UNIVERSITÄT STUTTGART

MSc. Air Quality Control, Solid Waste and Waste Water Process Engineering (WASTE)

Dynamic modeling of cost potential curves of captured CO_2

Master Thesis by

Eugenio Salvador Arellano Ruiz

Performed at
Deutsches Luft und Raumfahrtzentrum,
Institut für Vernetzte Energiesysteme
Stuttgart

Supervision:
Niklas Wulff

Evaluation:
Univ.-Prof. Dr. techn. Günter Scheffknecht

February 18, 2021

Abstract

The field of system modeling and scenario analysis has a demand for high resolution data to drive their models and validate their outputs, in some cases technology related data is not existent or sparse because of novelty or small widespread. This is the case for emission mitigation strategies such as Carbon Capture and Sequestration, which can play a role in emission reduction scenarios. Static CO_2 capture potential analysis as well as project-related cost assessments have been thoroughly reported in literature. These approaches have shortcomings. They assume constant boundary conditions and usually focus on a single target year and region. Energy system model analysis however require various geographical scopes and temporal horizons.

To address this need, a flexible model is developed to generate cost potential curves in a Geographically distributed manner. In order to test the model, a literature research is done to collect and harmonize input values such as cost and efficiency loss, these are then aggregated and statistically analyzed to create error ranges of the model, given the diversity of the sources the ranges are relatively wide. The geographical distribution is achieved using literature data for industrial processes and open source energy production databases whose missing values are completed using the available information within itself. To fill the gaps mainly in reported efficiency values a computational regression model is built with moderately reliable results of 5% standard error.

The cost potential distribution is calculated for the European union and their neighboring countries, examples maps for Germany, Italy and the Netherlands are built from this, insights on the spatial distribution of the sources are done. Curves for Germany are drawn showing a technical potential exceeding 25 Mt/y only from industrial sources wit ha max price of 45 €/t and only on demands of more than 70.7 Mt/y the fossil fuels become relevant with a potential of more than 200 Mt/y with a max price of 110 €/t.

Basic scenario developments are modeled for Germany and the EU27 countries using a naive plant closing approach that builds on the GECO scenarios also considering the German coal phase out, the industrial sector development is not modeled. There is a 310 Mt/y potential difference between the Reference and 1.5°C scenarios for the EU27 countries and no difference for Germany as the scenario emission goals are met by German coal phase-out by 2038.

A sensitivity analysis shows that higher efficiency losses imply a lower marginal cost of CO_2, the same can be said for the Capacity factor effect, reducing the capture efficiency from 90 % to 50% increases the cost from 80 €/t to 150 €/t and the potential is reduced from 1500Mt/y to 750 Mt/y .

To further validate the data, cross-validation of the amounts is done using Emission Reports compiled in the PRIMAP project but the comparison potential is limited given the high dependency of assumptions. Similarly with the aid of an alternative data source it was found that the whole structure of the curve is sensitive to the input by changing the max potential from 1800 Mt/y to 1600 Mt/y and the median cost from 75 €/t to 50 €/t. With the available data it is concluded that the production does not represent a bottleneck in the CO_2 supply chain.

Dedication

This work is dedicated to my family, who supported me all the time during my studies. My brother Aldo, my mother Lorenza and my father Fernando.

To all my professors, but especially those in the Air quality specialization, Professors Dr. Vogt, Dr. Reiser and Dr Scheffknecht for offering the challenges that helped me grow my expertise. A very special thank for Dr.Friedrich along Dorothea Schmid and Julia Neuhäuser as they inspired and pushed my interest in the field of simulation and modeling with a focus on environmental research.

I want to thank as well my working team and colleagues in DLR, especially Niklas Wulff for being an understanding but also a very challenging supervisor who really pushed my scientific thinking forward, Benjamin Fuchs for the technical knowledge and support that helped me develop my programming skills, Felix Nitsch and Ulrich Frey whose work was of a huge value for the validation of the developments in this thesis, Hans Christian Gils whose leadership style has been really inspiring for me and Patrick Jochem for managing a very promising and diverse department in which I was fortunate to take part of.

I extend thanks also to the members of the scientific community who supported me during the research and development process. Especially Henriette Naims and Lorenzo Cremonese for the insights and expertise in the field of CO_2 capture they shared. Tjerk Zitscher and Aniko Walther whose own input and advice pushed the early research in this thesis.

Contents

Abstract **i**

Dedication **ii**

1 Introduction **1**
1.1 Problem 1
1.2 Background 2
1.3 Purpose 2
1.4 Approach 3

2 Technical Background **4**
2.1 Carbon Capture Technologies 4
2.2 Measurements of Costs of CCUS 6
2.3 Power Plant Technologies 7
2.3.1 Coal based power plants 7
2.3.2 Natural gas based power plants 8

3 Methodology **9**
3.1 Materials 9
3.1.1 Software 9
3.1.2 Packages 9
3.1.3 Datasets 10
3.1.4 Literature 12
3.2 Reference Harmonization 12
3.2.1 Unit Harmonization 12
3.2.2 Cost Harmonization 14
3.2.3 Calculation of cost of carbon capture for individual plants 16
3.2.4 Assumptions for the harmonization of values 16
3.2.5 Statistical aggregation of harmonized literature values 17
3.3 Geographical Distribution of Potential and Cost 17
3.3.1 Conventional power plants distribution 18
3.3.2 Industrial processes 18
3.3.3 Strategies for the completion of efficiency data 20
3.3.4 Geographical aggregation 21
3.3.5 Assumptions for inclusion of Biomass and Biogas 21
3.3.6 Assumptions for the geographical distribution of cost 22
3.4 Cost Potential Curves 22
3.4.1 Scenario Development 23

4 Results **25**
4.1 Reference Harmonization 25
4.1.1 Key findings of harmonization 26
4.1.2 Aggregation of Harmonized values 28
4.2 Strategies for the completion of efficiency data 30
4.2.1 Random forest regression 30
4.3 Geographical distribution of potential and cost 32
4.4 Cost Potential Curves 37
4.5 Scenario Development 39
4.6 Sensitivity Analysis 42

4.6.1 Data input sensitivity . . . 43
4.7 Emission validation . . . 44

5 Discussion **47**

6 Conclusions **49**

Appendices **50**

Appendix A: Input Data **51**

Appendix B: Harmonized Values **57**

Appendix B: Correlations **62**

List of Figures

1.1 Energy supply scenarios 1

2.1 Carbon capture principles 4
2.2 CCUS Technologies Status 5
2.3 Flow diagram of MEA 5

3.1 Energy sector capital and operating cost indexes from the IHS [2020a,b] 10
3.2 Completeness of aggregated datasets 11
3.3 Unit harmonization general flow 13
3.4 Relative development after 2002 of cost indexes 15
3.5 General process of the creation of a geographical distribution of cost of carbon capture 19
3.6 General process of the production of cost potential curves 23

4.1 Cost distribution for the different power technology groups, the colors represent if the study is done considering the retrofitting of an existing plant or not. 25
4.2 Harmonized values of cost of carbon capture for different fuel categorizations . . . 26
4.3 Harmonized cost of carbon capture from different sources 27
4.4 CAPEX distribution values, the outer blue lines are the one standard deviation limits, the middle red line is the meand and the green line is the median 28
4.5 Distribution values, the blue lines are the one standard deviation limits, the red line is the mean and the green line is the median. 29
4.6 Heat rate gain regression 30
4.7 Distribution of error of the efficiency regression methods 31
4.8 Distribution of reported and random forest predicted efficiency values. 31
4.9 Distribution of reported and random forest predicted efficiency values, effect of commission year 32
4.10 General view of the potential and cost of carbon capture. 33
4.11 Germany view of the potential and cost of carbon capture, the values are €/tCO_2 33
4.12 Italy view of the potential and cost of carbon capture, the values are €/tCO_2 . . . 34
4.13 Netherlands view of the potential and cost of carbon capture, the values are €/tCO_2 34
4.14 Left: German natural gas power plant cost distribution for captured CO2 in 2019€/t. Right: Potential distribution in Mt/y. Here the potential is evenly distributed across the territory 35
4.15 Left: German coal power plant cost distribution for captured CO2 in 2019€/t. Right: Potential distribution in Mt/y. In this case, there is a formation of focal points in major mining territories 35
4.16 Left: German industrial process cost distribution for captured CO2 in 2019€/t. Right: Potential distribution in Mt/y. industry is sparser than natural gas plants yet has lower costs. 36
4.17 Left: German Bioenergy plant cost distribution for captured CO2 in 2019€/t. Right: Potential distribution in Mt/y. The data used for the production of these distributions does not allow strong conclusions 36
4.18 Cost potential curve for the countries in EU27, the cost is in €/tCO2 and the potential in Mt/y 37
4.19 Cost potential curve for example countries, the cost is in €/tCO2 and the potential in Mt/y 38
4.20 (a.) Cost potential curve for Germany for all the different sources. (b.) Cost potential curve for Germany for lower potential sources. 38

4.21 (a.)Germany reference scenario at 2050, Curve.(b.) Germany 1.5°C scenario at 2050, Curve . . . 39
4.22 (left) Captured CO2 cost distribution of Germany at 2050 in the reference scenario, [2019€/t].(right) Captured CO2 potential distribution of Germany at 250 in the reference scenario, [Mt/y]. . . . 39
4.23 (left) Captured CO2 cost distribution of Germany at 2050 in the 1.5C scenario, [2019€/t].(right) Captured CO2 potential distribution of Germany at 250 in the 1.5C scenario, [Mt/y]. . . . 40
4.24 European potential reference scenario at 2050. . . . 40
4.25 European potential 1.5°C scenario at 2050. . . . 41
4.26 Cost potential curve for the countries in EU27 at years 2020 and 2050 . . . 41
4.27 Cost potential curve for the countries in Germany at years 2020 and 2050 . . . 41
4.29 Effect of variables in the cost potential curves, values are in table 4.1 . . . 42
4.30 (left) PPM European cost potential Curve. (right) Internal dataset cost potential curve 43
4.31 (left) PPM German cost potential Curve. (right) Internal dataset cost potential curve 44
4.32 CO2 from combustion activities across the last century. HISTCR is red and HISTTP is blue. . . . 45
4.33 CO2 from combustion activities for years after 2010. HISTCR is red and HISTTP is blue. . . . 45
4.34 CO2 amounts based on the different points of the calculation reference and the values reported by PRIMAP, the stripped squares do not represent actual values. . . . 46

1 Correlation matrix of all the input and output variables from the harmonization process . . . 62

Acronyms

BEC	Bare Erected Cost.
CCGT	Combined Cycle Gas Turbine.
CCS	Carbon Capture and Secuestration.
CCU	Carbon Capture and Utilization.
CCUS	Carbon Capture, Utilization and Secuestration.
CF	Capacity Factor.
COC	Cost of Clinker.
COE	Cost of Electricity.
EPC	Engineering Procurement and Construnction cost.
ERE	Excess Renewable Electricity.
FCF	Fixed Charge Factor.
FLH	Full Load Hours.
FOM	Fixed Operation and Maintenance.
GECO	Global Energy and Climate Outlook.
HHV	High Heat Value.
HR	Heat Rate.
IGCC	Integrated Gasification Coal Combustion.
IPCC	International Panel on CLimate Change.
LCA	Life Cycle Assesment.
LCOE	Levelized Cost of Electricity.
LHV	Low Heat Value.
MEA	Post-combustion monoethanolamine.
O&M	Operation and Maintenance.
OCGT	Open Cycle Gas Turbine.
PPM	Power Plant Matching.
SCPC	Super Critical Pulverized Coal.
TCC	Total Capital Cost.
TCR	Total Capital Requeriment.
TOC	Total Overnight Cost.
TPC	Total Plant Cost.
USPC	Ultra Super Critical Pulverized Coal.
VOM	Variable Operation and Maintenance.

Introduction

1.1 Problem

Reports F20 [2020] that the achievement of the 1.5°C global average temperature rise goal set by the IPCC in 2018 is possible according to an advanced climate model published in the book Achieving the Paris Climate Agreement Goals (APCAG). This model considers a complete replacement of fossil fuels by renewable alternatives by 2050.

This model also considers a 6% share of synthetic fuels as part of the global final energy demand, this accounts for an amount of 6328 PJ per year of fuels that have to be used to satisfy the demand of the transportation sector in which electrification is expected to comprise up to a 50% share of it, the rest from which will be satisfied using hydrogen and renewable fuels.

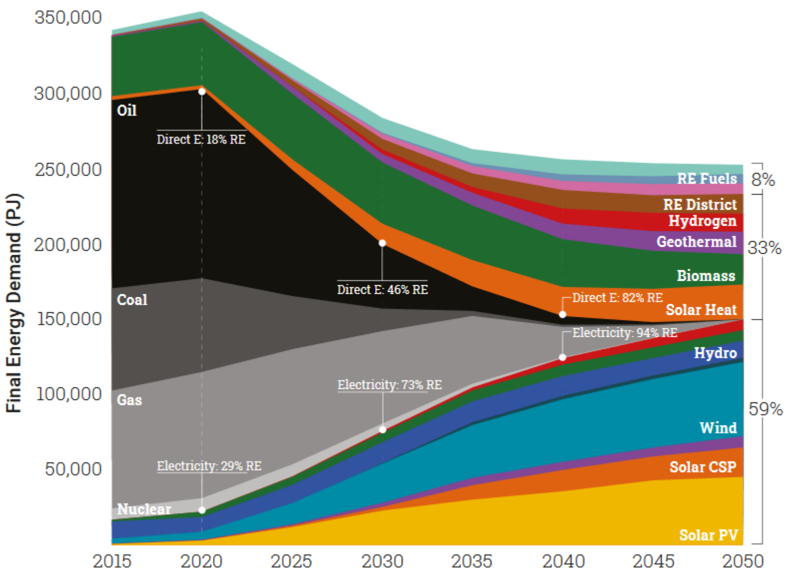


Figure 1.1: Energy supply scenarios for electricity, heat and mobility found in F20 [2020]

In order to be categorized as renewable, the production of synthetic fuels is highly dependent on the usage of a source of carbon atoms that is equally renewable [Sterner, 2009], which usual form is of a CO2 molecule. This raises the question of whether the availability of carbon from such sources can be accounted to in a future where the usage of CO2 emitting fuels is expected to be drastically reduced. Millinger et al. [2020] proposes a collection of scenarios in which the Excess Renewable Electricity (ERE) is used to produce synthetic fuels, there it tackles the question picking a robust set of assumptions using LCA based scenarios in which Carbon Capture and Utilization in the production of fuels is assessed as part of the CO2 stream and considering the German power mix goals.

