

DATA ANALYSIS ON THE IMPACT OF GAS EMISSIONS ON THE GREENHOUSE EFFECT

Gabrijela Dimić¹ Ljiljana Pecić²

Abstract: The study presented in the paper emphasises on tracking and investigating various aspects of the greenhouse effect through data analysis methods. Two datasets were used for the case study that is discussed in the paper. The Python libraries Pandas, NumPy, and Matplotlib were implemented, together with the Google Colab environment. The study determined which human activities had the biggest influence on the emissions of the gases that cause the greenhouse effect. With a focus on Serbia and its surrounding nations, a thorough assessment of the distribution of greenhouse gas concentrations across various geographic regions was carried out. Understanding the causes and effects of the greenhouse effect is the research's contribution, since it will help design strategies for lowering emissions and lessening the effects of climate change.

Key words: Data Analysis, Data Visualization, Greenhouse Effect, Emission of Gases, Python Libraries

1. INTRODUCTION

One of the biggest problems facing the world today is climate change, which has long-term and irreversible effects on many aspects of life, such as natural ecosystems, economic activity, and human health. Elevated levels of greenhouse gases in the atmosphere are the main contributors to climate change. Global warming is closely linked to the greenhouse effect, which causes the Earth to warm. [1]. The phrase "greenhouse effect" was coined to describe this process, which is comparable to what occurs in greenhouses when the glass walls retain heat inside the enclosed area. The phenomena happens in the lowest layers of the atmosphere when particular gases, like carbon dioxide, methane and nitrous oxide, trap solar heat that reaches the Earth's surface. These gases stop thermal energy from being completely discharged back into space because of their insulating qualities. Without these gases in the atmosphere, the Earth's surface temperature would be so low that most living forms could not survive. However, because of human activities like burning fossil fuels, deforestation, and intensive agriculture, the concentration of these gases has significantly increased during the past century. The average global temperature has significantly increased as a result, and the greenhouse effect has gotten stronger [2]. Researchers from all around the world are committed to determining and verifying empirically the impact of different elements that may help lower greenhouse gas emissions. Their objective is to find practical answers and create instruments that can guarantee long-term climate benefits. In [3], the authors described the application of the model to assess global warming by reducing greenhouse gas emissions and adapting to any warming that occurs. The authors of the study [4] present the process for determining indicators to evaluate greenhouse gas emissions. These indicators include national indicators, per capita indicators, GDP-based indicators, and indicators related to international trade. The analysis of the industrialized accumulated emissions per capita indicator shows that historical accumulated emissions per capita in industrialized countries such as the United Kingdom and the United States are higher than the world average and emissions in developing countries.

The research described in the paper focuses on the application of data analysis processes to investigate and monitor various aspects of the greenhouse effect. Two datasets were used for the case study that is detailed in the paper. The first dataset, which includes information on greenhouse gas emissions by nations worldwide from 1850 to 2023, was obtained via the Kaggle platform [5]. The Republic of Serbia's Statistical Office provided the second dataset [6]. For the years 2018–2022, it contains information on human activities (households, transportation, industry, agriculture, and building) that produce greenhouse gas emissions that contribute to the greenhouse effect. The Python programming language [7] and the Google Colab environment [8] were used to conduct the research. Python libraries [9] such as Pandas, NumPy, and Matplotlib greatly facilitated the application of various data

¹Higher lecturer, Academy of technical and art applied studies, Starine Novaka 24, Belgrade, Department School of Electrical and Computer Engineering, email:gdimic@viser.edu.rs

²Professor of applied studies, Academy of technical and art applied studies, Starine Novaka 24, Belgrade, Department School of Electrical and Computer Engineering, email: ljiljanap@viser.edu.rs

processing techniques, including analysis, visualization, and dataset cleaning. The study identified the gases most responsible for the greenhouse effect and the human activities that have the greatest impact on their emissions. A comprehensive evaluation of the distribution of greenhouse gas concentrations across different geographical regions was conducted, with an emphasis on Serbia and neighboring countries. The aim of the research is to establish a scientific foundation for understanding the causes and consequences of the greenhouse effect in order to help develop strategies for reducing emissions and mitigating the impacts of climate change.

