To what extent will the COVID-19 pandemic contribute towards reaching goals stated in the Paris Climate Agreement for Germany?

Final Result - Group 10

Florian Auinger 03659767
Florian Butsch 03663488
Florian Hölzl 03659355
Christoph Miller 03721920
Johannes Gensheimer 03721520
Johannes Kiechle 03735507
Johannes Gahr 03675464
Constantin Nowak 03676376

September 4, 2020

Motivation

We are currently in the middle of the COVID-19 pandemic, the biggest pandemic in the last 100 years. While all of our lives are rapidly affected by it, other issues are overshadowed. The changing climate is one of the most severe challenges for mankind on a long term basis. In recent years, initiatives to fight climate change and global warming have been raising awareness about the topic. There exist different scenarios by a variety of researchers which all point towards the same fact that greenhouse gas emissions have to be reduced to keep global warming in an acceptable range. Germany's goal is to reduce greenhouse gas emissions by 55 percent until 2030 compared to 1990. However, there are already predictions that see the achievement of the climate goals at risk.

The effects of the COVID-19 pandemic are easy to observe. Many countries have decided on lockdowns for both, the economy and individuals in order to slow down the spread of the virus. Consequences are among others a drastic decrease in industrial production due to closed factories and reduced demand as well as travel bans and even restrictions on all non-obligatory personal movement. Currently, a noticeable decrease in greenhouse gas emissions can already be measured, but so far, there is no certain forecast on how the numbers will develop during the time of this crisis and thereafter. In this project, we are going to try to tackle this problem by using data to make predictions how Germany's emissions could develop further down the crisis. These resulting insights and their significance will then allow us to discuss Germany's adherence to the EU Climate Goals for 2030 and if the COVID-19 pandemic will have any impact in reaching those goals.

Contents

1 Project Description

Recent data has shown, that the COVID-19 pandemic and associated measures such as a public lockdown resulted not only in a decrease in mobility and consumption but also in a reduction of greenhouse gas emissions in Germany. The scope of this project is to analyze the effects of the pandemic on the emissions in Germany and consequentially its impact on reaching the EU Climate Goals for 2030.
Germany, as the biggest economy and with the largest population of any member country in the European Union, is used here as a representative to highlight the consequences of the crisis and the associated measures imposed on the economy and the population for our current approach on combating climate change.
The project is split into three parts. First, a regression approach is used to get an estimate on how the emissions changed so far during the crisis by mapping recent data of available features which have an impact on the carbon output. Afterwards, we fit the numbers of COVID-19 infections to the predicted emissions which then allows us to make adjustable forecasts for the rest of the year. Lastly, we use a seasonal ARIMA model to develop a ground truth by fitting emission data not affected by the pandemic.

Research Question

Which effects will the COVID-19 pandemic have on Germany's greenhouse gas emissions in 2020 in the context of reaching the EU Climate Goals for 2030?

Goals

The goal of this project can be split into three parts. First, we want to get an estimate on how the greenhouse gas emissions have changed so far during the pandemic. This will give us a lower bound on the impact of the pandemic and a basis for fitting COVID-19 infections to emissions. In our project we will focus on CO_2 emissions as they make up the biggest share of man-made contributions to the greenhouse effect and consequentially climate change.
In order to better understand the long-term consequences of the crisis, our second goal is to forecast the emissions for the rest of 2020. Under the assumption, that with possible vaccines against COVID-19 and other counter-measures, the year 2020 will be the worst of the pandemic, we should be able to make an approximation on how the pandemic had an impact on reaching the climate goals for 2030. As the impact of the crisis highly depends on the number of infections in the future, our goal is to have a forecast which is adjustable depending on the COVID-19 cases.
Lastly, we want to compare our findings to a scenario without the pandemic in order to get a better approximate on its impact. We do so, by developing a ground truth and comparing it to the results of our COVID-19 affected scenario. Our goal is to the use the deviation to discuss the impact of the pandemic and put the findings into the long-term context of Germany reaching the EU Climate Goals for 2030.

2 Data basis

The collected data is a very important basis for the whole project and for machine learning applications is crucial for the quality of the results. For this purpose, at the beginning of the project extensive data was collected for the sectors mobility, economy, and energy and households, as well as ground truth data for the CO_2 values. These are all listed in Milestone 2. Starting the modeling part and constantly researching literature the database of the project was adjusted. The database creation is an iterative process and requires significant share of time and effort of the whole project.

In the modeling process we decided to remove the sector economy, because all our regression models have learned wrong correlations and a feature selection based on too few features did not make sense. In the sector energy and households we decided to focus only on the energy segment because feature mapping did not lead to reasonable results. Besides, features that are not available until June 2020 are not considered in the feature mapping. Nevertheless, one can still explore all the data collected in Milestone 2 in the GUI.

We have not only discarded data but also added new data and information sources compared to the second milestone. We have added various fuel sale data for different fuel types and corona infection numbers. Furthermore, we now use information sources that allow us to make certain conversions, namely from energy data to CO_2 equivalents, annual CO_2 values to monthly values and the conversion of CO_2 savings to other metrics.

In the following section we describe in more detail the sources we ultimately use, how we collect them and the preprocessing steps.

Sources

In this section we list all data and information sources that we ultimately use for our final machine learning model and which are relevant for the end result.