Millinger et al. proposes in its model a global approach for the development of the availability of carbon from captured sources, and it also considers a fixed price for the obtained carbon. At the same time, the availability of ERE is highly geographically constrained [Scholz, 2012]. At least during the energy transition, the cost and availability of carbon from captured sources are dependent on factors that are integral to the energy system. These facts pave the way to start questioning if the cost assumptions of studies like the one from Millinger et al. still holds true after taking the geographical distribution of the carbon sources into consideration. The first step to start addressing this doubt is to find a way of conceptualize the geographical distribution of cost and potential.

1.2 Background

Fröhlich [2019] has done an analysis of the potential of CO2 from captured sources distribution for the German territory to satisfy the demand from the transformation of ERE to gas or liquid fuel,the methodology is similar to the one in this work but focused on the emission reduction potential.

The Carbon Capture and Utilization (CCU) technology CO2 reducing potential at a macro scale is visited by Budinis et al. [2018], who makes a widely scoped analysis building differently configured scenarios up to year 2100; it makes considerations of fossil fuel usage developments and accomplishment of GHG reduction goals. It also has technology specific assumptions and builds up on other widely cited studies like GCCS [2011], this makes it very useful to get a top level idea of how the technology may develop; In terms of cost, it aggregates at a global level to reflect the total investment is needed at that scale to achieve the scenarios it analyses.

At a plant scale there is abundant literature assessing the costs, each has a wide set of assumptions that sets them apart from each other significantly, in the year 2005 the IPCC collected the information of several of such studies to generate the chapters 3 and 9 of their Carbon Capture and storage Report [IPCC, 2005], a very good collection of general sources for CCS studies to build upon, said report was updated in 2015 by one of its original authors Rubin et al. [2015] assessing the changes in assumptions that appeared in the time frame of 10 years. As an special mention, one of such reports, the IEAGHG [2014] is one of the most extensive and openly available reports on Coal plant CO2 capture performance. Rubin, in other work Rubin et al. [2013], proposed a general methodology to asses carbon capture technologies cost estimates for power plant applications, here he generates an extensive list of factors and terms that the different agencies utilize to generate their reports. This in order to harmonize future publications, such harmonization is not exhaustive but still able to make the results of said studies more easily comparable. This last paper is used as guide to do the harmonization of values in this thesis.

The best example of the methodology is sought to develop in this study comes from Naims [2016] where she makes a economic assessment of the potential supply and demand of CCS technologies, she collects a series of technologies and builds potential curves which she later uses to analyses different scenarios of demand from different industries in a global scale.

Von der Assen [2016] Does an analysis based on the emissions reported by the EEA [2019]. It research energy demands of carbon capture in industrial and energy processes, develops a LCA model to estimate a capture potential, calculates their GHG emission reduction and ranks their usage on a theoretical future demand by their environmental impact potential.

The shortcoming this work intends to compensate for are the lack of a point specific bottom up, source and technology sensitive estimation of the emission potential and capture cost of the European continent. With the exception of Fröhlich [2019], the analyses mentioned are either based on total reported emissions and are aggregated at a very high level (European, World). The approach in this thesis will bother less about emission reduction potential, as Fröhlich and Von der Assen do it already in a exhaustive way, and focus on the economical requirements, without considering transport or any kind of tax or bonus.

1.3 Purpose

The mentioned studies are very helpful for doing general reports regarding the future of carbon capture and utilization, is only when one has interest of analyzing the complexities of implementing it into an actual spatially distributed system when the shortcoming start to show up, one of such examples is the inclusion of this technology in a sector coupled model of the European energy system [Brown T., 2018], such a model also requires flexible implementation of the geographical distribution as they usually have various levels of aggregation; single node European level, countries, NUTS regions etc. Having a spatial and temporal resolution of cost implies a high number of variables; such as fuel availability and prices, return of investment of the projects and technological development of the capture technology are some among many others. At the date of the production of this work, there are no published attempts to create supply cost curves for CO2 from carbon capture sources in a geographically differentiated resolution. Because the existence of such curves would be of relative value for scenario analysis, the steps for their production and the initial attempts to do them will be addressed in this thesis.

The goal of this master thesis is first to develop and implement a methodology to produce a geographically and temporally distributed dataset of cost and availability of carbon dioxide from captured sources. Second, to exemplify its utilization by using it to produce an exemplary dataset using publicly available European data of high CO2 producing industries and fossil fuel power plants. And finally to use it to answer the following theoretical questions:

- What are the advantages and disadvantages of the different CO_2 sources?
- Can a demand of 25 Mt_{CO_2}/y by 2050 which corresponds to the highest demand in Millinger et al. for Germany be met?
- If yes, under what cost can this demand be met?
- Economically, are fossil fuels a viable source of CO_2?

1.4 Approach

The main characteristic of the methodology is that it's data driven, its bottom up, and the values are calculated as function of the industrial activity and not their reported emissions. It is highly sensitive to the input data quality and the assumptions are kept at a minimum and only done when it is absolutely necessary(Scenario analysis, for example). The idea is to make a data stream that can be updated whenever new information is available or a different data source is to be included. This is to allow the emergent inclusion of new sets of technological and economical assumptions. This characteristic is preferred over a more fixed model given the rapid changing nature of the field that is being worked with. Although the end product of this work is intended to include both power and industrial sources, the basis for the cost calculations will be build around the power plant sector because the area is already widely developed. This will be clarified in the theoretical framework and material sections 3.1. This said, a major part of this thesis is dedicated to solving the challenges of processing data from different sources and the production of a technically grounded homogenization. Another important part is related to the selection of optimal input structures to which the data has to be converted, analyzing the necessary fields and deciding what is relevant and what is not. Some work is dedicated to the geographical distribution and algorithmic production of the cost potential curves. And finally a rough implementation of scenario development is going to be addressed. An example scenario study is presented to demonstrate the approach effectiveness and its uncertainties as well to draw conclusions.

Technical Background

2.1 Carbon Capture Technologies

There is a number of ways to perform carbon capture based on different physical and chemical properties of the CO_2 molecule, to get general overview of some of these technologies, Figure 2.1 shows a classification in function of their active principles.

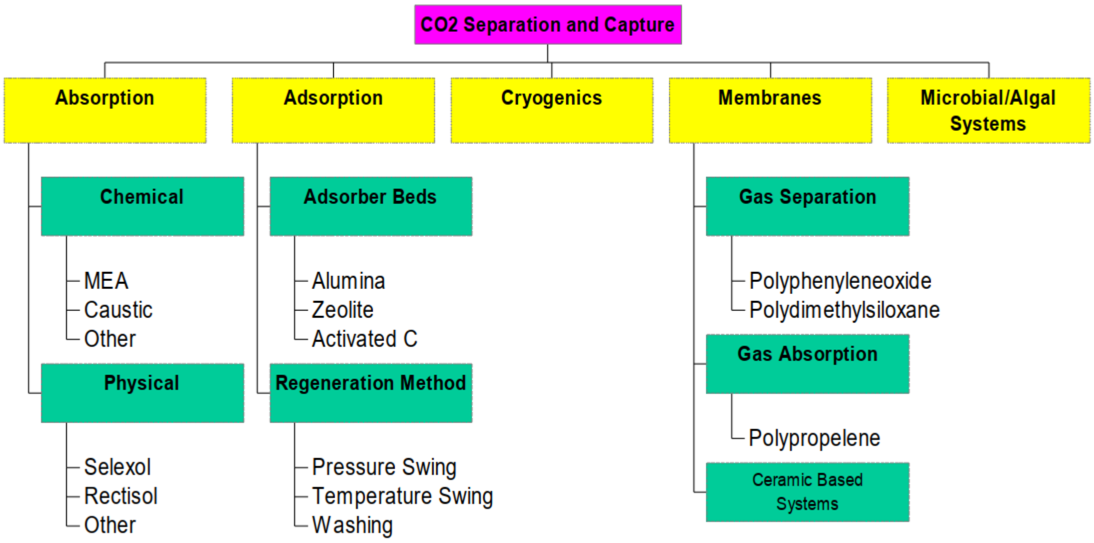


Figure 2.1: Clasification of carbon capture technologies according to their active principles, IECM [2019].

From a technical standpoint Carbon Capture, Utilization and Secuestration (CCUS) it's already in a commercial stage for Amine based chemical capture methods, their outputs are used in Enhanced Oil Recovery, however there is still a lot of them in development stages that can play an important role during the process of dacarbonising the industrial and energy sectors [Bui et al., 2018].The stages of development of different carbon capture related technologies can be seen in figure 2.2.

It can be noted that for power plants most of the capture technologies are in pilot or demonstration stages and only Post-combustion monoethanolamine (MEA) is at commercial stage, unfortunately it has not gone beyond demonstration projects for industrial processes. From now, we will concentrate in the MEA process for power plant gas streams, the understanding of this process is important for the cost measurement section (3.2.3)and the generation of cost assumption values for the power plant dataset. In order to get the technical background of the industrial process assumptions one can visit the industrial processes section fo the methodology chapter (3.3.2).

Post-combustion monoethanolamine (MEA) is considered to be the current state of the art process to remove CO_2 from gas steams. It is represented in Figure 2.3 and consist on two main operations, the absorption of carbon dioxide gas into a amine based sorbent and the stripping of CO_2 as a high concentration stream. [IECM, 2019]

The main source of cost of a amine based scrubber its the energy requirement of the sorbent recovery, in fact the optimization of the energy flows is one of the main subjects of the research and development efforts done around the technology [Lee et al., 2016]. Because of this, complexity

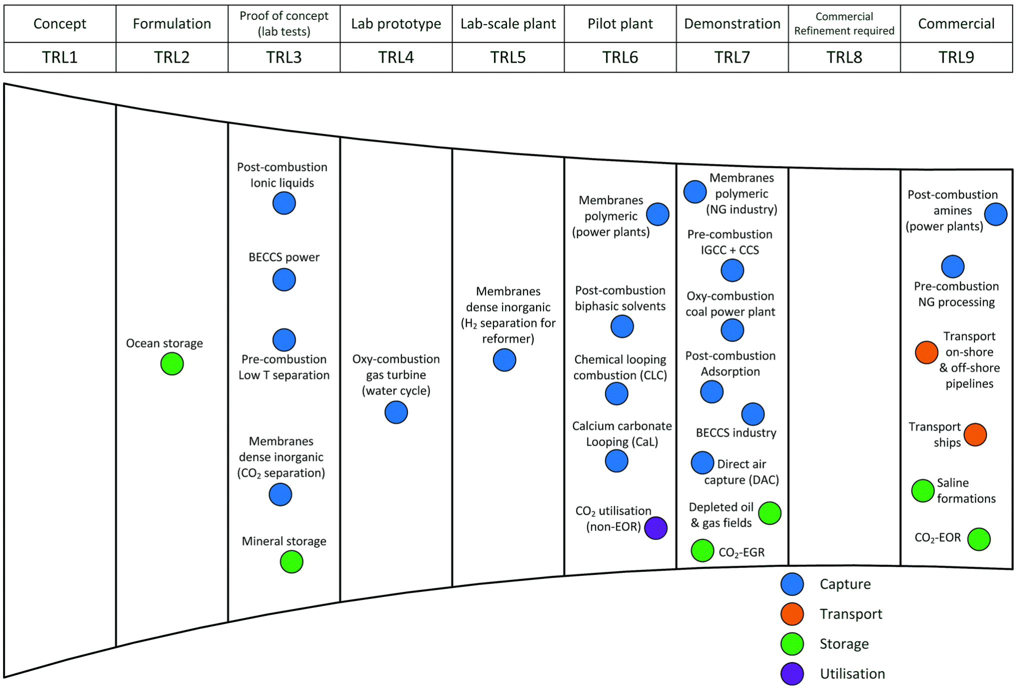


Figure 2.2: Current development progress of carbon capture, storage and utilisation technologies in terms of technology readiness level, Bui et al. [2018]

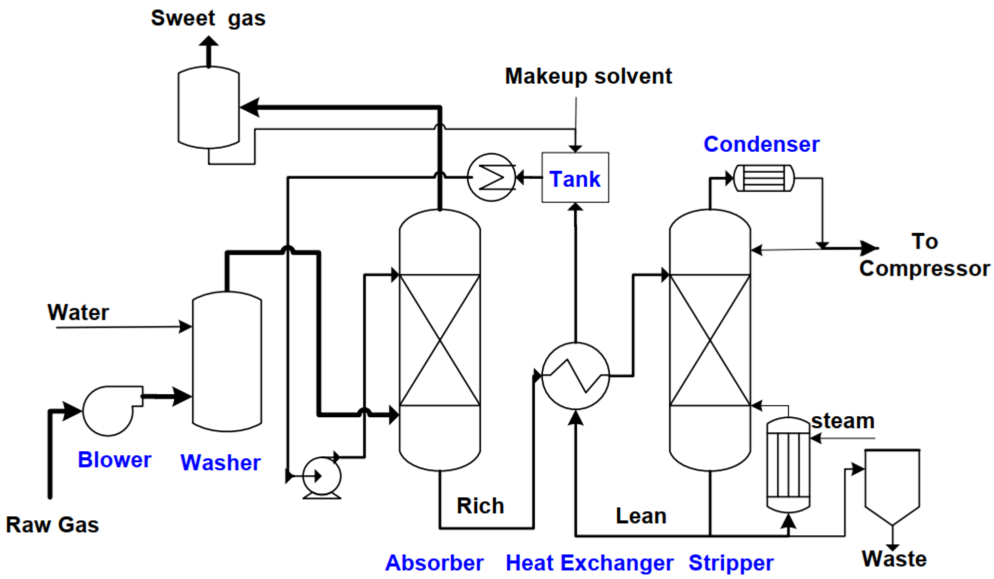


Figure 2.3: Basic flowchart of CO_2 removal using MEA absorpion, Roussanaly et al. [2013]

arises in the development of the plant configuration. The location of the CO_2 removing module in the process and the source of auxiliary energy to run it are very strong factors in the measurement of its ecological and economical impact, this can be better appreciated in Mantripragada et al. [2019] where models based on the two operating carbon capture power plants are compared yielding significantly different results. The names of these projects are Boundary Dam, which is found in Canada and it was retrofitted with an auxiliary gas power plant to provide energy for the capture process; and Petra Nova, located in the US, it was retrofitted using the existing power units as energy sources taking an efficiency penalty in the process. Both are amine based, although they use different solvents. This does not only affect the overall performance but also the cost, as each configuration implies also a very different business model to be used. And although these considerations have a significant weight in the final cost, they will be significantly simplified in this work, this will be made clear in the measurements of cost section(2.2).