2. THE LIFECYCLE OF THE DATA ANALYSIS PROCESS

Combining mathematics, statistics, programming, machine learning, data visualisation, and data mining, data science is an interdisciplinary field. Data science is one of the professions that is expanding the fastest across all industries due to the advancement of technology and the expansion of data sources. The data analysis process life cycle comprises seven essential phases [10]. Defining the research goal: Establishing the purpose of the study and figuring out the data required to meet these objectives is the definition of the research goal. Data collection: During this stage, information is gathered using a variety of techniques from all pertinent sources. Among the methods are online scraping, surveys, interviews, real-time data harvesting from gadgets, and more. Data processing: This stage entails cleaning the data, eliminating duplicate entries, formatting the data, and storing it. The format that was predetermined for the research is used to standardise the data. Data analysis: Several data analysis techniques are used at this phase [11]: descriptive analysis, diagnostic analysis, predictive and prescriptive analysis. Presentation of analysis findings: Following analysis, important insights are found and shared with the appropriate parties. Usually, reports or other types of data visualisation are used for this. Decision model selection and application: A suitable decision model is selected based on the analysis's findings. Monitoring model performance: Tracking the performance of the chosen model is the last phase in the data analysis process. Verifying if the model performs as anticipated and whether the intended outcomes were obtained within the allotted period is the aim of monitoring.

3. CASE STUDY

The dataset obtained from the Kaggle platform contained 41.238 rows and 6 columns. Using the relevant Pandas library functions, the dataset was imported into a *DataFrame* after being downloaded in CSV format. The following columns were found: *Entity*, *Code*, *Year*, *Annual CO₂ Emissions*, *Annual Methane Emissions*, and *Annual Nitrous Oxide Emissions in CO₂ Equivalents*. The *Code* column was eliminated during the data preprocessing phase since it was unnecessary for the study. *Drzave (Countries)* and *Godine (Years)* were substituted as the column titles for *Entity* and *Year*, respectively. No null values were found after using the *isna().sum()* method, and no duplicate rows were found using the *duplicated()* function. Data on Serbia's gas emissions were retrieved during the descriptive analysis phase. The annual emissions of greenhouse gases from 1850 to 2023 were plotted on a line graph. (Fig. 1).

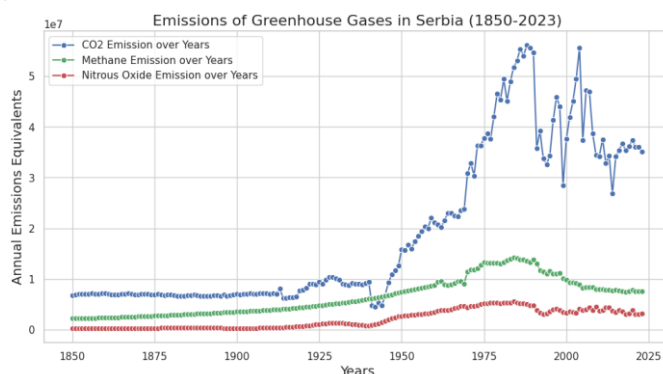


Figure 1 – The annual emissions of greenhouse gases in Serbia from 1850 to 2023.

Global CO₂ emissions in Serbia increased rapidly between 1925 and 1975 (Fig. 1), peaking in that year. The atmospheric CO₂ concentrations were shown to stabilise after 1975. Overall greenhouse gas contributions to the environment must be zero, though, because carbon sinks balance out any residual emissions. To do this, emissions must be drastically cut. Data extraction on Serbia's gas emissions from 2010 to 2023 was the next part of the descriptive analysis phase. Figure 2 shows the greenhouse gas emissions for the period from 2010 to 2023.

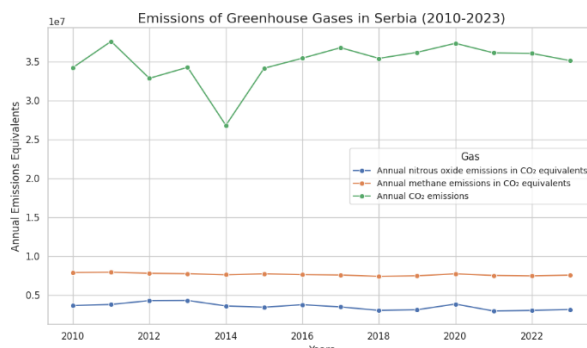


Figure 2 – The annual emissions of greenhouse gases in Serbia from 2010 to 2023.

It is clear that concentrations dropped from 2010 to 2014 before rising again after that year (Fig. 2). Numerous factors, such as changes in industrial operations, energy production adjustments, regulatory changes, or shifts in the global economy, could be responsible for this variance in emissions. Additional factors that could account for the increase after 2014 include increased energy use, increased transportation-related emissions, or a slowdown in the transition to renewable energy.

Data for the following countries were removed and presented for a more thorough analysis: Bosnia and Herzegovina, Croatia, Slovenia, Montenegro, Romania, Greece, Austria, and Bulgaria (Fig. 3).

