen in the air combine during high-temperature combustion, usually during the burning of fossil fuels such as oil, diesel, and gas. It is exceedingly harmful and causes horrible illnesses in tens of thousands of Dutch. Since nitrogen dioxide is easier to detect than soot or particle matter, it is typically used to evaluate the severity of traffic-related air pollution. (IQAir 2020.)

The air in Austria includes 75% of fine dust that is harmful to human health and is caused by transportation. According to the most current studies, industrial operations account for 45% of dust emissions, followed by traffic (33%), bulk goods (21%), and other sources (1%). Nevertheless, half of the harmful fine dust (PM_{10} and $PM_{2.5}$) comes from traffic. (IQAir 2021.)

The main causes of air pollution in Germany include vehicle traffic, power plant emissions, industrial activity, fossil fuel heating, agriculture, and waste management. The mining, chemical, and metal sectors are part of Germany's large industrial sector, which is responsible for a sizable portion of emissions. Transport contributes significantly to air pollution in Germany emitting NO_X, CO, and PM_X pollutants. (IQAir 2020.)

There are many personal and heavy-duty vehicles on the road in Greece, which causes a lot of exhaust fumes to be in the air. These cars frequently burn fossil fuels such as diesel, which tend to produce more pollution than fuels from cleaner or more environmentally friendly sources. During the winter, power plants typically burn fossil fuels such as coal in Greece's northern regions which generates heat pollution. (IQAir 2020.)

The major influences on Belgium's air quality are emissions from vehicles, the auto sector, and food processing. Burning wood or solid fuel in a stove or using chemicals in agriculture also have an impact on air quality. Two-thirds of nitrogen dioxide (NO_2) emissions are attributable to transportation-related combustion engines (such as those found in cars, ships, and airplanes). The smallest particulate matter, elemental carbon (EC), is also greatly influenced by traffic. (IQAir 2021.)

In Spain's urban areas, where most of the population lives, car emissions are the main source of pollution. Road traffic, particularly diesel cars, is the source of more than 50% of nitrogen dioxide (NO_2) emissions. In some manufacturing areas and near major coal and oil thermoelectric facilities air quality is significantly decreasing. (IQAir 2021.)

In Finland, small-scale wood burning, and car emissions are the two largest sources of air pollution. Concentrations of air pollutants are higher near such sources than the urban average. Moreover, fine particles and some gaseous compounds effectively enter residential areas. The number of particulate emissions produced by international shipping is unimaginable according to Finnish scientists who monitored the ship's smoke booms. On the mainland, the total amount of particle emissions from buildings, companies, and automobiles is comparable to that of shipping. (IQAir 2022.)

Automobile emissions are among the major sources of pollution in Switzerland. Cars are a significant producer of hazardous chemical compounds such as sulfur dioxide (SO_2) and nitrogen dioxide (NO_2). Nitrogen dioxide is the most prevalent, in addition to increasing the concentration of PM_{10} and other fine particulate matter. (IQAir 2020.)

The estimated 750,000 wood-burning stoves in Denmark are the primary source of air pollution. The particles from these stoves result in roughly 550 fatalities annually. Particle pollution, which is brought on by the exhaust emissions from cars, trucks, and ships as well as by coal-fired power plants and wood smoke, is a major hazard to public health in Denmark. A significant additional source, accounting for 20% of total $PM_{2.5}$ emissions, is road traffic. (IQAir 2020.)

A prominent source of pollution in Sweden is the combination of $PM_{2.5}$ and PM_{10} particles. Burning wood for heating is the major cause of $PM_{2.5}$ emissions. Other important sources are domestic transportation and industry. The main causes of particulate emissions from road traffic into the atmosphere are deterioration of the road, tire abrasion, and vehicle braking. (IQAir 2020.)

Before the Czech Republic was formally "established" in 1993, the nation historically had severe air pollution problems. It concentrated on growing its industrial network while using low-quality, inexpensive coal as a fuel source. This lignite coal produces a significant amount of sulphur dioxide (SO_2), which is extremely detrimental to the environment. Residential heating with wood, coal, waste, or emissions from industry and traffic are the main causes of PM pollution. (IQAir 2020.)

There are numerous pollutants in France's air, each of which has a unique primary source of emission. The high levels of nitrogen dioxide (NO_2) emissions from cars and trucks in the atmosphere are used to directly assess air pollution. Another reason for air pollution is heating which requires burning fossil fuels and other possibly "unclean" fuels to achieve the desired results. Most of France's pollution is brought on by residential heating and roadways. (IQAir 2019.)

The number of severely polluting vehicles and overall emissions in Italy have been declining annually, but private urban road transportation remains one of the main sources of pollution. 16% of all $PM_{2.5}$ emissions are caused by transportation-related vehicles. The percentages of nitrogen oxide (NO) emissions, sulphur oxide (SO), and non-methane volatile organic compounds (NMVOCs) emissions (mainly benzene) are 5,66%, 18%, and 20%, respectively. (IQAir 2020.)

The main factor contributing to most of Poland's air pollution is the country's reliance on coal for energy production. The local economy continues to rely heavily on the coal industry. The second-largest coal-mining nation in Europe is Poland. The overall use of coal has decreased, while the use of natural gas as a fuel source has gradually increased, since the 1980s. As in numerous other countries, vehicles are a major source of pollution in Poland. (IQAir 2022.)

In Norway, nitrogen dioxide (NO_2) and particulate matter ($PM_{2.5}$ and PM_{10}) are the two main causes of air pollution. In addition to tires and exhaust emissions, combustion in motors, stoves, and fireplaces can also produce airborne dust. Although studded tires are inferior to standard tires, they are frequently used in Scandinavian regions where it snows a lot throughout the winter. (IQAir 2022.)

For the past few years, large levels of $PM_{2.5}$ pollutants have been produced in Hungary by a variety of pollution sources. The pollution from the millions of cars is one of the sources. Older automobiles and motorcycles that burn fuel inefficiently emit significantly more smoke and chemical emissions than modern vehicles. The "chimney smoke" epidemic in rural Hungary, where hundreds of thousands of traditional homes burn wood and coal in their fireplaces and stoves is one of the major causes of pollution. Another significant source of air pollution in Hungary is the vast amount of coal that power plants and companies consume to supply energy to the numerous homes around the country. (IQAir 2020.)

Large trucks and buses, which can make major contributions to air pollution, are the primary sources of traffic pollution in Turkey. Domestic heating and cooking are other sources of pollution since they typically include the use of substances that emit a lot of smoke and soot, such as coal, wood, or other dead organic waste, which contributes to pollution and haze. (IQAir 2019.)

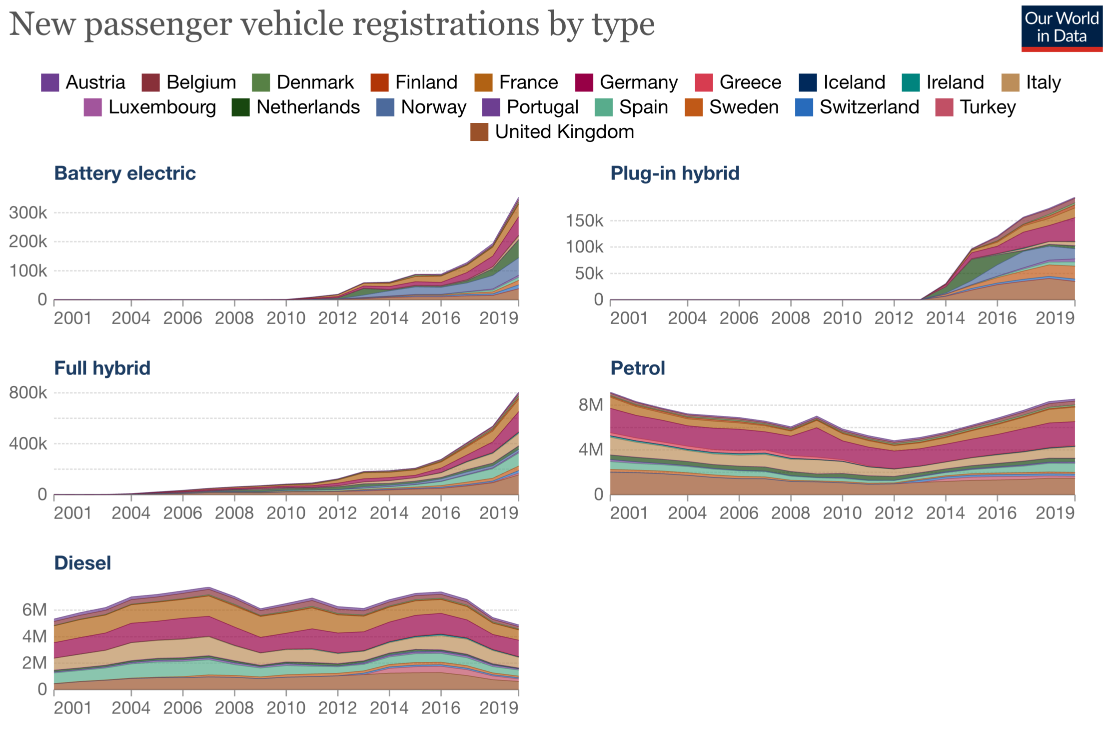
2.4 Engine type fraction per country

Concerns over the emissions from automobiles have grown in recent years. Due to the better fuel efficiency and less frequent maintenance requirements, the auto industry promotes diesel automobiles as being cleaner than petrol vehicles. Although diesel cars differ greatly from petrol cars in many ways, increasing the quantity of diesel cars at the expense of petrol cars might have significant effects on air quality, smog formation, global warming, and other environmental issues. (Airquality UK 2022.)

The interactive charts in Appendix 2 show the breakdown of new passenger vehicle registrations by type from 2001 to 2019 for 19 European countries (Ritchie, Roser & Rosado 2022).

Type is broken down by:
- petroleum.
- diesel.
- full hybrid (excluding plug-in hybrids).
- plug-in electric hybrids.
- fully electric battery vehicles.

Figure 21 represents the data about new passenger vehicle registrations by type from 2001 to 2019 for 19 European countries in the data.



Source: International Council on Clean Transport (ICCT) and European Environment Agency

OurWorldInData.org/transport • CC BY

FIGURE 21. New passenger vehicle registrations by type.

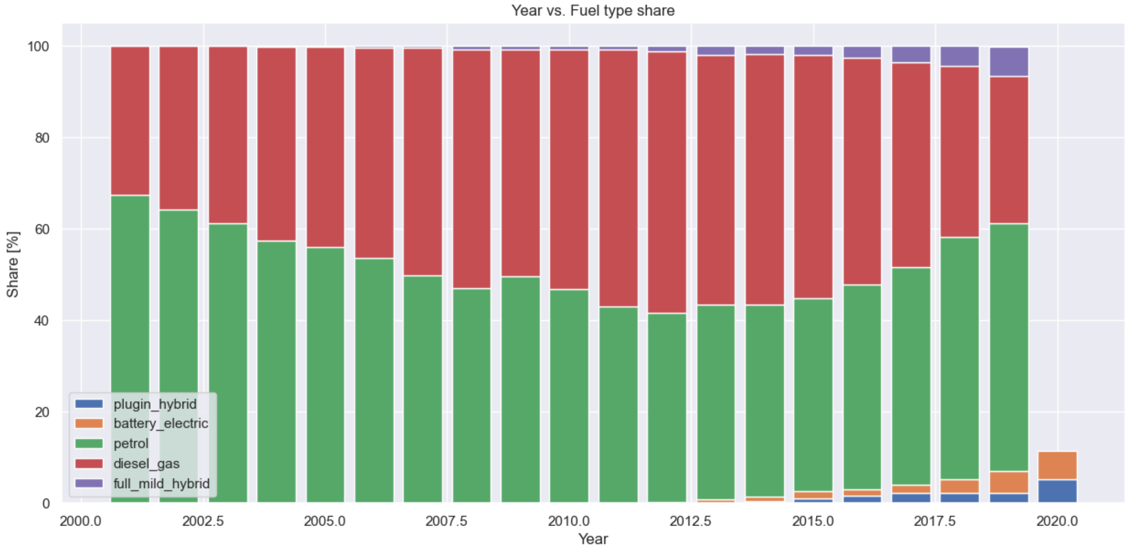


FIGURE 22. New passenger vehicle registrations fuel type share.

As seen in Figure 22, the amount of petrol vehicle types was more than 60% in 2000 out of all types of vehicles, then started to decline in 2012 was less than 50% but started to increase again with about 60% in 2019. The amount of diesel vehicle types was less than 40% in 2000 out of all types of vehicles, then started to increase and in 2012 was more than 50%, but started to decrease again, with less than 40% in 2019. The number of other types of vehicles (full hybrid (excluding plug-in hybrids), plug-in electric hybrids, and fully electric battery vehicles) in all countries was less than 10% in 2019 but is increasing since 2012.

2.5 Data cleaning and quality

Data cleaning is known as the process of removing inaccurate, flawed, poorly formatted, duplicate, or insufficient data from a dataset. There are multiple ways for data to be duplicated or incorrectly categorized while merging several data sources. Results and algorithms are unreliable despite appearing to be correct due to uncleaned data. The process of changing data from one format or structure to another is known as data transformation. Transformation activities involve the transition of data from one "raw" data type into another format for analysis. Although there are basic steps for organizing the data structure, the methods used to clean up the data may vary depending on the types of data the company is storing. (Tableu 2022.)

There are five characteristics of quality data (Tableu 2022):

1. Validity is the degree to which the data conforms to defined business rules or constraints.
2. Accuracy ensures the data is close to the true values.
3. Completeness is the degree to which all required data is known.
4. Consistency ensures the data is consistent within the same dataset and/or across multiple data sets.
5. Uniformity is the degree to which the data is specified using the same unit of measure.

Constraints must be applied during the data validation process to guarantee the data is accurate and consistent. The amount of data cleansing is reduced due to various data validation limitations (Scribbr 2022):

- Data-type constraints: values can only be accepted if they are of a certain type, such as numbers or text.
- Range constraints: values must fall within a certain range to be valid.
- Mandatory constraints: a value must be entered.

Data screening is known as the process of examining the dataset for inconsistent, false, omitted, or outlier data which can be done manually or by statistical techniques for example straightening up the dataset so that the data is organized and easy to understand, visually scan the data for possible discrepancies, use statistical techniques and tables/graphs to explore data. These are the general problems with “bad” data such as duplicate data, invalid data, missing values, and outliers. (Scribbr 2022.)

The thesis project uses several data sets from different sources which were merged into one: air quality annual statistics data gathered from AirBase & e-Reporting datasets, population, GDP per capita, fossil fuel consumption, fossil fuel consumption by type (gas, coal, oil), average near-surface temperatures in Europe.

The air quality annual statistics calculated by the European Environment Agency needed several cleaning steps to ensure high data quality for further analysis. Each column consisted of numeric or text types of values, which is acceptable. Some columns such as ‘Unit Of Air Pollution Level’ and ‘Air Quality Station Area’ showed inconsistency and needed to be fixed. The ‘Unit Of Air Pollution Level’ column contained values such as ng/m3 and ng.m-3. It was fixed by replacing wrong symbols in values with correct ones. The ‘Air Quality Station Area’ column contained values such as Urban and urban, Suburban and suburban, and Rural and rural. It was fixed by setting all the values to lowercase. The range of the numeric columns ‘Year’ and ‘'City Population' did not have any incorrectness in the data. The ‘Air Pollution Level’ numeric column contained values below zero, which is not acceptable, and it was fixed by removing rows with incorrect data. Mandatory constraints of the ‘Air Pollution Level’ column contained several NaN values, which is not acceptable, and it was fixed by removing such rows. Duplicates in the data were not found. The data was narrowed to six main pollutants data ($PM_{2.5}$, PM_{10}, O_3, CO, SO_2, NO_2) and the annual values of these pollutants. The data were grouped with the calculation of the mean ‘Air pollution level’ value for each year in each country for all air pollutants. Outliers were found in the carbon monoxide (CO) data with pollution levels above 100 mg/m^3. Rows with such values were removed from the data for better prediction. 'Unit Of Air Pollution Level' contained values in the 'mg/m3' (milli) unit for CO and '$\mu g/m^3$' (micro) unit for $PM_{2.5}$, PM_{10}, O_3, SO_2, NO_2. However, it was decided not to change it, but just convert the values of CO from milli to micro in conclusions.

The population data and GDP per capita by The World Bank did not have cleaning issues; however, they needed several changes to fit the main dataset before merging: the 1960-2021 columns were transformed into rows, the ‘Country Name’ column was renamed to ‘Country’ and 'Country Code’, and 'Indicator Name', 'Indicator Code' columns were dropped.

The fossil fuel consumption data and fossil fuel consumption by fuel type by Our World in Data did not have cleaning issues; however, they did not contain data about Lichtenstein, Andorra, Albania, Malta, Kosovo, Serbia, Bosnia and Herzegovina, Montenegro. The ‘Code’ column was dropped, the ‘Entity’ column was renamed to ‘Country’, and NaN values of 'Gas Consumption - TWh', 'Coal Consumption - TWh', and 'Oil Consumption - TWh' columns were dropped.

The temperature data by European Environment Agency did not have many cleaning issues. It was decided to drop the 'HadCRUT5', 'HadCRUT5max', 'HadCRUT5min', 'NOAAGlobalTemp', 'GISTEMP', 'ERA5', 'JRA-55' columns leaving only the ‘BerkeleyEarth’ column and rename 'Unnamed: 0' column to ‘Year’.

All datasets contained information about the same countries, but the names of the countries were used in different forms in each data table. To ensure that no country was overlooked in the analysis, the names of the countries were given in the same form. The countries' names in the population and GDP per capita data tables were taken as a basic form. Therefore, 'Lichtenstein’ was set as ‘Liechtenstein', 'Kosovo under UNSCR 1244/99' was set as 'Kosovo', 'Turkey' was set as 'Turkey', 'Slovakia' was set as 'Slovak Republic' for consistency.

After data merging the year range of data was 1973 – 2021 and the merged data contained information about 32 countries.

3 DESCRIPTION OF METHODS

3.1 Supervised learning

The term "machine learning" describes the computer algorithms which are used to learn from data and generate predictions about unknown future data. Machine learning can be divided into two categories: supervised learning and unsupervised learning. Supervised learning is when the algorithm learns the rule for converting the input variables into the target variable using labelled training data. Unsupervised learning is when the algorithm learns associative rules for the data using unlabelled training data. (Loy 2019, 7.)

The usual supervised learning approach is shown in Figure 23, where labelled training data are fed into a machine learning algorithm to form a predictive model that can forecast new, unlabelled data inputs. (Dzhulgakov et al 2022, 3.)

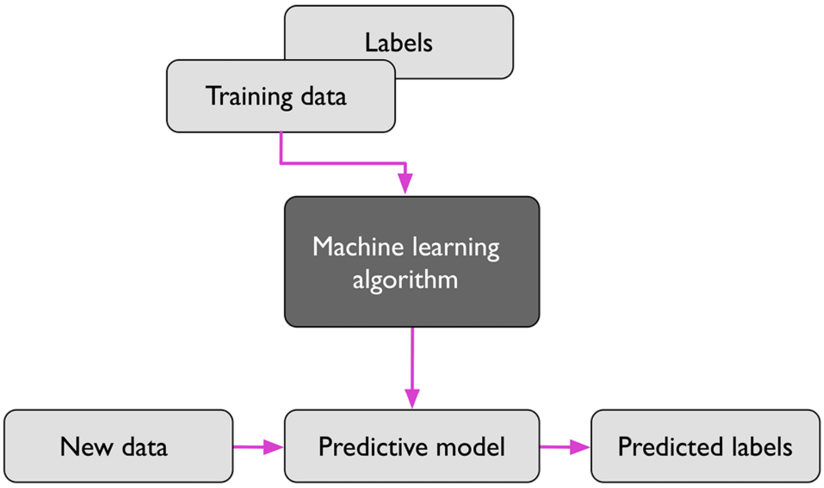


FIGURE 23. Supervised learning process.

3.2 Training, testing and validation sets

Splitting the data into training, testing, and validation sets prevents overfitting and offers an unbiased supply of data for evaluating model performance (Loy 2019, 59):

- Training set is used to train the neural network.

- Validation set is used to adjust the number of hidden layers by doing hyperparameter tuning.
- Testing set is used as the foundation for the neural network's final assessment.

The training and validation data are usually used to improve the model. The validation set can be utilized until the model performance on the validation set stops getting better. This enables the prevention of the neural network from overfitting. The neural network will never be trained using the testing set, and it is sometimes referred to as the holdout dataset. The model is using the testing set later to give a precise picture of how well the model performs. (Loy 2019, 59.)

Generally, the training data should be divided into 80% training and 20% testing, and the testing data into 80% training and 20% validation. This technique is shown in Figure 24 (Loy 2019, 60):

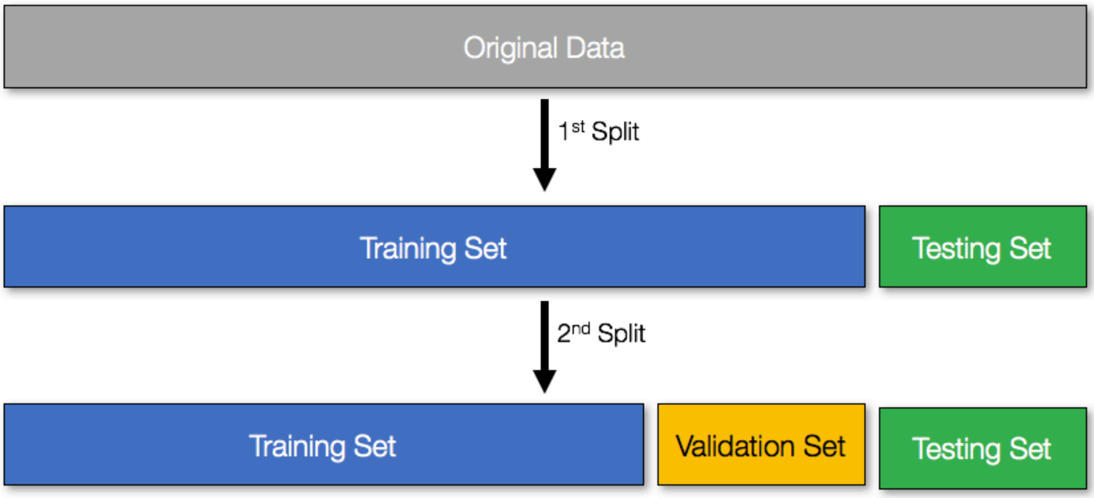


FIGURE 24. Split the data process.

3.3 Data standardizing

Standardizing the data in neural networks is crucial to ensuring that the algorithm performs as planned. Data normalization has the additional benefit of reducing the size of the variables and changing their scale to one that is more proportionate. (Loy 2019, 58.) Standard scaling is the process of changing the features' mean to 0 and standard deviation to 1, which transforms the features into a normal distribution. The following operation is used to do this, where each value of observation is deducted from its mean value before the result is divided by the standard deviation value (Galli 2022, 242.):

$$x_{scaled} = \frac{x - mean(x)}{std(x)}$$

3.4 Regression analysis

Modeling the relationship between one or more characteristics and a continuous target variable is the aim of linear regression. Regression analysis is a separate subsection of supervised learning, which forecasts outcomes on a continuous scale. Simple linear regression seeks to represent the relationship between a single feature (the explanatory variable, x) and a continuous-valued target (response variable, y). The equation of a simple linear regression is defined below, where the parameter (bias unit), b, represents the y-axis intercept and w_1 is the weight coefficient of the explanatory variable (Dzhulgakov et al 2022, 269):

$$y = w_1 x + b$$

The objective is to learn the weights of the linear equation that characterizes the relationship between the explanatory variable and the target variable to forecast the outcomes of additional factors that were not included in the training dataset. As seen in Figure 25, linear regression can be understood as locating the straight line that fits the training examples the best. (Dzhulgakov et al 2022, 269-270.)

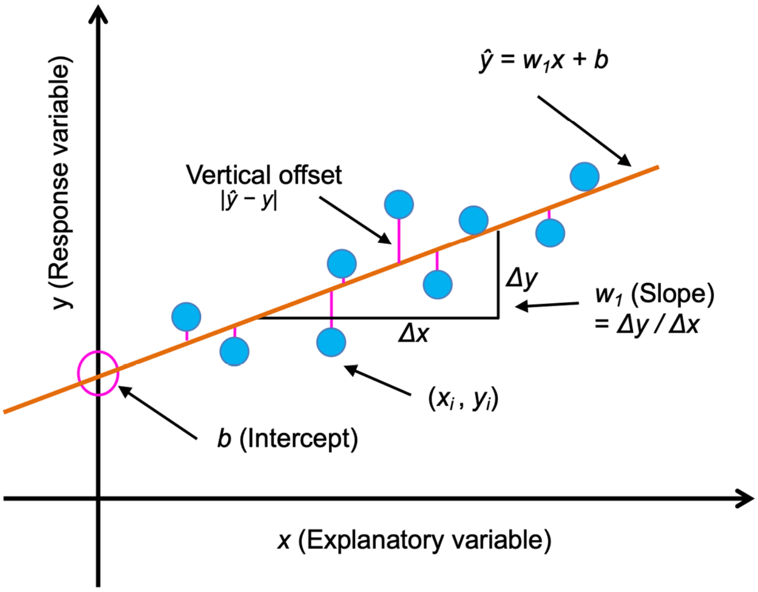


FIGURE 25. A simple one-feature linear regression example.