We won't go into detail in other technologies but it is relevant to mention that membrane technologies are depending on developments in material science and process engineering to become commercially viable [Zhai, 2019]. As for other industries, the general process is similar and it is adapted to each particular gas stream. The relevance of gas composition in the cost assumptions, when the flue gas contains too much sulfur or any other acidic substance, the performance of the absorption-stripping cycle for CO2 recovery will be greatly diminished by solvent degradation [Zhou et al., 2013], because of this, there might be cases in which the addition of a Flue Gas Desulphurization module has to be done, which in the case of retrofit projects implies an additional cost of around 1000 €2019 per KW of energy for the captured carbon. [Dillon et al., 2013].

2.2 Measurements of Costs of CCUS

In Thermal power plant analysis, there is no standard way to report cost of carbon capture technologies. There is various metrics that can be used depending on the boundary conditions of the analysis to be done. For example the energy penalty, which measures the change of performance of a reference energy plant with CCUS vs the same without the technology [Budinis et al., 2018], this metric is very useful when assessing the technology in a way that can be technically comparable to others. But to make a cost analysis in an economical frame work there has to be an inclusion of other dimensions of the project like plant erection costs, operational expenditures, fuel prices, etc. In order to do these inclusions there is a series of metrics based on the Levelized Cost of Electricity (LCOE) that are used across several sources in the field dating back up to 2005[IPCC, 2005]. LCOE is represented in Eq 2.1:

$$LCOE = \frac{(TCC)(FCF) + FOM}{(CF)(8760)(P)} + VOM + (HR)(FC) \tag{2.1}$$

Where TCC is the total capital cost of the power plant project in monetary units (€), often found in the literature with a basis per nominal capacity (€/MW); FCF is the fixed change factor used to transform the capital cost into annuities [Rubin et al., 2013]; FOM is the Fixed Operation and Maintenance (O&M) costs (€) often represented in the literature in a per nominal capacity basis (€/MW*a); CF is the average capacity factor through the lifetime of the plant, it is multiplied by 8760 to calculate the effective hours per year of the power plant, not so often one finds these hours directly reported in the literature as Full Load Hours (FLH), concept mostly used referring to renewables; P reports the net capacity of the power plant, or the effective capacity, it is related to the actual electricity output of the plant, it should be differentiated from gross capacity which is electricity production; VOM are all the O&M costs that vary in function of the generated energy(€/MWh); HR is the Heat rate of the plant (MJ/MWh), explained further in section 3.2 ; FC is the fuel energy cost(€/MJ), the studies report this value based either on Low Heat Value (LHV) or High Heat Value (HHV) basis, this has implications on the final cost as it is a function of the produced heat as seen in equation 3.15. The LCOE or often just labeled as Cost of Electricity (COE), consist of the net present cost of an energy unit generated by a power plant in a given timeframe, usually its lifetime, but sometimes can be adjusted to be used in repowering or retrofitting projects[Witta Ebel, 2019], aspect which will be important later in this study. From here this metric can be calculated to a series of reference and capture implemented plants to find the CO2 related metrics.

The first of these metrics is the cost of avoided CO2 which includes the whole CCUS chain, inclusive transport and storage; it is wide in its scope and mostly apt for comparing different types of projects like conventional power plants vs photovoltaic parks.

$$Cost\,of\,avoided\,CO_2 = \frac{LCOE_{CCUS} - LCOE_{ref}}{(t_{CO_2}/MWh_{ref}) - (t_{CO_2}/MWh_{CCUS})} \tag{2.2}$$

The subscript "ref" can be any reference scenario or project, not limited to the same power plant with CCUS. (tCO_2/MWh) refers to the amounts of CO2 in tons avoided per effective electricity generated. The second approach for measuring the cost of CO2 and the one preferred for this study refers to the cost of CO2 captured 2.3.

$$Cost\,of\,captured\,CO_2 = \frac{LCOE_{CCUS} - LCOE_{ref}}{t_{CO_2}/MWh_{captured}} \tag{2.3}$$

It is strictly on site as it ignores subsequent costs of transport and storage. The reason it is preferred in this study is because such assumption allows more flexibility at the time of using the output value as an input into other process analysis.

From Eq 2.1 we can discern how are the different components of the cost of electricity, and as consequence the cost of CO2 captured, related. The first component is attributed to the capital cost; in general it is the value attributed to the erection of a power plant, it can have different values depending on the analysis level that is being done, this varies among studies and can be expressed in its different layers as follows [Rubin et al., 2013]:

BEC : (Bare Erected Cost) It is the lowest layer and comprises of the equipment, the supporting facilities and the labor including material and sales taxes.

EPC (Engineering Procurement and Construction cost) It is the BEC plus engineering services.

TPC (Total plant cost)It is all of the above with the contingencies of the process and project added.

TOC (Total overnight cost)This includes a series of costs together called Owner's costs which are for example: Land, feasibility studies, incentives and initial materials.

TCR (Total capital requirement)This includes cost escalation and interests during the construction process.

For this thesis the assumptions to be done are on the level of TCR for the calculation of the LCOE in the different power plant projects to be analyzed. The next factor is the Operation and Maintenance (O&M) which has two main components, the Variable Operation and Maintenance (VOM) and the Variable Operation and Maintenance (VOM). The former consist on those associated to the production of electricity like the chemicals for flue gas cleaning and the later are the ones considered fixed for the plant, like labor. There is several ways to calculate them but for the simplicity of the estimations a percentage of the power plant erection costs will be utilized. And finally the last component is the one associated to the cost of fuel, this is in some cases like in the use of MEA the most influential factor in the cost of captured carbon as the process is energy intensive, this implies a methodological challenge since energy carrier prices fluctuate across time, the effect of these prices can be seen in the sensitivity analysis section 4.6.

2.3 Power Plant Technologies

The equations presented in the previous section are general for the power sector. In order to conduct the analysis, the different authors in the literature researched for this thesis differentiate among types of fuels and power plant technologies. This section is built in order to ensure the reader has a basic understanding the importance of this differentiation for the development of the cost potential curves.

2.3.1 Coal based power plants

Coal power represented in 2019 25.6% of the global primary energy consumption [Gaeldicke, 2019]. The dominant technology for the construction of new plants is Super Critical Pulverized Coal (SCPC) which is also a mature technology meaning it has very low room for improvement having a maximum efficiency of 46%(LHV) with the potential of reaching around 50% with the so called Ultra Super Critical Pulverized Coal (USPC) technology[IEA, 2010]. The configuration of these power plants consist on the feeding of a furnace with previously pulverized coal for direct combustion, the supercritical, ultra-supercritical, sub-critical definition comes from the thermodynamic conditions of the oxidized gas stream being above the water vapor critical pressure of around 220 bar [Sarkar, 2015].

Another coal technology is the Integrated Gasification Coal Combustion (IGCC). This is a developing technology consisting of the preheating of the coal in an inert atmosphere to release the combustible gases (CO and H_2) to then use them in a gas and steam turbines combination in a similar fashion to a natural gas combined cycle. Theoretically it can achieve higher efficiencies than the PC but there is questions raised regarding the fact it is not completely avoiding GHG emissions. [IEA, 2010]

The efficiency of the power plant is inversely proportional to the heat rate, CO2 concentration in the flue gas is inversely proportional to the efficiency. On a larger scope there is effect of the process steps like storage and transportation on the emissions of the power plant, a very important matter addressed by LCA analyses. However, the majority of the emissions come from the fuel burning in the flue gas stream, those are the ones that are to be considered in this work because they are the only capturable emissions. This said, there is a further differentiation among coal power plants that is to be addressed.

In relation to the type of fuel, coal power is divided into lignite and hard coal plants. The first consisting on lower quality fuels with more water, less specific heat and higher emissions which in place are associated with lower plant efficiency; lignite power plants are usually located next to the fuel source to avoid transportation and make up for the low return value they have, one example is Schwarze Pumpe operated by the LEAG group and found in Vattenfall, Germany. The second type has usually higher efficiencies, uses relatively less purer coal forms and is not necessarily located near the source of energy, an example of this type of power plant is the one operated by RWE and located in Neurath, Germany.

2.3.2 Natural gas based power plants

Natural gas was in 2019 the 22% of the global primary energy consumption [Gaeldicke, 2019]. There is two main technologies for natural gas power generation Open Cycle Gas Turbine (OCGT) and Combined Cycle Gas Turbine (CCGT). The first consist on a single compressor-turbine connected to a generator, the second has the same initial principle with the addition of a vapor based energy recovery cycle. The efficiencies of the OCGT are around 35% while for the CCGT can theoretically reach the 60% mark. While the coal is expected to be reduce in most scenarios in the short term, gas is often considered as a growing alternative given its lower specific emissions. OCGT is not considered as a candidate for carbon capture in the literature while the considerations of CCGT often lean towards the construction of new plants, not retrofitting existing ones.

Methodology

3.1 Materials

To achieve the objectives of this thesis a complementary tool was developed to systematically transform data, do calculations and generate outputs. This tool is primarily programmed in python 3.8 using a collection of resources that will be listed in the following sections.

3.1.1 Software

The code used in this work is completely written in python 3.8. There is various grounds to pick said language. The first being that there is a very rich collection of open source work on energy system analysis available in python, one outstanding example is PyPSA [Brown T., 2018] which is already referenced in several papers and used to build complex models. Another ground is the fact that python is very easy to write on and be read by other people, the simplicity of the language makes up for things like lower efficiency than other options. And lastly, given its popularity there is a wide collection of packages that can ease the execution of complex task like building regression models. Other software used in this work is:

- QGIS, Despite most of the geographical output is being done in python, QGIS is going to be used as a support tool for representation of input and intermediate geographical data.
- Git is the state of the art software for version control, it is used to manage the development.
- MS Excel, Some of the source material is in an excel format, so this software is necessary to adapt these, the code books for the input are also built as excel files for accessibility.
- Pycharm as main IDE software used to write the code.

3.1.2 Packages

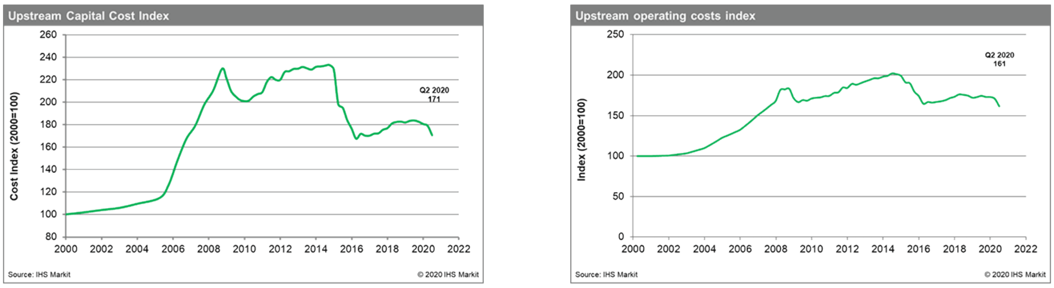
There is a number of third party python packages used in this work, the ones that are standard for python development like numpy are not listed here, instead we list the tools of specific intend:

- Pandas, a data library to perform database like operations. This is the main tool for the systematical calculation of values.
- Geopandas, a library built around pandas with geographical operations implemented, itself contains representation libraries like cartopy. This is used mainly for geographical aggregation.
- Scikit-learn, a machine learning library used in this work for creating regression models
- matplotlib, and a higher level package, seaborn, for the construction of plots.
- powerplantmatching, API to obtain the power plant matching dataset, more on this in the database section.
- missingno, library to visualize the completeness of data.

3.1.3 Datasets

Open source data is used in this thesis but it is not bound exclusively to these data sources; the method is generic and as long as the data source contains the input specified in the Data Flow section, it should work, in order to validate this assumption an internal database will be used.For the geographical distribution of data the official European Union data for the NUTS regions from 2016 is used. EUROSTAT [2020].

For cost and investment developments a series of indexes were used. The first of them is the CEPCI, sourced from the official website reports of the Chemical Engineering Magazine [ChemicalEngineeringOnline, 2019] for the values after 2016, and from Turton [2018] for the values before, this is a general indicator for the prices development of the chemical industry. For the price development of the operational and plant capital costs the UOCI/UCCI from the IHS [2020a] is used, the values however are not in a table form, they are openly as graphics in the company website, they are tabulated from there. The graphics are shown in Figures 3.1a and 3.1b.