Sector mobility:

Feature	Source	Characteristics
Mobility Trend Corona BAST	https://www.bast.de/BASt_2017/DE/Statistik/Verkehrsdaten/Verkehrsbarometer.html?nn=1820340	03/2020 - 07/2020, daily, numeric Format: CSV
Traffic count Germany	https://www.bast.de/BASt_2017/DE/Verkehrstechnik/Fachthemen/v2-verkehrszaehlung/zaehl_node.html	2003 - 2018, hourly, numeric Format: CSV
Traffic count Bavaria	https://www.baysis.bayern.de/web/content/verkehrsdaten/dauerzaehlstellen.aspx	02/2017 - 04/2020, monthly, numeric Format: CSV, XLS
MWV fuel sale data	https://www.mwv.de/statistiken/mineraloelabsatz/	01/2000 - 06/2020, monthly, numeric Format: CSV
Aviation traffic statistic in germany	https://www-genesis.destatis.de/genesis//online?operation=table&code=46421-0012&bypass=true&levelindex=0&levelid=1592905811215#abreadcrumb	01/2011-07/2020, monthly, numeric Format: CSV, XLSX
Scheduled flights Germany	https://www.oag.com/coronavirus-airline-schedules-data	01/2019 - 08/2019 and 01/2020 - 08/2020, weekly, numeric Format: XLSX

Corona related data:

Data	Source	Characteristics
Corona case numbers	https://de.statista.com/statistik/daten/studie/1094950/umfrage/entwicklung-der-weltweiten-fallzahl-des-coronavirus/#professional	01/2020 - 08/2020, daily, numeric Format: XLS

Target values (ground truth for sector mobility):

Data	Source	Characteristics
Target values of greenhouse gas emissions of Germany	https://www.oeko.de/fileadmin/oekodoc/Sektorale-Abgrenzung-Treibhausgasemissionen-Datenbasis-20191217.xlsx Contact person for target values at "Umweltbundesamt": Patrick Gniffke Tel: +49 (0)340 2103 2757 Patrick.Gniffke@uba.de	1990 - 2030, yearly, numeric 1990 - 2017: actually published data of "Umweltbundesamt" 2018 - 2030: currently estimated prognosis of "Umweltbundesamt" by Patrick Gniffke Format: XLSX
Temporal profile of CO2 data	https://www.che-project.eu/sites/default/files/2019-01/CHE-D2-3-V1-0.pdf	Conversion factors, numeric Format: PDF

Sector energy:

Data	Source	Characteristics
Carbon intensity of each type of power plant	https://www.ipcc.ch/site/assets/uploads/2018/02/ipcc_wg3_ar5_annex-iii.pdf#page=7 Page 1335	Conversion factors, numeric Format: PDF
Electric Energy Data of Germany starting from 2015 ordered by production type	https://www.smard.de/home/downloadcenter/download_marktdaten/726#!?downloadAttributes=%7B%22selectedCategory%22:1,%22selectedSubCategory%22:1,%22selectedRegion%22:%22DE%22,%22from%22:1559340000000,%22to%22:1622584799999,%22selectedFileType%22:%22CSV%22%7D	01/2015 - 06/2020, 15min steps, numeric Format: CSV, XLS, XML
Electricity generation, net heat generation, fuel input - Germany	https://www-genesis.destatis.de/genesis//online?operation=table&code=43311-0002&bypass=true&levelindex=0&levelid=1592908210645#abreadcrumb GENESIS-Tabelle: 43311-0002	01/2002 - 12/2014, monthly, numeric Format: CSV

Data Collection

The data collection is done as follows, we download or copy data from websites and save them as CSV or XLS. We use data on conversion coefficients directly in the code.

A consistent form of the data is beneficial to work more efficiently in the further process. Therefore, we adapt the downloaded data manually or via python code, so that we get a form where the first column contains a date-time format and each additional column represents a feature. The data has been updated regularly. For the final submission the goal is to have a actuality until June 2020 for as many features as possible.

Preprocessing

The overall goal of this first data preprocessing is to generate one uniform and standardized database as a JSON-file which works as the interface to the data models. Since the individual data sources are available in very different forms, they are preprocessed in separate Jupyter notebooks and stored in a defined format as CSV. In a further step, all these CSV files are merged in an additional notebook and stored in the above mentioned JSON.

To get as many training samples as possible, the annually ground truth CO_2 data is converted to monthly resolution using the conversion table listed in Sources.

In the same way, the features for mapping must be available on a monthly basis. The goal is to have them available until June 2020. Unfortunately, not all data sources cover the desired time range and therefore have to be merged in order to map a consistent feature. This is done for vehicle traffic, aviation and energy data. The procedure is described in detail in Milestone 3 (Section 2.2 Feature merging).

At this point, a common preprocessing like PCA or scaling does not make sense, because this is required differently for each data model and explained in detail in the Milestone 3 document.

3 Data Model

During the development journey many different regression models were investigated. These include the fields of feature to emission mapping, feature to emission forecasting, and emission to emission forecasting. For the final data model in total five models are used to estimate the impact of Corona on CO_2 emissions.

Remarks on WaveNet outlined in Milestone 3
Besides the final chosen approach, incorporating linear regression and SARIMA, a parallel method based on the WaveNet architecture was intensively pursued, assessed and evaluated. Originally, WaveNet was introduced by Google Deep Mind for the purpose of serving as a deep generative neural network for raw audio waveform [15]. However, as 1D-CNNs are increasingly used for general purpose time series processing, WaveNet is rising awareness as an auspicious technique for time series forecasting in numerous regression tasks these days, too [9],[10]. Due to elaborate techniques embraced in the WaveNet architecture, which allow for efficient processing of time series data, the overall complexity level of the resulting neural network is elevated too. The higher the complexity, the more knobs has to be tweaked in order to leverage the prediction performance. As the final WaveNet model, ready for hyperparamter optimization, had a considerable number of parameters, this had proven to be a very tedious exercise without taking advantage of any GPU support. Furthermore, as there were some troubles in obtaining reproducible results using this method, the aforementioned combination of a linear regression model and SARIMA was picked for the purpose of this project, which in addition allows an easier interpretation of the final results.