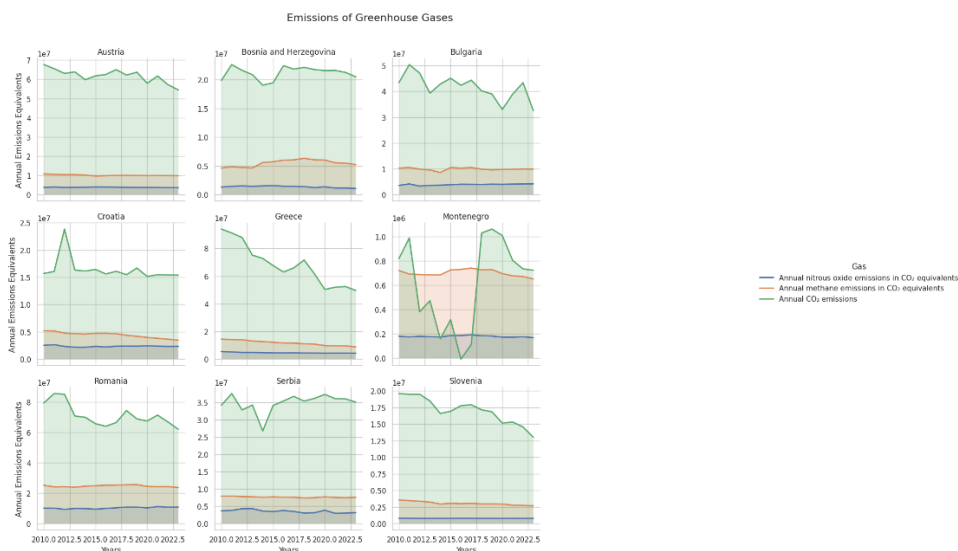


Figure 3 – The annual emissions of greenhouse gases in Serbia and the countries in the region from 2010 to 2023.

The study revealed various changes in the concentrations of greenhouse gas emissions among the countries in the region. For example, Montenegro experienced a dramatic decrease in emissions from 2010 to 2016, followed by a strong increase in CO₂ emissions. While the sharp rise after 2016 might suggest a recovery in energy consumption growth or a lack of effective emission control strategies, the

decrease in emissions may be due to a shift to cleaner energy sources or a reduction in industrial activity.

To visualise geographical data, a geographic map was created using the *scatter_geo* function from the *Plotly* package in *Python* (Fig.4). This function is useful for displaying geographic data points. To enable a more efficient display, a new column was created to show the total greenhouse gas emissions at specific locations.

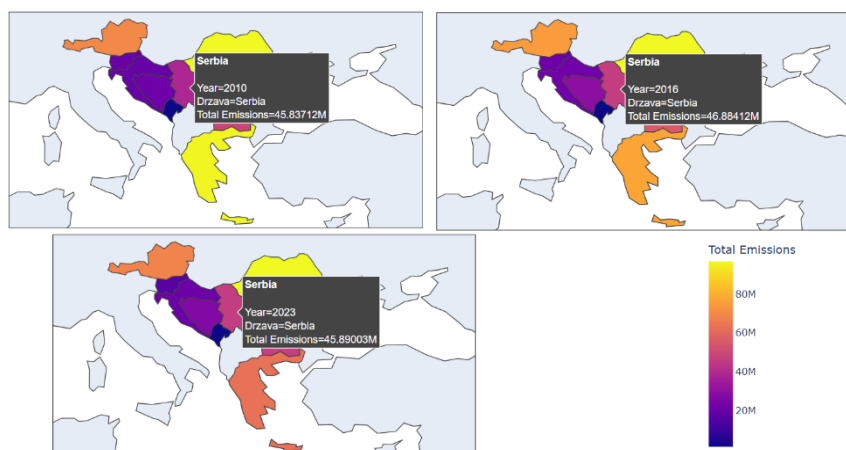
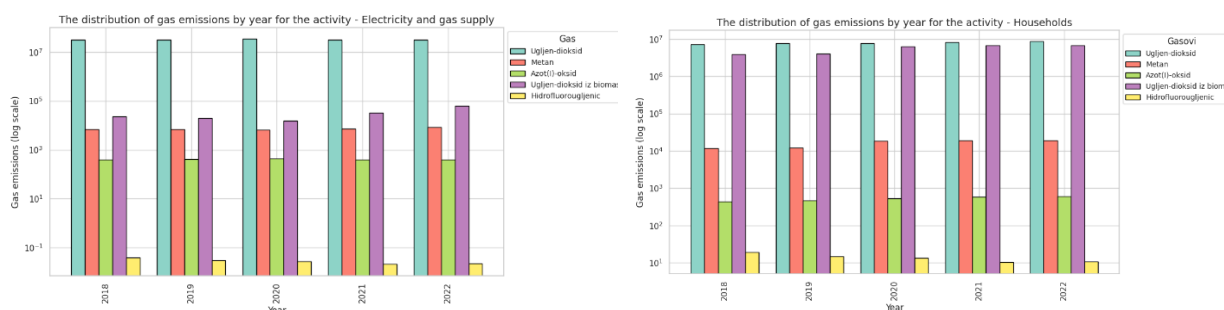