The linear regression model can be also generalized into multiple explanatory variables which is called multiple linear regression (Dzhulgakov et al 2022 271):

$$y = w_1 x_1 + \cdots + w_m x_m + b = \sum_{i=1}^{m} w_i x_i + b = w^T x + b$$

The same concepts and evaluation methods underlie simple and multiple linear regression. The two-dimensional, fitted hyperplane of a multivariate linear regression model with two features is shown in Figure 26:

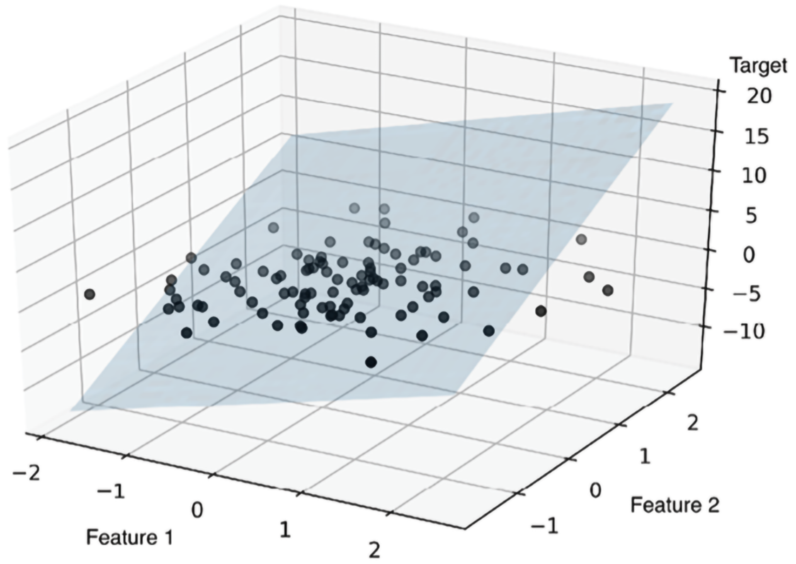


FIGURE 26. A two-feature linear regression model.

3.5 Neural network

The input layer, hidden layer, and output layer make up a basic neural network, as shown in Figure 27:

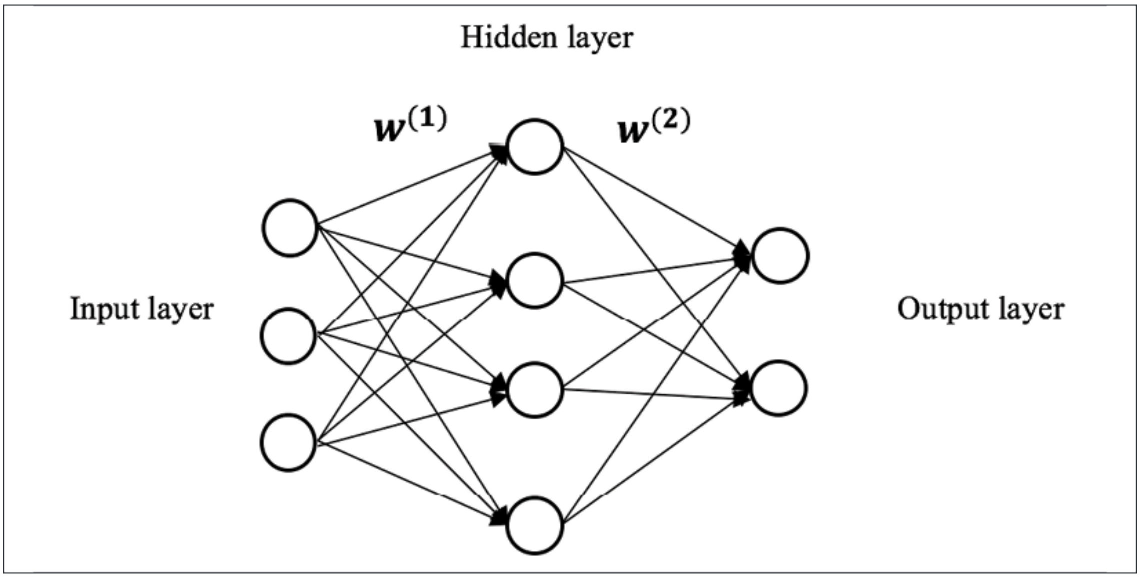


FIGURE 27. A simple shallow neural network.

A layer is a group of nodes, also known as units, that represent the neurons in a biological brain. Each node is a predictive feature, and the input layer represents the input features, x. The goal variable(s) is represented by the output layer. Regression uses a single node in the output layer, which value is the outcome of the prediction. Signals are passed from one neuron in one layer to another neuron in the following layer via conceptual edges. The model's weights, w, are used to parameterize the edges. (Liu 2020, 254.)

A neural network often has several hidden layers when it is used in applications. In Figure 27 $w^{(1)}$ links the input and hidden layers, while $w^{(2)}$ links the hidden and output layers. Data is sent from the input layer to the output layer through a hidden layer in a typical neural network(s). Because of that, it is called a feedforward neural type of network. The fundamental relationship between the input data and the target is easier to learn for neural networks with one or more hidden layers between the input and output layers. (Liu 2020, 255.)

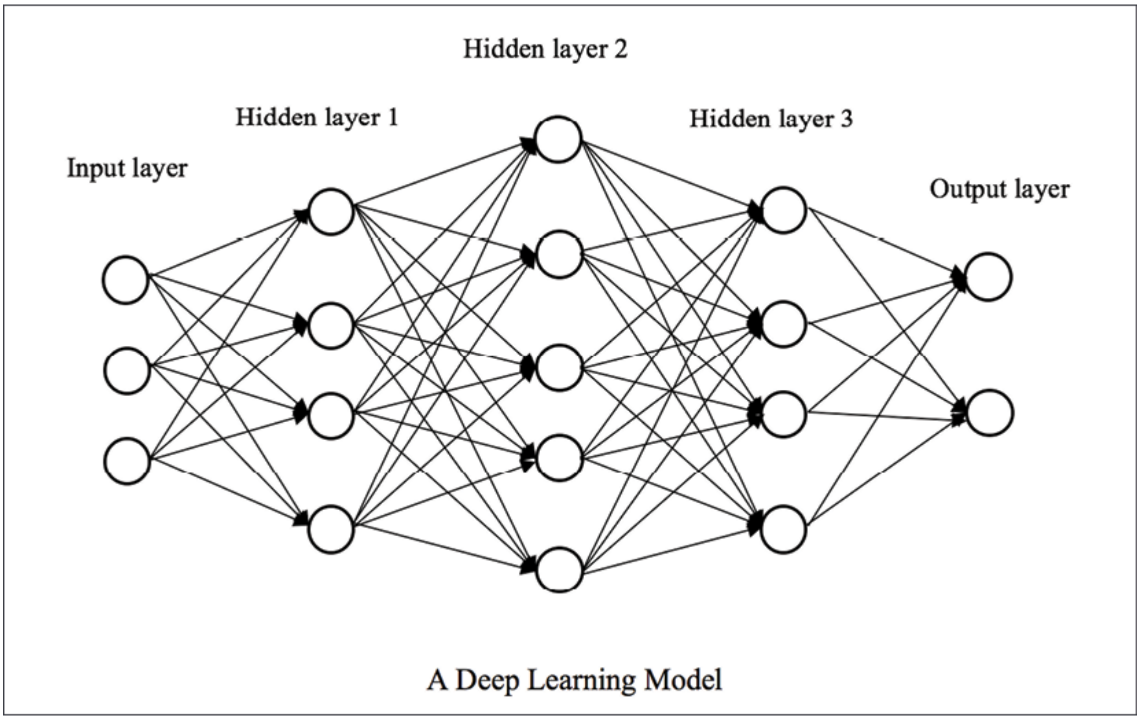


FIGURE 28. A deep neural network.

As shown in Figure 28, each hidden layer's input is the result of the layer before it in a stack of numerous hidden layers. Each hidden layer is extracted for features. Patterns from many levels are represented by features from various layers. Beyond shallow neural networks, which typically have one hidden layer, a deep learning model, which often has two or more hidden layers, can learn complicated non-linear correlations from data more effectively. (Liu 2020, 260.)

3.6 Multilayer perceptron analysis

A multilayer perceptron neural network has multiple layers, as the name suggests. The basic structure is unchanged - there must be a layer for accepting input data and another for producing output values. Hidden layers are those that model the data to distribute various weights and biases. MLP has three layers: the hidden layer, which modifies the values, the input layer, which accepts input values, and the output layer, which causes an output value. A set of input values are sent to the MLP model, which changes them and causes output values to be generated in the output layer. (Nokeri 2021, 158.)

The data training process for this regressor is iterative. To update the parameters, it calculates the partial derivatives of the loss function with respect to the model parameters at each step. To minimize the model parameters and prevent data overfitting, a regularisation term is introduced to the loss function. A multilayer perceptron regressor is implemented via the MLPRegressor() function. (Ciaburro 2022, 337.)

3.7 The coefficient of determination

The process of determining if the model used in the study is acceptable is the most crucial stage in any empirical investigation. While performing an empirical study, it is important to follow the procedures before the study and verify results once the regression is carried out. The first approach is to calculate the mean of the dependent variable. The second approach is based on regression, which can account for some of the variances in the dependent variable, in addition to the mean of the dependent variable. The sum of squares total (SST) provides the total variation of the observations for the dependent variable and is divided into two parts: the part that is explained by the regression model (sum of squares regression SSR) and the part that is not explained by the regression model (residual sum of squares or SSE for the sum of squares error). (Naghshpour 2016, 67.)

The coefficient of determination, or R^2, is a measure of how well the model matches the data. It is the ratio of the total sum of squares (SST), which represents the overall variation in the dependent variable, to the sum of squares of regression (SSR), which represents the fraction of variation explained by the regression (Naghshpour 2016, 76):

$$R^2 = \frac{SSR}{SST}$$

R^2 is typically stated as a percentage. Depending on the research, it falls between zero and one, or between 0% and 100%. Higher values of R^2 imply a model's greater ability to explain the data. A high R^2 does not always equal a good fit, there are some instances where it does not. R^2 has the drawback of not accounting for the number of variables used in its calculation, which causes it to rise as more variables are included in the model, sometimes even independent of their ability to explain dependent variables. (Naghshpour 2016, 76.)

3.8 Python tools

Python has grown in favour among data engineers, scientists, and Machine Learning (ML) admires over the past few years despite being a more general-purpose programming and scripting language. The building of ML models can be done interactively in well-designed programming environments such as Jupyter Notebook, which allows for speedy analysis of the data. Numerical data may be used effectively with the help of robust modules such as NumPy and Pandas. Using the SciPy library is best for scientific computation. In scikit-learn (also known as sklearn), several fundamental ML methods have been effectively implemented. These have persuaded many ML developers to choose Python as the language for data exploration, pattern recognition, model building, and model deployment to the production environment. (Swamynathan 2019, 3)

In 2019, Kaggle conducted an annual survey for data scientists and machine learning experts globally. This survey is one of our most trustworthy sources of information on the state of the sector. The percentage of usage for various machine learning software frameworks is shown in Figure 29. (Chollet 2021, 19.)

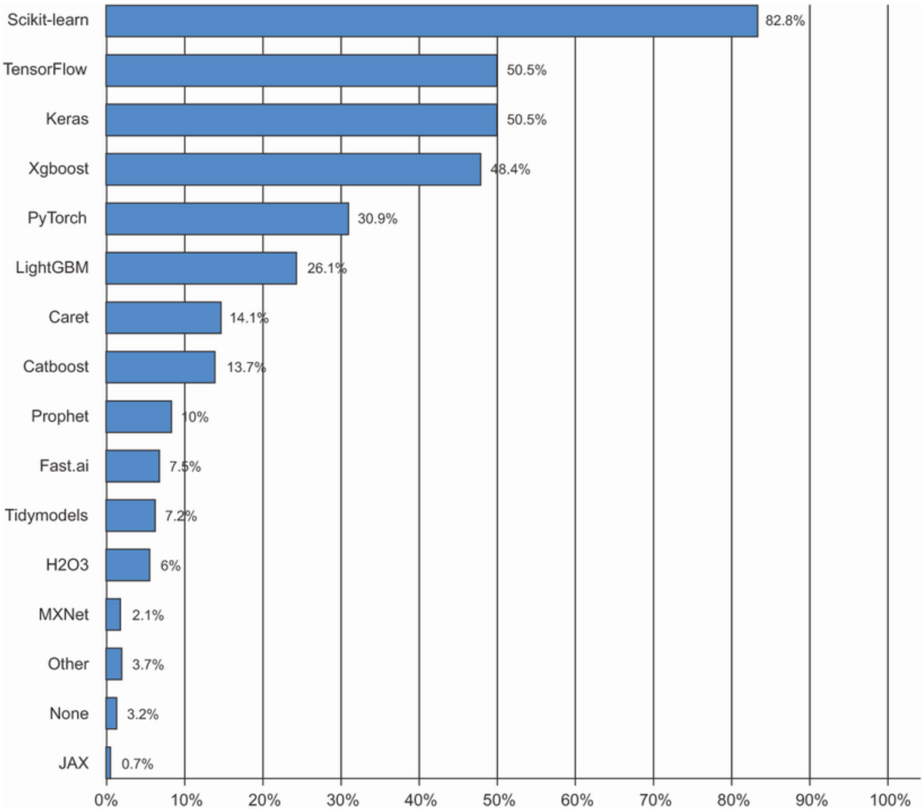


FIGURE 29. Tool usage across the machine learning and data science industry.

The thesis work used the scikit-learn software machine learning library to carry out predictive data analysis. It offers a variety of effective methods for statistical modelling and machine learning, including regression. This library is based on NumPy, SciPy, and Matplotlib and was written primarily in Python.

Two types of models were used in the thesis: LinearRegression for data with one target and one feature and MLPRegressor (neural network) for data with one target but multiple numbers of features. The data were scaled using sklearn.preprocessing.StandardScaler which standardizes features by removing the mean and scaling to unit variance. The convenient sklearn.model_selection.train_test_split method was used to randomly split the data to train and test subsets. The R^2 score was used to evaluate the model's performance and eli5.show_weights were used to explain the weights of the model features.

4 AIR POLLUTION MODELLING RESULTS

The objective of this thesis was to describe air pollution and make a forecast of pollutant levels in Europe for the next ten years from 2022 to 2032. The factors of air pollution were defined, and their levels were predicted using LinearRegression models. The main six air pollutants ($PM_{2.5}$, PM_{10}, O_3, CO, SO_2, NO_2) were selected, and their levels were predicted using MLPRegressor models based on the factor's levels.

LinearRegression model datasets were separated into 80% for train and 20% for test sets. MLPRegressor model datasets were separated into 80% for train and 20% for test sets, or 90% for train and 10% for test sets for a better prediction. The MLPRegressor models used four hidden layers with a size of 100 and an iteration size of 500 for each pollutant. The models were evaluated by the R^2 score to express how well the model explained the data and most of them predicted data with 80% accuracy.

4.1 Factors models

The models' settings and R^2 scores for each factor are presented in Table 10.

TABLE 10. Factors forecast settings.

Pollution factor	Test size	Model type	R^2 score train/test set
Population	0.2	LinearRegression	0.98/0.98
GDP per capita	0.2	LinearRegression	0.94/0.88
Fossil fuel consumption	0.2	LinearRegression	0.61/0.72
Gas consumption	0.2	LinearRegression	0.85/0.80
Coal consumption	0.2	LinearRegression	0.87/0.93
Oil consumption	0.2	LinearRegression	0.60/0.75
Earth temperature	0.2	LinearRegression	0.61/0.73

The population, GDP per capita, gas consumption, and coal consumption models have R^2 scores around 0.9 which means that models explain more than 90% of the factors data. The fossil fuel, oil consumption, and Earth temperature models have R^2 scores around 0.7 which means that models explain around 70% of the factors data and it is not as good as for other factors but still is reliable.

An example of visual evaluation is presented in Figure 30 for the population forecast. The charts show how well the predicted values fit the data.

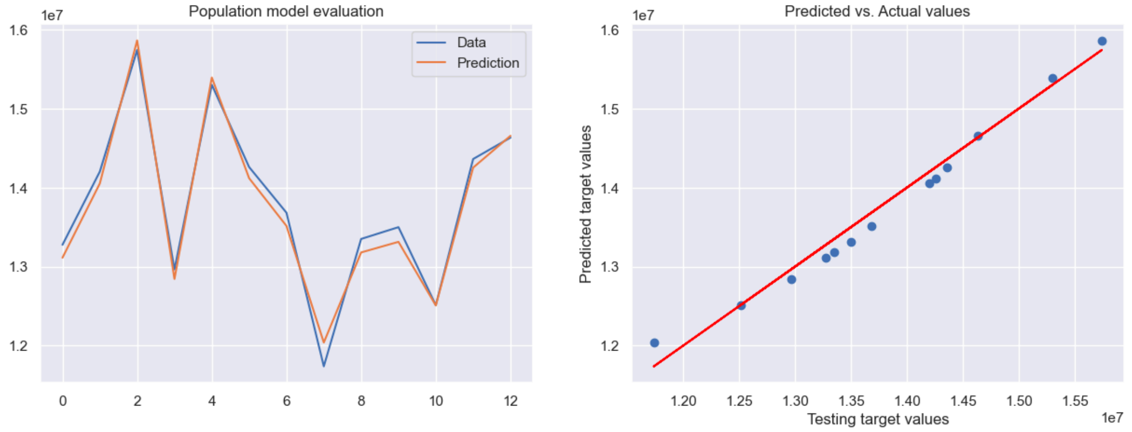


FIGURE 30. Population linear regression model evaluation.

The predicted values of the factors are presented in Table 11.

TABLE 11. Pollution factors prediction results.

Pollution factor	Trend	2022	2032
Population	Increasing	16000000	16700000
GDP per capita	Increasing	37000 (US$)	45000 (US$)
Fossil fuel consumption	Decreasing	490 (TWh)	460 (TWh)
Gas consumption	Increasing	190 (TWh)	220 (TWh)
Coal consumption	Decreasing	80 (TWh)	52 (TWh)
Oil consumption	Decreasing	240 (TWh)	220 (TWh)
Earth temperature	Increasing	1.9 (°C)	2.2 (°C)

Since fossil fuel consumption is a major contributor to pollution, it is important to note the decrease in total fossil fuel consumption, mainly coal and oil consumption. At the same time, the gas consumption will continue to increase. However, even with an increasing trend, the gas consumption level will still be below or equal to the level of oil consumption. The charts of gas, coal, and oil model prediction results are presented in Figures 31, 32, and 33.

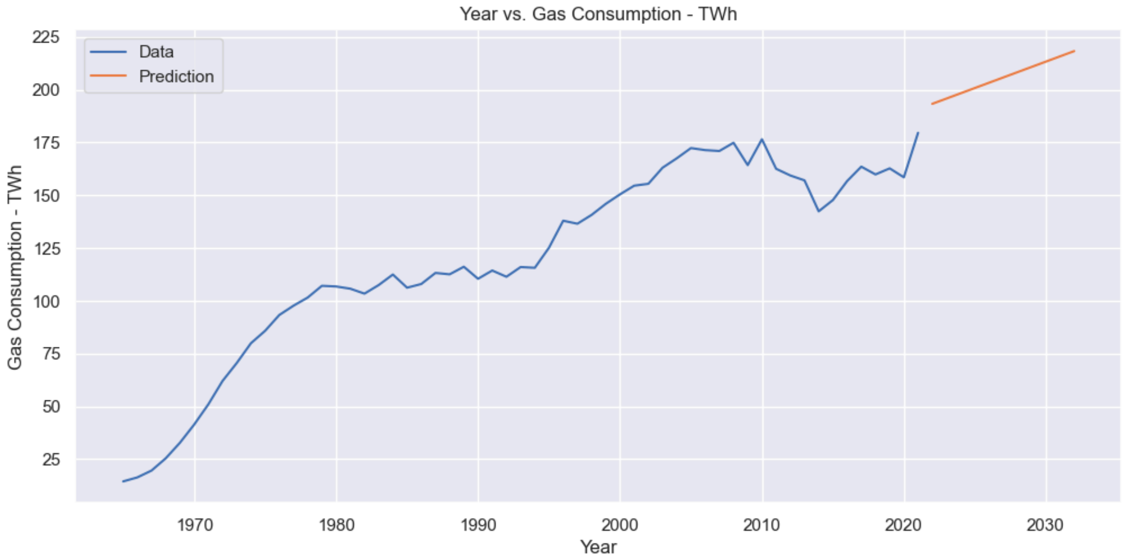


FIGURE 31. Gas consumption trend from 1960 to 2032.

FIGURE 32. Coal consumption trend from 1960 to 2032.

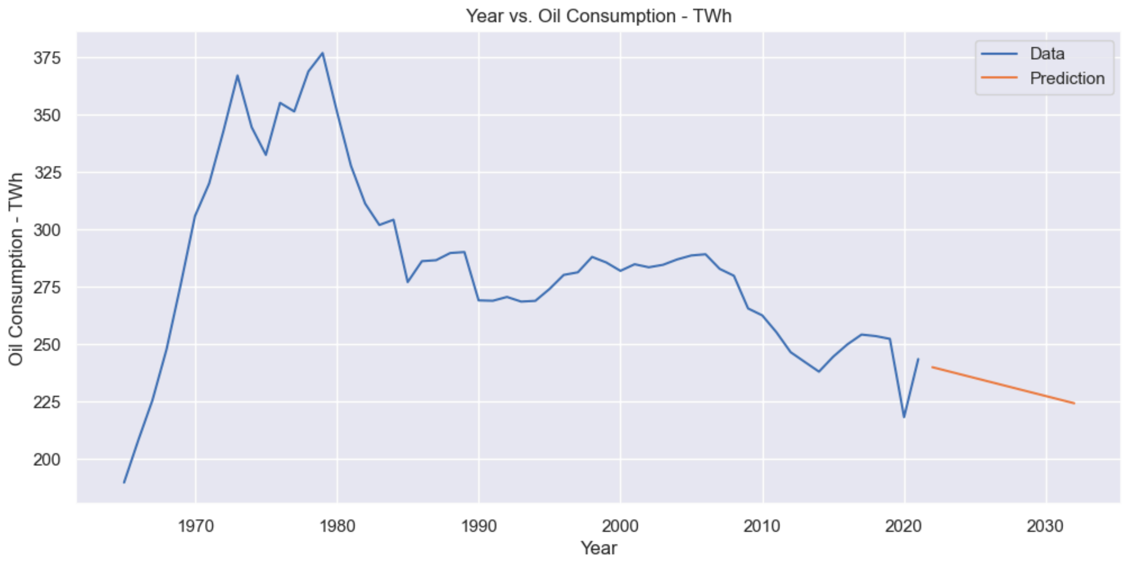


FIGURE 33. Oil consumption trend from 1960 to 2032.

4.2 Pollutants models

The model settings and R^2 scores for the pollutant are presented in Table 12.

TABLE 12. Pollutant forecast settings.

Pollutant	Test size	Model type	Hidden layers sizes	Iteration size	R^2 score train/test set
$PM_{2.5}$	0.2	MLPRegressor	(100,100,100,100)	500	0.82/0.67
PM_{10}	0.1	MLPRegressor	(100,100,100,100)	500	0.90/0.74
O_3	0.1	MLPRegressor	(100,100,100,100)	500	0.62/0.43
CO	0.2	MLPRegressor	(100,100,100,100)	500	0.87/0.34
SO_2	0.2	MLPRegressor	(100,100,100,100)	500	0.97/0.67
NO_2	0.2	MLPRegressor	(100,100,100,100)	500	0.88/0.57

Almost all the models have R^2 scores around 0.9 which means that models explain more than 90% of the factors data which makes the prediction reliable.

Different scenarios were created to test the stability of the models and how they work with different parameters. Firstly, the most probable pollutant values were calculated using previously predicted levels of GDP per capita, population, fossil fuel consumption, gas consumption, coal consumption, oil consumption, and temperature. These values were compared with the Year Average Common Air Quality Index (YACAQI). Then various scenarios were created in which these factors varied either increasing or decreasing in the range of 1 to 2. Average levels of each pollutant from 2022 to 2032 of each scenario were calculated and a range of results was determined to see how the models fluctuate. This method showed that:

- The $PM_{2.5}$ pollutant model seems to be stable in the range of 4 to 22 $\mu g/m^3$.
- The PM_{10} pollutant model seems to be stable in the range of 10 to 28 $\mu g/m^3$.
- The O_3 pollutant model seems to be stable in the range of 43 to 65 $\mu g/m^3$.
- The CO pollutant model seems to be stable in the range from 0.06 to 0.35 mg/m^3.
- The SO_2 pollutant model seems to be stable in the range from 1 to 6 $\mu g/m^3$.
- The NO_2 pollutant model seems to be stable in the range of 6 to 20 $\mu g/m^3$.

The thesis work predicted pollutants levels in 30 different scenarios:

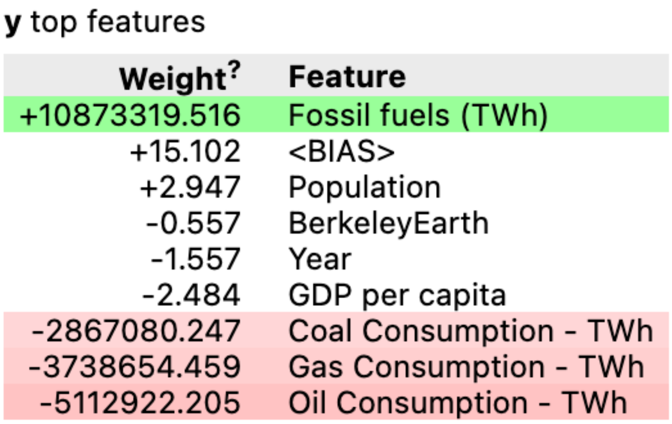
1. The population is decreasing, and gas consumption is increasing.
2. The population is increasing, and gas consumption is decreasing.
3. The population is decreasing, and coal consumption is increasing.
4. The population is increasing, and coal consumption is decreasing.
5. The population is decreasing, and oil consumption is increasing.
6. The population is increasing, and oil consumption is decreasing.
7. The GDP per capita is decreasing, and gas consumption is increasing.
8. The GDP per capita is increasing, and gas consumption is decreasing.
9. The GDP per capita is decreasing, and coal consumption is increasing.
10. The GDP per capita is increasing, and coal consumption is decreasing.
11. The GDP per capita is decreasing, and oil consumption is increasing.
12. The GDP per capita is increasing, and oil consumption is decreasing.
13. The Earth's temperature is decreasing, and gas consumption is increasing.
14. The Earth's temperature is increasing, and gas consumption is decreasing.
15. The Earth's temperature is decreasing, and coal consumption is increasing.
16. The Earth's temperature is increasing, and coal consumption is decreasing.
17. The Earth's temperature is decreasing, and oil consumption is increasing.