Approach

The investigated models were reduced to the usage of two SARIMA models, one LASSO Linear Regression, and two basic Linear Regressions.
As in section 2 described the sectors were also reduced to mobility, and energy. One SARIMA model for each sector is used to forecast the time series of CO_2 emissions for the hypothetical situation of no Corona pandemic in 2020. In the sector mobility the LASSO Linear Regression model is used to build a connection between the mobility features and the monthly CO_2 emission estimations by the Umweltbundesamt using TNO-MACs temporal profiles. One basic Linear Regression model is used in each sector to map the monthly Corona infection rates in Germany on the percentage deviation of the CO_2 emissions of one month in 2020 to the corresponding month in 2019. Using this mapping approach the seasonality of CO_2 emissions is considered. For the time span July to December 2020 the Linear Regressions are used to estimate the effect of monthly Corona infections, which the user can adjust, on CO_2 emissions in each sector.
Flow charts in figure 1 visualize the above explained approach.

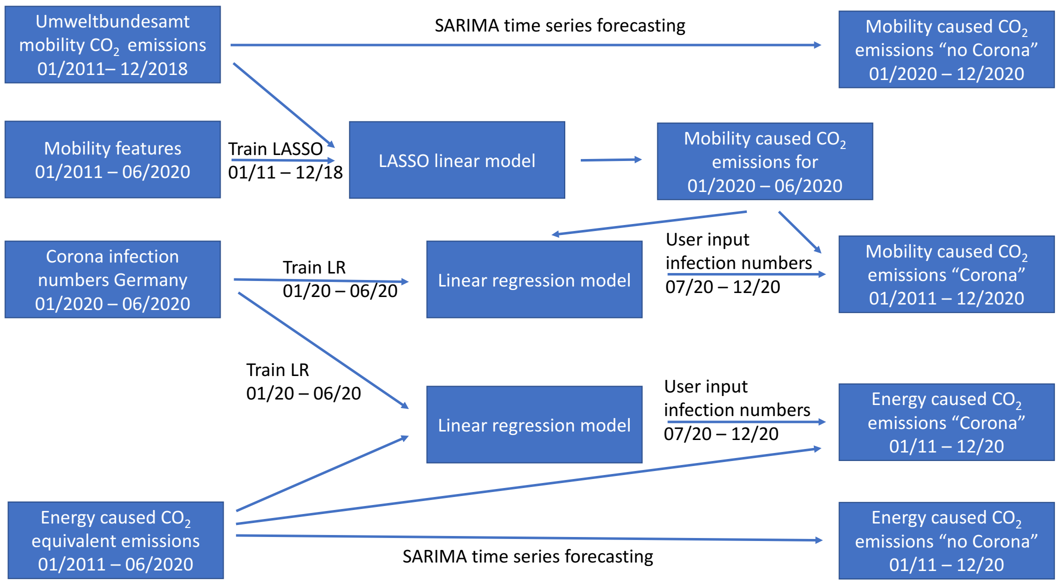


Figure 1: Flow chart visualization for modeling mobility and energy caused CO_2 emissions

Training

The training of all five models is part of the initialization procedure of the web interface. All models, except of the Linear Regression of the Corona to emission mapping, are optimized. To optimize the SARIMA models, the pmdarima [14] implementation is used. It aims to minimize the AIC of the model, by doing a something like a grid search with possible p and q values, as well as different configurations for the seasonal data. With the historic data from the sectors in the time span 01/11 - 12/18 the models are trained, to complete the ground truth for 24 months, until the end of year 2020. The best parameter for alpha in the LASSO implementation is searched using a grid search with build in k-fold cross validation. As metrics the the mean squared error is used. Data of the time span 01/11 - 12/18 is used as the emission estimations until 2018 are approved values. The Linear Regressions of the Corona to emission regression are trained with all six values for the months 01/20 - 06/20 for both sectors individually using the Corona case numbers as input and the percentage emission reduction of one month related to the corresponding month in 2019 as ground truth.

Evaluation

SARIMA models for the sector mobility and energy

To forecast emissions based on historic emission data, ARIMA/SARIMA models and a MLP network were evaluated in Milestone 3. The RSME of the SARIMA model was better in comparison to the MLP, as displayed in figure 1. In addition to that, the forecast curve looks smoother.

Model	CO2 Emissions RMSE train	test
MLP	2.32	1.74
SARIMA	0.57	

Table 1: RMSE in million tons CO2

The auto SARIMA approach described in section Training is used to find the best parameters to forecast the sectors mobility and energy. The configuration with the lowest AIC score is then selected. This value is relative, therefore negative as well as positive scores are possible. The SARIMA configuration for the two sectors are shown in figure 2.

	SARIMA configuration				
Sector	p	d	q	seasonal component	AIC
Mobility	0	1	1	2,1,2,12	-324.736
Energy	0	1	1	0,1,1,12	328.204

Table 2: Configuration found by AUTO SARIMA model for the sectors.

The forecasts of the SARIMA models are shown in figure 2.