Figure 4 – Geographic representation of Annual Emissions by Country in Europe.

Many human activities that release gases into the atmosphere are the cause of the greenhouse effect. These include manufacturing, transportation, industry, agriculture, and construction. The gases primarily impacted by the greenhouse effect are carbon dioxide, methane, nitrous oxide, carbon dioxide from biomass, hydrofluorocarbons, and others. The dataset, which was acquired from the Republic of Serbia's Statistical Office, contained data on greenhouse gas emissions brought on by human activity. The information was gathered between 2018 and 2022. It has been determined that the activities that contribute most to greenhouse gas emissions are households, manufacturing, various industries, agriculture, and the provision of gas and electricity. A study of gas concentrations by year was presented for the selected activities (Fig. 5). The data was analysed using a histogram bar chart, which shows the distribution of gas emissions over different time periods. Histograms are helpful for showing frequency across categories, while bar charts provide a simple way to compare various values or changes over time. By improving data interpretation and analysis, this method aids in spotting important long-term patterns in gas emissions.



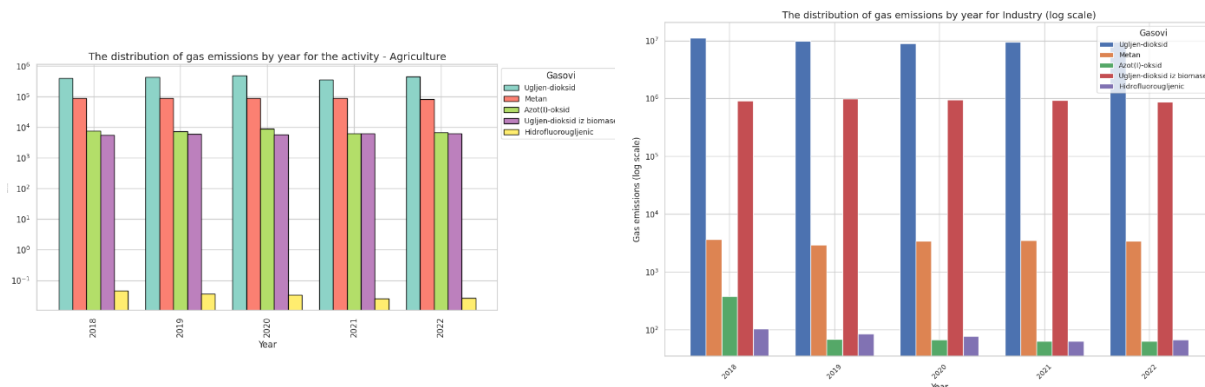


Figure 5 – The distribution of gas emissions by year for the activity

The study demonstrates that carbon dioxide, or CO₂, is by far the most common gas for all processes that were examined. Carbon dioxide concentrations from biomass are noticeably higher for houses. Along with increased levels of carbon dioxide, nitrous oxide, and carbon dioxide from biomass, methane concentrations are also noticeably elevated in agriculture. Methane [12] is a naturally occurring gas that helps to warm the surface of our planet. The decomposition of organic matter from plants and animals, geological processes including rock breakup, and volcanoes are some of the most frequent natural sources of methane emissions into the atmosphere. Human activities including burning fossil fuels, raising livestock, and landfills can also release methane into the atmosphere. The last stage of the breakdown of organic matter, which can originate from either plants or animals, in an oxygen-deficient environment results in the formation of methane. During this stage, biological matter reverts to its most basic form, which consists of molecules like methane that are discharged into the atmosphere. An important contributor to agricultural gas emissions is cattle. The stomach of cattle is distinct from that of humans in that it contains four chambers. Partially digested food is kept and fermented in one of these compartments. According to [13], this atmosphere, which lacks enough air, produces the conditions for methane generation. Nitrous oxide, or N₂O, is also widely used in agriculture. Fertilisation is one of the primary sources of nitrous oxide production in agriculture. About half of the fertiliser that is applied to the soil is absorbed by it, with the other half remaining on the soil's surface. The leftover fertiliser starts to change into nitrous oxide if it doesn't have enough oxygen. Compared to other activities, the presence of methane and nitrous oxide is much lower in the case of manufacturing. Hydrofluorocarbon gases, which contain fluorine, are the least represented of all the activities that have been recorded.