18. The Earth's temperature is increasing, and oil consumption is decreasing.
19. Coal consumption is decreasing, and gas consumption is increasing.
20. Coal consumption is increasing, and gas consumption is decreasing.
21. Oil consumption is decreasing, and gas consumption is increasing.
22. Oil consumption is increasing, and gas consumption is decreasing.
23. Oil consumption is decreasing, and coal consumption is increasing.
24. Oil consumption is increasing, and coal consumption is decreasing.
25. The population is increasing, and the GDP per capita is decreasing.
26. The population is decreasing, and GDP per capita is increasing.
27. The population is increasing, and the Earth's temperature is decreasing.
28. The population is decreasing, and the Earth's temperature is increasing.
29. The GDP per capita is increasing, and the Earth's temperature is decreasing.
30. The GDP per capita is decreasing, and the Earth's temperature is increasing.

These scenarios show to which direction the level of pollution will go, depending on the level (1 to 2) of change in the factors.

4.2.1 Particulate matter and nitrogen dioxide models

Figure 34 shows an example of the factor weights for the PM$_{2.5}$, PM$_{10}$, and NO$_2$ pollutants prediction models. The fossil fuel factor has the greatest weight, while coal, gas, and oil consumption factors have less weight. This suggests that the most difference in pollution comes from the difference in total fossil fuel consumption.



y top features

Weight?	Feature
+10873319.516	Fossil fuels (TWh)
+15.102	<BIAS>
+2.947	Population
-0.557	BerkeleyEarth
-1.557	Year
-2.484	GDP per capita
-2867080.247	Coal Consumption - TWh
-3738654.459	Gas Consumption - TWh
-5112922.205	Oil Consumption - TWh

FIGURE 34. Factors weights for PM$_{2.5}$ pollutant prediction.

The most probable scenario forecasts for $PM_{2.5}$, PM_{10}, and NO_2 pollution are presented in Figures 35, 36, and 37 respectively. The level of $PM_{2.5}$ is going to decrease to 10 µg/m^3, the level of PM_{10} is going to decrease to 13 µg/m^3, and the level of NO_2 is going to decrease to 12.5 µg/m^3. These results can be explained by the fact that fossil fuel consumption which is the main factor for these models is predicted to decrease as well. Based on the Year Average Common Air Quality Index (YACAQI) levels of $PM_{2.5}$, $PM_{10,}$ and NO_2 pollutants are below the annual mean.

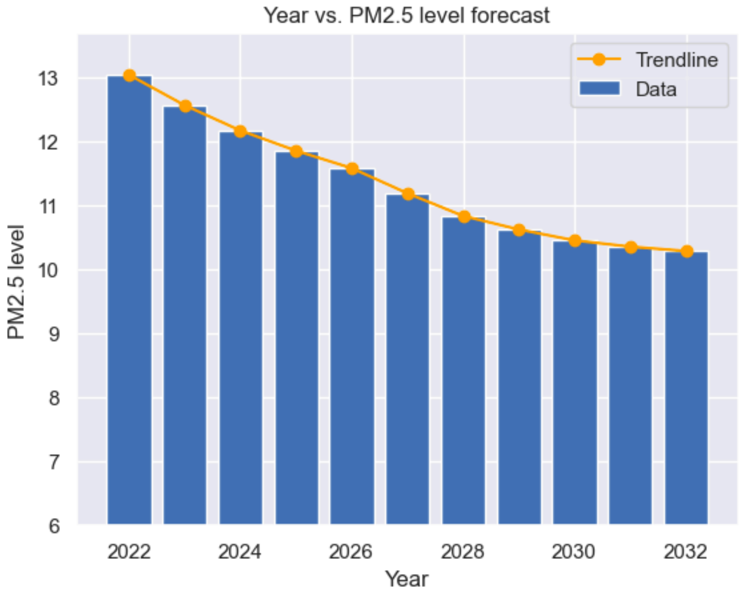


FIGURE 35. $PM_{2.5}$ most probable scenario forecast from 2022 to 2032.

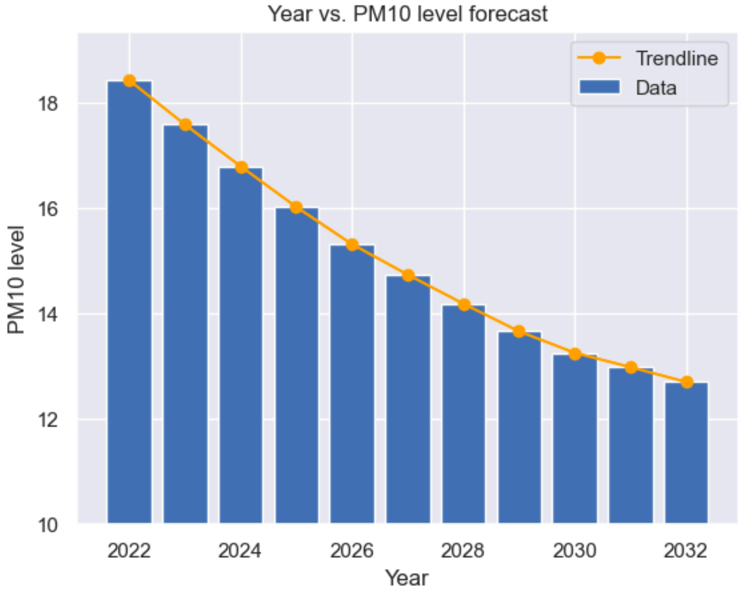


FIGURE 36. PM_{10} most probable scenario forecast from 2022 to 2032.

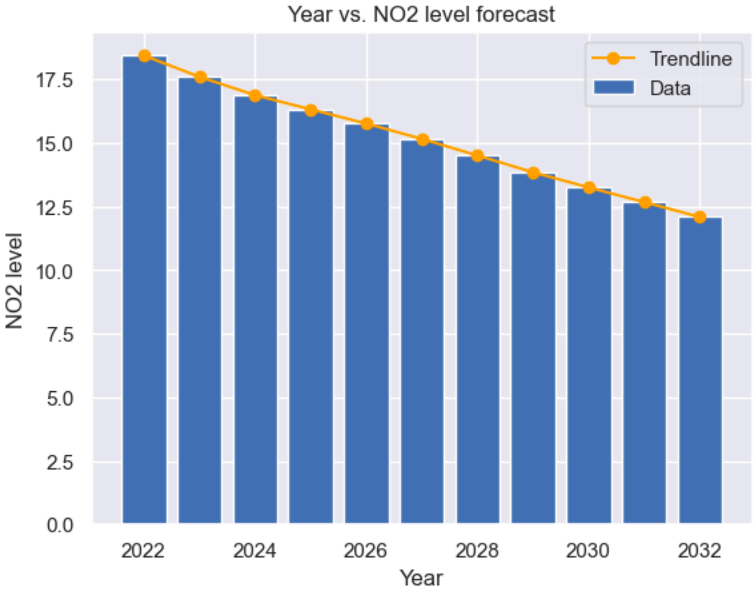


FIGURE 37. NO_2 most probable scenario forecast from 2022 to 2032.

The results of different scenarios for $PM_{2.5}$, PM_{10}, and NO_2 pollution are presented in Tables 13, 14, and 15 respectively.

TABLE 13. $PM_{2.5}$ pollution scenarios results.

Scenario/modifier	1	1.25	1.5	1.75	2
1	11.35	13.70	15.81	17.56	19.40
2	11.35	9.40	8.03	7.14	6.65
3	11.35	12.16	12.65	12.96	13.21
4	11.35	10.35	9.43	8.45	7.69
5	11.35	11.36	10.99	10.27	9.41
6	11.35	10.79	10.02	9.21	8.50
7	11.35	14.76	17.95	20.42	22.36
8	11.35	9.12	8.00	6.92	6.08
9	11.35	13.56	16.07	18.53	20.54
10	11.35	10.41	10.44	11.07	11.45
11	11.35	12.84	15.02	16.83	17.73
12	11.35	10.91	11.22	11.39	10.68
13	11.35	12.73	13.39	13.54	14.12
14	11.35	9.68	6.90	5.29	4.67

15	11.35	11.65	12.22	12.48	12.72
16	11.35	10.75	8.20	6.04	5.19
17	11.35	11.02	10.56	9.99	9.50
18	11.35	11.39	8.62	6.4	5.59
19	11.35	12.49	13.84	15.33	16.66
20	11.35	10.22	9.45	9.03	8.78
21	11.35	13.31	15.34	17.47	19.85
22	11.35	9.81	8.76	8.14	7.88
23	11.35	11.60	11.45	11.28	11.06
24	11.35	10.49	9.70	8.94	8.11
25	11.35	12.41	13.06	13.18	13.04
26	11.35	11.43	11.72	11.98	10.58
27	11.35	10.67	10.08	9.38	8.74
28	11.35	11.79	8.91	6.36	5.41
29	11.35	10.19	9.87	9.85	9.23
30	11.35	11.70	9.11	7.34	6.75

As seen in Table 13, the $PM_{2.5}$ pollution level in these scenarios ranges from 4.67 to 22.36 $\mu g/m^3$. The baseline level of $PM_{2.5}$ pollution is 11.35 $\mu g/m^3$, which is the average of the most probable values from 2022 to 2032.

The pollution level in the first scenario significantly increases up to 19.50 $\mu g/m^3$ when the population is decreasing, and gas consumption is increasing. And in the second scenario, the pollution level significantly decreases to 6.65 $\mu g/m^3$ when the population is increasing, and the gas consumption is decreasing. These results suggest that gas consumption plays a major factor in $PM_{2.5}$ pollution.

The maximum pollution level of 22.36 $\mu g/m^3$ shows the seventh scenario when the GDP per capita is decreasing and gas consumption is increasing. The eighth scenario shows how pollution decreases to 6.08 $\mu g/m^3$ when GDP per capita is increasing and gas consumption is decreasing. The ninth scenario shows how pollution increases up to 20.54 $\mu g/m^3$ when GDP per capita is decreasing and coal consumption is increasing. The eleventh scenario shows how pollution increases up to 17.73 $\mu g/m^3$ when GDP per capita is decreasing and oil consumption is increasing. These results suggest that the GDP per capita plays a major role in $PM_{2.5}$ pollution.

The minimum pollution level of 4.67 μg/m^3 shows the fourteenth scenario when the Earth's temperature is increasing, and gas consumption is decreasing. The sixteenth scenario shows how pollution decreases to 5.19 μg/m^3 when Earth's temperature is increasing, and coal consumption is decreasing. The eighteenth scenario shows how pollution decreases to 5.59 μg/m^3 when Earth's temperature is increasing, and oil consumption is decreasing. These results suggest that even with increasing temperatures, decreasing fossil fuel consumption is the major factor in $PM_{2.5}$ pollution.

The nineteenth scenario shows how pollution increases up to 16.66 μg/m^3 when coal consumption is decreasing, and gas consumption is increasing. The twenty-first scenario shows how pollution increases up to 19.85 μg/m^3 when oil consumption is decreasing, and gas consumption is increasing. The twenty-second scenario shows how pollution decreases to 7.88 μg/m^3 when oil consumption is increasing, and gas consumption is decreasing. These results suggest that gas consumption plays a role in $PM_{2.5}$ pollution.

The twenty-eighth scenario shows how pollution decreases to 5.41 μg/m^3 when the population is decreasing, and the Earth's temperature is increasing. This result suggests that the decreasing population also plays a major factor in $PM_{2.5}$ pollution.

The thirteenth scenario shows how pollution decreases to 6.75 μg/m^3 when the GDP per capita is decreasing, and the Earth's temperature is increasing. This is an interesting finding because decreasing GDP per capita and increasing Earth's temperatures should, in theory, increase pollution level. However, decreasing GDP per capita may imply that the economy is shrinking, and fewer people want to live in such places, which leads to lower energy consumption and therefore less pollution.

TABLE 14. PM_{10} pollution scenarios results.

Scenario/modifier	1	1.25	1.5	1.75	2
1	15.06	16	16.61	16.96	18.07
2	15.06	14.44	15.02	16.22	17.56
3	15.06	15	14.37	14.38	13.98
4	15.06	14.81	14.77	15.14	15.74
5	15.06	16.22	17.40	18.51	19.54

6	15.06	14.91	15.93	16.94	17.84
7	15.06	16.92	18.59	21.58	24.57
8	15.06	15.89	14.31	13.44	12.91
9	15.06	17.20	20.07	21.70	23.15
10	15.06	17.53	17.62	15.25	13.17
11	15.06	16.77	19.55	22.77	28.01
12	15.06	16.71	17.49	16.03	14.16
13	15.06	19.56	20.71	19.14	17.70
14	15.06	12.80	12.70	13.02	12.34
15	15.06	17.13	16.25	15.27	15.10
16	15.06	12.50	12.80	13	12.36
17	15.06	18.6	18.02	17.34	17.36
18	15.06	11.98	12.27	12.56	10.46
19	15.06	15.83	16.21	16.13	16.26
20	15.06	14.35	14.37	14.58	14.77
21	15.06	15.04	15.16	15.16	14.90
22	15.06	15.27	15.31	16	16.57
23	15.06	14.37	14.41	14.45	14.47
24	15.06	15.99	16.70	17.31	17.99
25	15.06	17.91	20.80	22.63	23.59
26	15.06	18.22	18.26	16.15	13.54
27	15.06	17.35	17.12	16.84	17.36
28	15.06	12.58	12.58	12.78	10.40
29	15.06	20.78	19.00	15.66	14.47
30	15.06	14.43	14.75	15.31	14.80

As seen in Table 14, the PM_{10} pollution level in five scenarios ranges from 10.40 to 28.01 $\mu g/m^3$. The baseline level of PM_{10} pollution is 15.06 $\mu g/m^3$, which is the average of the most probable values from 2022 to 2032.

The fifth scenario shows how pollution increases up to 19.54 $\mu g/m^3$ when the population is decreasing, and oil consumption is increasing. The seventh scenario shows how pollution increases up to 24.57 $\mu g/m^3$ when the GDP per capita is decreasing, and gas consumption is increasing. The ninth scenario shows how pollution increases up to 23.15 $\mu g/m^3$ when the GDP per capita is decreasing, and coal consumption is increasing. The maximum pollution level of 28.01 μg/m^3 shows the eleventh scenario when the GDP per capita is decreasing, and oil consumption is increasing. These results suggest that increasing fossil fuel consumption is a major factor in PM_{10} pollution.

The eighteenth scenario shows how pollution decreases to 10.46 μg/m^3 when Earth's temperature is increasing, and oil consumption is decreasing. The minimum pollution level of 10.40 μg/m^3 shows the twenty-eighth scenario when the population is decreasing, and the Earth's temperature is increasing. These results suggest that decreasing oil consumption and decreasing population are major factors in PM_{10} pollution.

TABLE 15. NO_2 pollution scenarios results.

Scenario/modifier	1	1.25	1.5	1.75	2
1	15.14	15.39	16.07	16.88	17.67
2	15.14	14.61	13.48	12.39	10.98
3	15.14	15.54	15.85	16.08	15.76
4	15.14	14.19	12.96	11.65	10.00
5	15.14	16.64	17.85	18.81	19.39
6	15.14	13.26	11.33	9.93	8.92
7	15.14	14.62	14.31	14.03	13.92
8	15.14	13.95	14.34	13.44	12.76
9	15.14	14.78	14.58	15.03	15.33
10	15.14	14.64	14.77	15.49	15.85
11	15.14	15.58	16.19	16.61	16.49
12	15.14	13.62	14.68	16.34	17.23
13	15.14	12.96	11.23	11.09	11.58
14	15.14	13.37	10.78	10.80	12.66
15	15.14	12.84	9.78	7.87	6.57
16	15.14	12.95	10.58	10.73	12.69
17	15.14	13.30	10.76	9.27	8.54
18	15.14	11.95	9.21	9.84	11.71
19	15.14	14.77	14.65	15.11	15.94
20	15.14	15.42	15.56	15.58	15.08

21	15.14	13.80	12.82	12.95	13.40
22	15.14	16.27	17.10	17.86	18.57
23	15.14	14.14	13.24	12.43	11.74
24	15.14	15.87	16.40	16.63	16.45
25	15.14	14.14	13.41	12.73	11.64
26	15.14	15.77	16.41	15.32	14.42
27	15.14	10.91	7.04	6.09	5.79
28	15.14	13.37	11.32	11.87	13.69
29	15.14	13.70	11.85	7.65	7.13
30	15.14	13.25	12.20	11.65	12.14

As seen in Table 15, the NO_2 pollution level in five scenarios ranges from 5.79 to 19.39 $\mu g/m^3$. The baseline of the NO_2 pollution is 15.14 $\mu g/m^3$, which is the average of the most probable values from 2022 to 2032.

The second scenario shows how pollution decreases to 10.98 $\mu g/m^3$ when the population is increasing, and gas consumption is decreasing. The fourth scenario shows how pollution decreases to 10.00 $\mu g/m^3$ when the population is increasing, and coal consumption is decreasing. The maximum pollution level of 19.39 $\mu g/m^3$ shows the fifth scenario when the population is decreasing, and oil consumption is increasing. The sixth scenario shows how pollution decreases to 8.92 $\mu g/m^3$ when the population is increasing, and oil consumption is decreasing. These results suggest that decreasing overall fossil fuel consumption and increasing oil consumption are major factors in NO_2 pollution.

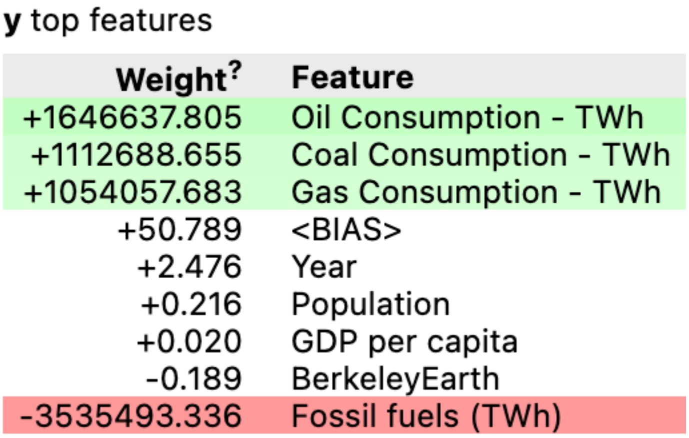
The fifteenth scenario shows how pollution decreases to 6.57 $\mu g/m^3$ when the Earth's temperature is decreasing, and coal consumption is increasing. The seventeenth scenario shows how pollution decreases to 8.54 $\mu g/m^3$ when the Earth's temperature is decreasing, and oil consumption is increasing. These results suggest that decreasing Earth's temperature is a major factor in NO_2 pollution.

The minimum pollution level of 5.79 $\mu g/m^3$ shows the twenty-seventh scenario when the population is increasing, and the Earth's temperature is decreasing. The pollution level twenty-ninth scenario decreases to 7.13 $\mu g/m^3$ when the GDP per capita is increasing, and the Earth's temperature is decreasing. These results suggest that increasing GDP per capita and decreasing Earth's temperature are major factors in NO_2 pollution.

In conclusion, these results of different scenarios show that the GDP per capita, population, and fossil fuel consumption are important factors of $PM_{2.5}$, PM_{10}, and NO_2 pollution. The decreasing Earth's temperature also plays a major role in decreasing the level of NO_2 pollution.

4.2.2 Ozone, carbon monoxide, and sulfur dioxide models

Figure 38 shows the factors' weight for O_3, CO, and SO_2 pollutants prediction models. Oil, coal, and gas consumption factor has the greatest weight, while fossil fuel consumption has less weight. This suggests that the most difference in pollution comes from the difference in oil consumption.



y top features

Weight?	Feature
+1646637.805	Oil Consumption - TWh
+1112688.655	Coal Consumption - TWh
+1054057.683	Gas Consumption - TWh
+50.789	<BIAS>
+2.476	Year
+0.216	Population
+0.020	GDP per capita
-0.189	BerkeleyEarth
-3535493.336	Fossil fuels (TWh)

FIGURE 38. Factors weights for Ozone pollutant prediction.

The most probable scenario forecasts for O_3, CO, and SO_2 pollution are presented in Figures 39, 40, and 41 respectively. The level of O_3 is going to increase to 54 $\mu g/m^3$ assumingly because the gas consumption is predicted to increase as well. The results cannot be compared to annual limits because O_3 concentration is usually calculated in hours. The level of CO is going to decrease to 0.22 mg/m^3 or 220 $\mu g/m^3$, and the level of SO_2 is going to decrease to 1.5 $\mu g/m^3$. These results can be explained by the fact that oil consumption which is the main factor for these models is predicted to decrease as well. Based on the Year Average Common Air Quality Index (YACAQI) levels of CO and SO_2 pollutants are below the annual mean.

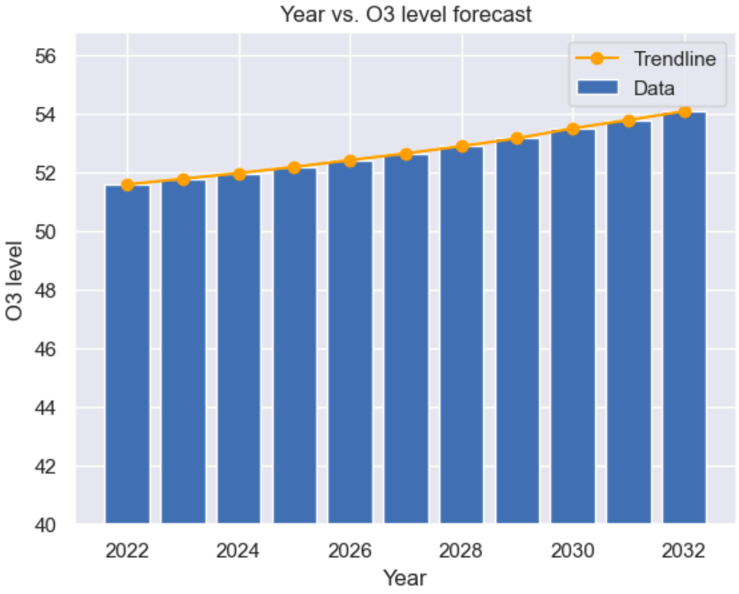


FIGURE 39. O_3 most probable scenario forecast from 2022 to 2032.

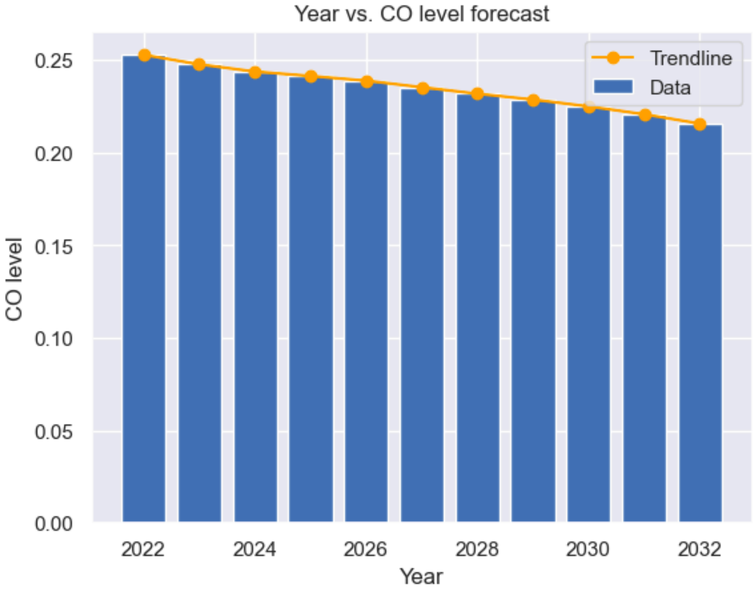


FIGURE 40. CO most probable scenario forecast from 2022 to 2032.

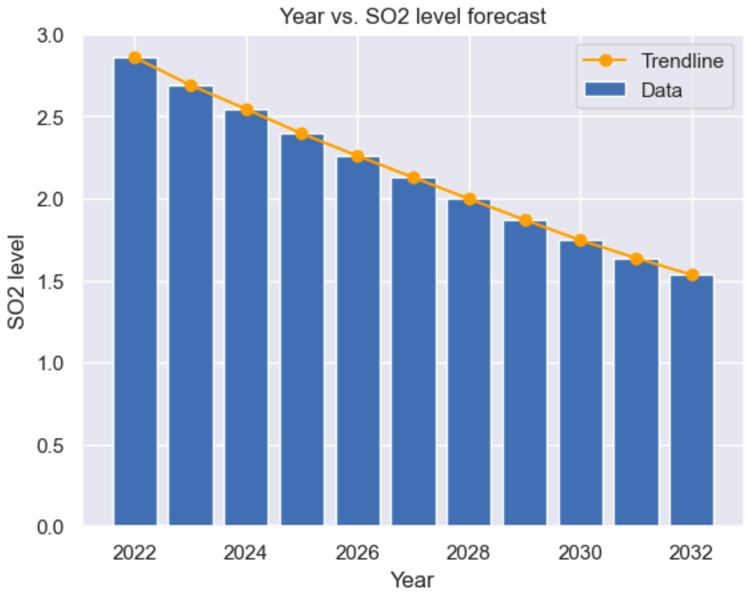


FIGURE 41. SO_2 most probable scenario forecast from 2022 to 2032.

The results of five different scenarios for O_3, CO, and SO_2 pollution are presented in Tables 16, 17, and 18 respectively.

TABLE 16. Ozone (O_3) pollution scenarios results.

