**Research Report**

Greenhouse Gas Emissions

Suhainsly Martha, student number 2602068

Tomas van Daalen, student number 2598668

Rob Schilder, student number 2572418  
  
Digitization: from Object to Data  
Vrije Universiteit Amsterdam

October 2018

**Introduction**

In 2018, it is impossible to ignore the reality of climate change; from droughts in Syria, to national weather changes. Even universities such as the Vrije University allows people the option to study climate change with courses available in the Humanities. In political news, the European Union (EU) member states have set goals in various agreements to limit the increase of global temperatures. Countries globally agreed under the United Nations Framework Convention on Climate Change (UNFCCC) to limit 2°C compared to the average temperature in pre-industrial times to prevent the most severe impacts of climate change and possibly catastrophic changes in the global environment. To achieve this, the world must stop the growth in greenhouse gas emissions by 2020, and reduce it by 60% by 2050 compared with 2010. The new goals have been set in the Paris agreements in 2015. These were to decrease global average temperature to well below 2 °C above pre-industrial levels and to limit the increase to 1.5 °C, since this would substantially reduce the risks and effects of climate change. The rest of this paper will focus on this issue.

**The research**

The goals of this research is to find out which industry needs to minimize the growth of emissions of greenhouse gasses and to find out how this growth can be explained, subsequently, our research question is:

*In which type of industry did the greenhouse emissions grow the fastest between 1990 and 2016 and how can this be explained?*

To answer this question, this research will look at the data set on greenhouse gas emissions and removals sent by EU member states to UNFCCC and the EU Greenhouse Gas Monitoring Mechanism. This data set reflects the Greenhouse Gas (GHG) inventory data for 2018 as reported under the UNFCCC. In the database, you can find the emission values from 1990 until 2016 reported by each country individually.

To find the answer to the above question, the first objective is to find the main polluting industry. Therefore, we ask two sub-questions:

*Which type of industry has the biggest growth in the emissions of greenhouse gasses in the period 1990-2016 following the dataset of the UNFCCC?*

We are trying to answer this question by analyzing the nature of the dataset, the dataset itself and the tools provided by the European Environment Agency (EEA) website. After finding out which type of industry has the biggest growth in emissions, the next step will be to explain why this specific industry has the biggest growth. This will be done following our second sub-question:

*What are the reasons for the growth of emissions of greenhouse gasses in the biggest growing polluting industry?*

To answer this last sub-question, we will look at secondary literature about the subject.

**Relevance of the research**

The emission of greenhouse gasses, with global warming as a result, is a subject that affects everybody. The global sea level rising and the complete disappearance of vegetation are just some reasons as to why global warming should concern everybody. This research is focussed on the industry that has the biggest increase in emission of greenhouse gasses in the period of 1990-2016. The information that will be retrieved from this research can be used to tackle the problem of the emissions of greenhouse gasses and affect actual policy discussions. We will also look for explanations for the increase in the emission of greenhouse gasses. Because of this, it is possible to look at the problem from a new perspective, and with that, try to come up with new solutions to tackle a small part of the problem of the pollution of the earth.

**Data analysis**

This dataset is massive and is structured and given in csv format, although the initial download is of a zip file on the EEA website. The dataset and subsequently the data is copyrighted by this agency. However it is open, linked and available for the public to view as long as the user complies with the EEA standard reuse policy. The dataset and eventual visualization materials are available for commercial or non-commercial purposes free of charge, provided that the source is acknowledged.[[1]](#footnote-0) Further metadata is found on the Eionet website. Eionet is a governmental open big data source. With Eionet EEA is trying to promote and further environmental research and policies via big data. This data is gathered through continuous monitoring and reporting. On this website the viewer can find an extensive overview of not only the general information but also legislative papers, the common reporting obligations papers and the original reporting papers in several languages of the 28 EU member states over several years, just to name a few. Metadata is important for the interpretation and searching of particular data especially if the data comes from various sources (Janssen, Charalabidis & Zuiderwijk, 2012). In the case of this dataset, the metadata is extensive.