(a) Upstream Capital Cost Index

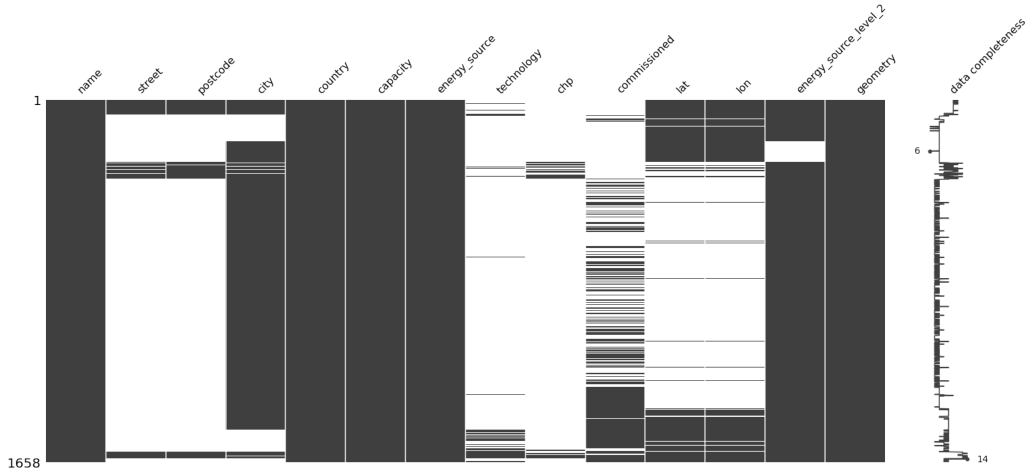
(b) Upstream Operating Cost Index

Figure 3.1: Energy sector capital and operating cost indexes from the IHS [2020a,b]

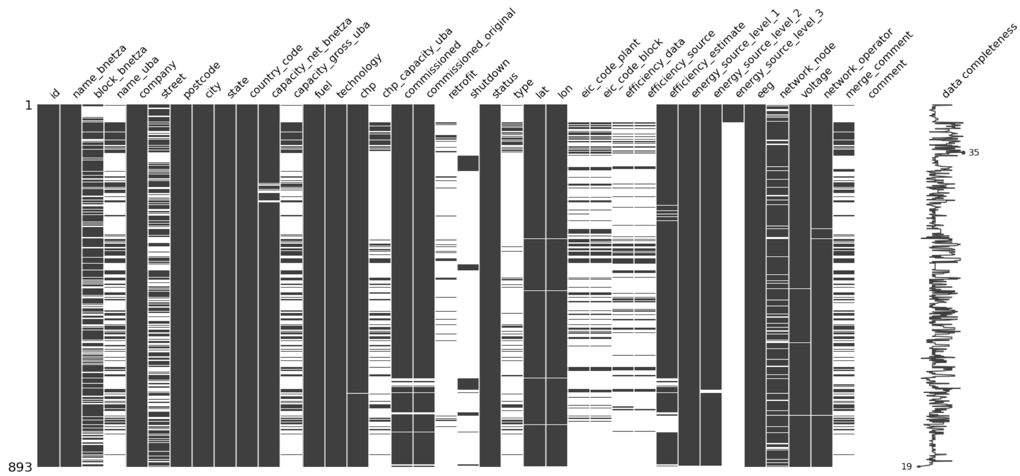
For the transformations of fuel costs, commodity price indexes were obtained from the World Bank Group [2020] for coal costs and the Henry Hub index [EIA, 2020a] for natural gas. An exchange rate for Euro vs Dollar transformations was extracted from Macrotrends [2020]. The details of the use of these indexes will be shown in the Harmonization section, all the index values used can be seen in table 3.4

Thanks to a collection of private and open source institutions, there is a rich collection of data related to power plants and industrial activity. The information ranges from basic like geographical location to detailed analysis and annual emission reports. The openmod [2020] initiative, for example, runs a wiki page that collects models electricity markets, demand and networks across different geographical and temporal scopes; data of energy demand, production and available technologies . All of this information is open to the public and have some sort of open license. Among their sources there is also tools that harmonize different sources to generate one that is richer and more complete, one of the most significant to this work is the Power Plant Matching (PPM) F. Gotzens and Hofmann [2019]. The advantages and disadvantages of different datasets are outlined in Table 3.1. Other sources exist but are not listed as they are all mostly already processed and contained in either Open Power System Data (OPSD) [Jens Weibezahn, 2018] or PPM.

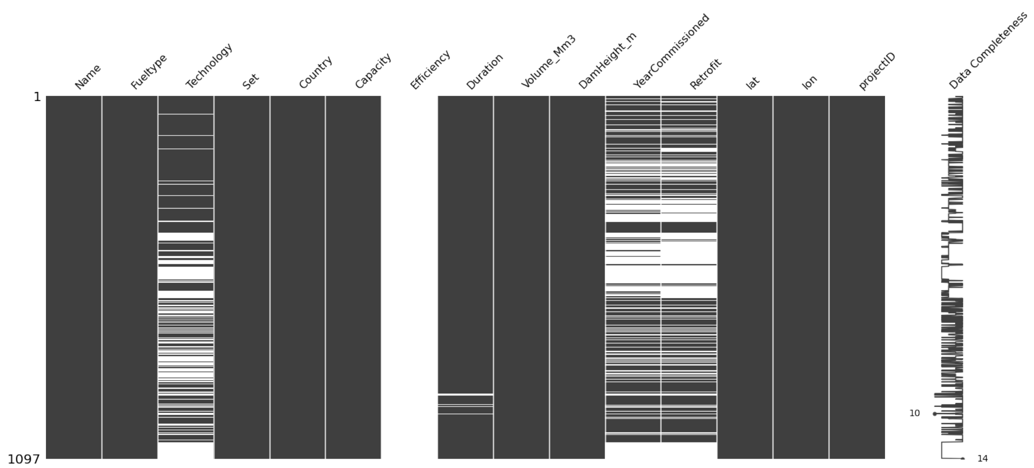
For further comparison the sources, a package for visualization of data is used [Bilogur, 2018]. In Figures 3.2b, 3.2a and 3.2c we can appreciate in color black the data that is available and in white the data that is not, each column represents a column in the source and the rightmost column summarizes the general shape of the data completeness, the number at the bottom left is the size of the dataset. In order to be made comparable, both datasets were filtered to include only the information regarding conventional power plants. It is important to remark how the dataset geographical information of the OPSD is represented mostly in addresses while in the PPM they are all already parsed into geographical points. The reason why PPM seems to be a smaller data set for almost the same region is because it aggregates energy units while OPSD has them separated. The efficiency data available from OPSD is only for the German network.



(a) Completeness of European OPSD Dataset



(b) Completeness of German OPSD Dataset



(c) Completeness of PPM Dataset

Figure 3.2: Completeness of aggregated datasets

Dataset	Advantages	Disadvantages
Open Power System Data [Jens Weibezahn, 2018]	It has an extensive coverage of individual power units. It is the one with the better commissioning year completeness. It contains some points with efficiency data. Data is almost completely processed.	Its geographical details like longitude and latitude are limited to an extent. It's limited to the European Union. One major weakness is that the classification of plants according to their energy source is not very specific in some cases.
Industry About [industryabout, 2019]	It has an extensive coverage of number of units, capacities and technologies. It contains information from countries .across the world	The data has to be heavily processed in order to be used in any kind of model, needs the development of a web parsing tool. Some of the sources are not updated.
Power Plant Matching [F. Gotzens and Hofmann, 2019]	It has a wide coverage in the European union It is very well processed, especially on the geographical side. As it has almost 100 % of the points	Although it's supposed to report efficiency, at the time of writing this work the feature was not working. It is also limited to the European union.

Table 3.1: Advantages and Disadvantages of using different power plant data sets.

3.1.4 Literature

To do significant cost estimations an harmonization of scientific and industrial values is to be done, the details of this methodic will be explained in the next section, for now is important to know that a large number of literature sources for carbon capture cost values. These sources and their extracted values are listed in the appendix 6

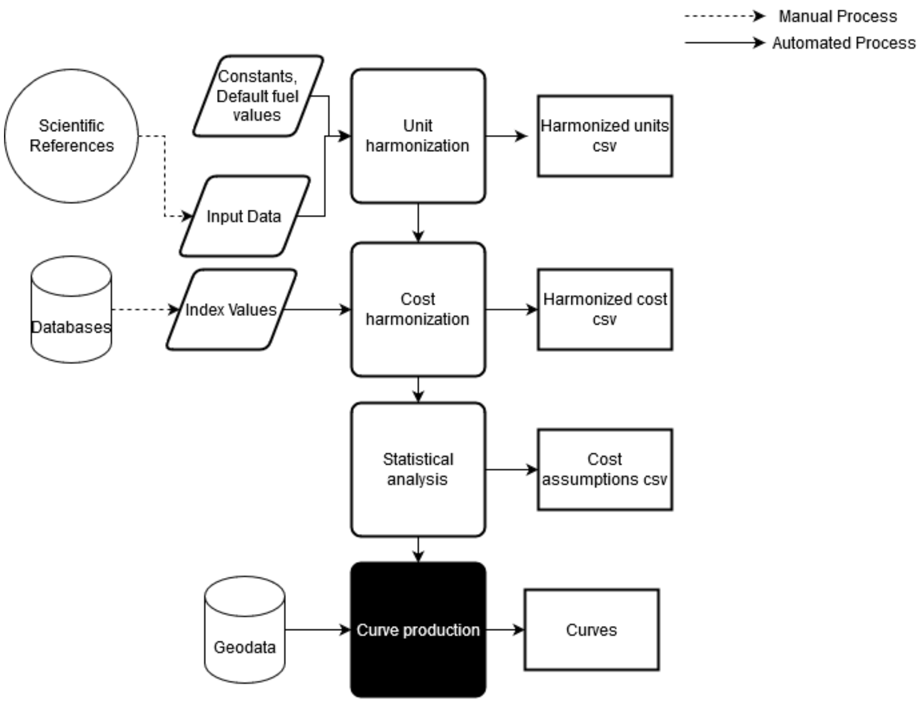
3.2 Reference Harmonization

There are two different approaches to input the cost assumptions to the model. The first of it is taking reference values from existing data collections built for the goal of doing such kind of analyses. The second being producing a set of assumptions aggregating statistically a series of researched of literature values. By doing either of those approaches there is a need of an harmonization of values to make them comparable among each other. The harmonization consist mainly in making sure the units used are the same and the cost values are in the same reference year. The general process of harmonization is represented in figure 3.3.

For the aggregation, a manual extraction of the data values was done, the general structure, data types and explanation of each input row can be found in appendix 6

3.2.1 Unit Harmonization

The harmonization of units is needed given the wide range of reference reporting formats. Each one of the sources that are being worked with has, depending on their objectives, a different structure of their reported values; some only report final LCOE and their compositions, some do the cost in a yearly basis and some as a whole for the project. There are also differences between the



!ht

Figure 3.3: Unit harmonization general flow

measurement units, for example, American studies tend to report in imperial units and some studies change the order of magnitude to achieve better clarity. Other ground for this need is that some studies don't report some values, but these can be obtained by using the given values, an example from this is calculating the specific fuel emissions; some reports only report on the emissions of the power plant itself, but they also report heat rates or efficiency so a transformation can be done very easily.

Each numeric value associated with a dimension with units or series of them was transformed into a set of general units, which are represented in table 3.2.

Value	**Units**
Heat	kJ GJ
Electricity	kWh
Mass	$kgCO_2$ $tonCO_2$

Table 3.2: Reference units

In general, mass units will be represented in Kg or tons depending of the context, heat will be represented in multiples of Joule and electricity Kilowatt-hour/Megawatt hour out of convention. These are going to be used consistently and if the input is in Imperial units, the necessary transformations will be done.

Cost values are reported in different formats, Capital cost is often represented as capital required per unit of nominal capacity (€/MW) but there is cases where only the total cost is given, Because of this a transformation is done to have values in the first form. Similarly, Fixed Operational and Maintenance(FOM) values are often represented as a total for the project, but it is not rare to find yearly representations or even per nominal capacity, Variable Operational and Maintenance(VOM) values are presented more consistently as costs per energy unit, given the high variability among the FOM representations, the harmonization is done by adding together VOM and FOM, sometimes FOM has to be transformed from the reported units into the same of VOM. In this part of the homogenization however, there is no currency value transformation, the homogenization is done in a generic currency unit (CURR). The value transformation happens in the cost harmonization module of the process.

There is no clear convention on how to represent the effectiveness of the heat conversion. Some studies report heat rates(HR) and some others Efficiencies. These are easily interchangeable by matching the energy units of the electricity generated with the necessary heat. Heat Rate is KJ of heat per kWh of electricity so to transform the value to efficiency one muss simply invert and multiply by 3600 seconds that correspond to an hour. To get the Heat rate from the Efficiency one should just do the opposite (Equation 3.1).

$$HR = \frac{3600}{\eta} \tag{3.1}$$

There is also differences among the heat value used, some of the studies report on High calorific value and others in Low calorific values; given that the price of the Fuels is consistent with this rates in each study there is no need of homogenization in that ground.

Despite doing the calculation themselves, some of the studies do not report directly the Fuel Emission Factors (EF). These are associated with the type of fuel used and the calorific value basis. However, there is a very simple work around this. Most studies report on the plant specific emissions (SE) and heat rates (HR) or efficiencies. Using these values we can calculate the utilized fuel specific emission factors (Equation 3.2).

$$EF = \frac{SE}{HR} \tag{3.2}$$

The fixed cost factor is an adjustment done to the capital cost to account the effect of inflation in the cost of the project across the length of it. It is a function of the inflation rate (r) and the length of the project life in years (T) (Equation 3.3) [Rubin et al., 2013].

$$FCF = \frac{r * (1+r)^T}{(1+r)^T - 1} \tag{3.3}$$

Some Studies report fuel cost of the fuels as per mass or volume unit. In these cases it is necessary to transform the values in to energy related costs. To do so, the Heat Value basis should be known. This is either High Heat Value or Low Heat Value. Most studies report this basis. To convert the Cost from one basis to another the following calculation has to be done(Equation 3.4).

$$FC_{Heat} = \frac{FC_{Mass}}{HV} \tag{3.4}$$

3.2.2 Cost Harmonization

In order to make the comparison of the costs significant a series of prices and cost conversions has to be done. These conversions are related to the type of cost, namely if it is a capital investment, if it is related to operation of the assets or if it is related to a commodity. These different costs have different developments that can be tracked using indicators that are reported periodically. There is another transformation to be done which is the currency exchange, it takes place at various steps of the process to ensure consistency. The indicators used are obtained from the sources listed in the datasets section, they and the component they are related to can be seen in table 3.3.