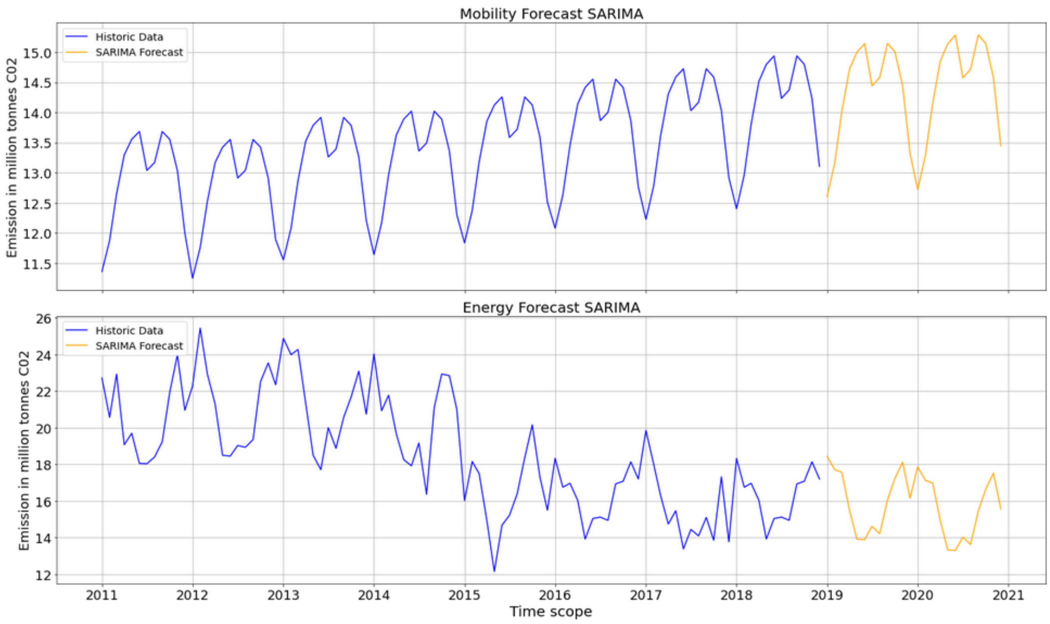


Figure 2: SARIMA predictions for the two sectors

Usage of LASSO linear regression for the mobility sector

The were Linear Regression, Partial Least Squares Regression (PLSR), LASSO linear regression, Ridge linear regression, and Neural Network regression were investigated to map the mobility features on the mobility CO_2 emissions. All of them, except of the Neural Network regression, showed similar behavior and similar metrics. The Neural Network regression was not considered due to given results. As LASSO linear regression has the best mean RMSE of the cross validation and the adventage of a build in feature selection using a tuning parameter alpha it was chosen for the feauture to emission mapping in the mobility sector.

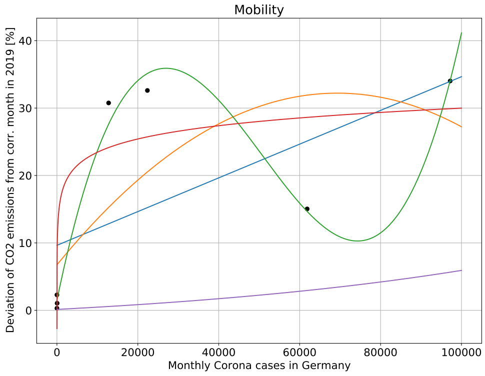
	Mobility RMSE	
Model	train	test
LR	0.43	0.45
NN	0.88	0.88
PLSR	0.35	0.46
LASSO	0.35	0.44
Ridge	0.36	0.44

Table 3: Mean RMSE of 10 fold cross validation in million tons CO_2 of the best performing hypterparameter.

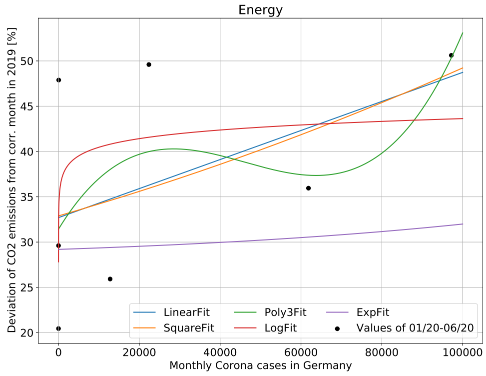
Linear Regression models for the Corona to emission regression

In the field of Machine Intelligence human judgment of the results is an important aspect, especially at a case with very few data points. Figures 3(a) and 3(b) show the regression of the monthly Corona infections on the the percentage deviation of CO_2 emissions of one month to the same month in 2019. Table 4 implies that the polynomial of degree one is one of the worst performing regarding its metrics r^2 and RMSE. Anyway a **linear fit** was chosen which is reasoned with figures 3(a) and 3(b). The **square fit** and **polynomial of degree 3** have a negative gradient at an area of very high infection numbers which is illogical as with rising infection numbers the governmental restrictions rise and therefor the mobility and energy CO_2 emissions decrease. The **logarithmic fit** is also illogical as the decrease of CO_2 emissions rises about 15% with a very small difference in infection numbers in both sectors. As it can be observed in both figures the **exponential fit** does not fit the data points well enough as the gradient is too low. Basically the curves (especially the polynomial of degree 3) show the behavior of overfitting and lose the real correlation of the features to the output.

Approaches for better results could be to adjust the square fit, so that the vertex is at the highest infection number or to set lockdown thresholds to produce a step function. We decided not to do it as this has a big impact on the overall curve fitting. Also the intercept of the linear fit could be set to zero, but we state that even the infection number is zero the pandemic still has effects on mobility and energy. For example flights are still at an extremely low level even though the infection numbers are very low in Germany in comparison to April 2020. We decided to keep this offset as a long term impact of the COVID-19 pandemic on the CO_2 emissions in 2020, due to governmental restrictions and human behavior to keep the infection numbers low. In contrast to the observed infection to emission correlation we assume an infinite fast reaction of the government and society on rising or decreasing infection case numbers. We do not state that this is fully scientifically correct.