Diving into the dataset itself, there are several things that are apparent. First of all the level of standardization. From the dataset it is clear that the the Intergovernmental Panel on Climate Change (IPCC) Common Reporting Format has been used. The data collected for this dataset has strict regulations for every country. The regulations are available on Eionet in the form of EU policy papers at which in the end there are approximately 50 tables that should be electronically submitted by every country annually (see appendix table 1). The tables represent the specific sectors and parameters set. So at EU report submission levels, the data collected is highly standardized. This reporting process is also a massive undertaking consisting of several teams for every country collecting the data, curating the data and submitting the data. However, both the extensiveness of the metadata and the openness of the data mentioned previously have their limitations. In the case of this particular dataset, the original submitted reports with tables of the EU countries are not available. They are seen and searchable on Eionet but when it comes to accessing them the viewer would need login access. To get an username, you would have to email the Eionet helpdesk and when done so there is no response. This leads us to the issues of data collection and provenance. Because the original data are not available for the public there is no say into how the data was changed to fit the final dataset. Moreover, the country’s collection processes are not made public.

Looking into the dataset itself, there appears to be no dumbing down of data once one looks at the sectors and values given. All tables submitted by countries are found back in the dataset. However, looking closely at the dataset brings up questions of quality. The data set is annotated. Where there no values for the emission, there are annotations “C”, “IE”, “NA”, “NO” and a few more. What these notations mean are not immediately clear in the dataset. However, in the previously mentioned IPCC reporting guidelines, the notations are defined. For example, for the notation “NA”, not estimated, the countries have to report why the estimation has not been made. This is however not made clear to the public because of the provenance issue. So maybe the countries have been this clear in the required tables, but this information did not make it back to the data set. So there is dumbing down of information once you look deeper. So given that there is no original data available, we researched further into the obligations that the countries’ have to stick to while collecting and writing down the data. One of these obligations is to not only make clear in the given tables that there is a level of uncertainty with values, but they have to make this clear in the form of completely different tables from the required 50. It is clear, that not all data gathered can be said to be 100% conclusive. However, this shred of uncertainty is not made clear anywhere in the data set. The issue of provenance plays a role here. We can assume that the given information per country is correct in the dataset used, what the regular public can not assume is that the data reported is factual. So the data is public in name only but is still private in practice, which makes curating the dataset nearly impossible (Janssen, Charalabidis & Zuiderwijk, 2012).

**Methodology**

The data in the dataset is complete enough to fully answer the research question. With a few exceptions, the ‘base year’ for tabulating greenhouse gas emissions has been set as 1990. Therefore, the ‘base’ emissions of the year 1990 still must be found. These can be found on the internet for each specific industry reported by the European Union. Due to the data being very detailed, no other steps of enriching the data was necessary to query the data.

The size of the document is 86,2 MB in the .csv format. The research has been conducted in the program OpenRefine. The data is in the format required for the research, thus no conversion was necessary. Due to the large nature of the dataset, it works best on a computer with a rather good processor.

To answer the first sub-question, some parts of the data are more relevant than other parts. The categories relevant for the research are: Country, Pollutant\_name, Year, Sector\_name and emissions. For example, consider the category ‘Pollutant\_name’ and ‘emissions’ from the individual countries. This database concerns the annual emissions of different GHGs. All these GGHs are expressed in Gigagrams CO2-equivalent (Gg CO2-eq). This relates the other GHGs to CO2 (EEA, 2018). The time period we consider is 1990 until 2016.

As previously mentioned it was not necessary to query the data but we did work around in Openrefine to be do an accurate data analysis. and to find individual values To find specific values for instance we opened the dataset in openrefine and we clicked the downwards arrow the relevant column, then went to facet and text facet to see all the text available in this column. Once the texts were available we were able to click on specific gasses and values to do some comparative analysis to answer the sub- and research questions. From this, we made the conclusions towards the total gas emissions in the several industries. Likewise, to make conclusions the multiple graphs and visualization provided by EEA website was used to make conclusions on the research and sub-questions.

**Findings and interpretations**

As said before, we did not really make use of the programs in order to understand and edit the data and available visual material provided by EEA was used. For instance, figure 1 in the appendix shows the greenhouse gas emissions for the main sectors of the European industries. Fuel combustion for transport is the second most important sector in 2016, with 24% of the total share. It has increased its contribution significantly since 1990. In table 3 in the appendix, the information of the pie chart is also explained in numbers. From this table, it is visible that the aviation industry in particular is growing really fast. In fact, it is the only industry that is growing in the period 1990-2016. For this reason our answer to the first sub-question is that aviation is the biggest growing industry in the period 1990-2016.

Out of the research conducted with the dataset came the result that the aviation industry has the biggest growth in emissions. This seems logical once one thinks about it. Flights are widely available and are extremely cheap depending on where you are going. If you want to travel outside the peak season, it is possible to fly to a wide variety of locations for under 50 euros, which is often even cheaper than travelling by train or car. The low pricing makes it very tempting to travel by plane, which might be one of the reasons why the emissions of greenhouse gasses of the aviation has grown that much in the period of 1990-2016. However, it is probably not the only reason.