In order to have a consistent effect of the indicators, before the transformations, all the values are homogenized into USD at their given years. This is mostly because the indicators are developed taking in account the american market and take into account inflation of the USD. Once they are all in USD, the general formula to transform a value is shown in Equation 3.5 .

$$V_{new} = V_{ref} * \frac{I_{new}}{I_{ref}} \tag{3.5}$$

The cost development indexes have reported values in a yearly basis that can and should be taken as they are. For the fuel development indexes that it is not the case, the reports go all the way down to daily granularity, so they should be aggregated to calculate an annual average. Same for the currency exchange rates. The development of the used index can be seen in figure 3.4 and

Indicator	Source	Cost Component
Upstream Operational Cost Index (UOCI)	IHS Markit	Costs associated to Operation and Maintenance
Upstream Capital Cost Index (UCCI)	IHS Markit	Capital costs
Henry Hub Index	U.S. Energy Information Administration (EIA)	Natural Gas Prices
Coal Commodity Prices	The World Bank Organization	Coal Prices
Historical Euro-USD Exchange Rate	Macrotrends	Development of the Exchange rates
Chemical Engineering Plant Cost Index (CEPCI)	Chemical Engineering Magazine	Optional replacement for IHS indexes in the differential capital cost of carbon capture.

Table 3.3: Assigned uses of the indicators

table 3.4. These values are used based on a reference Year corresponding to the earliest date of the literature, the values were converted to 2019 due to data availability for 2020 not being complete for all the sources. After doing this harmonization, the output values are used to calculate newly the cost of carbon capture.

Year	UOCI	UCCI	CEPCI	Coal	EUR-USD	Natural Gas
2002	1.00	1.00	1.00	1.00	1.00	1.00
2005	1.50	1.07	1.18	1.88	1.32	2.62
2008	1.80	2.10	1.32	2.84	1.47	1.17
2009	1.70	1.90	1.39	3.91	1.40	1.30
2010	1.60	1.90	1.45	5.02	1.56	2.63
2011	1.90	2.19	1.43	3.34	1.40	1.11
2013	1.85	2.00	1.48	4.80	1.47	1.19
2017	1.65	1.62	1.43	3.50	1.19	0.89
2019	1.70	1.72	1.53	3.08	1.20	0.76

Table 3.4: Index values with 2002 reference

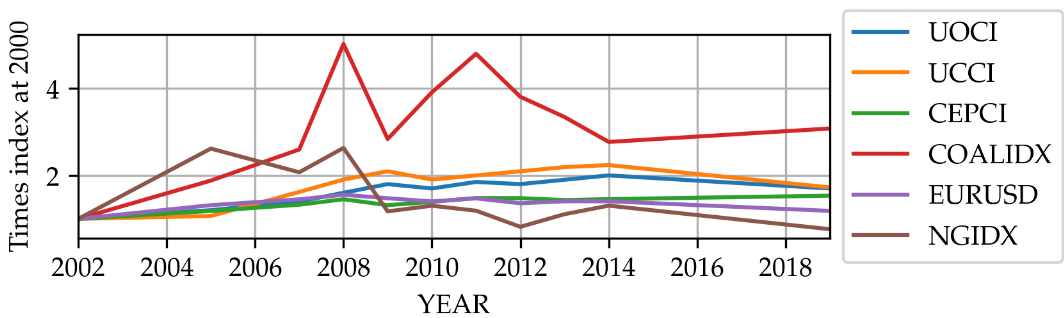


Figure 3.4: Relative development after 2002 of cost indexes

3.2.3 Calculation of cost of carbon capture for individual plants

The calculation of the harmonized cost of carbon capture is done based on Equations 2.1 and 2.3, Considering the TCC and FOM reported per net capacity units (TCCP and FOMP) the COE for the reference case can be written as Equation 3.6.

$$COE_{ref} = \frac{(TCCP_{ref})(FCF) + (FOMP_{ref})}{(CF)(8766)} + VOM_{ref} + (HR_{ref})(FC) \tag{3.6}$$

Same for the captured case, Equation 3.7.

$$COE_{cc} = \frac{(TCCP_{cc})(FCF) + (FOMP_{cc})}{(CF)(8766)} + VOM_{cc} + (HR_{cc})(FC) \tag{3.7}$$

We can also consider the capture case as a function of the base case and the additional values from adding CC. Equation 3.8.

$$\theta_{i,cc} = \theta_{i,ref} + \delta\theta \quad , \theta_i \in \{TCCP, FOMP, VOM, HR\} \tag{3.8}$$

With 3.8 we can rewrite 3.7 as 3.9

$$\begin{aligned} COE_{cc} = & \frac{(TCCP_{ref} + \delta TCCP)(FCF) + (FOMP_{ref} + \delta FOMP)}{(CF)(8766)} \\ & + VOM_{ref} + \delta VOM + (HR_{cc} + \delta HR)(FC) \end{aligned} \tag{3.9}$$

With 3.9 we calculate the numerator of 2.3 yielding 3.10 which represents the differential cost of electricity purely in terms of the added carbon capture module.

$$COE_{diff} = \frac{(\delta TCCP)(FCF) + (\delta FOMP)}{(CF)(8766)} + \delta VOM + (\delta HR)(FC) \tag{3.10}$$

The denominator of 2.3 cn be simplified in a similar fashion with equations 3.11 and 3.12 and 3.10 yielding 3.13

$$t_{CO_2}/MWh_{emitted} = HR * EF \tag{3.11}$$

$$t_{CO_2}/MWh_{captured} = \eta * (t_{CO_2}/MWh_{emitted}) \tag{3.12}$$

$$Cost\,of\,captured\,CO_2 = \frac{\frac{(\delta TCCP)(FCF) + (\delta FOMP)}{(CF)(8766)} + \delta VOM + (\delta HR)(FC)}{(HR_{ref} + \delta HR) * EF * \eta_{capture}} \tag{3.13}$$

Using the assumption of unified values of VOM and FOM, 3.14.

$$\delta OM = \frac{\delta FOMP}{(CF)(8766)} + \delta VOM \tag{3.14}$$

We can yield equation 3.15 which will be used to calculate the harmonized and end model cost of carbon capture.

$$Cost\,of\,captured\,CO_2 = \frac{\frac{(\delta TCCP)(FCF)}{(CF)(8766)} + \delta OM + (\delta HR)(FC)}{(HR_{ref} + \delta HR) * EF * \eta_{capture}} \tag{3.15}$$

3.2.4 Assumptions for the harmonization of values

For the creation of a cost of carbon capture distribution using equation 2.3 the following technical and economical assumptions were taken for each of the points in the analyzed sources:

- The costs correspond to the state of the art method of carbon capture, which, in the case of the used literature corresponds to a variation of the MEA process.

- The reference currency for all the cost values is 2019€. The order of the indexes for the transformation are those mentioned in table 3.4 and their values are those in table 3.3
- The efficiencies(heat rates) of the power plants, the costs and emission factors of the fuels were not modified in regard to their heat value basis. The ground is that it is expected that they are consistent within each study and doing transformatios in such regards would imply having to create a particular set of assumptions for each of the studies which can bring unnecessary uncertainty that may skew the comparation among values. The option to harmonize the fuel costs is however implemented in the code with transformations of High calorific values to Low calorific values based on reference fuels but not considered in this thesis.
- The life of the plants, as well as the return rate are harmonized to 25 years and 10% respectively (as it was the most common value pair used) and they are used to calculate the FCF using 3.3. An option for not doing so is implemented however.
- Given the capture efficiency from the studies is consistently around the value of 90% only with very few exceptions (see appendix 6) this value is used, this parameter affects significantly the output, so it is analyzed during the sensitivity analysis 4.6.

3.2.5 Statistical aggregation of harmonized literature values

To do an statistical aggregation of the data obtained two main groups will be identified, for each one of the variables related to the addition of a carbon capture module a series of analysis will be done based on the potential characteristics of the power plant that are available from the data sources. From this analysis the aggregation will be done in the form of a single average with standard deviation calculated from the values or a linear regression whenever a trend is found.

Some of the sources worked at explicitly report for Retrofit of power plants, since retrofitting will be the main focus of this study, ruling out that we can do an aggregation based on all the values obtained or just the ones that represent retrofit studies, which are way less, is important. In order to consider them independently or not we have to test if they could belong to the mix of all the values. For doing so, we perform a Kolmogorov-Smirnov test on the distributions with the following hypotheses. This test the goodness of fit of two distributions as a whole and it is used to find out if the two samples are likely from the same distribution or not [SPR, 2008]. The hypotheses used are the following:

$$H_0 : \text{"The 2 independent samples are drawn from the same continuous distribution"} \tag{3.16}$$

$$H_1 : \text{"The no retrofit sample has a lower cumulative distribution than the retrofit "} \tag{3.17}$$

After this test, depending on the result, values will be analyzed for their mean, median, standard deviation and other distribution properties. If a regression trend against the power plant related variables is found a line will be fit and used as an assumption in the analysis.

3.3 Geographical Distribution of Potential and Cost

For the creation of a geographical distribution of potential cost of carbon capture for a determined process there are two main parts that have to come together. First a geographically distributed indicator of the activity of a process and secondly a function that calculates emissions as function of this activity (Equation 3.18) and another that calculates the cost of capturing these emissions.

$$e_{g,s} = f(a_{g,s}) \tag{3.18}$$

$$c_{g,s} = f(e_{g,s}) \tag{3.19}$$

where:

g : geographical point

s : source type

e : emission rate

a : activity rate

c : cost of carbon capture

For the case of conventional power plant technologies these two last functions are combined in 3.15, in this case the cost is also affected by the activity itself given the energy for the process is obtained from the same power plant, this is not necessarily the case for other industrial processes or for power plant configurations that use external sources of energy.

3.3.1 Conventional power plants distribution

To apply equation 3.15 to every power plant some values have to be generalized and others have to be assigned depending on the characteristics of the point, this process is dependent on the data source used. It was earlier mentioned that there were various options from data sources to pick. and after doing a comparison using figure 3.2 and a reading of the documentation it was decided that the power plant matching (PPM) dataset was to be used.

The PPM dataset has the following columns that are relevant to this work: Capacity, Location, Type of Fuel, Type of power technology, Year of Commission, year of last retrofit and labels to associate to other data sources. Most of them have no missing values so we are good to take them to the next step. Besides this, there is technical information that it is still needed before adding the cost parameters; a value for heat rate, fuel heat value and emission factors are needed. To deal with the fuel part an harmonization for each type of fuel can be done, this will be detailed in the assumptions part of this section, for the efficiency there is two approaches that can be taken, the first being taking a set of assumptions from another source and the second a novel method developed for this thesis explained in the next subsection.

With the technical information assigned the values from adding carbon capture can be matched, for doing so there is two possible approaches: One is to use values from reference data sources like Agency [2016] and Klaus Görner [2016]. The second is done by using the statistical aggregation of the values obtained from the harmonization section. Once matched with values the calculations are done using 3.15. The steps described in this section for the production of an unified geographical data are shown in Figure 3.5.

3.3.2 Industrial processes

For the geographical distribution and activity values of industrial processes, from which in the case of this thesis only Cement and Steel production will be considered, the industryabout [2019] data compiled by Hu [2019] will be used. The cement plant calculations will be done based on values and equations reported by Gardarsdottir et al. [2019]. In the case of steel, the cost and assumptions will be taken from Kuramochi et al. [2012]. The cost from both sources will be transformed using the indexes reported in the harmonization section.

To calculate the cost of CO_2 from the production process specific emission an capture rates have to be calculated. These are obtained using equations 3.20 where the specific emissions are added to the indirect emissions associated with energy production of the process and 3.21 where energy emissions are calculated as a function of emission factor and required power.

$$e_{clk,eq} = e_{clk} + e_{el,clk} \tag{3.20}$$

$$e_{el,clk} = e_{el} * P_{el,clk} \tag{3.21}$$

The cement data has production values of tons of cement production per year, however emissions are associated to clinker production, the clinker/cement ratio varies across product specifications, for simplification, the average value reported by Hu [2019] will be used. Based on the clinker production a total emission is calculated based on values reported in Gardarsdottir et al. [2019]. Using this value and an efficiency value we can obtain the captured carbon. With the captured

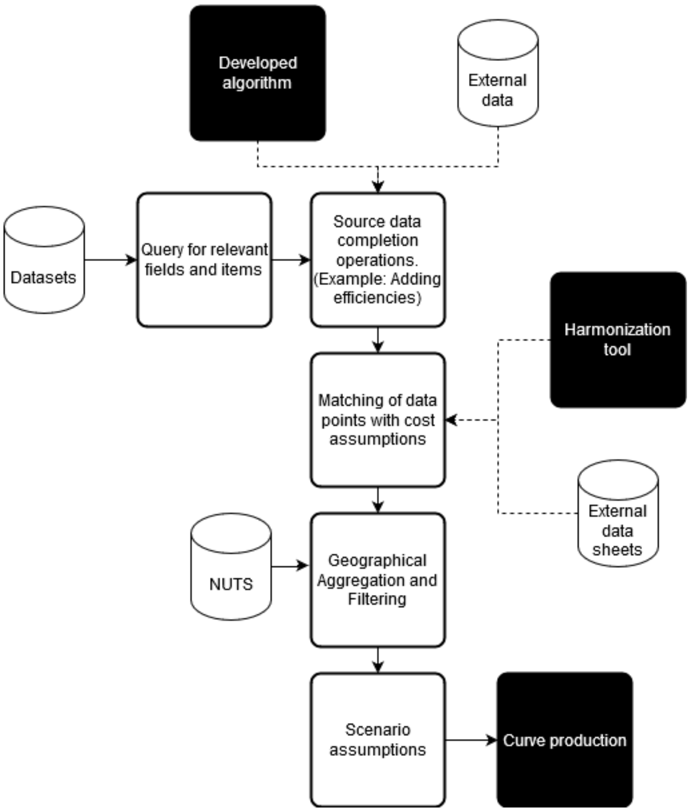


Figure 3.5: General process of the creation of a geographical distribution of cost of carbon capture

values and the cost reported by Gardarsdottir et al. [2019], the cost of carbon capture can be reported as in equation 3.22 where COC means Cost of Clinker.