((a)) Curve fittings for the sector Mobility



((b)) Curve fittings for the sector Energy

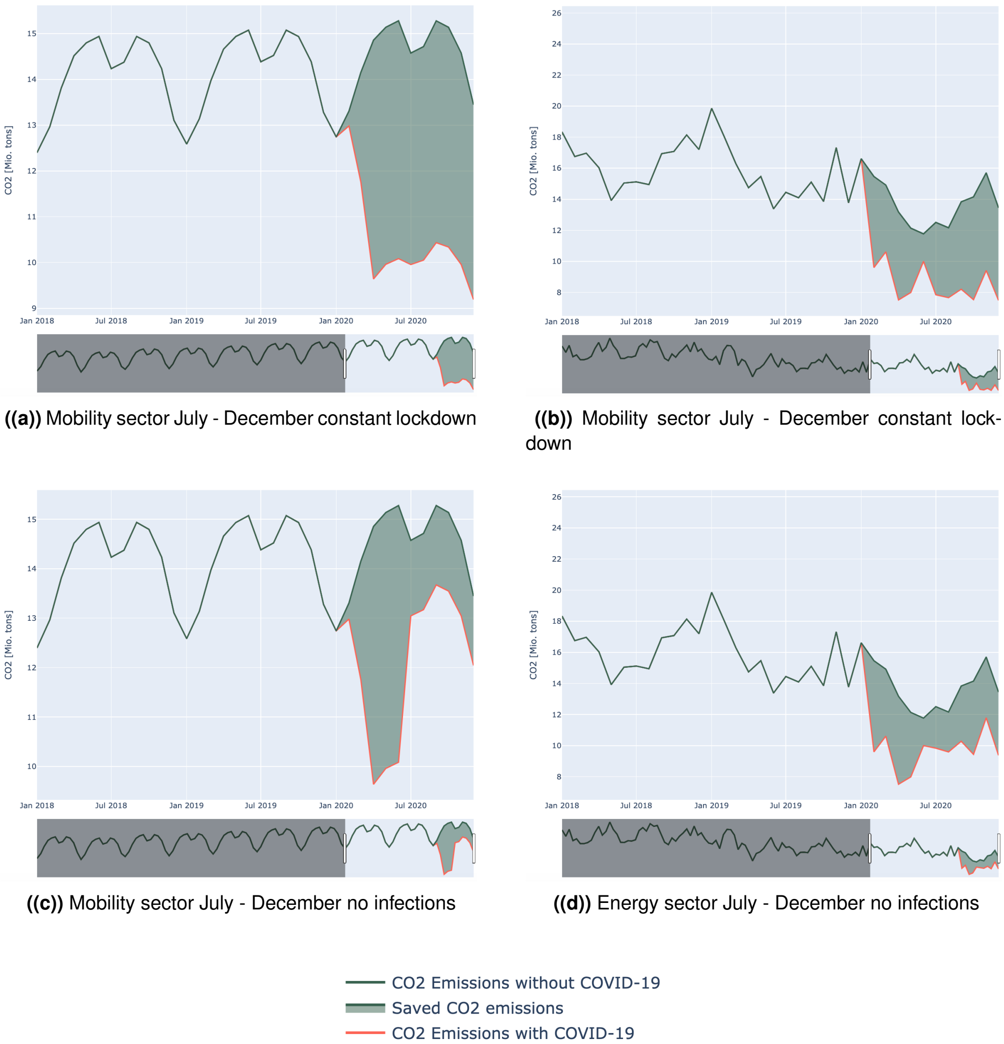
Figure 3: Different models applied for the Corona on Emission mapping.

	Mobility		Energy	
Model	r^2	RMSE	r^2	RMSE
Linear fit	0.36	0.12	0.24	0.1
Square fit	0.46	0.11	0.24	0.1
Poly 3 fit	0.99	0.02	0.31	0.1
Exponential fit	0.32	0.12	0.29	0.1
Logarithmic fit	0.76	0.07	0.24	0.1

Table 4: Metrics results of Corona infection numbers on CO_2 emission curve fits.

4 Results

For both investigated sectors the difference of the estimated CO_2 emissions of 2020 to the forecasted CO_2 emissions of the hypothetical situation where no Corona pandemic has happened is calculated. This absolute value is divided by the average emission per day in 2019 in order to estimate how much time of emissions we have saved due to the Corona pandemic.
To estimate the global impact of this difference FAIR v1.3 [13] is used to calculate the caused decrease in global temperature rise and global CO_2 concentration.



((a)) Mobility sector July - December constant lockdown

((b)) Mobility sector July - December constant lockdown

((c)) Mobility sector July - December no infections

((d)) Energy sector July - December no infections

Figure 4: Difference of the situation with the Corona pandemic and the hypothetical situation of no Corona pandemic of 2020.

Figure 4 shows the extreme cases of the two sectors mobility and energy. Figures 4(a) and 4(b) show a constant lockdown situation from July to December 2020 with monthly infection number greater than 90000. Figures 4(c) and 4(d) show the case of no infections from July to December 2020. The user of the web interface can adjust the infection numbers for July to December 2020 and the red curve will adjust for that time period accordingly. The adjustments of the user also result in the calculation of CO_2 savings, equivalent emission days, global concentration and temperature impact. Table 5 shows the corresponding values of the cases in figure 4.

	Emissions [Mio. tons]	Equivivalent emission days	Δ Concetration [ppb]	Δ Temperature [μK]
Constant lockdown				
Mobility	-48.4	103.3	-6.5	-11.7
Energy	-57.9	113.5	-7.4	-12.9
Sum	106.3	-	-13.9	-24.6
No new infections				
Mobility	-29.8	63.6	-4.1	-7.5
Energy	-45.7	89.6	-5.9	-10.2
Sum	-75.5	-	-10	-17.7

Table 5: Numeric results of extreme cases of constant lockdown and no infections in July to December 2020

Observations

As it can be observed in figure 4 and table 5 the impact of a second lockdown is in our calculation not as high as the impact of the first lockdown. This has the reason of the offset of the Corona to emission mapping explained in section 3. Even though there are no new infections the decrease of CO_2 emissions related to the previous year is still at -10% for mobility and even -32.5% for energy. We argue that it models the impact of the Corona pandemic on the long term and stay with the linear fit.

Depending on the infection numbers in July to December 2020 we've saved 63.6 to 103.3 days of emissions in the sector mobility and 89.6 to 113.5 days in energy.

The impact of the saved CO_2 emissions on the global concentration and temperature rise seems minor. The concentration change is in the range of 10ppb to 13.9ppb, which means only 10 to 13.9 particels per billion less than without the Corona pandemic. Currently there is a mean CO_2 concentration of about 400 particles per million (ppm) in the earths atmosphere [4]. The global temperature difference is in the range of micro degree Kelvin. Using Germany as a case study and scaling it to the world emissions this reduction can have an impact on the course of climate change, like explained in the global study by Forster et al. [11].