Hartmut Grassl (2003), a scientist who is specialized in climate change, did research on the issue of the aviation industry. According to him, the big increase of the emissions of greenhouse gasses in the aviation industry is a direct result of three factors; the low pricing of flights, the wide availability of flights and finally the large demand for flights. These three factors are closely connected, because low pricing causes a large demand, a large demand causes more availability and more availability causes lower prices. It is clear that the factors are very interdependent on each other, which might be a clear indication why the aviation industry is growing this hard.

Besides the fact that the three factors for the growth of the aviation industry are very interdependent, every factor has its own explanation. Let's first look at the wide availability of flights nowadays. Because of the increase of the demand for flights, a lot of new airlines started up. For example, in 2017 only, 78 new airlines opened their business.[[2]](#footnote-1) 29 of these 78 airlines are set in Europe. This high increase of airlines results directly in the growth in availability of flights.

Another important factor is the large demand for flights. The large demand can be declared from the decrease of the prices of flights. However, according to Grassl (2003), this is not the only reason. The increase of the demand for flights is also a direct result of the growing globalization. In short, globalization is the process in which the different countries world becomes more connected through faster communication and travelling. He also states that the globalization affected our lifestyle. Because it has become easier to go to other countries, people are much more prepared to travel further. This increasing willingness to travel further results in the increasing demand for flights.

The final factor is the low pricing of flights. The reason why the prices of flights are lower differs for every continent, following the research the focus is on Europe. From the moment that the aviation industry started up in Europe until the present day, it was largely supported by the governments of the European countries (Koning, 2017). Following a research by the European Commission it was discovered that the governments of the members of the EU gave 4,7 billion euros to the aviation industry (Koning, 2017). Another factor that influences the low pricing of the flights, is the fact that there are a lot of airports in Europe (Noack, 2017). The smaller airports have to compete with the big and famous airports to keep functioning. In order to compete with the bigger airports, the smaller airports pay airlines to come to their airport or they do not charge the airlines with landing fees. Lastly kerosene is free of excise duties, which means that the profits made from the absence of excise duties enables the airlines to lower their prices with about 20 percent (Koning, 2017).

**Conclusion**

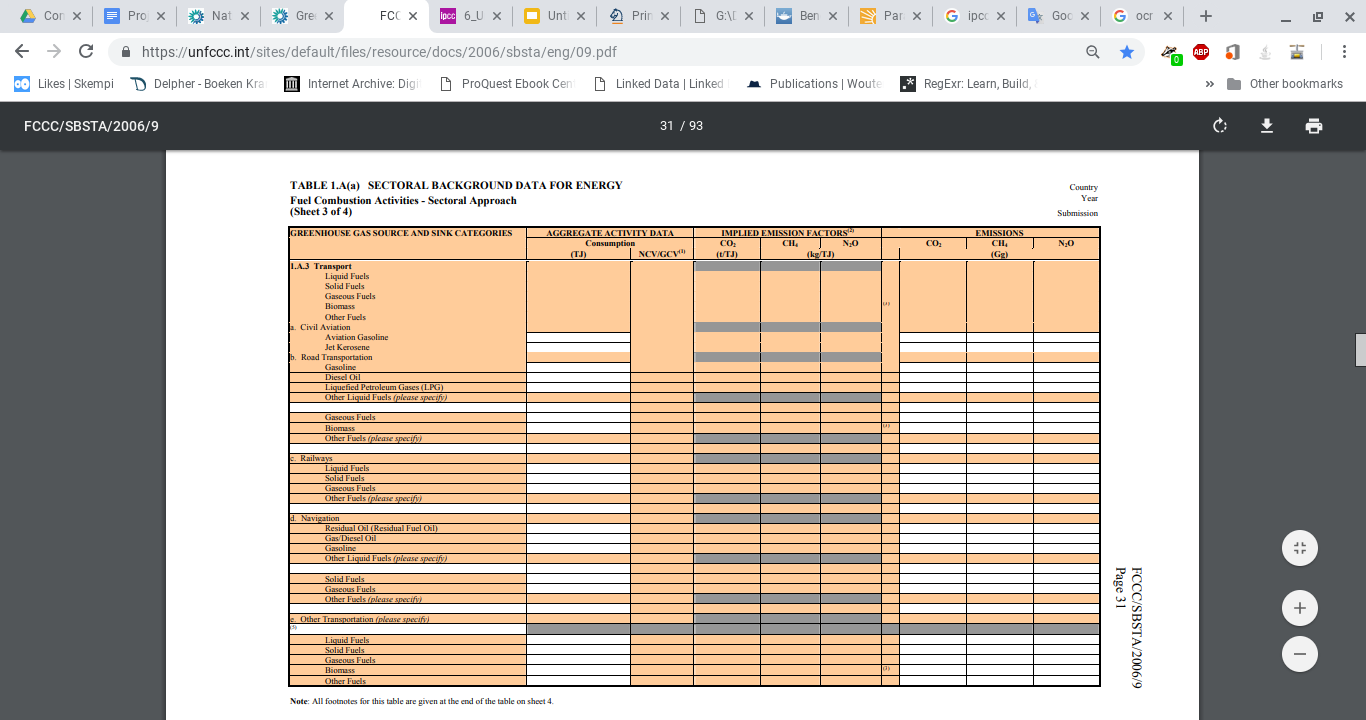
This research looked at which industry had the biggest growth in the emissions of greenhouse gasses in the period of 1990-2016. This research is conducted with a dataset that is provided by the United Nations Framework Convention on Climate Change (UNFCCC), the graphs provided with this dataset and external literature about the industry that had the biggest growth. The results of the research is that the Aviation industry had the biggest growth in the emissions of greenhouse gasses in the period 1990-2016. Further research showed that there were three factors that influenced the growth of the aviation industry, and with that the growth of the greenhouse emissions of this industry. These three factors are very interdependent and they form a certain loop that causes the aviation industry to keep growing vastly. Out of these three factors, one is especially important and probably was the main trigger for the big growth of the aviation industry. This most important factor is the pricing. The research showed that airlines were and are still able to offer flights for very low prices because of the large financial support of the governments, trough direct funding or special benefits in the taxes. For this reason, the pricing of the flights should be one of the most important things to keep in mind when attempting to tackle the emission problem of the aviation industry.