$$Cost\, of\, captured\, CO_2\, Cement = \frac{COC_{capture} - COC_{ref}}{\eta * e_{clk,eq}} \tag{3.22}$$

In the case of Iron and steel plants, the costs and capture values were taken directly from Kuramochi et al. [2012]. All the reported values were reported as function of the plant production(p), the values of captured CO_2 are calculated with 3.23. The cost values are scaled through an scaling index with equation 3.24. The cost of carbon capture in a particular plant is calculated with 3.25 and to transform it to a ratio of captured carbon equation 3.26 is used.

$$CO_{2captured,steel} = p_{steel} * c_{steel} \tag{3.23}$$

$$Cost_{production,specific} = \frac{p_{steel}}{p_{steel,ref}}^{sf} \tag{3.24}$$

$$Cost_{CO_2capture} = Cost_{production,specific,capture} - Cost_{production,specific,reference} \tag{3.25}$$

$$Cost\, of\, captured\, CO_2\, Steel = \frac{Cost_{CO_2capture}}{CO_{2captured,steel}} \tag{3.26}$$

The values for cement cost are collected in table 3.5 and the ones for steel are found in 3.6

Name	Value	Units
Clinker emission factor	0.5	kg_{CO_2} / $kg_{Clinker}$
Clinker cement ratio	0.5	$kg_{Clinker}$ / kg_{Cement}
Cost of clinker reference plant	62.6	2014€/$t_{Clinker}$
Cost of clinker Capture plant	107.4	2014€/$t_{Clinker}$
Specific power, reference	15.88	MW/$Mt_{Clinker}$
Specific power, capture	29.5	MW/$Mt_{Clinker}$
Electricity emission factor	850	Kg / MWh

Table 3.5: Assumption collection for Cement power plants [Gardarsdottir et al., 2019]

Name	Value	Units
Specific capture, Blast furnace	890	kg_{CO_2} / $t_{Pig\ Iron}$
Cost of production with capture	420	2007€ /t_{Roll}
Pig Iron to Steel Ratio	1	
Cost of production, reference	300	2007€ /t_{Roll}
Scaling factor	-0.2	
Reference scale	4	Mt/y

Table 3.6: Assumption collection for Steel power plants, [Kuramochi et al., 2012]

3.3.3 Strategies for the completion of efficiency data

Given the relevance of plant efficiency for the calculation of emissions and cost having values for efficiency is of great importance. Power plant projects are extremely diverse as they are built to fit the particular needs and characteristics of the region they are built in. This means that, assigning an efficiency value is not a simple task. To perform this compensation one can use different approaches, the simplest of them is a naive estimation using literature values for average efficiencies, the second is to associate it to other plant properties such as erection year, capacity and main fuel. From this second approach there is also the possibility to find values in the literature.

For this work we exercise a methodology that uses available information to complete the missing values using open data. One problem with this approach is that, currently, the data sources have limited availability of efficiency information as only the OPSD dataset for Germany has such characteristic with around 115 reported values. The amount and distribution are rich enough to be used as a source for an estimation to feed the other data points. Using this information we can develop a regression model that can be trained using the existing data and then apply it to estimate the missing efficiency values.

There is in the state of the art, many ways of completing missing data from a dataset. Some examples are: Probabilistic classification like Naive Bayes, linear regression, support vector machines(SVM) and decision trees. The tool to be used depends strongly on the type of dependent and independent variables to be analyzed. By looking at the selection of variables used they can be divided in numerical and categorical (Table 3.7). Given the fact that there is this mixture, using a

Categorical variables	**Numerical variables**
	Nominal capacity
Type of fuel	Year of plant comission
	Year of last retrofit

Table 3.7: Variables for regression of efficiency

regression can be tricky and may lead to misleading results. The dependent variable is continuous, so using a probabilistic approach is not considered. Ruling out a support vector machine is harder but justified by the fact that categorical and numerical independent variables are being used. This lets us boil down to using a decision tree structure, however a single tree, despite the small amount of independent variables is prone to error because the degree of effect for each variable is unknown. This problem can be worked around very easily by having an ensemble of trees that have different input independent variable vectors, that go through the process of being built and updated based on a loss function; the model composed of the ensemble of trees is called a Random Forest [Breiman, 2001].

The input of the trees has to be conditioned to get a better result. First, the categorical variables must be encoded, and because they are a small amount the best way to go is one-hot vector encoding [Geron, 2017]; That means to turn the columns of categories into series of columns associated with each characteristic and values of 0 if the data point does not have the characteristic and 1 if it does. For the numerical values, one option is to scale the values using a minmax normalization; turning the maximum number into 1 and the minimum to 0, but this is not necessary for the tree algorithms given their comparative nature, however, the scaling is done to open the possibility of using the same data input to perform a different kind of regression if it is necessary. When the training data is transformed, also the data to be predicted has to be transformed so that adds computations to be done, but given the scale of the data sets that is not currently an issue.

A last feature conversion is being done. It is related to the year and retrofit variable. In the data sets, these variables are being reported in the form of years. In the case of the retrofit variable its availability is intermittent, naturally, as not every plant has been retrofitted. But it creates a complex relationship with the effect of age in the efficiency. It is to be expected that the retrofitted plants have been updated used technology of the time they were retrofitted, so the relationship between the year they were built and the efficiency is expected to be less strong, but at the same time it is not expected that the correlation is as strong as if the plant was completely build that year, because the retrofitted plant may have old units still operating or parts that are deteriorated but are not worth being replaced yet. In order to take in account this complexity, a feature engineering is done: First, if the power plant has a retrofit year, this year will replace the commissioning year variable to take in account the latest year they are related to in a technological sense; second, the retrofit variable will be turned into a binary variable, so the decision trees can assign some weight to the effect of having the plants retrofitted and not.

In the prediction step, similar transformations have to be done, with some modifications. If the commissioning year information is not available in a data point, the average year is being used as a dummy. In the case of the capacity, in these datasets all the points have a value so this kind of transformation is not being done.

For comparison of the performance of this approach, a linear regression with the comissioning year and a single fixed value assumption dataset are created. These will be tested with random selections from the source points to asses the error.

3.3.4 Geographical aggregation

In general the geographical aggregation consist in taking a resolution, namely NUTS level; load the polygons at this resolution and then find all the points in the dataset that are found within each one of these polygons. Next, values are aggregated by summing them and the cost values by averaging them. From the polygon then a centroid is calculated and this is assigned as the new geographical representation of the point. The aggregated values can then be relabeled to a lower NUTS level to generate curves using these new points. In a practical sense, having this characteristic helps a potential user of the data to have choices in which resolution of data to use.

3.3.5 Assumptions for inclusion of Biomass and Biogas

The power plant matching contains data on bioenergy power plants but it does not specify the type of fuel used. To have a general idea of the distribution of the bioenergy sourced Captured CO_2 these values are matched just as its done with the fossil fuels but with costs and emission values sourced from Möllersten et al. [2003] where black liquor from pulp mills is considered as energy source. It is to be noted that analyzing the complexity of biomass and biogas cost and potential can be an independent project by itself as it can be seen in Zhang et al. [2020].

3.3.6 Assumptions for the geographical distribution of cost

For the power plant value cost calculation the following assumptions are considered:

- The life time of the powerplants with retrofitted carbon capture is homogenized to 25 years and their return rate is fixed to 10% per year.
- The capture efficiency is assumed to be 90%.
- The capacity factors of the plants are generalized in table 3.8.
- For fuels, the values assumed are shown in table 3.9.

Energy source	Capacity Factor
Lignite	50%
Hard Coal	50%
Natural Gas	50%
Bioenergy	30%

Table 3.8: Capacity factors assumed for the different energy production processes [Morales Pedraza, 2019]

Type	**Name**	**HHV** GJ/ton(m^3)	**LHV** GJ/ton(m^3)	**Cost** 2019USD/ton(m^3)	**EF** Kg_{CO_2}/GJ
Hard Coal	Illinois 6	25.35	24.12	34.47	94
Lignite	Powder River Basin	11.86	10.07	10.43	110
Natural Gas		40	36	84.3	56
Bioenergy	Black Liquor	21	19.3	63	71

Table 3.9: Fuel value assumptions, [EIA, 2020c], [EIA, 2020b], [Möllersten et al., 2003]

As boundary condition of this thesis a state of the art approach to capture technologies will be considered, this means that only well documented technologies are to be considered in the potentials; this is however not hard coded and can be easily expanded to include more technologies.

3.4 Cost Potential Curves

A cost potential curve for a determined area is created using the formula established by Naims [2016]. Which sorts the values from the lowest available cost to the highest in the collection (Equation 3.27). In the case of this thesis, however, the sorted indexes will correspond to a geographical point for a determined technology which could either be aggregated or not, this instead of the totals for the carbon producing processes.

$$p(q) = p_i \forall q \in]\, q_{1-1}; q_i] \tag{3.27}$$

For the space U as the union of all the quantity intervals sorted :

$$U_{i=1}^{n}(q_{i-1}; qi) \tag{3.28}$$

i : Rank of the CO_2 emitting source geographical point.

n : Number of ranked sources

p_i : Captured cost €/tCO_2.

q_i : Total emissions, Mt/y.

The process of producing a curve will be done for every geographical region given as input, and will be performed for all the points within the given region. After all the curves are produced, the values will be displayed in a single graph. The general process for the curve production is represented in figure 3.6

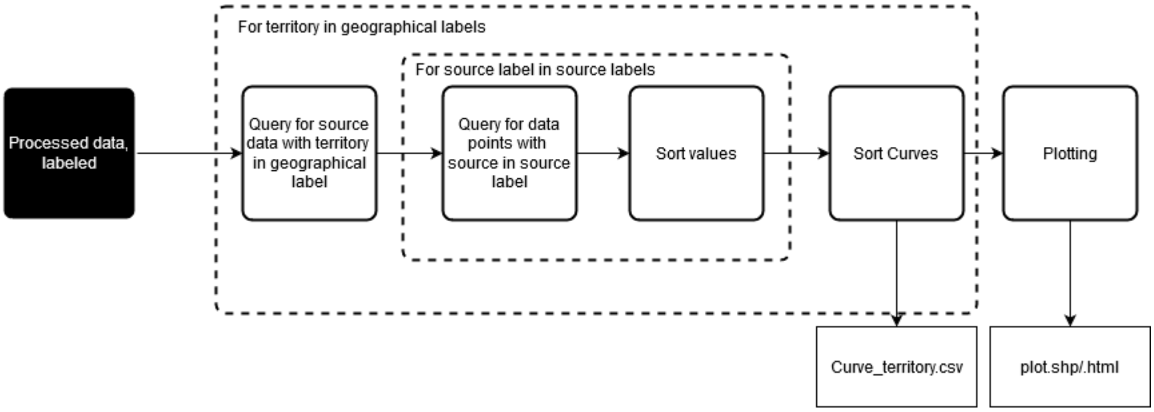


Figure 3.6: General process of the production of cost potential curves

3.4.1 Scenario Development

The analysis of detailed scenario development is not within the scope of this work. Despite that, a rudimentary estimation based on energy system and environmental emission scenarios is going to be applied for demonstration porpoises.

The Global Energy and Climate Outlook (GECO) 2019 edition [Keramidas K., 2021] analyses development pathways of the global energy system. It considers macroeconomic and technological assumptions that represent the set of decisions that would have to be taken in order to achieve the 2°C temperature rise above preindustrial levels goal. This is compared with a reference scenario and a more ambitious 1.5°C scenario. These scenarios consider decarbonisation of the energy system as one of the key elements of the transition process. They can give an insight of the future availability of CO2.

Another reason to chose the GECO scenarios is that they have 5 year intervals which helps to have a discrete time distribution that makes sense in terms of power plant operation.

The sheer amount of factors (which could be technological, political, economical etc.) that could affect the cost of carbon capture technologies makes it almost pointless to try to predict the absolute cost at future developments. It is possible, instead, to have a relative calculation as function of the Potentials which can be more easily modeled. Under this assumption, all the reports on cost in future scenarios should be considered relative to the current year (2020) values.

To estimate potential, there is two main approaches that were considered. The first of them and the one represented in the scenario results consists in having a projected amount of energy production in a goal year. With given value an algorithm is built that "closes" the power plants based on their age(i) and capacity(j), closing always first the oldest and smaller. The process is represented in eq 3.29, here H represents the open power plant park and T is the total capacity at the given scenario year.

$$f(H_{i,j},T) = \begin{cases} H_{i,j} & \text{if } T < \sum j \\ f(I_{i,j},T), I_{i,j} = \{h \in H_{i,j} / g(H_{i,j})\} & \text{if } T >= \sum j \end{cases} \tag{3.29}$$

$$g(H_{i,j}) = \begin{cases} H_{n,j}, n = \min i & if|H_{n,j}| = 1 \\ H_{n,m}, n = \min i, m = \min j & if|H_{n,j}| > 1 \end{cases} \tag{3.30}$$

The second approach consist in creating nodes using the geographical aggregation tool and change these nodes capacities until they meet the total energy requirements of the scenario. The only component of the scenarios which was hard coded in every case was the "kohleausstieg" which is supposed to happen by 2038 in Germany [Bundesministerium der Justiz und für Verbraucherschutz, 2020].