Scenario/modifier	1	1.25	1.5	1.75	2
1	52.74	52.34	52.08	51.40	50.87
2	52.74	53.34	53.50	52.51	51.75
3	52.74	52.87	52.47	52.51	50.53
4	52.74	52.68	52.64	52.29	51.30
5	52.74	54.68	56.20	56.91	57.37
6	52.74	51.60	50.67	49.39	48.12
7	52.74	54.06	55.09	55.53	56.27
8	52.74	51.99	51.51	49.98	48.59
9	52.74	54.10	55.21	54.88	54.46
10	52.74	50.98	49.98	47.77	45.58
11	52.74	55.92	58.89	61.08	63.04
12	52.74	50.51	47.54	44.91	43.01
13	52.74	51.03	50.04	49.36	48.58
14	52.74	54.49	58.62	62.18	65.08

15	52.74	51.56	50.68	49.78	49.05
16	52.74	53.90	57.76	61.58	64.56
17	52.74	52.94	53.72	54.44	55.40
18	52.74	52.59	56.79	61.78	65.25
19	52.74	52.20	52.39	52.67	52.68
20	52.74	53.48	53.98	54.42	54.85
21	52.74	51.07	50.29	49.85	49.69
22	52.74	55.07	57.42	59.49	61.05
23	52.74	51.59	50.79	49.88	48.98
24	52.74	54.36	56.21	57.96	59.49
25	52.74	53.95	54.56	54.37	53.98
26	52.74	49.91	47.66	45.04	42.87
27	52.74	51.69	51.35	51.09	50.61
28	52.74	54.39	56.72	59.54	63.06
29	52.74	49.25	47.59	45.83	46.01
30	52.74	55.17	56.32	56.62	58.55

As seen in Table 16, the Ozone (O_3) pollution level in five scenarios ranges from 42.87 to 65.25 $\mu g/m^3$. The baseline level of Ozone (O_3) pollution is 52.74 $\mu g/m^3$, which is the average of the most probable values from 2022 to 2032.

The fifth scenario shows how pollution increases up to 57.37 $\mu g/m^3$ when the population is decreasing, and oil consumption is increasing. The tenth scenario shows how pollution decreases to 45.58 $\mu g/m^3$ when the GDP per capita is increasing, and coal consumption is decreasing. The eleventh scenario shows how pollution increases up to 63.04 $\mu g/m^3$ when the GDP per capita is decreasing, and oil consumption is increasing. The twelfth scenario shows how pollution decreases to 43.01 $\mu g/m^3$ when the GDP per capita is increasing, and oil consumption is decreasing. These results suggest that fossil fuel consumption and GDP per capita are major factors in Ozone (O_3) pollution.

The fourteenth scenario shows how pollution increases up to 65.08 $\mu g/m^3$ when the Earth's temperature is increasing, and gas consumption is decreasing. The sixteenth scenario shows how pollution increases up to 64.56 $\mu g/m^3$ when the Earth's temperature is increasing, and coal consumption is decreasing. The maximum pollution level of 65.25 $\mu g/m^3$ shows the eighteenth scenario when Earth's temperature is increasing, and oil consumption is decreasing. These results suggest increasing Earth's temperature is a major factor in Ozone (O_3) pollution.

The twenty-second scenario shows how pollution increases up to 61.05 $\mu g/m^3$ when the oil consumption is increasing, and gas consumption is decreasing. The twenty-fourth scenario shows how pollution increases up to 59.49 $\mu g/m^3$ when the oil consumption is increasing, and coal consumption is decreasing. These results suggest increasing oil consumption is a major factor in Ozone (O_3) pollution.

The minimum pollution level of 42.87 $\mu g/m^3$ shows the twenty-sixth scenario when the population is decreasing, and GDP per capita is increasing. The twenty-eighth scenario shows how pollution increases up to 63.06 $\mu g/m^3$ when the population is decreasing, and the Earth's temperature is increasing. These results suggest that increasing GDP per capita and increasing Earth's temperature are major factors in Ozone (O_3) pollution.

The twenty-ninth scenario shows how pollution decreases to 46.01 $\mu g/m^3$ when the GDP per capita is increasing, and the Earth's temperature is decreasing. The thirtieth scenario shows how pollution increases up to 58.55 $\mu g/m^3$ when the GDP is decreasing, and the Earth's temperature is increasing. These results suggest that GDP per capita is a major factor in Ozone (O_3) pollution.

TABLE 17. CO pollution scenarios results.

Scenario/modifier	1	1.25	1.5	1.75	2
1	0.234	0.221	0.215	0.208	0.201
2	0.234	0.250	0.265	0.277	0.286
3	0.234	0.223	0.215	0.208	0.203
4	0.234	0.248	0.262	0.275	0.287
5	0.234	0.220	0.207	0.197	0.190
6	0.234	0.247	0.260	0.275	0.289
7	0.234	0.256	0.275	0.287	0.303
8	0.234	0.203	0.164	0.128	0.096
9	0.234	0.255	0.273	0.286	0.295
10	0.234	0.205	0.169	0.122	0.085

11	0.234	0.262	0.291	0.317	0.339
12	0.234	0.216	0.192	0.159	0.129
13	0.234	0.227	0.241	0.265	0.281
14	0.234	0.249	0.229	0.187	0.155
15	0.234	0.224	0.232	0.239	0.235
16	0.234	0.237	0.209	0.180	0.145
17	0.234	0.213	0.224	0.233	0.229
18	0.234	0.229	0.200	0.164	0.134
19	0.234	0.231	0.229	0.228	0.221
20	0.234	0.234	0.231	0.228	0.223
21	0.234	0.228	0.218	0.203	0.184
22	0.234	0.231	0.220	0.209	0.200
23	0.234	0.231	0.227	0.225	0.223
24	0.234	0.229	0.221	0.216	0.210
25	0.234	0.269	0.298	0.325	0.347
26	0.234	0.195	0.150	0.097	0.061
27	0.234	0.236	0.243	0.246	0.244
28	0.234	0.228	0.197	0.160	0.116
29	0.234	0.187	0.173	0.200	0.210
30	0.234	0.256	0.245	0.231	0.223

As seen in Table 17, the CO pollution level in five scenarios ranges from 0.061 to 0.347 mg/m^3 or from 61 to 347 μg/m^3. The baseline level of CO pollution is 0.234 mg/m^3, which is the average of the most probable values from 2022 to 2032.

The fifth scenario shows how pollution decreases to 0.190 mg/m^3 when the population is decreasing, and oil consumption is increasing. The sixth scenario shows how pollution increases up to 0.289 mg/m^3 when the population is increasing, and oil consumption is decreasing. These results suggest that population is a major factor in CO pollution.

The seventh scenario shows how pollution increases up to 0.303 mg/m^3 when the GDP per capita is decreasing, and gas consumption is increasing. The eighth scenario shows how pollution decreases to 0.096 mg/m^3 when the GDP per capita is increasing, and gas consumption is decreasing. The ninth scenario shows how pollution increases up to 0.295 mg/m^3 when the GDP per capita is decreasing, and coal consumption is increasing. The tenth scenario shows how pollution decreases to 0.085 mg/m^3 when the GDP per capita is increasing, and coal consumption is decreasing. The eleventh scenario shows how pollution increases up to 0.339 mg/m^3 when the GDP per capita is decreasing, and oil consumption is increasing. The twelfth scenario shows how pollution decreases to 0.129 mg/m^3 when the GDP per capita is increasing, and oil consumption is decreasing. These results suggest that fossil fuel consumption is a major factor in CO pollution.

The thirteenth scenario shows how pollution increases up to 0.281 mg/m^3 when the Earth's temperature is decreasing, and gas consumption is increasing. The fourteenth scenario shows how pollution decreases to 0.155 mg/m^3 when the Earth's temperature is increasing, and gas consumption is decreasing. The sixteenth scenario shows how pollution decreases to 0.145 mg/m^3 when the Earth's temperature is increasing, and coal consumption is decreasing. The eighteenth scenario shows how pollution decreases to 0.134 mg/m^3 when the Earth's temperature is increasing, and oil consumption is decreasing. These results suggest that even with increasing or decreasing Earth's temperature, fossil fuel consumption plays a major factor in CO pollution.

The twenty-first scenario shows how pollution decreases to 0.184 mg/m^3 when the oil consumption is decreasing, and gas consumption is increasing. The maximum pollution level of 0.347 mg/m^3 shows in the twenty-fifth scenario when the population is increasing, and the GDP per capita is decreasing. The minimum pollution level of 0.061 mg/m^3 shows in the twenty-sixth scenario when the oil consumption is decreasing, and gas consumption is increasing. The twenty-eighth scenario shows how pollution decreases to 0.116 mg/m^3 when the population is decreasing, and the Earth's temperature is increasing. These results suggest that oil consumption, GDP per capita, and population are major factors in CO pollution.

Table 18. SO_2 pollution scenarios results.

Scenario/modifier	1	1.25	1.5	1.75	2
1	2.15	2.00	2.03	2.09	2.16
2	2.15	2.44	2.72	3.02	3.34
3	2.15	1.94	1.92	1.92	1.91
4	2.15	2.47	2.81	3.13	3.44
5	2.15	2.00	2.05	2.11	2.17
6	2.15	2.43	2.75	3.05	3.36

7	2.15	2.27	1.62	1.09	0.68
8	2.15	1.99	2.02	2.07	2.11
9	2.15	2.71	2.57	2.53	2.52
10	2.15	2.02	2.04	2.00	1.95
11	2.15	2.47	1.97	1.78	1.69
12	2.15	1.97	2.01	1.98	1.92
13	2.15	2.12	1.85	1.59	1.37
14	2.15	1.97	2.07	1.97	2.09
15	2.15	2.25	2.16	2.13	2.13
16	2.15	1.99	2.06	2.08	2.42
17	2.15	2.25	2.16	2.08	1.98
18	2.15	1.98	2.06	2.05	2.27
19	2.15	2.18	2.16	2.11	2.00
20	2.15	2.11	2.08	2.04	1.98
21	2.15	2.13	2.14	2.13	2.06
22	2.15	2.14	2.13	2.13	2.14
23	2.15	2.10	2.04	1.99	1.94
24	2.15	2.19	2.22	2.25	2.25
25	2.15	2.79	2.61	2.85	4.61
26	2.15	2.02	2.17	2.23	2.22
27	2.15	2.46	2.58	2.88	3.08
28	2.15	1.97	2.03	2.04	2.21
29	2.15	1.82	1.68	1.44	1.21
30	2.15	2.85	3.73	4.95	6.35

As seen in Table 20, the SO_2 pollution level in five scenarios ranges from 0.68 to 6.35 $\mu g/m^3$. The baseline level of SO_2 pollution is 2.15 $\mu g/m^3$, which is the average of the most probable values from 2022 to 2032.

The second scenario shows how pollution increases up to 3.34 $\mu g/m^3$ when the population is increasing, and gas consumption is decreasing. The fourth scenario shows how pollution increases up to 3.44 $\mu g/m^3$ when the population is increasing, and coal consumption is decreasing. The sixth scenario shows how pollution increases up to 3.36 $\mu g/m^3$ when the population is increasing, and oil consumption is decreasing. These results suggest that increasing population is a major factor in SO_2 pollution even when fossil fuel consumption is decreasing.

The minimum pollution level of 0.68 $\mu g/m^3$ shows the seventh scenario when the GDP per capita is decreasing, and gas consumption is increasing. This is an interesting finding, since a decrease in GDP per capita and increased gas consumption should, in theory, increase pollution levels. However, just as it was for $PM_{2.5}$ pollution, decreasing GDP per capita may imply that the economy is shrinking, and fewer people want to live in such places, which leads to lower energy consumption and therefore less pollution.

The thirteenth scenario shows how pollution decreases to 1.37 $\mu g/m^3$ when the Earth's temperature is decreasing, and gas consumption is increasing. The sixteenth scenario shows how pollution increases up to 2.42 $\mu g/m^3$ when the Earth's temperature is increasing, and coal consumption is decreasing. These results suggest that Earth's temperature is a major factor in SO_2 pollution.

The twenty-fifth scenario shows how pollution increases up to 4.61 $\mu g/m^3$ when the population is increasing, and the GDP per capita is decreasing. The twenty-seventh scenario shows how pollution increases up to 3.08 $\mu g/m^3$ when the population is increasing, and the Earth's temperature is decreasing. The twenty-ninth scenario shows how pollution decreases to 1.21 $\mu g/m^3$ when the GDP per capita is increasing, and the Earth's temperature is decreasing. The maximum pollution level of 6.35 $\mu g/m^3$ shows the thirtieth scenario when the GDP per capita is decreasing, and the Earth's temperature is increasing. These results suggest that GDP per capita, population, and Earth's temperature are major factors in SO_2 pollution.

In conclusion, these results of different scenarios show that fossil fuel consumption, especially oil consumption, GDP per capita, and Earth's temperature are important factors of Ozone (O_3), and CO pollution. The population, Earth's temperature, and GDP per capita are important factors in SO_2 pollution.

5 DISCUSSION AND CONCLUSIONS

The thesis work focused on the air pollution situation in Europe and forecasting it using regression models. The main questions were:

- What are the main factors and main pollutants?
- What will be the most probable level of main pollutants level in Europe and what will it be in different scenarios?
- How factors are affecting the level in these scenarios?
- Do the results exceed the Year Average Common Air Quality Index (YACAQI)?

The data about air pollutant levels were gathered from European Environment Agency for 41 countries in Europe each year from 1969 to 2022. The European Environment Agency allowed downloading 100000 rows at a time for each country per year. There were 4350951 rows in total with a worth of 2.16 GB, but during data cleaning some information was omitted due to unimportance. The factors levels data was gathered from Our World in Data and The World Bank for the European region from 1960. Data cleaning was also necessary for these data. The air pollution and pollution factors levels data were merged, pre-processed, and fed to the regression models for prediction. There are a lot of data sources on different aspects of air pollution; however, some data is difficult to gather in a short time and still needs cleaning even though it is reported by large organizations.

The LinearRegression was selected to predict the pollution factors because there is only one target (pollution factor level) and one feature (year). The MLPRegressor was selected to predict the pollutants levels because one target (an air pollutant level) has numerous features (pollution factors levels). The model's performance was good with the highest R^2 score of 0.9 and the lowest of 0.7. These Python tools were an excellent way to forecast for several years ahead.

It was decided to make a 10-year forecast to show the overall air pollution trend in Europe and analyse it. The main pollutant factors are population, GDP per capita, fossil fuel consumption, gas consumption, coal consumption, oil consumption, and Earth temperature. The levels of these factors were predicted to be fed to regression models as features. The significant finding was that fossil fuel consumption, coal consumption, and oil consumption will decrease in the next ten years.

As fossil fuel is the main cause of the pollution this result suggested that the pollution level would decrease.

The predicted factor levels were taken as they were and fed to the pollutant's regression models to predict the most probable pollutant levels. It was found that the levels of most pollutants will decrease, with only Ozone (O_3) pollutant level increasing from 2022 to 2032. Assumptions about the causes of such outcomes were described. The model stability was proved by creating different scenarios with modified factor levels.

Various scenarios were created in which predicted factor levels were changed by the random number from 1 to 2, either increasing or decreasing it. The average results of these scenarios were compared to conclude which factors affected the level of pollution the most. At the same time, the weights of the factors of each model were defined by the Python tool before being compared with the scenarios' results. The total fossil fuel consumption factor weighted the most for $PM_{2.5}$, PM_{10}, and NO_2 pollutants models, while the oil consumption factor weighted the most for Ozone (O_3), CO, and SO_2 pollutants models. These results suggested that some pollutant levels depend on total fuel consumption, while others depend mainly on oil consumption.

By analysing different scenarios, it was found that the decreasing GDP per capita level is an important factor for almost all pollutants. The decreasing population and fossil fuel consumption, especially gas and oil, are the main factors of $PM_{2.5}$, PM_{10}, and NO_2 pollution levels. The decreasing Earth's temperature has an impact on decreasing levels of NO_2 pollution. Fossil fuel consumption, especially oil, is the main factor of Ozone (O_3) and CO pollution. Population is one of the main factors for CO and SO_2 pollution and temperature plays a major role in Ozone (O_3) and SO_2 pollution. An interesting conclusion is that as GDP per capita decreases, $PM_{2.5}$ and SO_2 pollution decreases, which can be explained by less energy consumption by industry and population and therefore less pollution.

The pollution limits usually are described in hours or days, which was not suitable for predicted data of this thesis. The only metric that worked for the thesis was YACAQI, which describes year limits for $PM_{2.5}$, PM_{10}, CO, SO_2, and NO_2, but not for Ozone (O_3) which is usually measured hourly. Levels of five pollutants were below the annual average, suggesting there would be no health impact.

REFERENCES

Air-quality UK 2022. Motor Vehicle Emission Controls: Fuel Types. Search date 07.05.2023 https://www.air-quality.org.uk/26.php

Callen, Tim 2012. Gross domestic product: An economy's all. International Monetary Fund. Search date 18.02.2023 https://www.imf.org/en/Publications/fandd/issues/Series/Back-to-Basics/gross-domestic-product-GDP

Center for Science Education 2021. How Weather Affects Air Quality. Search date 20.02.2023 https://scied.ucar.edu/learning-zone/air-quality/how-weather-affects-air-quality

Chollet, Francois 2021. Deep Learning with Python. Second Edition. United States, NY: Manning Publications. Search date 22.03.2023 https://learning.oreilly.com/library/view/deep-learning-with/9781617296864/

Ciaburro, Giuseppe 2022. Hands-On Simulation Modeling with Python: Develop Simulation Models for Improved Efficiency and Precision in the Decision-making Process. United Kingdom: Packt Publishing. Search date 28.03.2023 https://learning.oreilly.com/library/view/hands-on-simulation-modeling/9781804616888/

Dovers, Stephen, & Butler, Colin 2019. Population and environment: a global challenge. Australian Academy of Science. Search date 18.02.2023 https://www.science.org.au/curious/earth-environment/population-environment

Dzhulgakov, Dmytro, Mirjalili, Vahid, Liu, Yuxi (Hayden), Raschka, Sebastian 2022. Machine Learning with PyTorch and Scikit-Learn: Develop Machine Learning and Deep Learning Models with Python. India: Packt Publishing. Search date 22.03.2023 https://learning.oreilly.com/library/view/machine-learning-with/9781801819312/

EPA 2022. Air Pollution and Cardiovascular Disease Basics. Search date 07.05.2023 https://www.epa.gov/air-research/air-pollution-and-cardiovascular-disease-basics

European Environmental Agency 2020. What is air pollution. Search date 30.01.2023 https://www.eea.europa.eu/signals/signals-2020/infographics/what-is-pollution

European Environmental Agency 2021. Air quality standards. Search date 30.01.2023 https://www.eea.europa.eu/themes/air/air-quality-concentrations/air-quality-standards

European Environmental Agency 2021. Emissions of air pollutants from transport. Search date 27.12.2022 https://www.eea.europa.eu/data-and-maps/indicators/transport-emissions-of-air-pollutants-8/transport-emissions-of-air-pollutants-8

European Environmental Agency 2022. Air pollutant emissions data viewer (Gothenburg Protocol, LRTAP Convention) 1990-2020. Search date 27.12.2022 https://www.eea.europa.eu/data-and-maps/dashboards/air-pollutant-emissions-data-viewer-4

European Environmental Agency 2022. Air pollution country fact sheets 2022. Search date 27.12.2022 https://www.eea.europa.eu/themes/air/country-fact-sheets/2022-country-fact-sheets

European Environmental Agency 2022. Air pollution is the biggest environmental health risk in Europe. Search date 26.03.2023 https://www.eea.europa.eu/themes/air/air-pollution-is-the-single

European Environmental Agency 2022. Air pollution sources. Search date 02.01.2023 https://www.eea.europa.eu/themes/air/air-pollution-sources-1

European Environmental Agency 2022. Air Quality e-Reporting (AQ e-Reporting). Search date 27.12.2022 https://www.eea.europa.eu/data-and-maps/data/aqereporting-9

European Environmental Agency 2022. Air Quality Health Risk Assessments. Search date 27.12.2022 https://www.eea.europa.eu/data-and-maps/data/air-quality-health-risk-assessments

European Environmental Agency 2022. Emissions of the main air pollutants in Europe. Search date 27.12.2022 https://www.eea.europa.eu/ims/emissions-of-the-main-air

European Environmental Agency 2022. European Air Quality Index. Search date 27.12.2022 https://airindex.eea.europa.eu/Map/AQI/Viewer/

European Environmental Agency 2022. Global (left) and European land (right) average near-surface temperatures relative to the pre-industrial period 1850-1900. Search date 13.02.2023 https://www.eea.europa.eu/data-and-maps/figures/global-left-and-european-land-1

European Environmental Agency 2022. Heavy metal emissions in Europe. Search date 27.12.2022 https://www.eea.europa.eu/ims/heavy-metal-emissions-in-europe

European Environmental Agency 2022. Industrial pollutant releases to air in Europe. Search date 27.12.2022 https://www.eea.europa.eu/ims/industrial-pollutant-releases-to-air

European Environmental Agency 2022. National Emission reductions Commitments (NEC) Directive emission inventory data. Search date 27.12.2022 https://www.eea.europa.eu/data-and-maps/data/national-emission-ceilings-nec-directive-inventory-19

European Environmental Agency Signals 2013. Every breath we take Improving air quality in Europe. Search date 02.01.2023 https://www.eea.europa.eu/publications/eea-signals-2013

Galli, Soledad 2022. Python Feature Engineering Cookbook: Over 70 Recipes for Creating, Engineering, and Transforming Features to Build Machine Learning Models, 2nd Edition. United Kingdom: Packt Publishing, Limited. Search date 26.03.2023 https://learning.oreilly.com/library/view/python-feature-engineering/9781804611302/

IQAir 2019. Air quality in France. Search date 27.02.2023 https://www.iqair.com/france

IQAir 2019. Air quality in Turkey. Search date 27.02.2023 https://www.iqair.com/turkey

IQAir 2020. Air quality in Czech Republic. Search date 20.02.2023 https://www.iqair.com/czech-republic

IQAir 2020. Air quality in Denmark. Search date 20.02.2023 https://www.iqair.com/denmark

IQAir 2020. Air quality in Germany. Search date 20.02.2023 https://www.iqair.com/germany

IQAir 2020. Air quality in Greece. Search date 27.02.2023 https://www.iqair.com/greece

IQAir 2020. Air quality in Hungary. Search date 27.02.2023 https://www.iqair.com/hungary

IQAir 2020. Air quality in Italy. Search date 27.02.2023 https://www.iqair.com/italy

IQAir 2020. Air quality in Sweden. Search date 20.02.2023 https://www.iqair.com/sweden

IQAir 2020. Air quality in Switzerland. Search date 27.02.2023 https://www.iqair.com/switzerland

IQAir 2020. Air quality in the Netherlands. Search date 27.02.2023 https://www.iqair.com/netherlands

IQAir 2021. Air quality in Austria. Search date 27.02.2023 https://www.iqair.com/austria

IQAir 2021. Air quality in Belgium. Search date 27.02.2023 https://www.iqair.com/belgium

IQAir 2021. Air quality in Spain. Search date 27.02.2023 https://www.iqair.com/spain

IQAir 2021. Air quality in United Kingdom. Search date 20.02.2023 https://www.iqair.com/uk

IQAir 2022. Air quality in Finland. Search date 20.02.2023 https://www.iqair.com/finland

IQAir 2022. Air quality in Norway. Search date 20.02.2023 https://www.iqair.com/norway

IQAir 2023. Air quality in Poland. Search date 20.02.2023 https://www.iqair.com/poland

IQAir 2023. Interactive global map of 2022 PM2.5 concentrations by City. Search date 23.03.2023 https://www.iqair.com/us/world-air-quality-report

Liu, Yuxi (Hayden) 2020. Python Machine Learning By Example: Build Intelligent Systems Using Python, TensorFlow 2, PyTorch, and Scikit-learn, 3rd Edition. United Kingdom: Packt Publishing. Search date 22.03.2023 https://learning.oreilly.com/library/view/python-machine-learning/9781800209718/

Loy, James 2019. Neural Network Projects with Python: The Ultimate Guide to Using Python to Explore the True Power of Neural Networks Through Six Projects. India: Packt Publishing. Search date 26.03.2023 https://learning.oreilly.com/library/view/neural-network-projects/9781789138900/

McHugh, Christine, Clayton, James, Kourti, Ioanna, & Warren, Ben 2014. The Greater London Authority. Comparison of Air Quality in London with a Number of World and European Cities. AMEC Environment & Infrastructure UK Limited 2014. Search date 23.02.2023 https://www.london.gov.uk/sites/default/files/comparison_of_air_quality_in_world_cities_study_final.pdf

Naghshpour, Shahdad 2016. Regression for Economics, Second Edition. United States: Business Expert Press. Search date 27.03.2023 http://ebookcentral.proquest.com/lib/oamk-ebooks/detail.action?docID=4518807

Nokeri, Tshepo Chris 2021. Implementing Machine Learning for Finance: A Systematic Approach to Predictive Risk and Performance Analysis for Investment Portfolios. United States: Apress. Search date 28.03.2023 https://learning.oreilly.com/library/view/implementing-machine-learning/9781484271100/

Ritchie, Hannah, Roser, Max & Rosado, Pablo 2020. CO_2 and Greenhouse Gas Emissions. Our World in Data. Search date 23.02.2023 https://ourworldindata.org/co2-and-greenhouse-gas-emissions

Ritchie, Hannah, Roser, Max & Rosado, Pablo 2022. Fossil Fuels. Our World in Data. Search date 09.02.2023 https://ourworldindata.org/fossil-fuels

Ritchie, Hannah, Roser, Max & Rosado, Pablo 2022. Transport. Our World in Data. Search date 02.03.2023 https://ourworldindata.org/transport

Scribbr 2022. Data Cleaning. A Guide with Examples & Steps. Search date 15.02.2023 https://www.scribbr.co.uk/research-methods/data-cleaning/

Swamynathan, Manohar 2019. Mastering Machine Learning with Python in Six Steps: A Practical Implementation Guide to Predictive Data Analytics Using Python. Germany: Apress. Search date 27.03.2023 https://learning.oreilly.com/library/view/mastering-machine-learning/9781484249475/

Tableu 2022. Guide To Data Cleaning: Definition, Benefits, Components, And How To Clean Your Data. Search date 15.02.2023 https://www.tableau.com/learn/articles/what-is-data-cleaning

The World Bank 2022. GDP per capita (current US$) – European Union. Search date 08.02.2023 https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=EU

The World Bank 2022. Population, total – European Union. Search date 08.02.2023 https://data.worldbank.org/indicator/SP.POP.TOTL?locations=EU

UK Air 2021. UK Air Quality Limits. Search date 07.05.2023 https://uk-air.defra.gov.uk/air-pollution/uk-eu-limits

Vallero, Daniel A. 2014. Fundamentals of air pollution. Fifth edition. Amsterdam, Netherlands: Elsevier.