**Possibilities for future research**

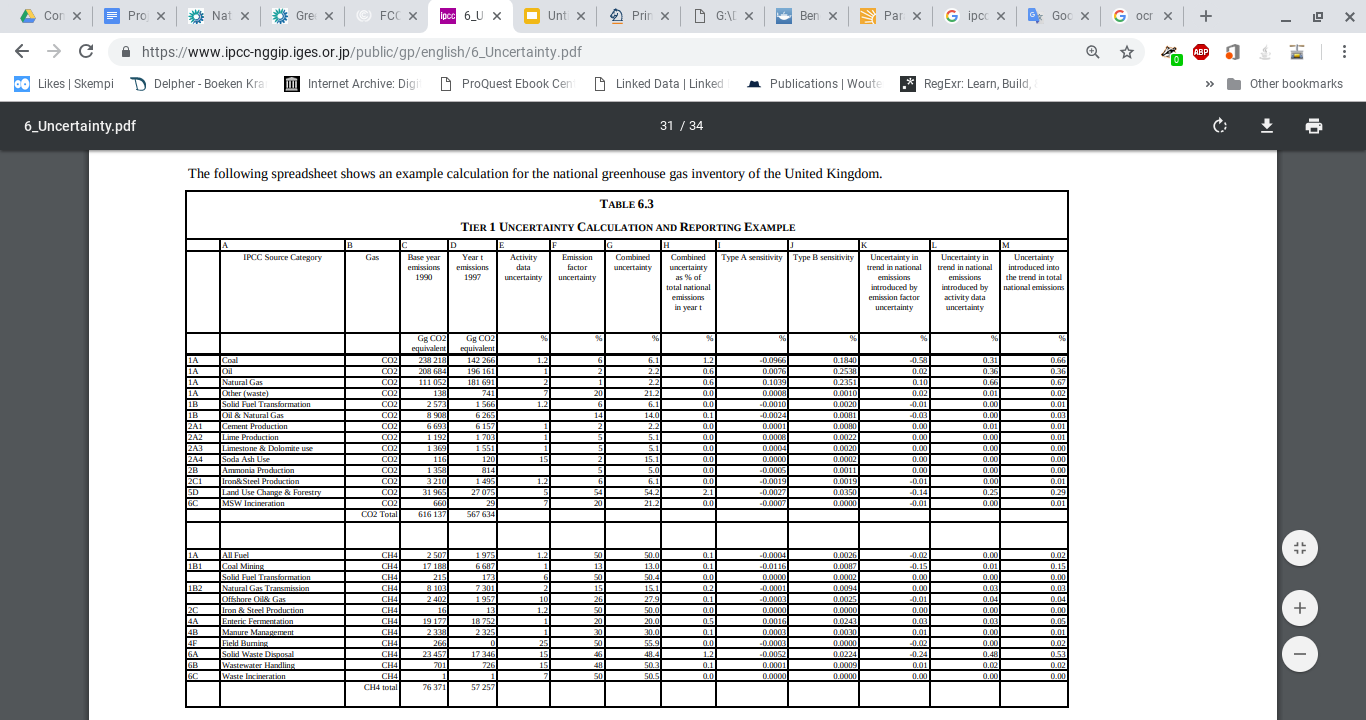
The result of the research of the dataset made clear that aviation has the biggest growth in the emissions of greenhouse gasses in the period of 1990-2016. This research only focussed on why there was such a big growth in the emissions of greenhouse gasses in the aviation industry. For further research, it might be interesting to come up with solutions for this problem. These solutions can be based on this research, but it is also possible to come up with another kind of research. It might be possible that the pollution of the aviation industry is vastly reduced in the future, because of improved planes or a more sustainable type of fuel. In this case, it might be possible to do further research on new threats for the environment with the same method as this research. This will probably give new results, which can be investigated. Finally, as made clear by the data analysis, it would be prudent to research the issues of provenance and data collection of these submitted values of greenhouse gas emissions.