Results

4.1 Reference Harmonization

The general distribution of the harmonized cost of carbon capture from different the sources can be appreciated in figure 4.3. It can be noticed at first glance that the variation is already very high, in order to make major conclusions one has to look at the different characteristics of the studies and how the variables affect the final cost. An overview of the cost distribution for the different technology groups can be viewed in figure 4.1 where ultra super critical and supercritical pulverized coal boilers are bundled into the label SCPC, subcritical and fluidized bed are bundled into SUBC and natural gas combined cycle is labeled as NGCC.

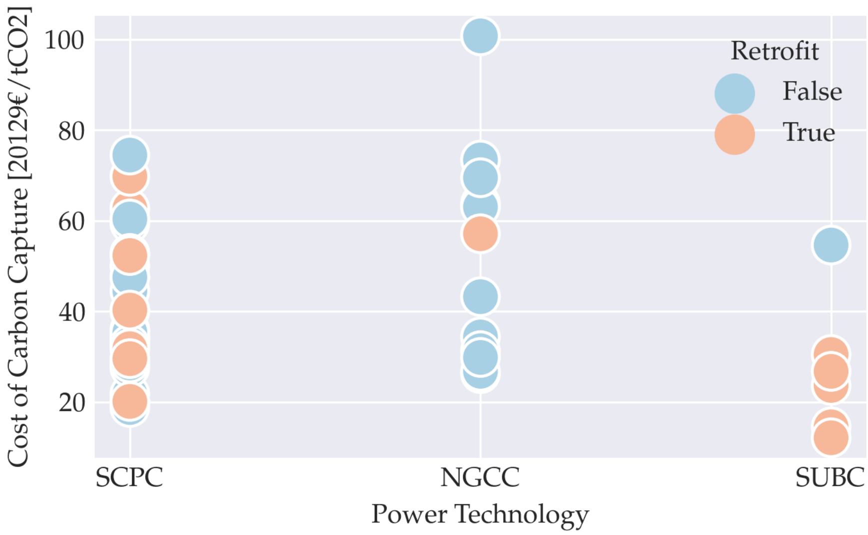


Figure 4.1: Cost distribution for the different power technology groups, the colors represent if the study is done considering the retrofitting of an existing plant or not.

To asses the effect of the variables a regression analysis was done, a general overview of the absolute correlations between the variables can be observed in the correlation matrix (Figure 1) found in the appendix 6 where the expected relationships like the effect of Heat rate on the emissions and captured amounts of CO_2 can be seen. This representation ignores all of the categorical variables, it is however a good start to find places where the variables may be related. A further analysis was done and it was found that the effect of power plant capacity on the marginal cost is negligible, the case is the same for efficiency and publication year. There was however an interesting finding where the effect of the efficiency of the base plant was strongly correlated with that of the capture plant, meaning that the better the power plant efficiency the less harsh is the penalty of adding capture modules.

Despite the difference among power plant technologies being significantly higher for coals, the classification is not useful in relation to the datasets available for this study, because of this it was opted instead to differentiate among fuel types. The value distribution among fuel types can be seen in figure 4.2. This classification is easier to map for individual power plants when their technological information is not available. However cost values are not to be used directly, instead we use their components, this is done to consider the effect of the emitted carbon amounts on the general cost. But before doing the reference value generalization of the cost components they should be statistically analyzed.

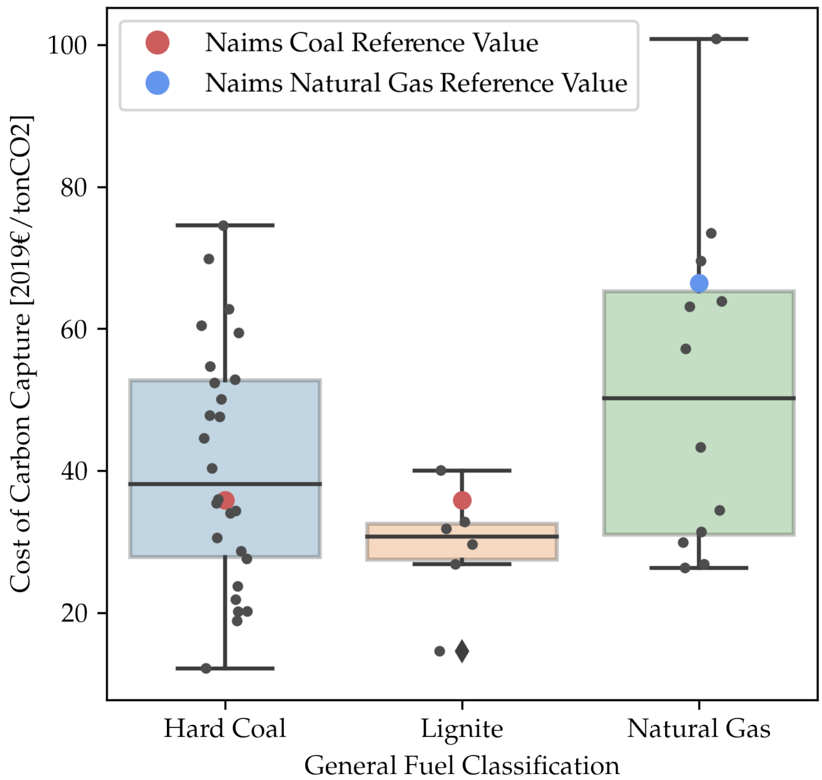


Figure 4.2: Harmonized values of cost of carbon capture for different fuel categorizations

4.1.1 Key findings of harmonization

- We can already rule out most of the costs of the reference plant and focus on the differential effects of adding Carbon capture
- There is not a strong case to differentiate among power plant Net capacities for calculating costs
- Fuel types are a very important factor for the cost calculation so differentiation of these values is important.
- Efficiency has a very slight effect, this effect can be accounted easily as the LCOE calculation already includes this value.
- The considered data shows a tendency for the values of capex to increase over time and operational values to decrease.
- The only outstanding region is China as it shows to have lower costs in general. This is because of the differences on costing frameworks of the country. [Hu and Zhai, 2017]

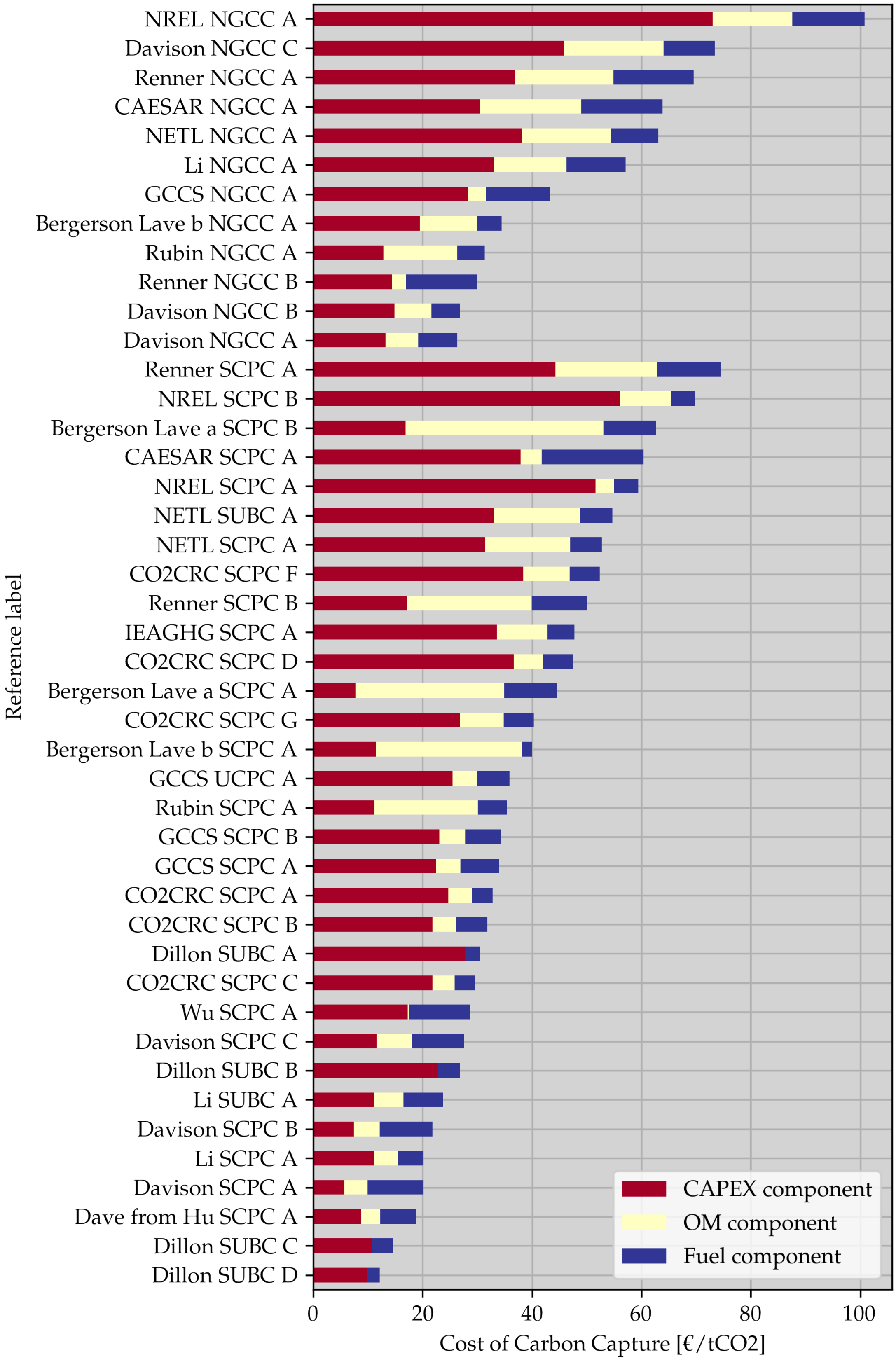


Figure 4.3: Harmonized cost of carbon capture from different sources

4.1.2 Aggregation of Harmonized values

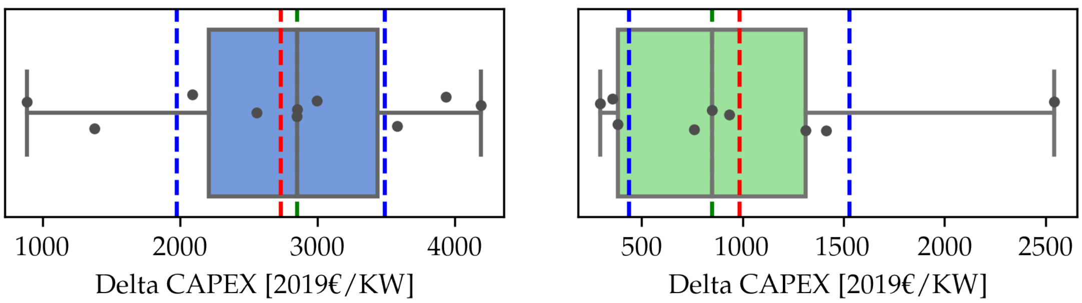
To use the data from the harmonization tool to calculate emissions of power plants the values were aggregated, but before, a series of validations had to be performed using the methodology mentioned in the corresponding section (3.2.5). The implementation of the Kolomogorov-Smirnov test can be seen in the following code.

```
from scipy.stats import ks_2samp
retro_true = df_capex[(df_capex.retrofit == True) & (df_capex.fuel_type ==
    "coal")]["delta_capex"].dropna().values
retro_false = df_capex[(df_capex.retrofit == False)& (df_capex.fuel_type ==
    "coal")]["delta_capex"].dropna().values
ks_test = ks_2samp(retro_true, retro_false, alternative = "lower")
p = 0.05
print("The ks statistic is {}, the p-value is: {}".format(*ks_test))
if ks_test[1] < p:
print("H0 is rejected, the samples are likely to come from different populations")
else:
print("H0 is accepted, the samples are likely to come from the population")
```

The output for the coal delta capex values are the following:

```
"The ks statistic is 0.5125, the p-value is: 0.026723848234145718
H0 is rejected, the samples are likely to come from different populations"
```

This means that the capital cost component of our retrofit data can not be clustered together with the non retrofit data to get an averaged value. The final distribution can be seen in figure 4.4a. By analyzing each of the points of the distribution it was found out that the retrofit values in the interquartil range were all having similar technical assumptions, namely they all considered the additional coupling of a Flue Gas Desulphurization and Selective catalytic reduction [CO2CRC, 2017], [Dillon et al., 2013]. Lower values, are from the same studies but without sulfur treatment. Higher values are from Veatch [2012] which reports higher costs in general, from circulating fluidized bed [Dillon et al., 2013] plants which has a different configuration than the usual SCPC and the advanced solvent configuration from [CO2CRC, 2017] which is different enough to keep out. The mean values and the intervals were calculated assuming a t-student distribution with a confidence interval of 95%, this is because the sample size of reports is smaller than 30 making the use of statistical scores like z not applicable.



(a) Distribution of delta CAPEX of adding carbon capture to a coal power plant project.

(b) Distribution of delta CAPEX of adding carbon capture to a natural gas power plant project.

Figure 4.4: CAPEX distribution values, the outer blue lines are the one standard deviation limits, the middle red line is the meand and the green line is the median

There is no explicit report of retrofit projects for CCS on Natural gas plants, because of this, the reference values to be taken are going to be aggregated from all the data points. The distribution can be found in figure 4.4b. From this distribution only the values in the interquartil range will be

Fuel	Value Name	Lower Limit	Upper Limit	Reference Value	Units
Natural Gas	delta capex	431.723	1504.038	967.881	2019€/KW
	delta om	0.004	0.010	0.007	2019€/KWh
	delta heatrate	675.315	2166.751	1421.033	KW/KWh
Coal	delta capex	1943.219	3433.327	2688.273	2019€/KW
	delta om	0.006	0.013	0.009	2019€/KWh

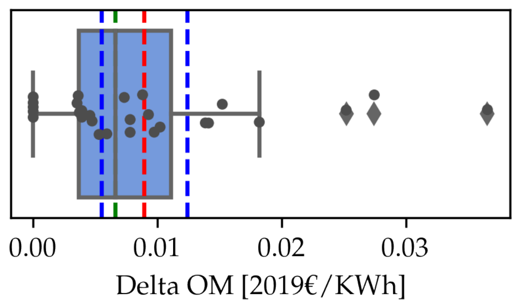
Table 4.1: Assumption map for power plant cost data obtained from the source harmonization.

considered for the aggregation.