Trends

Based on the used CO_2 data from the Umweltbundesamt [7] the total amount of emitted CO_2 is decreasing in Germany. We can also observe this trend in the energy data from [2] and [5].

Based on our Corona infection numbers to emission mapping in the sectors mobility and energy we observe a long term decrease in CO_2 emissions caused by the Corona pandemic, even without any new infections. The impact of Corona on CO_2 emissions can also be observed in current papers, e.g. [11] or [8].
Even though the emissions of many countries is decreasing the emissions of other countries and in total is constantly rising [4]. The long life time of CO_2 in the atmosphere and the rising emissions result in a rising concentration of this and other greenhouse gases.

5 Discussion

In principle, results must always be scrutinized critically, especially for this task, as the results cannot be checked for incorrectness or correctness. Rather, one has to think about the meaningfulness and logic of the prediction. Ultimately, the change in behavior of people and political decisions are decisive for the future development of CO_2 emissions and therefore no model can accurately predict the future trend. That is why we have created a possibility to calculate the minimum CO_2 savings and to run through various scenarios for the rest of the year, to be able to get an idea about a possible trend.

Interpretation of the results

Considering the results summarized in the previous section, the question rises how to interpret the results. The drastic drop in CO_2 emissions is obvious and matched the findings of recent publications [11] [8]. According to Patrick Gniffke from the Umweltbundesamt, the approach of inferring actual CO_2 emissions from indicators is a common way of predicting CO_2 values. Therefore, the emission values from the feature emission mapping can be viewed as valid and realistic results. Nevertheless, one has to be aware that false correlations can also be learned, which applies to the economy sector in our case.This is why this was left out, but the data can still be explored in the GUI. The prediction from July till December 2020 is an interactive feature to estimate a realistic scenario in which the seasonality is mapped with the emissions. For this regression, the assumption was made that the consequences of a lock down would have immediate effects. This means that society reacts to it infinitely quick. Therefore the result is only a rough estimation. However, since it cannot be said how the CO_2 emissions will develop, this is also a valid approach to simulate different scenarios.

Critical assessment of the results and assumptions

In order to be able to achieve this result, several assumptions had to be made and of course these bring uncertainties with them, some more and less. First, the conversion from annually to monthly CO_2 values for the ground truth, by TNO-MAC III [6] creates inaccuracies. However, these inaccuracies cannot be avoided and must be accepted, as no monthly CO_2 is available and these are necessary to have sufficient samples for training and accurate prediction. This uncertainty is mainly due to the poor data situation. However, data availability is also a general problem with regard to the features. In order to be able to estimate the impact of Corona on CO_2 emissions, the data must be as current as possible. We have defined availability until June as a minimum requirement. Since not all features meet this criterion, they are sorted out in the mapping model and not taken into account in the regression. This may cause important information to be lost.
It must also be said that machine learning or deep learning is not the general solution for all problems. Since the aim of this project is to answer the question with the help of machine learning algorithms, a sensible use of it was aimed at, although it was actually not always necessary. A locally more scientific accurate approach can be implemented using fuel sale data and emission factors for the mobility sector instead of a machine learning regression model but using the machine learning model makes the approach more scalable to other countries. The availability of fuel

sale data worldwide is generally unsatisfactory. For example [12] uses national yearly fuel sale data to estimate the traffic in San Francisco, Bay Area, California. Yearly emission values can be extracted from emission inventories like Edgar [1] or TNO-MAC III [6]. Features can differ from locations depending on which data is locally available. This approach shows the usage of machine learning to create a scalable approach independent on the local feature data availability.
Regarding the mobility mapping result it need to be said, that the features are not weighted in advance. The result may indicate that Aviation is too heavily weighted. This feature has suffered a particular decline, but at 3% it only accounts for a small part of the total emissions compared to road traffic [3] .
Based on the difficult question, inaccuracy of forecasting, especially longer forecastings and the data availability, unfortunately, it was only possible to calculate the emissions savings for 2020 and not, as originally planned, until 2030. Scenarios such as shutting down the coal-fired power stations makes a long-term forecast much more difficult or the information content would be very questionable. Nevertheless, from our point of view, the overall result is respectable and represents a possible scenario. In the end, however, it must be pointed out that the economy sector could not be estimated and that there are other smaller sectors [7], the effects of which were not examined.

Proposed Answer to the Research Question

To figure out the effects of the COVID-19 pandemic on Germany's greenhouse gas emissions in 2020 in the context of reaching the EU Climate Goals for 2030 was a ambitious research task we asked ourselves. In the end we had to realize that this question is very difficult to answer. Nevertheless, we were able to predict the CO_2 savings for the Mobility and Energy sectors. Depending on the COVID-19 case number development, the CO_2 savings for Mobility range between 29.8 and 48.4 million tons of CO_2. This corresponds to a time gain for the achievement of the climate targets of 63.6 days up to 103.3 days. For Energy the emission savings range between 45.7 million tons of CO_2 and 57.9 million tons of CO_2. That equates to 89.6 to 113.5 days of time saved. Beyond the research question, we calculated their impact on decreasing the global CO_2 concentration and global temperature increase, which are detailed in the conclusion.

In summary, it can be said that the research question could be answered in principle, even if the information content of the result is not a concrete statement as to whether we will achieve the climate goals or not. However, we were able to clearly demonstrate that the COVID-19 crisis has saved considerable amounts of CO_2 for the mobility and energy sectors. The number of days saved with average emissions shows how high the savings are. The COVID-19 crisis has given us a time lead to achieve our climate goals. Unfortunately, it cannot be predicted whether this time advantage will be sufficient or whether it can even be expanded.