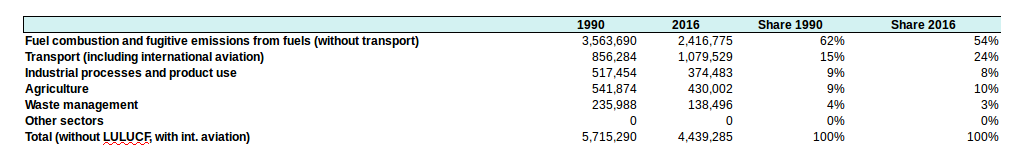
**Appendix**

**Table 1: Example Emission Reporting Table**

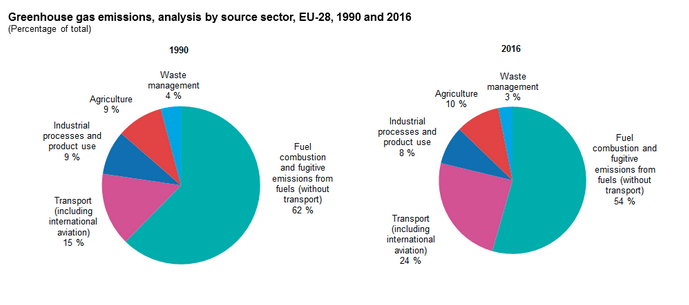
****

**Table 2: Example IPCC Uncertainty Reporting Table**

****

**Table 3: Total greenhouse gas emissions (including international aviation and indirect CO2, excluding LULUCF), 1990–2016 (million tonnes of CO2-equivalents)**

**Figure 1: Greenhouse gas emissions, analysis by source sector, EU-28, 1990 and 2016**



**Bibliography**

Aero, A. ‘79 new airlines begin life in 2017; 25 go out of business – Europe has the highest number of start-ups, but also the most failures’ [Airline Network News and Analysis]. (18 october 2017). Retrieved from: <<https://www.anna.aero/2017/10/18/78-new-airlines-begin-life-2017-25-go-out-business/>>

Grassl, H. (2003). Aviation and Atmosphere In: Meyers R.A. (eds) Encyclopedia of Sustainability Science and Technology. Springer, New York, NY. <<https://link-springer-com.vu-nl.idm.oclc.org/referenceworkentry/10.1007/978-1-4419-0851-3_555>>

Janssen, M., Charalabidis, Y. & Zuiderwijk, A. (2012) Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management, 29*(4), 258-268. DOI: 10.1080/10580530.2012.716740

De Koning, B. (15 july 2017). Vliegen is veel te goedkoop en dat is onhoudbaar. *De Volkskrant,* Retrieved from: <<https://www.volkskrant.nl/columns-opinie/-vliegen-is-veel-te-goedkoop-en-dat-is-onhoudbaar-~b17d4836/>>

Noack, R. (12 october 2017). Why are flights so much cheaper in Europe than in the U.S.?. *The Washington Post*, Retrieved from: <<https://www.washingtonpost.com/news/worldviews/wp/2017/10/12/why-are-flights-so-much-cheaper-in-europe-than-in-the-u-s/?utm_term=.578fd332ddd5>>

1. <https://www.eea.europa.eu/legal/copyright> [↑](#footnote-ref-0)
2. Anna Aero; airline network news and analysis, 2017. [↑](#footnote-ref-1)