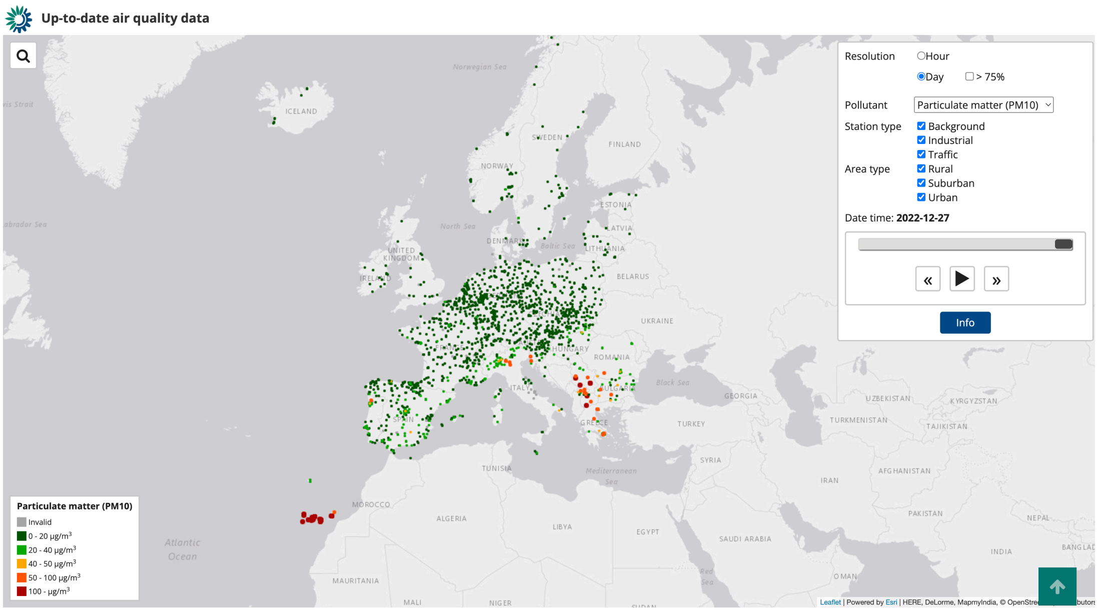
World Health Organization 2019. Health consequences of air pollution on populations. Search date 07.05.2013 https://www.who.int/news/item/15-11-2019-what-are-health-consequences-of-air-pollution-on-populations

World Health Organization 2022. Air Quality Database. Search date 27.12.2022 https://www.who.int/data/gho/data/themes/air-pollution/who-air-quality-database

World Health Organization 2022. Concentrations of fine particulate matter (PM2.5). Search date 27.12.2022 https://www.who.int/data/gho/data/indicators/indicator-details/GHO/concentrations-of-fine-particulate-matter-(pm2-5)

World Health Organization 2022. National Air Quality Standards. Search date 27.12.2022 https://whoairquality.shinyapps.io/AirQualityStandards/

1. Air quality live: Up-to-date air quality measurements show the latest measurements from the European air quality monitoring network on five air pollutants (Particles less than 2.5 µm, Particles less than 10 µm, Nitrogen dioxide, Ozone, Sulphur dioxide



Up-to-date air quality data

Resolution ○Hour ●Day ☐ > 75%

Pollutant: Particulate matter (PM10)

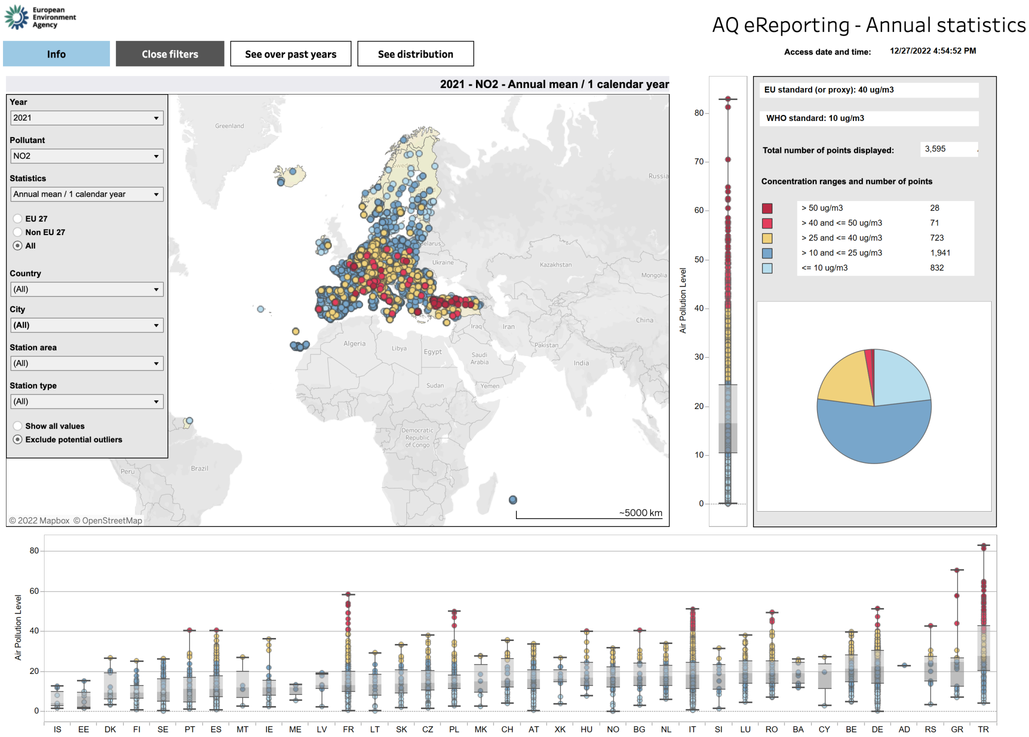
Station type: ☑ Background ☑ Industrial ☑ Traffic

Area type: ☑ Rural ☑ Suburban ☑ Urban

Date time: **2022-12-27**

Particulate matter (PM10)
- ▨ Invalid
- 🟩 0 - 20 µg/m³
- 🟩 20 - 40 µg/m³
- 🟧 40 - 50 µg/m³
- 🟧 50 - 100 µg/m³
- 🟥 100 - µg/m³

2. Key air quality statistics for the main air pollutants: map viewer. The map was made for the years 2013 to 2021.



AQ eReporting - Annual statistics

Access date and time: 12/27/2022 4:54:52 PM

2021 - NO2 - Annual mean / 1 calendar year

Year: 2021

Pollutant: NO2

Statistics: Annual mean / 1 calendar year

○ EU 27
○ Non EU 27
● All

Country: (All)

City: (All)

Station area: (All)

Station type: (All)

○ Show all values
● Exclude potential outliers

EU standard (or proxy): 40 ug/m3

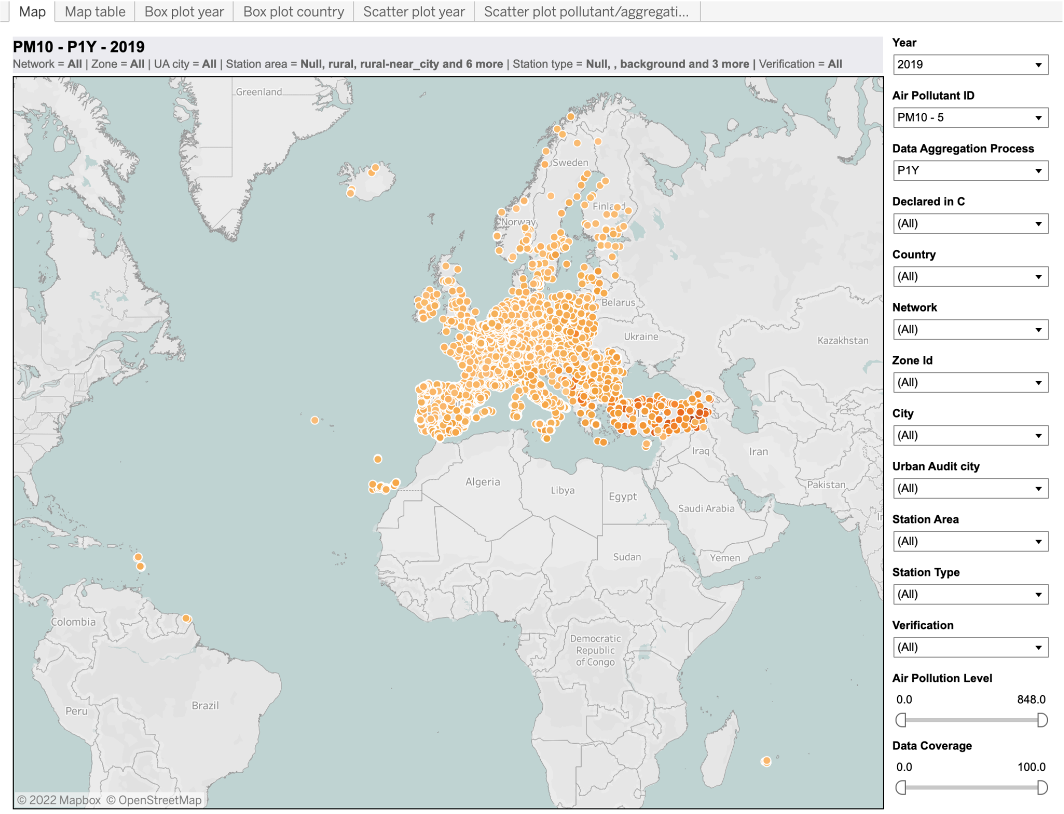
WHO standard: 10 ug/m3

Total number of points displayed: 3,595

Concentration ranges and number of points

🟥	> 50 ug/m3	28
🟥	> 40 and <= 50 ug/m3	71
🟨	> 25 and <= 40 ug/m3	723
🟦	> 10 and <= 25 ug/m3	1,941
🟦	<= 10 ug/m3	832

3. Detailed air quality parameters: advanced map viewer. The map was made for the years 1969 to 2021.



PM10 - P1Y - 2019

Network = **All** | Zone = **All** | UA city = **All** | Station area = **Null, rural, rural-near_city and 6 more** | Station type = **Null, , background and 3 more** | Verification = **All**

Year: 2019

Air Pollutant ID: PM10 - 5

Data Aggregation Process: P1Y

Declared in C: (All)

Country: (All)

Network: (All)

Zone Id: (All)

City: (All)

Urban Audit city: (All)

Station Area: (All)

Station Type: (All)

Verification: (All)

Air Pollution Level: 0.0 – 848.0

Data Coverage: 0.0 – 100.0

4. Assessment methods meta-data reported by countries (data flow D) are data tables that describe techniques for measuring or modelling air pollutants. Measurements descriptions were made for the years 2017 to 2022. Models and objective estimations descriptions were made for the years 2014 to 2021.

5. Air quality modeling results processed by the EEA (data flow E1b) data are annual aggregated values from air quality modeling processed by EEA to a common format and projection based on the data uploaded by countries into CDR. The EEA provides an Air Quality Modeling Viewer for checking air pollution concentrations modeling in different countries. The calculations were made for the years 2015 to 2021.

6. Air quality time series (E1a & E2a data sets) data has been exported from EEA's SQL database of primary validated assessment data and primary up-to-date assessment data reported by countries. The calculations were made for the years 2014 to 2021.

7. Air quality zone information reported by countries (data flow B and B preliminary) data table describes the delineation and type of zones and agglomerations in which air quality should be assessed and managed. The descriptions were made for the years 2012 to 2023.

8. Air quality zone geometries reported by countries (data flow B) are available in several spatial formats: ESRI Shapefile, SpatialLite, GeoPackage, and GeoJSON.

9. Air quality assessment regimes reported by countries (data flow C and C preliminary) is a dataset that describes air quality assessment regimes and lists all the individual sampling points and/or models used for each pollutant within individual zones and agglomerations. Air quality assessment regime descriptions were made for the years 2012 to 2023. Air quality assessment regimes – methods descriptions were made for the years 2013-2023.

10. Attainments of (air quality) environmental objectives reported by countries (data flow G) is a dataset that shows the attainment status of the environmental air quality goals established in European legislation, as done by countries for each pollutant within individual zones and agglomerations. It also lists all the individual sampling points and/or models used to assess each of the attainment statuses. Air quality attainments calculations were made for the years 2013-2021. Air quality attainments – methods descriptions were made for the years 2013-2021.

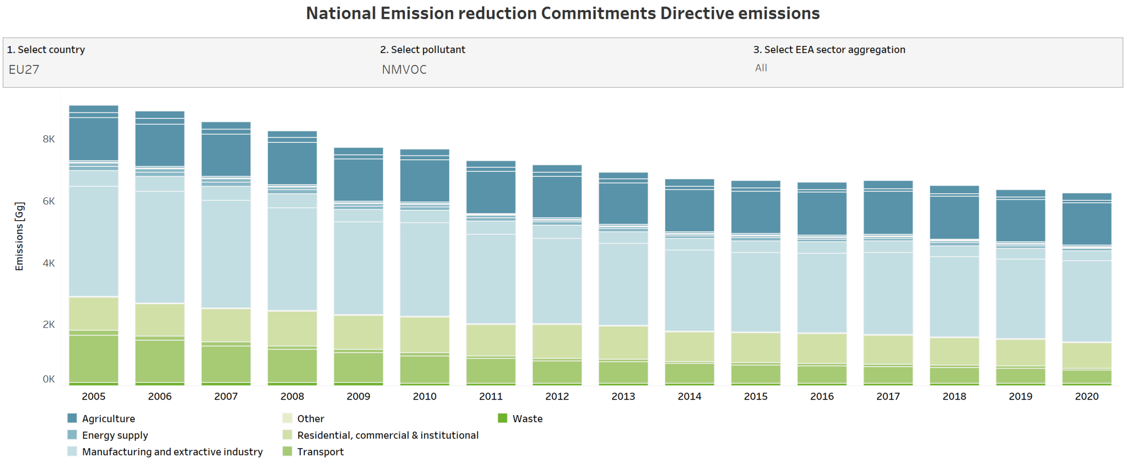
11. Air quality plans reported by countries (data flow H) is a dataset that provides information on air quality plans developed by countries for areas and agglomerations where environmental goals have not been met. Air quality plan descriptions were made for the years 2012-2019.

12. Source apportionments reported by countries (data flow I) is a dataset that shows the information on the distribution of sources by the country for each air pollutant considered in the air quality plans (see data flow H, air quality plans). The descriptions were made for the years 2012 to 2019.

13. Air quality scenarios reported by countries (data flow J) is a dataset that provides information on air quality scenarios submitted by countries and supporting air quality plans (see data flow H, air quality plans). The descriptions were made for the years 2012 to 2019.

14. Air quality measures reported by countries (data flow K) is a dataset that shows the information on air quality measures reported by countries and implemented as part of air quality plans (see data flow H, attainment reports). The descriptions were made the for years 2012 to 2019.

15. The national air pollutant emissions data viewer provides access to the latest air pollutant emission inventories submitted to the EEA by EU Member States under the National Emission Reduction Commitments (NEC) Directive. The description was made for the years 2005-2020.



National Emission reduction Commitments Directive emissions

1. Select country
EU27

2. Select pollutant
NMVOC

3. Select EEA sector aggregation
All

Member States emission reduction committments for 2020 and 2030

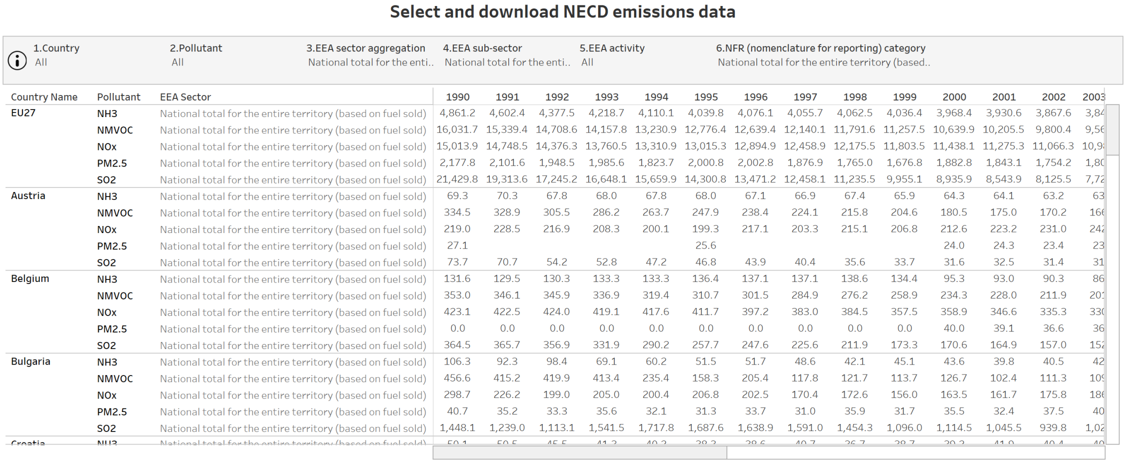
1. Select country
Finland

2. Select pollutant
NH3

🟦 Reduction commitment

		NH3	NMVOC	NOx	PM2.5	SO2
Finland	2020	▽	▽	▽	▽	▽
	2030	▽	△	▽	▽	▽

Current emission levels below the emission reduction commitment	▽
Emission reduction needed by less than 10 % from current levels	△
Emission reduction needed by 10 % to 30 % from current levels	△
Emission reduction needed by 30 % to 50 % from current levels	△
Emission reduction needed by more than 50 % from current levels	△



Select and download NECD emissions data

1.Country
All

2.Pollutant
All

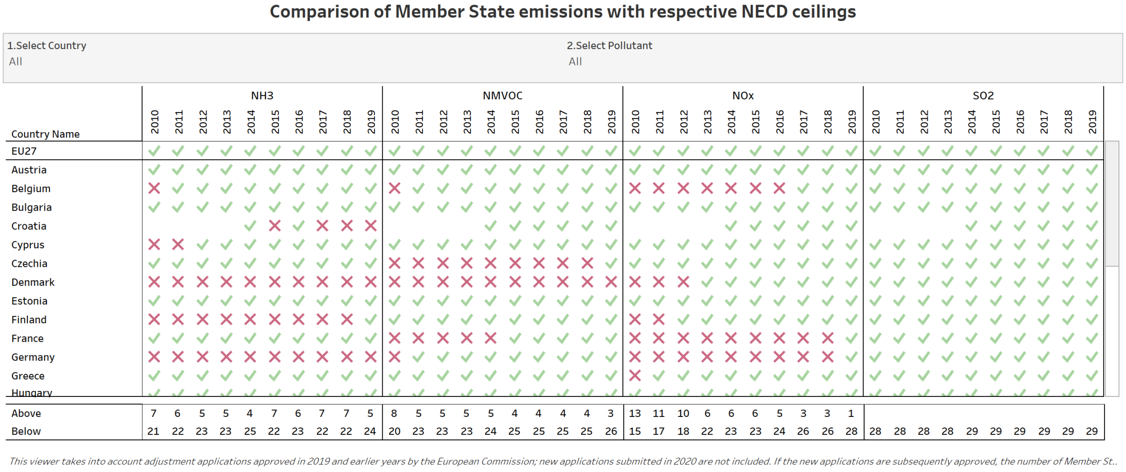
3.EEA sector aggregation
National total for the enti..

4.EEA sub-sector
National total for the enti..

5.EEA activity
All

6.NFR (nomenclature for reporting) category
National total for the entire territory (based..

Country Name	Pollutant	EEA Sector	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
EU27	NH3	National total for the entire territory (based on fuel sold)	4,861.2	4,602.4	4,377.5	4,218.7	4,110.1	4,039.8	4,076.1	4,055.7	4,062.5	4,036.4	3,968.4	3,930.6	3,867.6
	NMVOC	National total for the entire territory (based on fuel sold)	16,031.7	15,339.4	14,708.6	14,157.8	13,230.9	12,776.4	12,639.4	12,140.1	11,791.6	11,257.5	10,639.9	10,205.5	9,800.4
	NOx	National total for the entire territory (based on fuel sold)	15,013.9	14,748.5	14,376.3	13,760.5	13,310.9	13,015.3	12,894.9	12,458.9	12,175.5	11,803.5	11,438.1	11,275.3	11,066.3
	PM2.5	National total for the entire territory (based on fuel sold)	2,177.8	2,101.6	1,948.5	1,985.6	1,823.7	2,000.8	2,002.8	1,876.9	1,765.0	1,676.8	1,882.8	1,843.1	1,754.2
	SO2	National total for the entire territory (based on fuel sold)	21,429.8	19,313.6	17,245.2	16,648.1	15,659.9	14,300.8	13,471.2	12,458.1	11,235.5	9,955.1	8,935.9	8,543.9	8,125.5
Austria	NH3	National total for the entire territory (based on fuel sold)	69.3	70.3	67.8	68.0	67.8	68.0	67.1	66.9	67.4	65.9	64.3	64.1	63.2
	NMVOC	National total for the entire territory (based on fuel sold)	334.5	328.9	305.5	286.2	263.7	247.9	238.4	224.1	215.8	204.6	180.5	175.0	170.2
	NOx	National total for the entire territory (based on fuel sold)	219.0	228.5	216.9	208.3	200.1	199.3	217.1	203.3	215.1	206.8	212.6	223.2	231.0
	PM2.5	National total for the entire territory (based on fuel sold)	27.1					25.6					24.0	24.3	23.4
	SO2	National total for the entire territory (based on fuel sold)	73.7	70.7	54.2	52.8	47.2	46.8	43.9	40.4	35.6	33.7	31.6	32.5	31.4
Belgium	NH3	National total for the entire territory (based on fuel sold)	131.6	129.5	130.3	133.3	133.3	136.4	137.1	137.1	138.6	134.4	95.3	93.0	90.3
	NMVOC	National total for the entire territory (based on fuel sold)	353.0	346.1	345.9	336.9	319.4	310.7	301.5	284.9	276.2	258.9	234.3	228.0	211.9
	NOx	National total for the entire territory (based on fuel sold)	423.1	422.5	424.0	419.1	417.6	411.7	397.2	383.0	384.5	357.5	358.9	346.6	335.3
	PM2.5	National total for the entire territory (based on fuel sold)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	40.0	39.1	36.6
	SO2	National total for the entire territory (based on fuel sold)	364.5	365.7	356.9	331.9	290.2	257.7	247.6	225.6	211.9	173.3	170.6	164.9	157.0
Bulgaria	NH3	National total for the entire territory (based on fuel sold)	106.3	92.3	98.4	69.1	60.2	51.5	51.7	48.6	42.1	45.1	43.6	39.8	40.5
	NMVOC	National total for the entire territory (based on fuel sold)	456.6	415.2	419.9	413.4	235.4	158.3	205.4	117.8	121.7	113.7	126.7	102.4	111.3
	NOx	National total for the entire territory (based on fuel sold)	298.7	226.2	199.0	205.0	200.4	206.8	202.5	170.4	172.6	156.0	163.5	161.7	175.8
	PM2.5	National total for the entire territory (based on fuel sold)	40.7	35.2	33.3	35.6	32.1	31.3	33.7	31.0	35.9	31.7	35.5	32.4	37.5
	SO2	National total for the entire territory (based on fuel sold)	1,448.1	1,239.0	1,113.1	1,541.5	1,717.8	1,687.6	1,638.9	1,591.0	1,454.3	1,096.0	1,114.5	1,045.5	939.8



Comparison of Member State emissions with respective NECD ceilings

1.Select Country
All

2.Select Pollutant
All

Country Name	NH3									
	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
EU27	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Austria	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Belgium	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓
Bulgaria	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Croatia					✓	✗	✓	✗	✗	✗
Cyprus	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓
Czechia	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Denmark	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
Estonia	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Finland	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓
France	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Germany	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
Greece	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Above	7	6	5	5	4	7	6	7	7	5
Below	21	22	23	23	25	22	23	22	22	24

Country Name	NMVOC									
	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
EU27	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Austria	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Belgium	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓
Bulgaria	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Croatia					✓	✓	✓	✓	✓	✓
Cyprus	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Czechia	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓
Denmark	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
Estonia	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Finland	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
France	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓
Germany	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓
Greece	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Above	8	5	5	5	5	4	4	4	4	3
Below	20	23	23	23	24	25	25	25	25	26

Country Name	NOx									
	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
EU27	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Austria	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Belgium	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓
Bulgaria	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Croatia					✓	✓	✓	✓	✓	✓
Cyprus	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Czechia	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Denmark	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓
Estonia	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Finland	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓
France	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓
Germany	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓
Greece	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓
Above	13	11	10	6	6	6	5	3	3	1
Below	15	17	18	22	23	23	24	26	26	28

Country Name	SO2									
	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
EU27	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Austria	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Belgium	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Bulgaria	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Croatia					✓	✓	✓	✓	✓	✓
Cyprus	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Czechia	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Denmark	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Estonia	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Finland	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
France	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Germany	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Greece	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Above										
Below	28	28	28	28	29	29	29	29	29	29

This viewer takes into account adjustment applications approved in 2019 and earlier years by the European Commission; new applications submitted in 2020 are not included. If the new applications are subsequently approved, the number of Member St..

16. Convention on Long-range Transboundary Air Pollution (LRTAP) is an emission data reported by EEA member countries and used in the European Union's annual submission to the UNECE LRTAP Convention. This data set is used by the EEA in its air pollution indicators and reports.

17. National emissions reported to the Convention on Long-range Transboundary Air Pollution (LRTAP Convention) are data on emissions of air pollutants submitted to the LRTAP Convention and copied to EEA. The description was made for the years 1990-2020.

18. Industrial Reporting under the Industrial Emissions Directive 2010/75/EU and European Pollutant Release and Transfer Register Regulation (EC) No 166/2006 is a data that contains location and administrative data on Europe's largest industrial complexes, emissions, and transfers of regulated substances to all environments, waste transfers reported from 2007 to 2020. As well as more detailed energy consumption and emissions data for large combustion plants reported from 2016 to 2020. The EEA has removed access to these files since the data problems were discovered.