4. CONCLUSION

A case study that applies a data analysis procedure on how gas emissions affect the greenhouse effect is described in the paper. The Statistical Office of the Republic of Serbia's database and the Kaggle platform provided the dataset for analysis. Preparing the data for analysis was one of the necessary steps completed during the preprocessing phase. The data collection, which was acquired from Kaggle, included country-by-country records of greenhouse gas emissions from 1850 to 2023. Data for the years 2010–2023 pertaining to Serbia and the neighbouring countries were retrieved. Records of human activities that result in gas emissions that cause the greenhouse effect were included in the dataset, which was taken from the Statistical Office of the Republic of Serbia's database and covered the years 2018–2022. Visualisations were made once the concentrations of petrol emissions were noted.

What the entire globe has been warning about has only been validated by data analysis pertaining to the greenhouse effect and the gas emissions that affect it. Every year, the issue gets exponentially worse, and it's getting harder to ignore how quickly emissions are rising. Similar to industry, which is a major source of pollution, agriculture has a big influence on gas emissions in some nations, including Serbia. The examination of emissions from human activities reveals which activities are mainly responsible for the release of particular gases. Decisions regarding the best course of action to cut emissions can be made in light of these findings. It's critical to take into account every option for lowering or stopping gas emissions and to determine how much money will be needed to deal with this

worldwide problem. It's also important to consider the disparities in industrial and agricultural growth as well as the financial capacities of various nations.

To find the best answers, the decision-support system should take into account a lot of variables. Data on emissions from prior years and changes in gas emissions are important components. Future projections that can be produced with sophisticated machine learning algorithms should also be included. To determine how lowering petrol emissions will affect the world temperature, data on temperature anomalies should also be incorporated. The existing laws and regulations play a significant role in this process as well. Every suggested remedy needs to abide by current laws and global accords. The decision-support system should use all of this information to offer practical solutions that would lower gas emissions, speed up the financial return on investments, and lessen temperature anomalies—all while making sure that the solutions comply with current laws and regulations.

5. REFERENCES

- [1] Jeffry L., Ong M. Y., Nomanbhay S., Mofijur M., Mubashir M., Show P. L.: *Greenhouse gases utilization: A review*, Fuel, 301, 121017, 2021.
- [2] Miśkiewicz, R. *The Impact of Innovation and Information Technology on Greenhouse Gas Emissions: A Case of the Visegrád Countries*, Journal of Risk and Financial Management, 14(2), 59, 2021, <https://doi.org/10.3390/jrfm14020059>.
- [3] Hope C., Anderson J., & Wenman P.: *Policy analysis of the greenhouse effect: an application of the PAGE model*, Energy Policy, 21(3), 327-338, 1993.
- [4] Zhang Z., Qu J., & Zeng J.: *A quantitative comparison and analysis of the assessment indicators of greenhouse gas emissions*. Journal of Geographical Sciences, 18, 387-399, 2008.
- [5] Kaggle, "CO2 and Greenhouse Gas Emissions", <https://www.kaggle.com/datasets/imtkaggleteam/co-and-greenhouse-gas-emissions?resource=download>, December 2024.
- [6] <https://data.stat.gov.rs/Home/Result/25040101?languageCode=en-US>, December 2024.
- [7] <https://www.python.org/>, December 2024.
- [8] <https://colab.research.google.com/>, December 2024.
- [9] VanderPlas, J.: *Python Data Science Handbook*, O'Reilly Media, 2016.
- [10] Dimić G., Pecić Lj., Bogojević B.: Implementation of the data analysis process to identify the most common causes of air pollution in the Republic of Serbia, *International Multidisciplinary Conference "Challenges of Contemporary Higher Education" - CCHE 2024*, Volume 1, pp. 537-542.
- [11] Grus, J.: *Data Science from Scratch: First Principles with Python*, O'Reilly Media, 2019.
- [12] <https://www.carboncollective.co/sustainable-investing/methane-ch4>, January 2024.
- [13] CLEAR Centre, "How Do Cattle Produce Methane?", <https://clear.ucdavis.edu/explainers/how-do-cows-produce-methane>, January 2024.

Please save your paper as Surname-ACCHE-2025.doc written according to the provided guidelines and submit it using the online upload form no later than **January 20, 2025. at e-mail: conference@acche.rs** .