6 Conclusion

The global COVID-19 pandemic in the year 2020 has shown to impact not only social aspects but also the environment. Governmental restrictions and the economic and social lockdown has shown to reduce CO_2 emissions in the sector mobility and energy strongly. Social changes can also lead to environmental changes, like reduced consumption, using video calls for business meetings or reduce the vacation distance, also on the long term.

Summary of the Results

A strong decrease in CO_2 emissions of up to 32.5% in the mobility and about 50% in the energy sector can be observed. This equals a range of saved CO_2 emissions from 28.8 to 48.4 million tons in the mobility and 45.7 to 57.9 million tons in the energy sector for the year 2020. Both sectors combined the impact on the global temperature varies from 17.7μK to 24.6μK and in global concentration from 10ppb to 13.9ppb. The absolute numbers do not imply a big impact, but this is only the impact of German emissions on the whole worlds environment. Scaling the impact up up on global emissions and the fact that the COVID-19 pandemic impacted the whole world would increase the absolute numbers.
On the implementation side the results show that the used machine learning approaches for the emission forecast for July to December 2020 struggle strongly with suddenly and unexpected occurring events, why we state that they are not suited for this case. The observed feature value ranges and seasonality during the Corona pandemic did not occur in the time series, resulting in not reliable and realistic results.
Machine learning is not always suited in the area of environmental science as human behavior, especially unexpected behavior, can not be easily modelled. We did only forecast to the end of 2020, as we state that a further forecast can result in too big uncertainties caused by social behavior and governmental decisions.

Future Work

Further work has to be done in order to make this project even more accurate and more scientifically valuable. These include improvements and expansions and are

- Higher temporal resolution in order to get more data samples and increase the reliability of the regression methods and the feature selection process, also for the Corona on Emission mapping.
- Find more accurate features for the economic and industrial sector and all other CO_2 emissions sectors, to investigate the full impact of COVID-19 on CO_2 emissions.
- Scale the approach to Europe or the world using worldwide data or methods for emission upscaling (e.g. using the share of worldwide emissions of Germany).
- Include other meaningfull greenhouse gas emissions like CH_4 or N_2O.

7 Comments to the Group Work Experience

This years semester project felt quite challenging for a bunch of reasons. First of all, in comparison to other courses at university, the ultimate goal of this project left a rather big open space for our own creativity. This is not a circumstance which students conducting a university project are used to, however showcases the mundane life in research or in industry. Starting with formulating an own original research question, where there is no correct answer to be taken out of the shelf, was the very first and rather big barrier to come over, as we guess only the minority of students ever have been starting any project work this way. Along with the broad spectrum which accompanies with the project, the organizational aspect (e.g. communication, subdivision of the work, taking account the various backgrounds of every group member, assignment of responsibilities) is not to be neglected. As a consequence, the following is about general remarks on the group work experience, where we intend to convey our learning outcomes to the reader.

Remarks on the organizational structure and learning outcomes: This section gives an insight on the soft-skill handling during the entire project execution. This was even more important this semester, as any meetings in person and after work get-togethers for improvements of bonding's among every team member (i.e. after work beer) had to be conducted online. In order reduce the workload for every individual, a smart subdivision of the tasks during the entire project handling was the very core. As a consequence, subgroups for the different main tasks have been established (e.g. mapping-team, forecast-team, GUI-team). This is especially beneficial, as the group workload is split across many shoulders, discussions are becoming more efficient (i.e. time-wise), setting up informal meetings is much easier as less group individuals are involved. Nevertheless, all-hands group meetings are indispensable all along the total project time, however essentially require a person within the group who takes the lead in the sense of a anchorman/moderator role. This was especially perceptible during our weekly "Zoom meeting", as otherwise discussions would not have been that constructive within a given time scope of 1,5 hours. Nevertheless, online meetings come along with crucial benefits too, as everybody is uncommitted to location for instance. During our meetings we have made use of a collaborative online platform (i.e. Miro board) which extensively contributed for collecting agenda items, discussion agreements, next steps to be followed and milestone arrangements in a diligent way . In order to do so, we stringently followed the structure of a predefined template for documentation purposes of every all-hands team meeting (see Figure 5).

Remarks on the development process, group size and and learning outcomes: As every team individual is equipped with a different background and owns skills in diverse areas of profession, in order to make valuable contributions to the project progress, the subdivision of the tasks was likewise based on the skill level of every team member involved. However, especially comprehensive projects, incorporating vibrant task lists, always leave the door open for exploring into different areas of interest of every participant. Therefore, anyone was given the opportunity to gain knowledge in a broad spectrum, which might be helpful in feature endeavours. Especially, being part of a rather big project group is valuable, for both, hard- and soft-skill aspects, however is also more challenging for some reasons. Although there was no additional barrier in a sense of the main communication language (i.e. everybody's mother tongue was German), discussions often drag on because everyone tried to express his own opinion and sell at the best "price" possible. Furthermore, as much as splitting the workload into smaller packages is essential, this is always a representation for possible pitfalls in a sense of redundant work. Moreover, one has to be able to rely on all other team members and accept results that one may have done differently.