19. Reported information on large combustion plants under the Energy Community Treaty is a dataset that contains location and administrative data for large internal combustion power plants in Energy Community member countries, as well as more detailed data on energy consumption and air emissions. These data are reported by the EEA under the Energy Community Treaty. The descriptions were made for the years 2018 to 2021.

20. NECD policies and measures database contains policies and measures (PMs) reported by the EU Member States following the European Commission Implementing Decision (EC) establishing a common format for national air pollution control programs under Directive (EC) of the European Parliament and the Council on the reduction of national emissions of certain atmospheric pollutants.
    1. The PaMs Identification sheet contains information on the Identification and Characterization of a single or a package of PaMs. Implementation Start for years 2004-2030. Implementation Finish for years 2017 – 9999.
    2. The emission Reductions sheet contains information on the expected quantified emission reduction reported by the EU Member States for a single and/or a package of PaMs. The calculations were made for the years 2020, 2025, and 2030.
    3. AQ Impacts sheet contains information on reported impacts on air quality where available. Missing information does not mean that impact on air quality is not occurring.
    4. The costs sheet contains information on the expected costs related to a single or a package of PaMs. This information was not mandatory reporting. Missing data does not mean that cost assessment isn't available.

5. The agriculture sheet contains information on the reported additional details concerning the measures targeting the agricultural sector to comply with the reduction commitments.
6. The adoption sheet contains information only related to the PaMs selected for adoption. The description was made for the years 2004-2030.

21. Estimated effects of increased RES consumption since 2005 on fossil fuels and GHG emissions is a dataset with the most recent set of EEA estimates on the amounts of gross avoided fossil fuels and GHG emissions, per Member State and for the EU as a whole, due to the increase in gross final RES consumption since 2005. The estimates are presented as annual data for the years 2005-2017.

22. Approximate estimates of the share of gross final consumption of renewable energy in the 2021 dataset refer to the EEA 2021 renewable energy (RE) share. The EEA and its European Topic Centre for Climate Change Mitigation and Energy (ETC/CME) produce each year a set of early estimates concerning the RES shares achieved by the countries and the EU as a whole in the previous year. The cut-off date for most data sources incorporated in the calculation of the approximated RES shares is 31 July of the publication year. The calculations were made for the years 2019-2021.

23. Approximate estimates of primary and final energy consumption in the 2021 dataset refer to the EEA 2021 primary and final energy consumption figures. The EEA and its European Topic Centre for Climate Change Mitigation and Energy (ETC/CME) produce each year a set of early estimates concerning the consumption of primary and final energy in the previous year, across the EU as a whole and in each Member State. These estimates are compatible with the scope of the energy efficiency targets for 2020 and 2030. The calculations were made for the years 2017-2021.

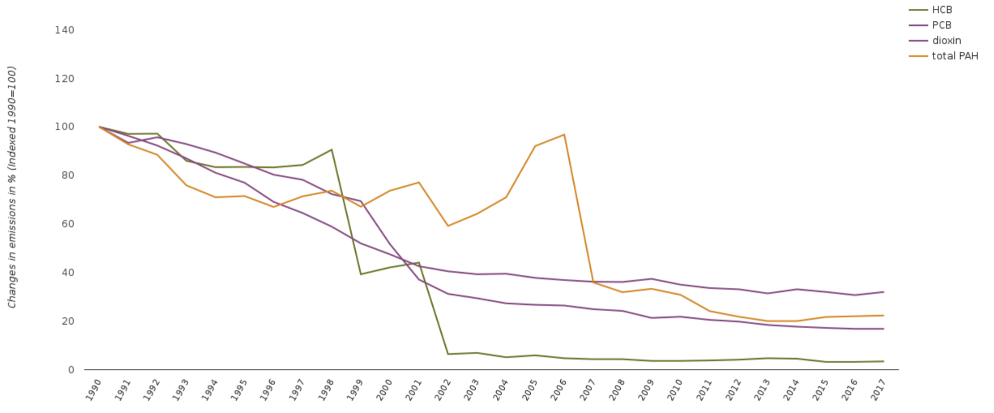
24. The European Pollutant Release and Transfer Register (E-PRTR), Member States reporting is a database that contains facility-by-facility data on releases and transfers of pollutants and waste covered by Regulation (EC) on the establishment of a European Pollutant Release and Transfer Register (EPRTR Regulation) for the years 2007-2017.

25. European Pollutant Release and Transfer Register (E-PRTR) diffuse air emission datasets is a set of 32 maps that shows the distribution of air pollutant emissions from diffuse sources. The data is based on a scale of 5 by 5 km grid and includes details of nitrogen oxides (NO_x), Sulphur oxides (SO_x), carbon monoxide (CO), ammonia (NH_3), carbon dioxide (CO_2), and particulate matter (PM_{10}). The timestamp of the data is unknown.

26. Interpolated air quality data. Interpolated maps showing air quality in Europe. Maps are derived primarily from e-Reporting database station monitoring data, and additional EMEP stations monitoring data, supplemented with additional data such as instance, altitude, meteorological ECMWF data, and EMEP concentration modeling data. The calculations were made for the years 2006- 2019.

27. Persistent organic pollutant (POPs) emissions. This indicator tracks trends in anthropogenic POP emissions since 1990. PAH emissions are currently described, but emissions of other POP compounds will be added in the future. The indicator also provides information on emissions by sector: Energy Production and Distribution; Industrial Energy Use; Industrial Processes; Automotive Transportation; Off-Highway Transportation; Commercial, Institutional, and Households; Solvent and Product Use; Agriculture; Waste; and Other.

Chart — Emission trends of persistant organic pollutants

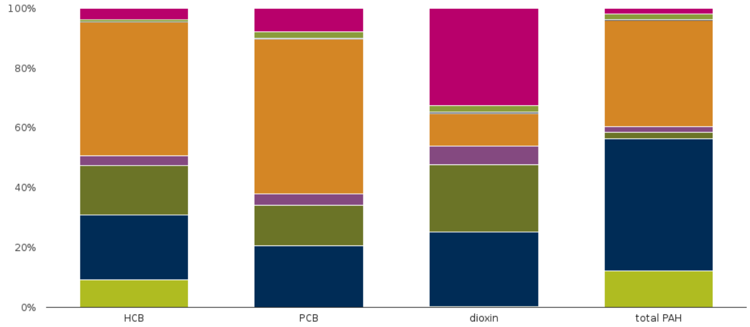


Notes:
Not all EEA member countries reported data for the single pollutants. Total PAHs trend is affected by reported emission by Portugal and Bulgaria related to asphalt blowing activities.
- HCB: hexachlorobenzene
- PAH: polycyclic aromatic hydrocarbon
- PCB: polychlorinated biphenyl.

Data sources:
National emissions reported to the Convention on Long-range Transboundary Air Pollution (LRTAP Convention) provided by European Environment Agency (EEA)

European Environment Agency

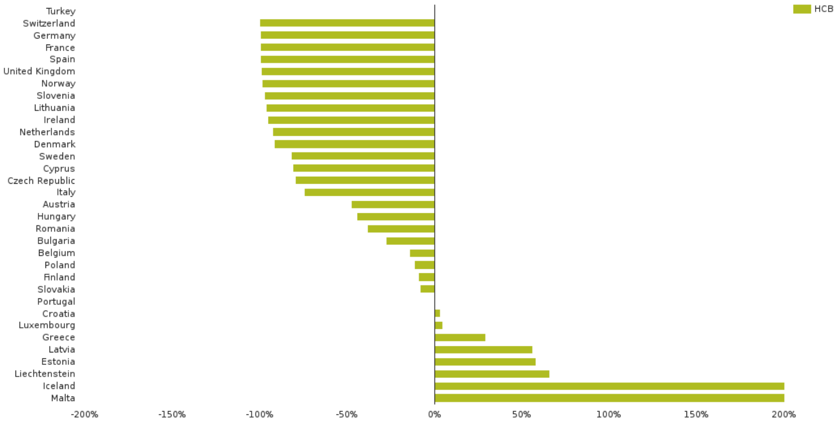
Chart — Emissions of selected persistent organic pollutants by sector



Notes:
Not all EEA member countries reported data for single pollutants.
HCB: hexachlorobenzene;
PAH: polycyclic aromatic hydrocarbon;
PCB: polychlorinated biphenyl.

Data sources:
National emissions reported to the Convention on Long-range Transboundary Air Pollution (LRTAP Convention) provided by European Environment Agency (EEA)

Chart — Change in emissions of persistant organic pollutants

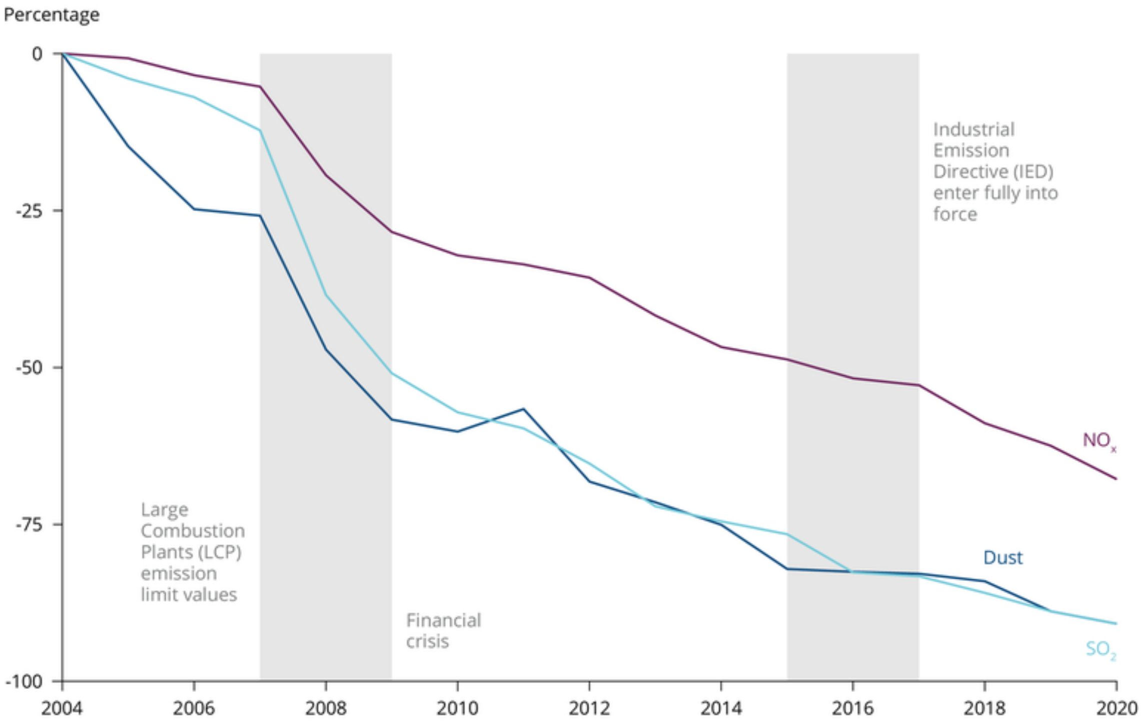


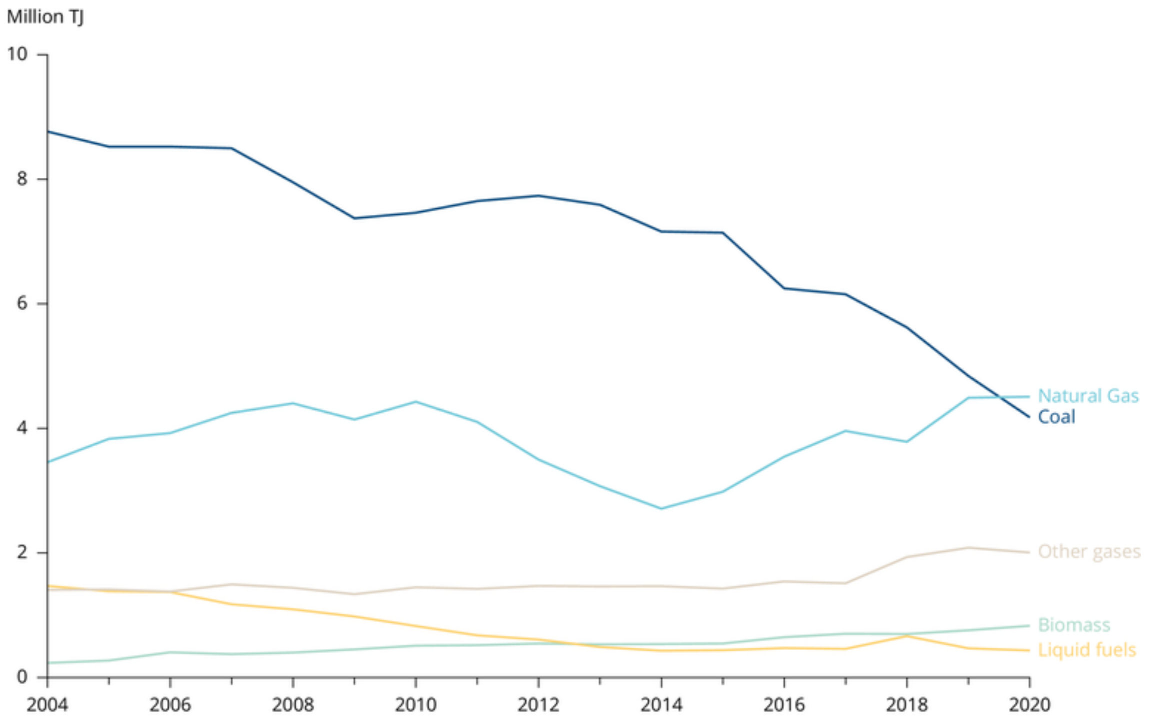
Turkey did not reported emission on POPs. Liechtenstein and Switzerland did not reported PCB emissions.

Data sources:
LRTAP Data

28. Emissions of air pollutants from large combustion plants. This indicator tracks trends for the so-called Large Combustion Plants (LCPs) emissions of SO_2, NOx, and dust, as well as the evolution of the energy mix used in these plants since 2004. LCPs comprise industrial combustion plants with a total rated thermal input equal to or greater than 50 MW. The

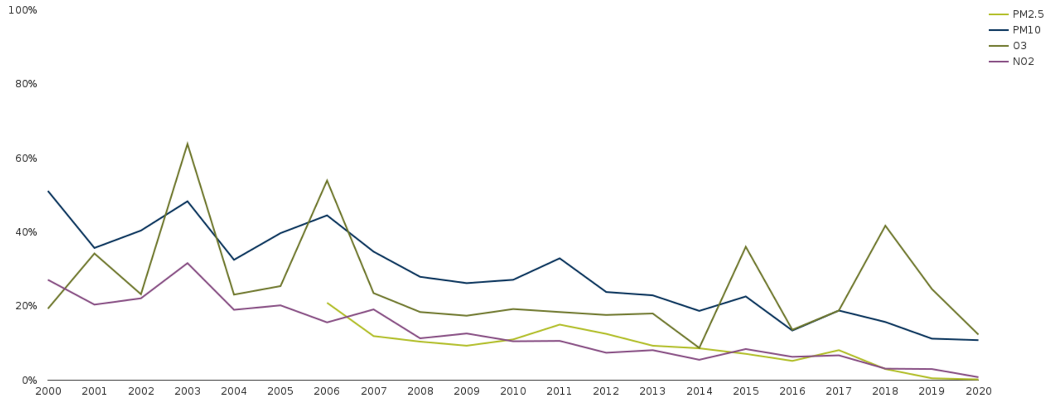
calculations were made for the years 2004-2020. From 2004 to 2020 emissions in Europe decreased: sulphur dioxide (SO_2) and dust by 91%, and nitrogen oxides (NOx) by 68%. Declines in emissions and improvements in environmental performance were caused by European policy, e.g., legally binding emission limit values for large combustion plants. Fossil fuel usage decreased by 26%, as energy production shifted to environmentally friendly sources and coal is no longer the most used fuel in LCP. In the upcoming years, it is anticipated that stricter emission limit values and policies intended to promote the use of renewable or cleaner fuels would result in additional reductions in combustion plant emissions.





29. Exceedance of air quality limit values in urban areas. This indicator shows the fraction of the EU-27 urban population that is potentially exposed to ambient air concentrations of six key pollutants ($PM_{2.5}$, PM_{10}, O_3, NO_2, SO_2, and BaP) that are over the EU limit or target values (EU, 2004, 2008) set for the protection of human health, and to concentrations of these pollutants over the WHO Guidelines. The calculations were made for the years 2000-2020.

Chart — Urban population exposed to air pollutant concentrations above selected EU air quality standards, EU-27



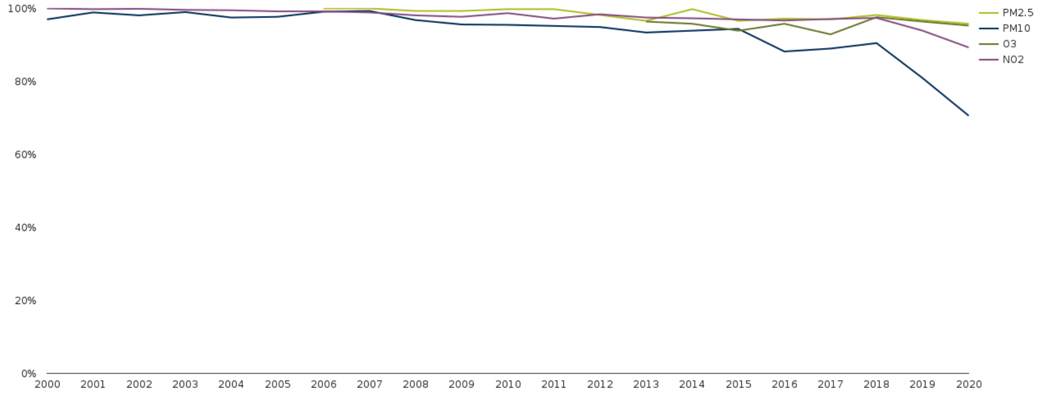
Notes:
$PM_{2.5}$ — particulate matter with a diameter of less than 2.5 µm; PM_{10} — particulate matter with a diameter of less than 10 µm; O_3 — tropospheric ozone; NO_2: nitrogen dioxide.

Criteria for each pollutant: $PM_{2.5}$ — population exposed to concentrations above 25µg/m^3; PM_{10} — population exposed to daily concentrations exceeding 50µg/m^3 for more than 35 days per year; O_3 — population exposed to maximum daily 8-hour mean concentrations exceeding 120µg/m^3 for more than 25 days per year; NO_2 — population exposed to annual concentrations above 40µg/m^3. These criteria are based on the limit values (for PM and NO_2) and target values (for O_3) set in the Ambient Air Quality Directive 2008/50/EC.

Different numbers of countries and stations are included in the illustrated data sets over time.

Data sources:
Air Quality e-Reporting (AQ e-Reporting) provided by European Commission
Gisco - Urban Audit 2020 provided by Statistical Office of the European Union (Eurostat)
Population on 1 January by age group and sex - cities and greater cities (urb_cpop1) provided by Statistical Office of the European Union (Eurostat)

Chart — Urban population exposed to air pollutant concentrations above 2005 WHO air quality guidelines, EU-27



Notes:
$PM_{2.5}$ — particulate matter with a diameter of less than 2.5 µm; PM_{10} — particulate matter with a diameter of less than 10 µm; O_3 — tropospheric ozone; NO_2 — nitrogen dioxide.

Criteria for each pollutant: $PM_{2.5}$ — population exposed to annual concentrations above 5µg/m^3; PM_{10} — population exposed to annual concentrations above 15µg/m^3; O_3 — population exposed to maximum daily 8-hour mean concentrations exceeding 100µg/m^3 for more than 3-4 days per year; NO_2 — population exposed to annual concentrations above 10µg/m^3. These criteria are based on the air quality guideline values defined by the World Health Organization in 2021.

Different numbers of countries and stations are included in the illustrated data sets over time.

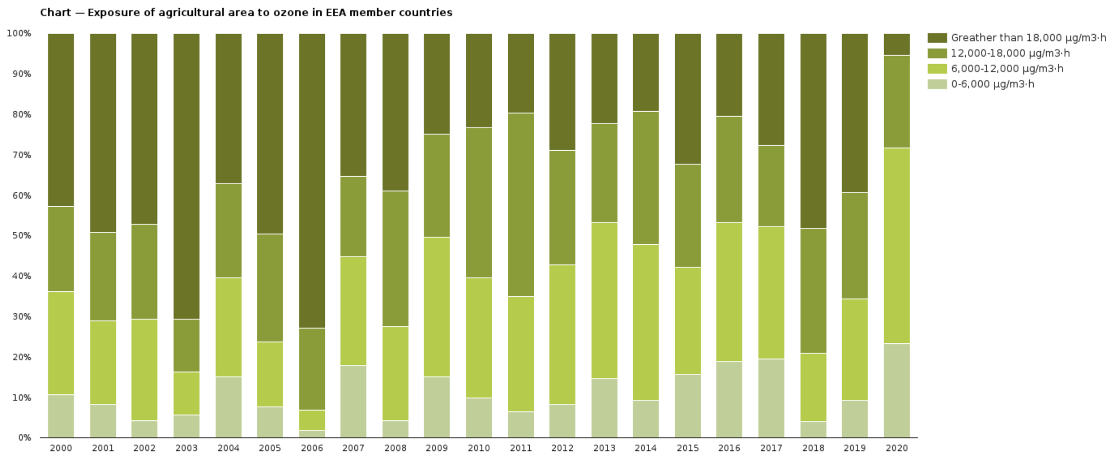
Data sources:
Air Quality e-Reporting (AQ e-Reporting) provided by European Commission
Gisco - Urban Audit 2020 provided by Statistical Office of the European Union (Eurostat)
Population on 1 January by age group and sex - cities and greater cities (urb_cpop1) provided by Statistical Office of the European Union (Eurostat)

30. Exceedances of air quality objectives due to traffic. This indicator compares concentrations of pollutants at background stations with those at traffic stations. This comparison provides an estimate of the increased levels of air pollution to which the population is exposed in

areas with relatively high levels of road traffic. It also provides a measure of the impact of the technical and non-technical measures adopted to reduce the road transport sector's contribution to observed pollutant concentrations. The calculations were made in 2017.

31. Air pollution due to ozone: health impacts and effects of climate change. The indicator presents an overview of ozone concentrations over Europe in recent years, their effects on human health, and an estimate of the changes in these concentrations due to the effect of climate change. It presents the annual mean of the maximum daily eight-hour mean ozone concentrations by station type, the modelled projected change, due to climate change, in summertime surface ozone concentrations over Europe in the middle and at the end of the 21st century, the relative effect of climate change on ozone concentrations in the middle of the 21st century, compared to other contributions, a selection of meteorological parameters that might increase under future climate change and their impact on ozone levels. The calculations were made from 2003-2012.

32. Exposure of ecosystems to acidification, eutrophication, and ozone. This indicator shows the exposure of areas covered with specific vegetation (crops and forests) to ground-level ozone. The calculations were made from 2000-2020.

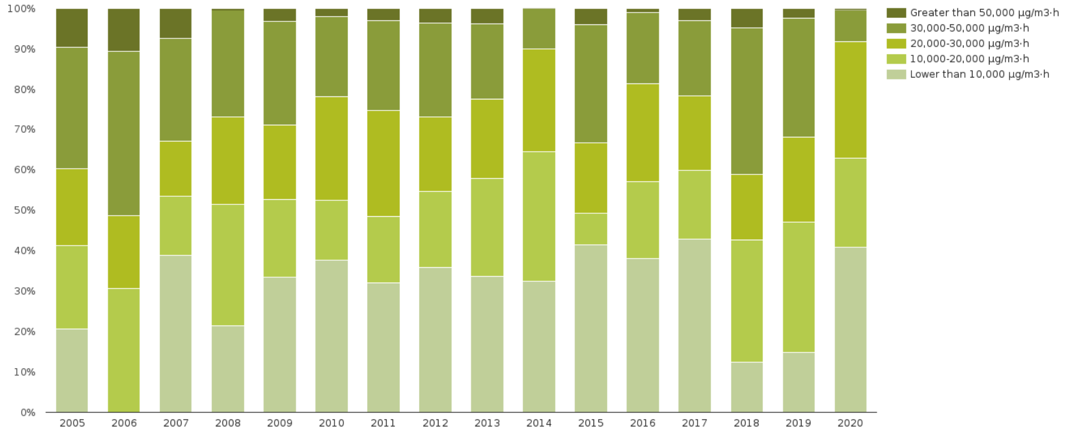


Notes:
- EU long-term objective for the protection of vegetation: 6,000 µg/m^3 · h
- EU target value for the protection of vegetation: 18,000 µg/m^3 · h (averaged over five years)

More informationData sources:
Iceland and Norway are included from 2007 onwards; Switzerland from 2008 onwards; and Türkiye from 2016 onwards. Apart from that, the current EEA-32 member countries are included during the entire period. Since 2018 maps, CLC2018 land cover data has been used (up to 2010: CLC2010; 2011-2015: CLC2006; 2016-2017: CLC2012).