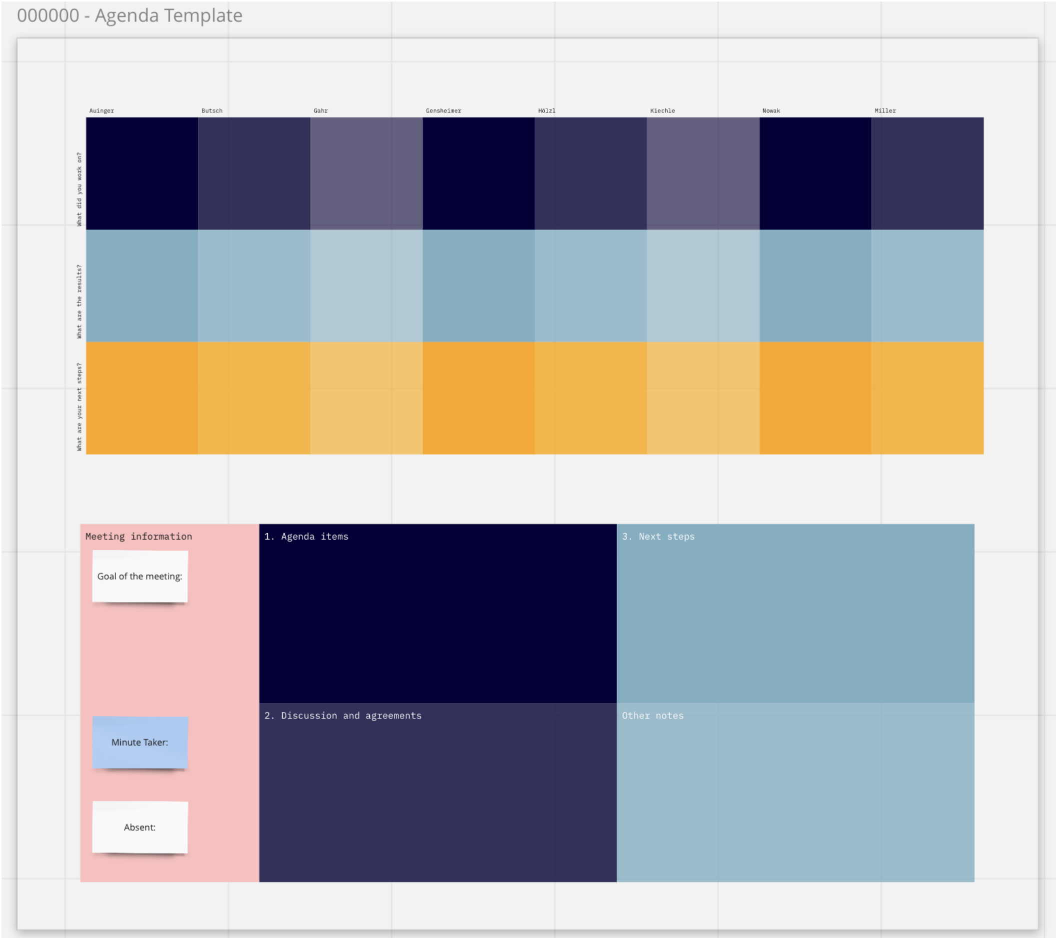


Figure 5: Template in Miro for all-hands team meetings.

References

[1] EDGAR emission inventory. https://edgar.jrc.ec.europa.eu/l. Accessed: 2020-09-01.

[2] GENSIS energy data. https://www-genesis.destatis.de/genesis//online?ope. Accessed: 2020-08-14.

[3] Klimaschutzreport 2019. https://www.bdl.aero/de/publikation/klimaschutzreport/.

[4] Nature, The hard truths of climate change — by the numbers. https://www.nature.com/immersive/d41586-019-02711-4/index.html. Accessed: 2020-09-01.

[5] SMARD energy data. https://www.smard.de/home/downloadcenter/download_marktdaten/726#!?downloadAttributes=%7B%22selectedCategory%22:1,%22selectedSubCategory%22:1,%22selectedRegion%22:%22DE%22,%22from%22:1559340000000,%22to%22:1622584799999,%22selectedFileType%22:%22CSV%22%7D. Accessed: 2020-08-14.

[6] TNO-MAC III high resolution European emission inventory. https://essd.copernicus.org/preprints/essd-2017-124/essd-2017-124.pdf. Accessed: 2020-09-01.

[7] Umweltbundesamt CO2 Emissionen. https://www.umweltbundesamt.de/indikator-emission-von-treibhausgasen#die-wichtigsten-fakten. Accessed: 2020-08-11.

[8] Temporary reduction in daily global co2 emissions during the covid-19 forced confinement. *Nature Climate Change*, 10(7):647–653, 2020.

[9] Anastasia Borovykh, Sander Bohte, and Cornelis W Oosterlee. Conditional time series forecasting with convolutional neural networks. *arXiv preprint arXiv:1703.04691*, 2017.

[10] Anastasia Borovykh, Sander Bohte, and Cornelis W Oosterlee. Dilated convolutional neural networks for time series forecasting. *Journal of Computational Finance, Forthcoming*, 2018.

[11] Piers M. Forster, Harriet I. Forster, Mat J. Evans, Matthew J. Gidden, Chris D. Jones, Christoph A. Keller, Robin D. Lamboll, Corinne Le Quéré, Joeri Rogelj, Deborah Rosen, Carl-Friedrich Schleussner, Thomas B. Richardson, Christopher J. Smith, and Steven T. Turnock. Current and future global climate impacts resulting from covid-19. *Nature Climate Change*, 2020.

[12] Brian C. McDonald, Zoe C. McBride, Elliot W. Martin, and Robert A. Harley. High-resolution mapping of motor vehicle carbon dioxide emissions. *Journal of Geophysical Research: Atmospheres*, 119(9):5283–5298, 2014.

[13] C. J. Smith, P. M. Forster, M. Allen, N. Leach, R. J. Millar, G. A. Passerello, and L. A. Regayre. Fair v1.3: a simple emissions-based impulse response and carbon cycle model. *Geoscientific Model Development*, 11(6):2273–2297, 2018.

[14] Taylor G. Smith et al. pmdarima: Arima estimators for Python, 2017–. [Online; accessed <today>].

[15] Aäron van den Oord, Sander Dieleman, Heiga Zen, Karen Simonyan, Oriol Vinyals, Alex Graves, Nal Kalchbrenner, Andrew W. Senior, and Koray Kavukcuoglu. Wavenet: A generative model for raw audio. *CoRR*, abs/1609.03499, 2016.