More informationData sources:
Air Quality e-Reporting (AQ e-Reporting) provided by European Commission
Interpolated air quality data provided by European Environment Agency (EEA)
Air quality in Europe 2022 provided by European Environment Agency (EEA)

Chart — Exposure of forest area to ozone in EEA member countries

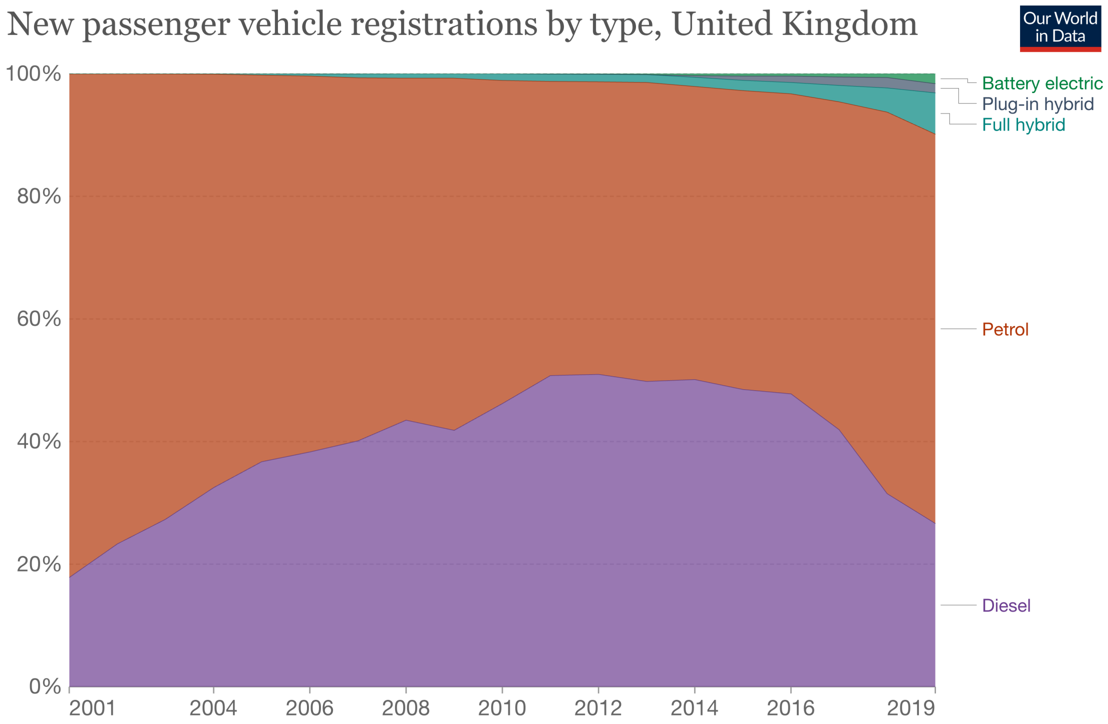


More informationData sources:
Iceland and Norway are included from 2007 onwards; Switzerland from 2008 onwards; and Türkiye from 2016 onwards. Apart from that, the current EEA-32 member countries are included during the entire period. Since 2018 maps, CLC2018 land cover data has been used (up to 2010: CLC2010; 2011-2015: CLC2006; 2016-2017: CLC2012).

More informationData sources:

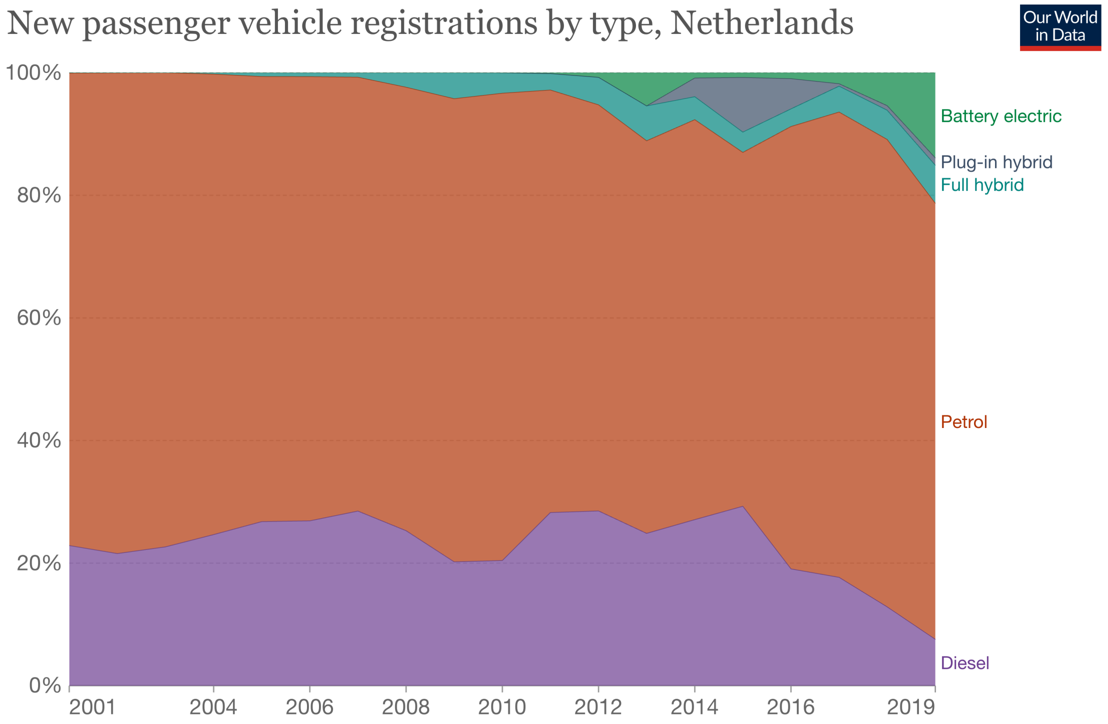
Air Quality e-Reporting (AQ e-Reporting) provided by European Commission
Interpolated air quality data provided by European Environment Agency (EEA)
Air quality in Europe 2022 provided by European Environment Agency (EEA)

1. United Kingdom



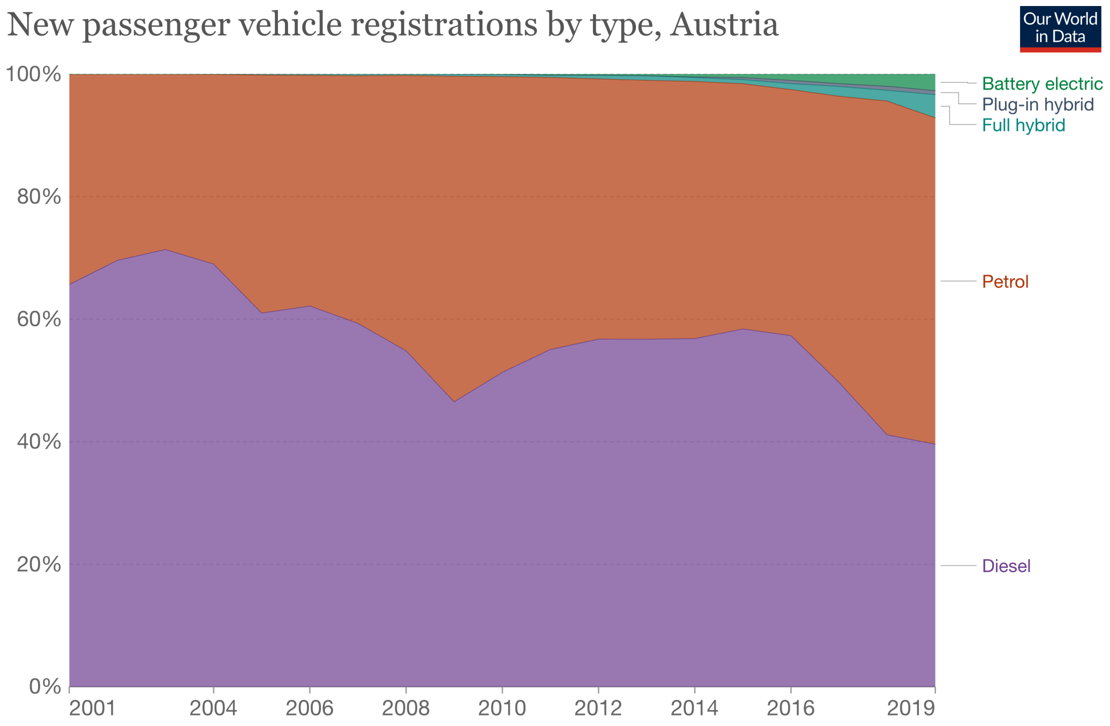
Source: International Council on Clean Transport (ICCT) and European Environment Agency OurWorldInData.org/transport • CC BY

2. Netherlands



Source: International Council on Clean Transport (ICCT) and European Environment Agency OurWorldInData.org/transport • CC BY

3. Austria

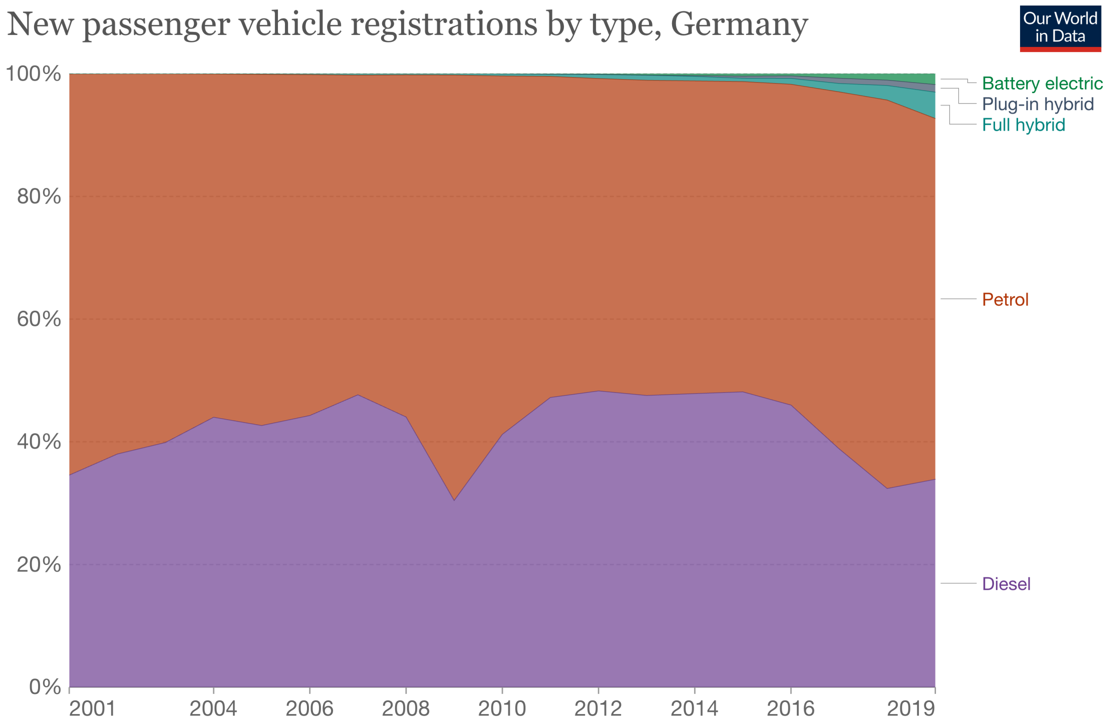


New passenger vehicle registrations by type, Austria

Source: International Council on Clean Transport (ICCT) and European Environment Agency

OurWorldInData.org/transport • CC BY

4. Germany

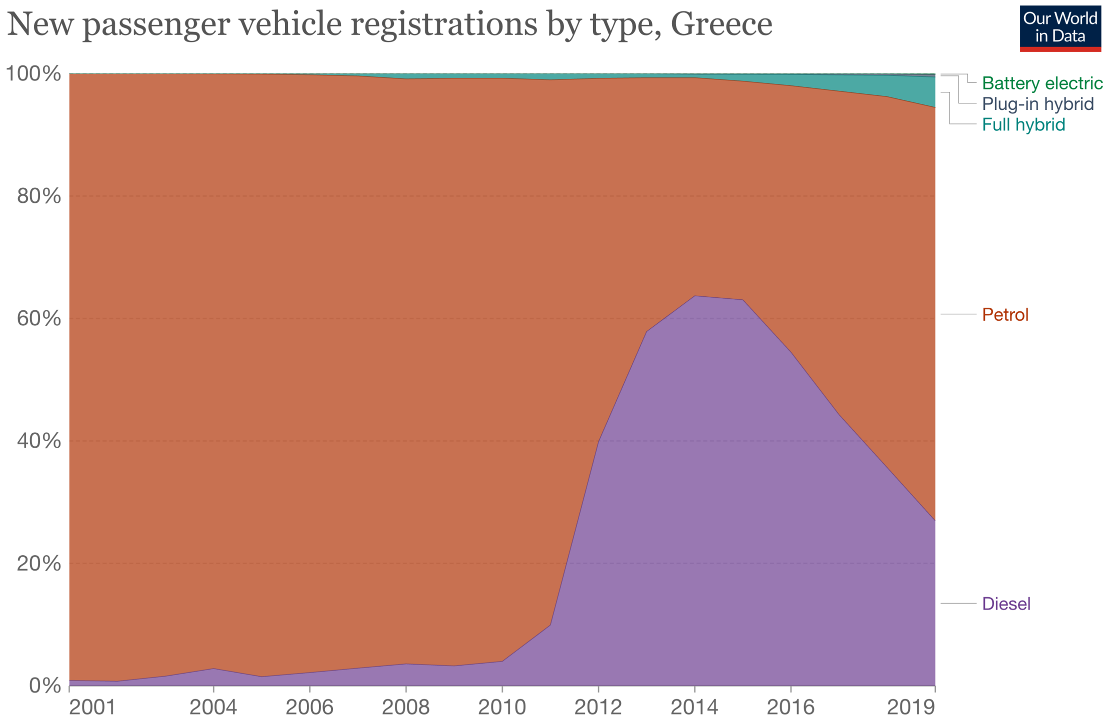


New passenger vehicle registrations by type, Germany

Source: International Council on Clean Transport (ICCT) and European Environment Agency

OurWorldInData.org/transport • CC BY

5. Greece

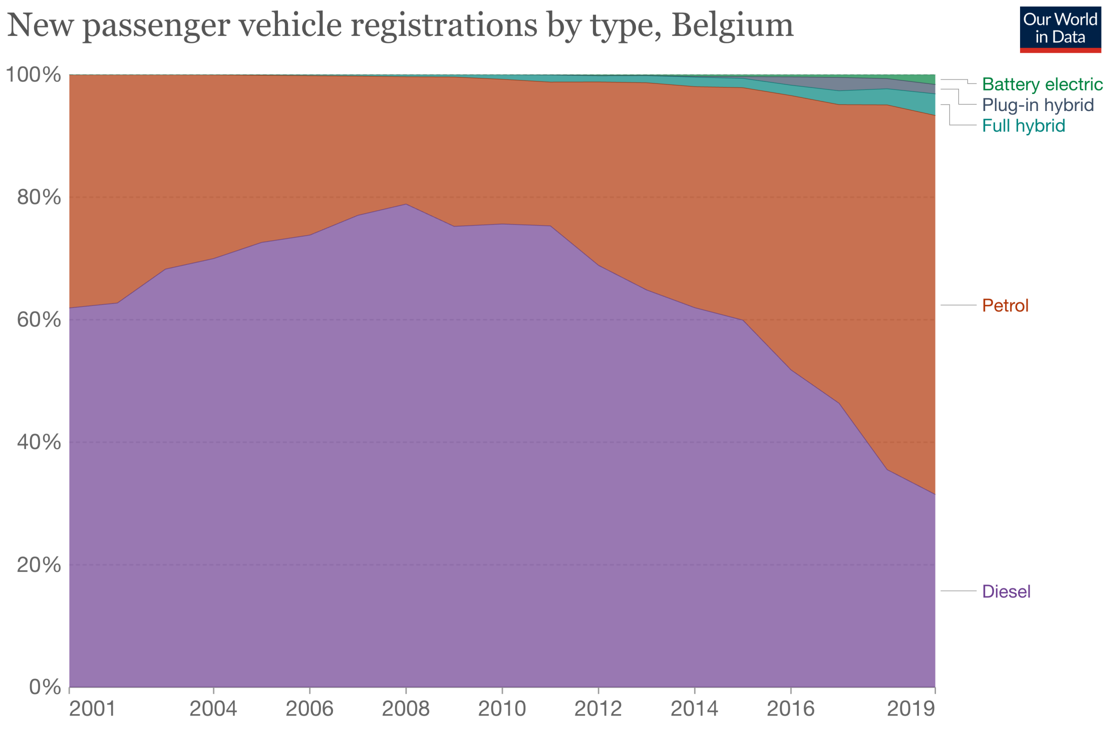


New passenger vehicle registrations by type, Greece

Source: International Council on Clean Transport (ICCT) and European Environment Agency

OurWorldInData.org/transport • CC BY

6. Belgium

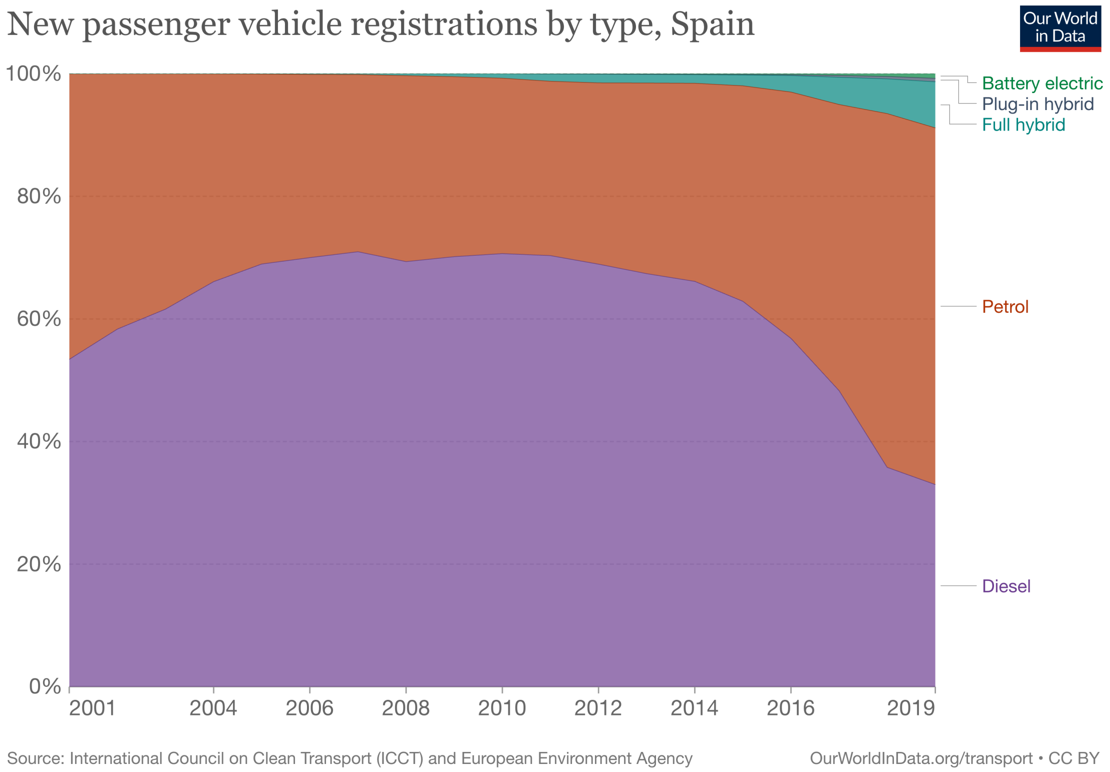


New passenger vehicle registrations by type, Belgium

Source: International Council on Clean Transport (ICCT) and European Environment Agency

OurWorldInData.org/transport • CC BY

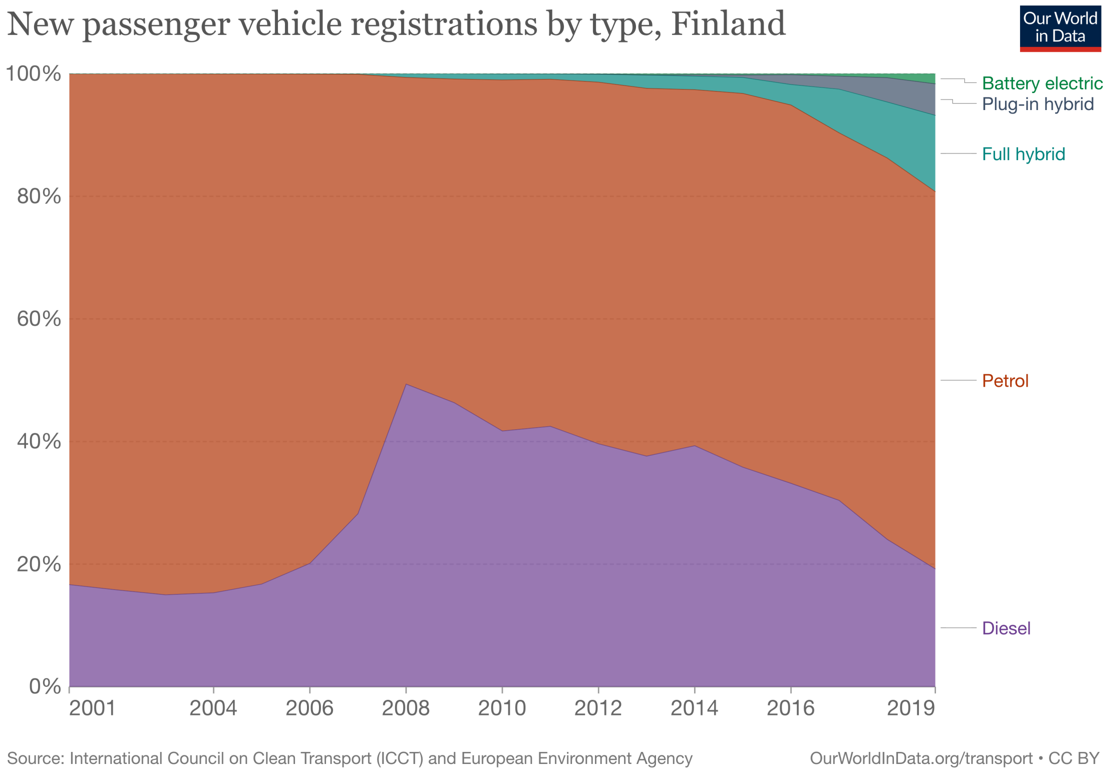
7. Spain


New passenger vehicle registrations by type, Spain

Source: International Council on Clean Transport (ICCT) and European Environment Agency

OurWorldInData.org/transport • CC BY

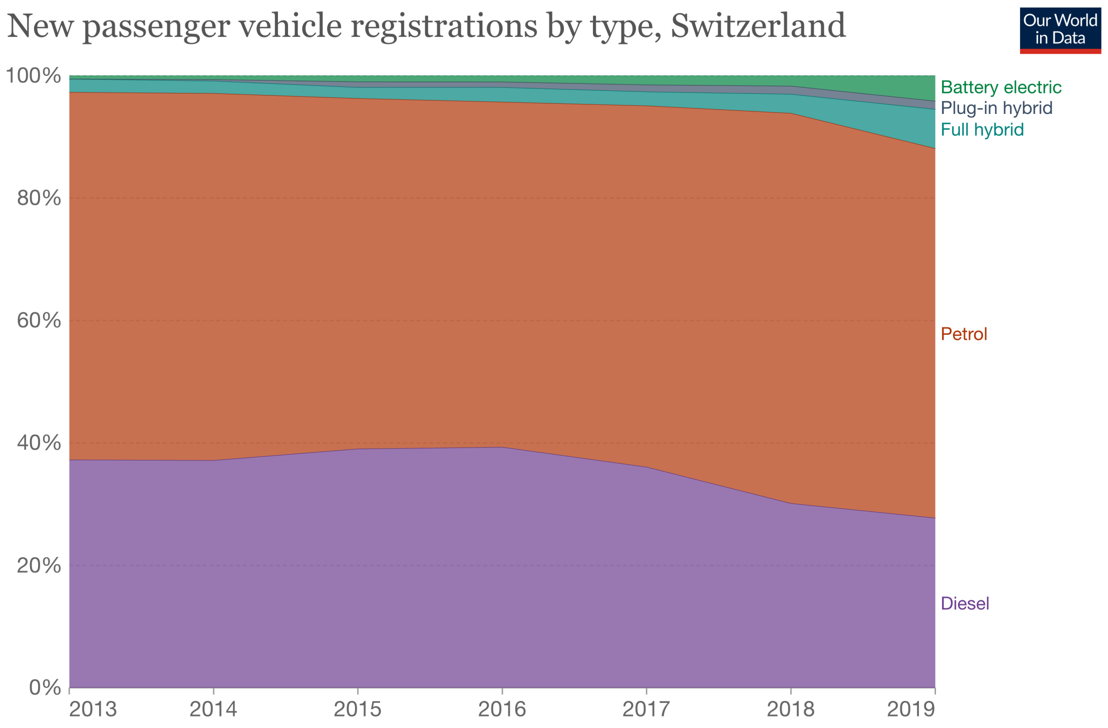
8. Finland


New passenger vehicle registrations by type, Finland

Source: International Council on Clean Transport (ICCT) and European Environment Agency

OurWorldInData.org/transport • CC BY

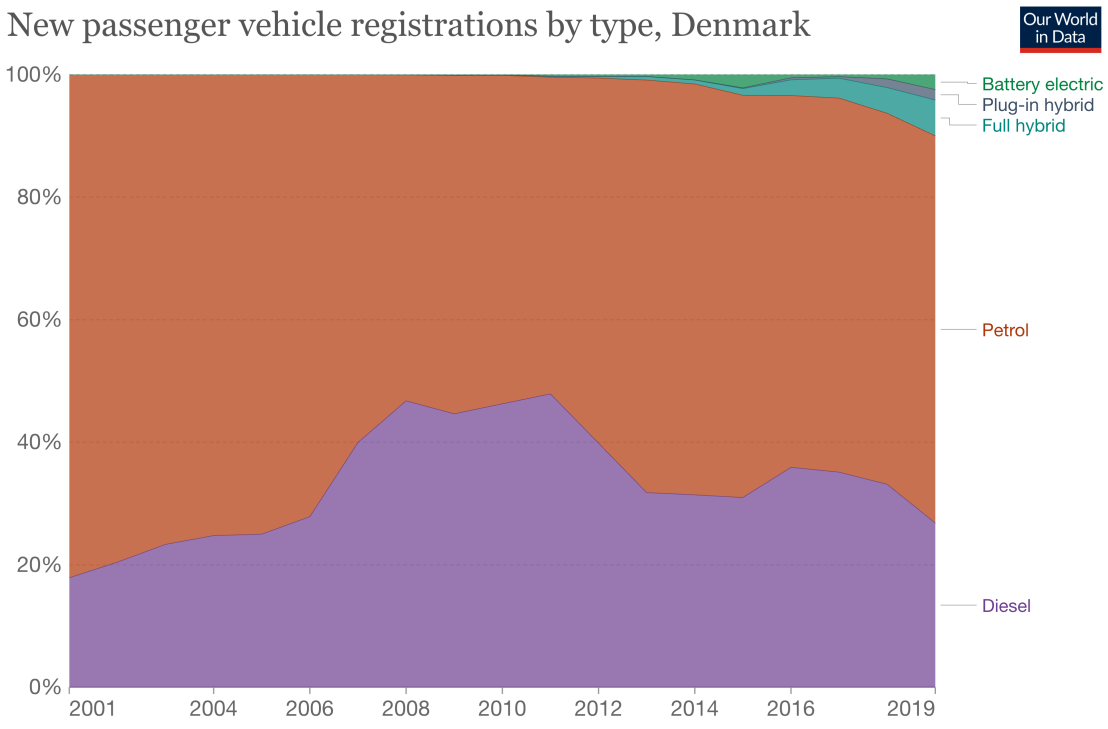
9. Switzerland


New passenger vehicle registrations by type, Switzerland

Source: International Council on Clean Transport (ICCT) and European Environment Agency

OurWorldInData.org/transport • CC BY

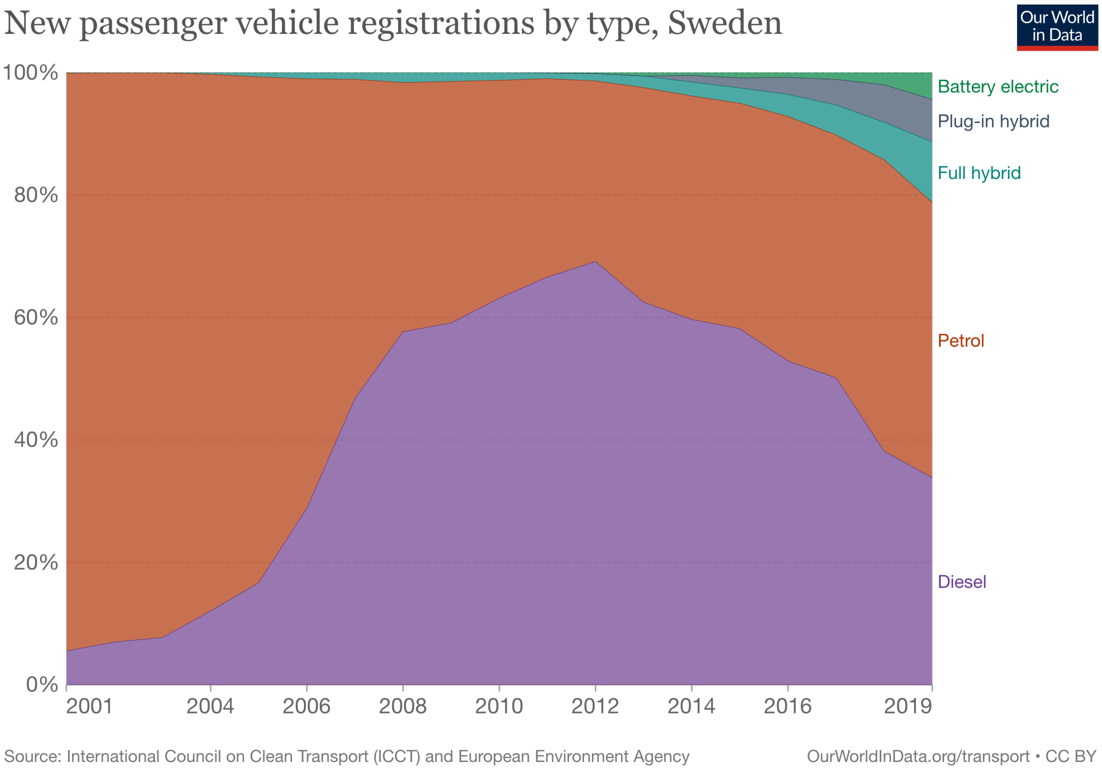
10. Denmark


New passenger vehicle registrations by type, Denmark

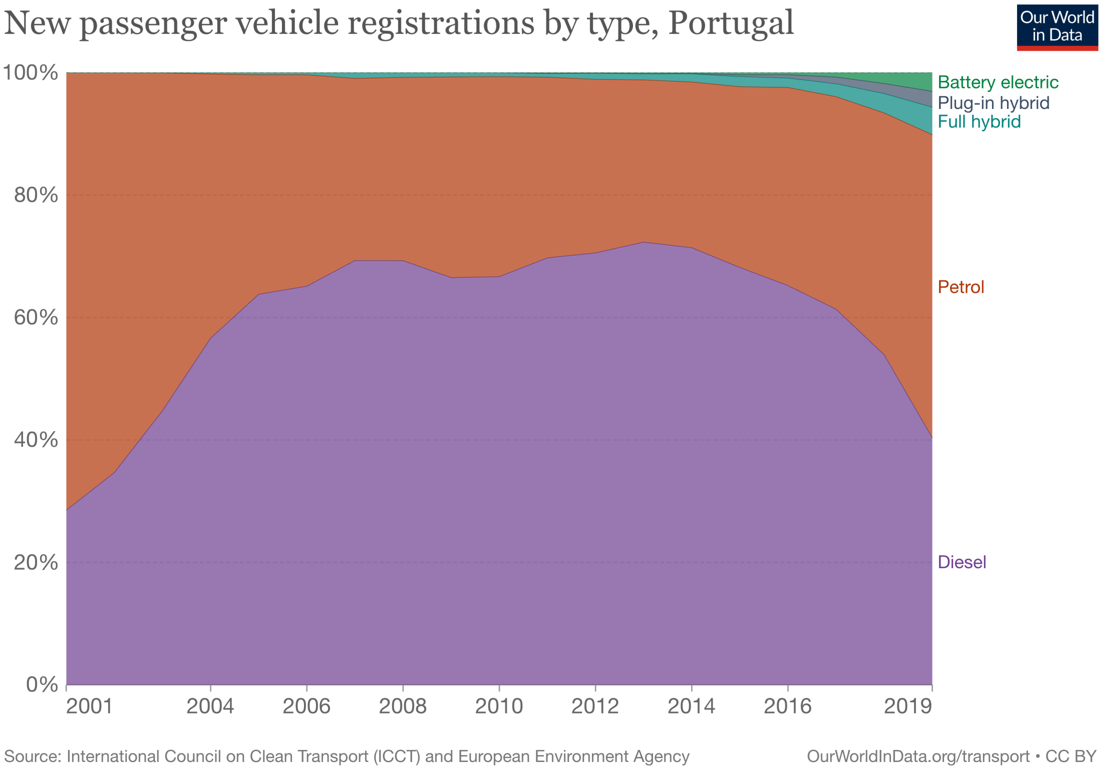
Source: International Council on Clean Transport (ICCT) and European Environment Agency

OurWorldInData.org/transport • CC BY

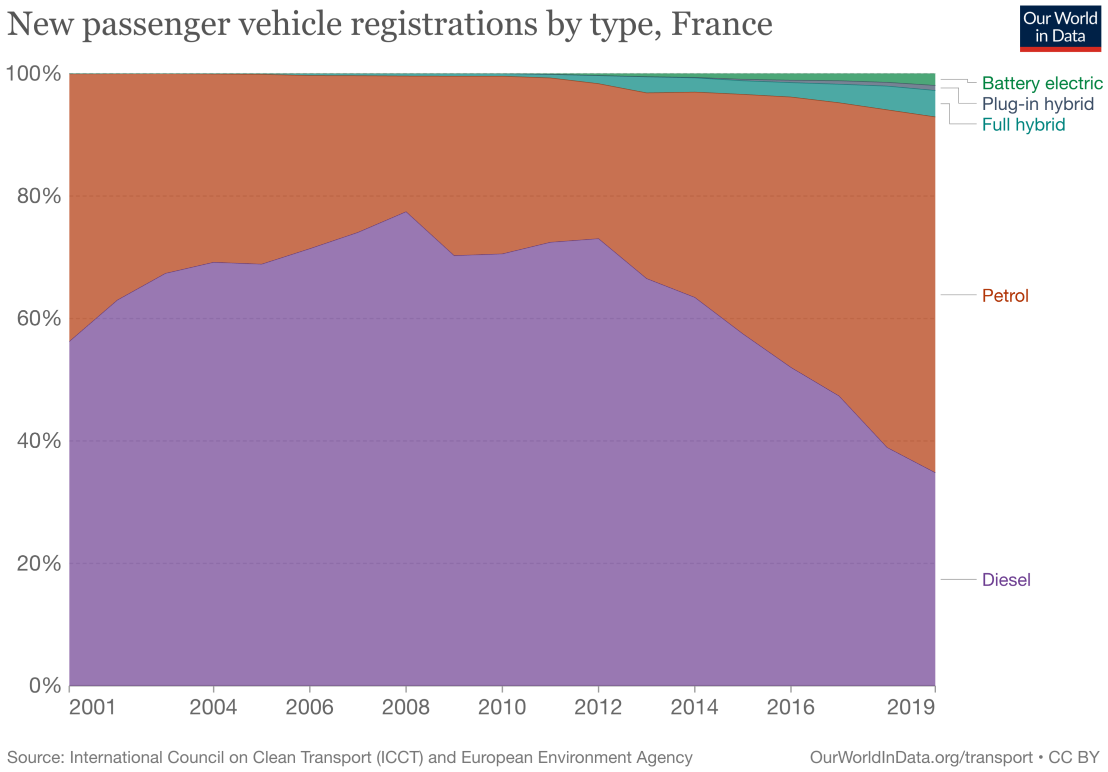
11. Sweden



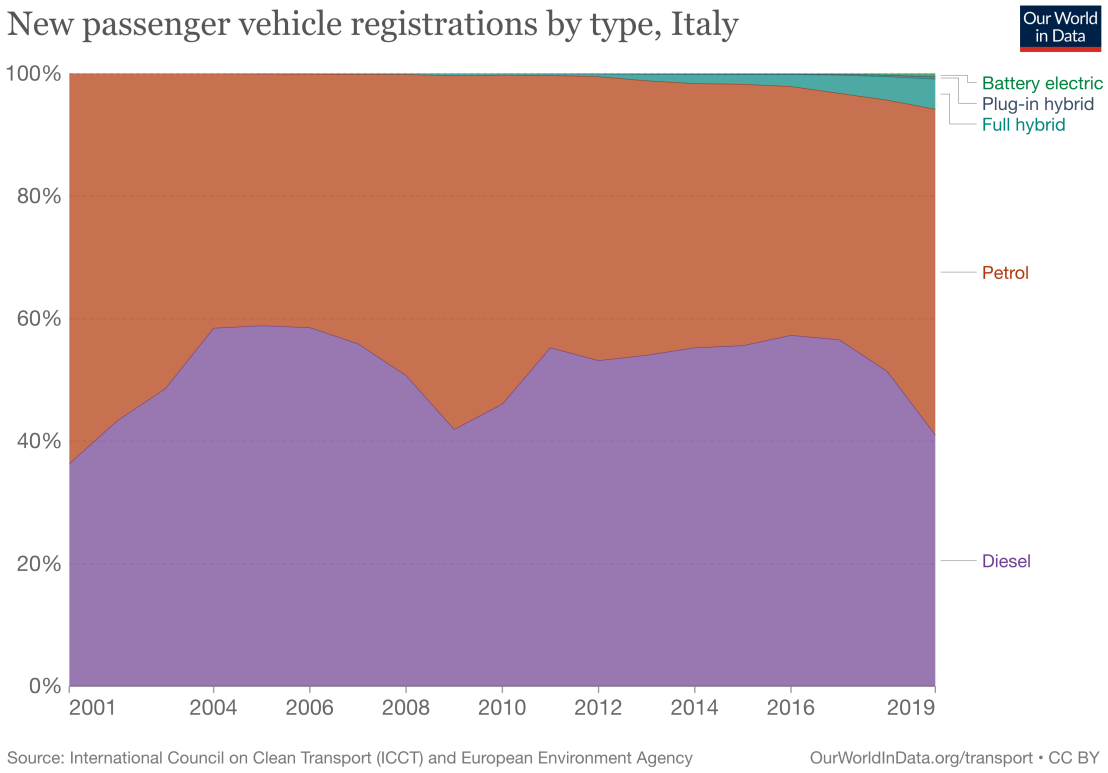
12. Portugal



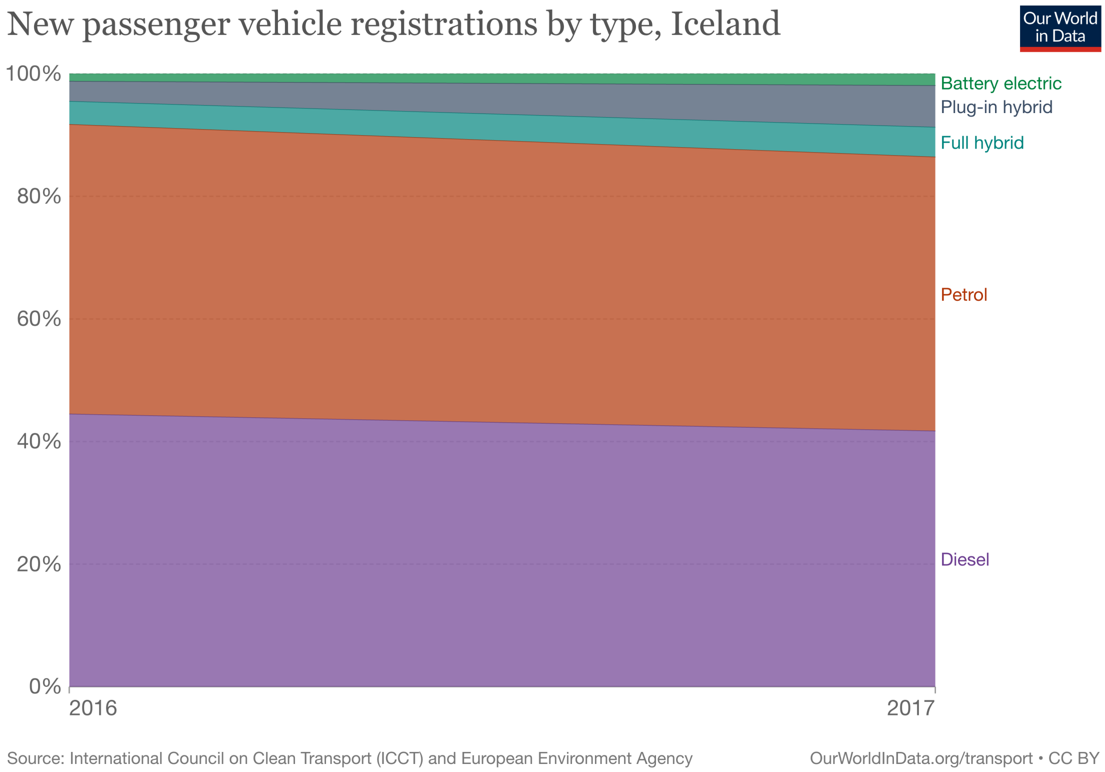
13. France



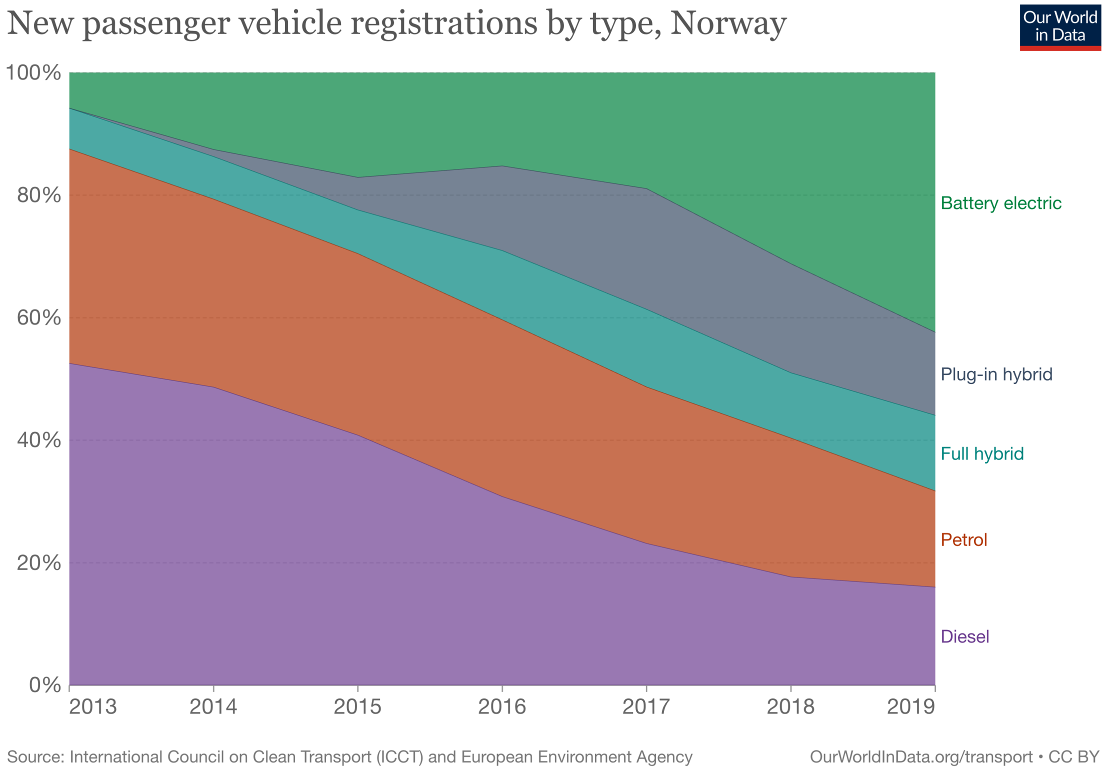
14. Italy



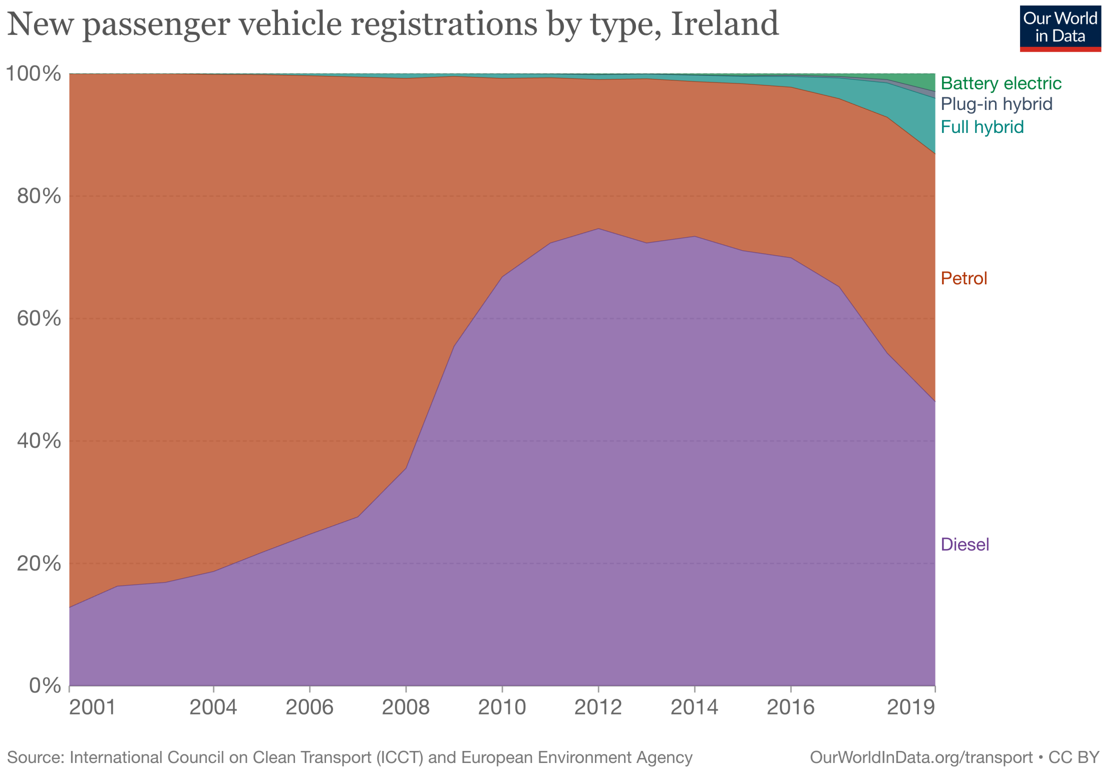
15. Iceland



16. Norway



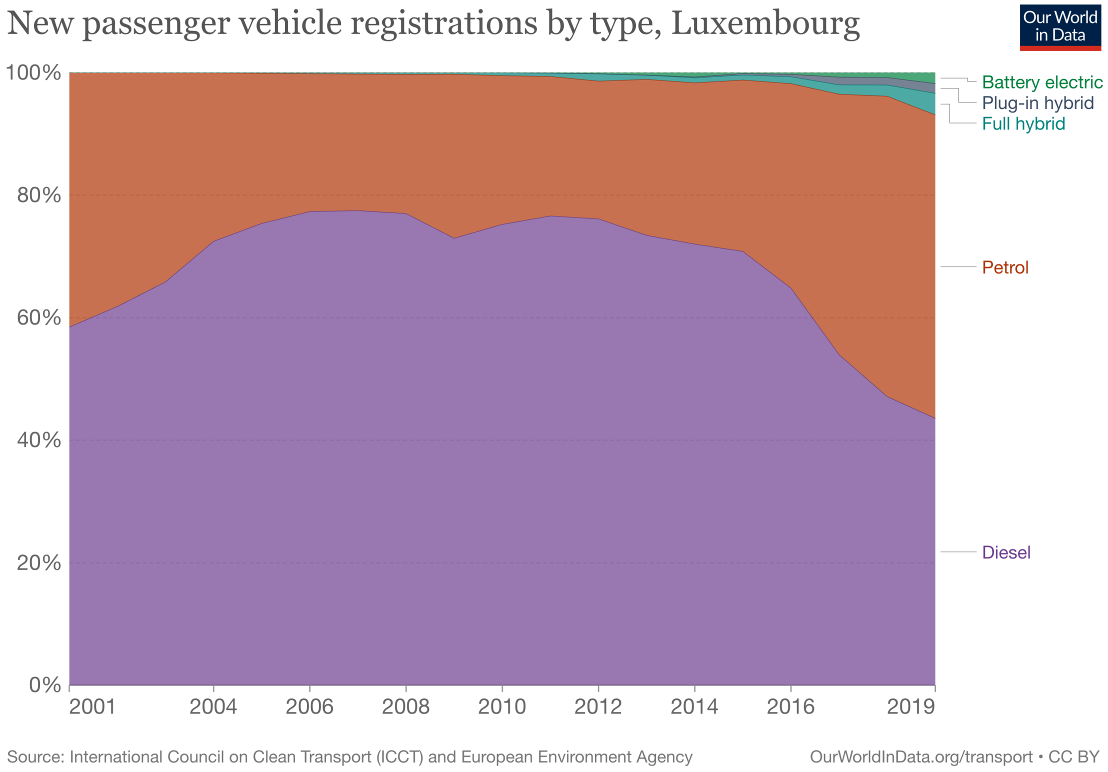
17. Ireland



Source: International Council on Clean Transport (ICCT) and European Environment Agency

OurWorldInData.org/transport • CC BY

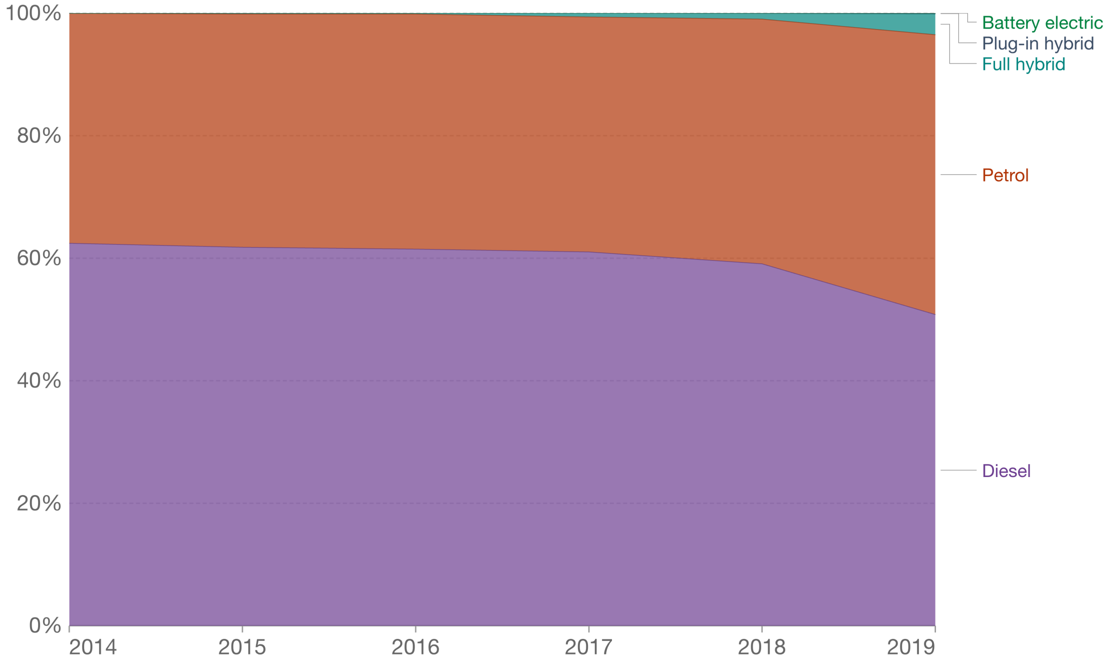
18. Luxembourg



Source: International Council on Clean Transport (ICCT) and European Environment Agency

OurWorldInData.org/transport • CC BY

19. Turkey



New passenger vehicle registrations by type, Turkey

Our World in Data

Source: International Council on Clean Transport (ICCT) and European Environment Agency

OurWorldInData.org/transport • CC BY