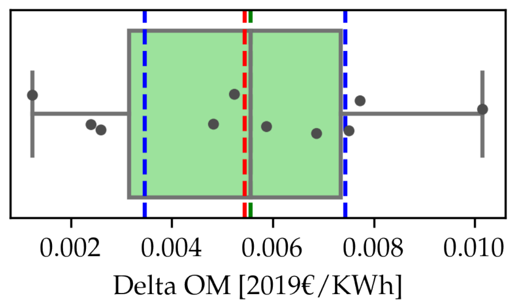
For the operational and maintenance values a similar approach will be used:

```
"The ks statistic is 0.45454545454545453, the p-value is: 0.09206684492725692
H0 is accepted, the samples are likely to come from the same population"
```

In this case, the retrofit values for coal are likely to be of the same population of the non retrofit values so we can aggregate them together. The distribution can be seen in figure 4.5a. Same procedure is done for natural gas seen in figure 4.5b. The collection of the assumptions extracted from this analysis can be seen in table 4.1



(a) Distribution of delta O&M of adding carbon capture to a coal power plant project.



(b) Distribution of delta O&M of adding carbon capture to a natural gas power plant project.

Figure 4.5: Distribution values, the blue lines are the one standard deviation limits, the red line is the mean and the green line is the median.

For the heat rate values a similar method as for the cost values was followed for natural gas. For coals given the correlation found between the efficiency and the efficiency penalty, a linear regression represented by equation 4.1 will be used for the calculations. The parameters of the regression can be seen in table 4.2 and the plot in figure 4.6.

$$\delta HR = c * HR_{ref} + b \quad \{HR_{ref} | MinHR <= HR_{ref} <= MaxHR\} \tag{4.1}$$

Fuel	Slope(c)	Intercept(b)	Min HR	Max HR	Intercept error	Slope error
coal	1.418061	-9696.578989	7912	14400	2050	0.20444

Table 4.2: Regrssion parameters for the calculation of delta heat rate.

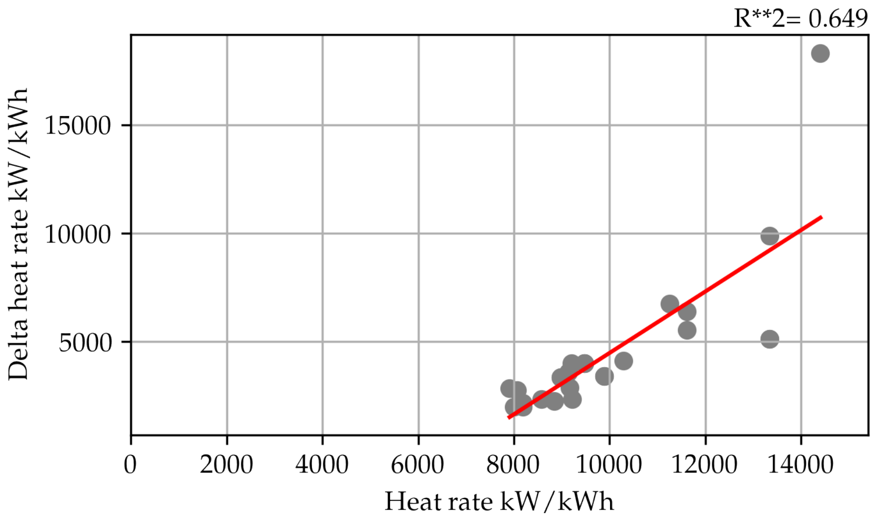


Figure 4.6: Heat rate gain regression

4.2 Strategies for the completion of efficiency data

The approaches of proxy efficiency were executed for fifty different random combinations of train-test pairs to measure how uncertain each approximation is, for comparison a random value between the maximum and minimum of each fuel efficiency is given. The average standard error is shown in table 4.3 and the plots of the different runs can be seen in figure 4.7. It can be observed that both regression approaches are similar in almost all three cases,but it is slightly better for natural gases. This behavior in natural gases is used to justify the usage of the Random Forest proxy method.

Fuel	Method	Standard error (Efficiency points)
Hard Coal	Random Forest	2.3%
	Linear Regression	2.2%
	Naive Approach	4.6%
	Random Values	3.7%
Lignite	Random Forest	1.5%
	Linear Regression	1.6%
	Naive Approach	4.1%
	Random Values	3.2%
Natural Gas	Random Forest	6.2%
	Linear Regression	7.2%
	Naive Approach	12.8%
	Random Values	9.3%

Table 4.3: Standard errors of .

4.2.1 Random forest regression

The dataset resulting from the random forest regression is shown in figures 4.8a, 4.9 4.8b. There the stars represent the real data and the dots represent the approximations of the algorithm, this data can now be used to match cost values in the next step.

The power plant efficiency along the temporal axis has a constant bottom behavior followed by a linear increase of different slopes for each fuel type with gas having the most pronounced slope, this slope could be probably related to the higher incidence of combined cycle power plants in later

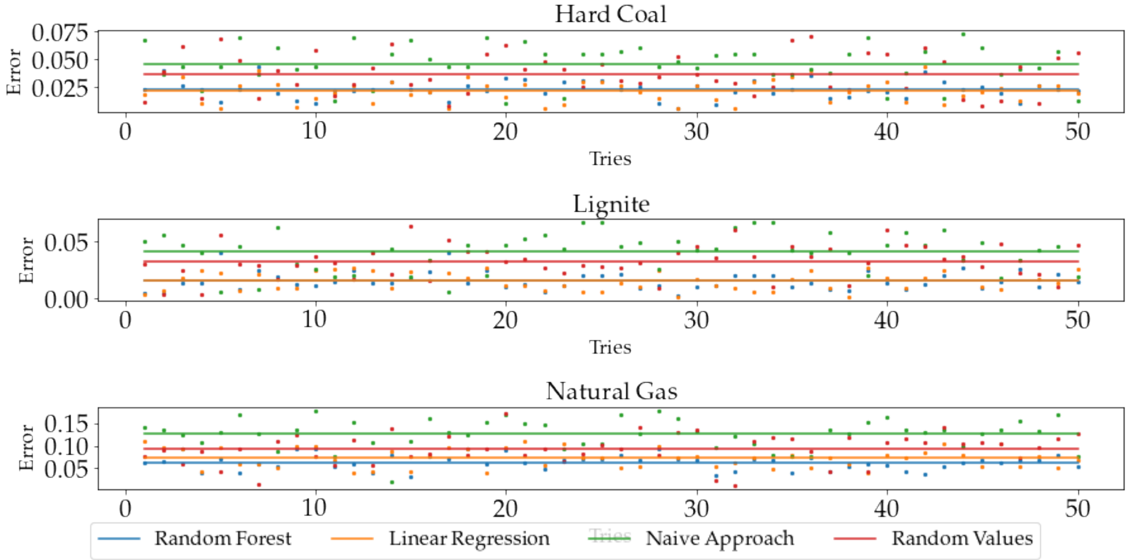
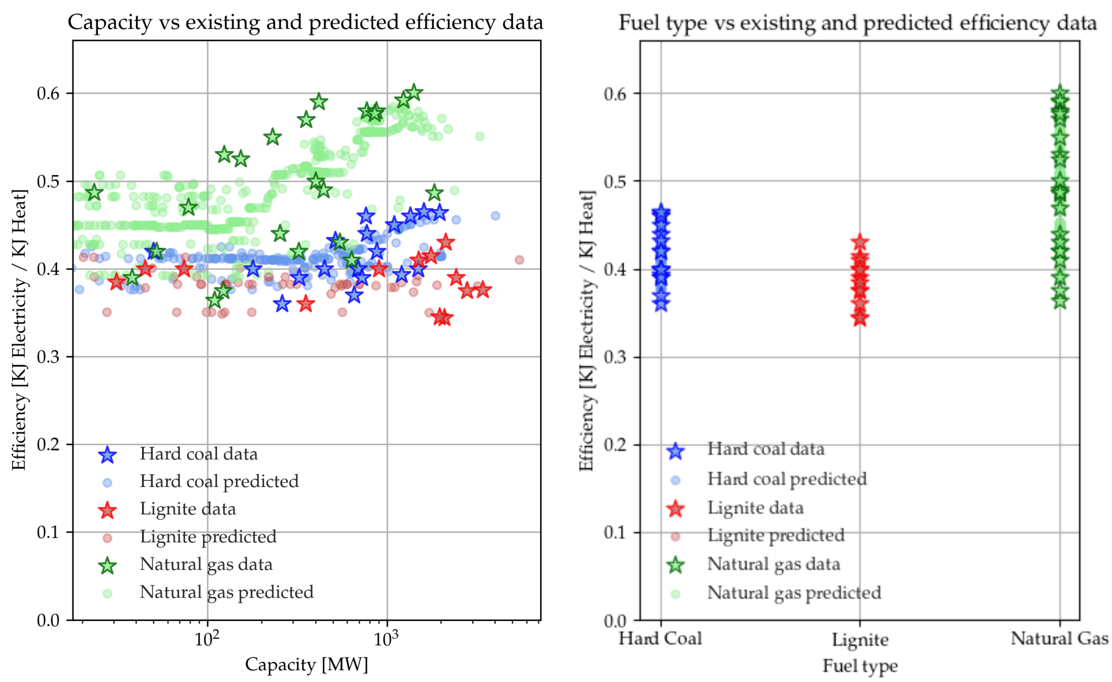


Figure 4.7: Distribution of error of the efficiency regression methods

years. From the capacity axis we see a interesting shape that out of experience from the author is assumed to be a logarithmic increase, this for natural gas and slightly for hard coal; lignite has no such behavior having an apparent 0 slope.

Looking at the fuel raw distribution we can see that natural gas has the higher spread, this is probably because the different parameters used in the regression have a stronger effect.



(a) Effect of power plant capacity, logarithmic notation

(b) Effect of fuel type

Figure 4.8: Distribution of reported and random forest predicted efficiency values.

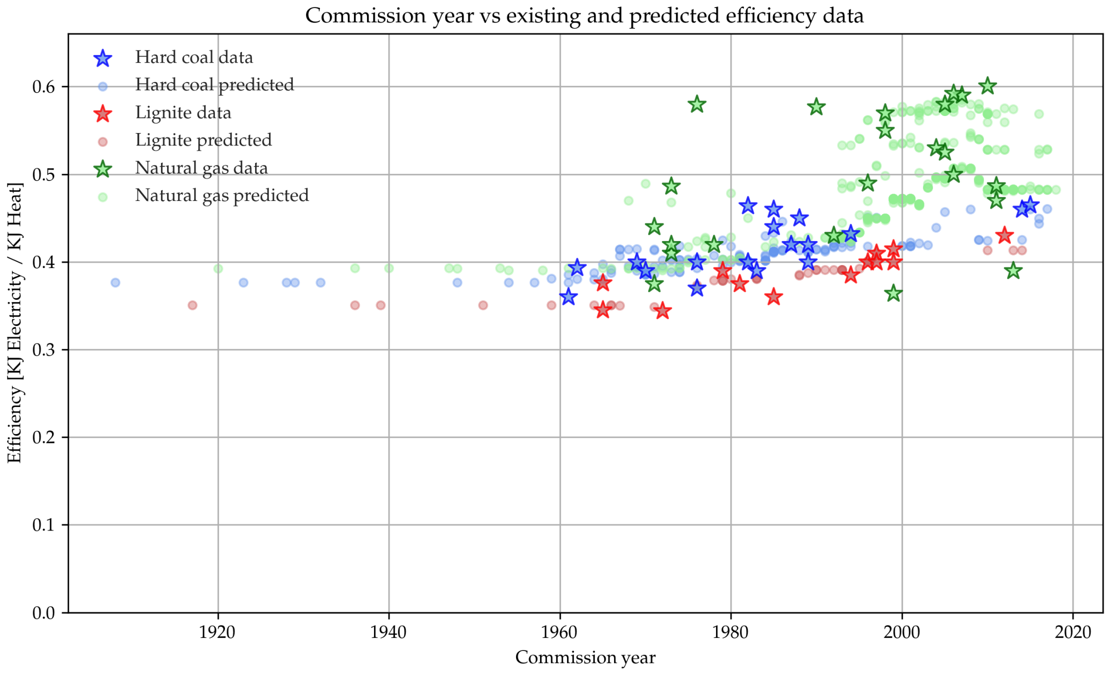


Figure 4.9: Distribution of reported and random forest predicted efficiency values, effect of commission year

4.3 Geographical distribution of potential and cost

To explore the geographical distribution a scattered Pie chart representation was chosen. The size of the pies is to give a general idea of the magnitude of the potential amount, it does not translate directly into actual predicted amounts; to do quantitative analysis of these use of the curves is recommended. The composition of the chart reflects the source of the captured CO_2. From this view the highest potentials can be attributed to Germany, Poland, Italy and Spain. In Germany and Poland, the largest source corresponds to Coal sources. In France, Spain and Italy industrial processes have a relatively higher share. Only in Finland the role of Bioenergy is significantly higher than the other sources. This overview of the European territory can be found in figure 4.10.

The finer details can be better appreciated looking at particular countries. To get a better view, in figures 4.11, 4.13 and 4.12 we can see the distribution for Germany, Italy and Netherlands. In figure 4.11 some locations like the German Rheinisches Revier among other high energy producing and industrial areas start to show up. One can also notice how Netherlands in figure 4.13 has low potential for industrial processes and the fossil sources are sparse but well spread across the country. Italy (4.12) a well distributed potential for all sources, Natural Gas sources are specially abundant.

The distribution of different sources can be viewed individually for comparison in figures 4.14 through 4.17. One can see that in general the natural gas sources are well spread across the country, and the areas surrounding the most important industrial areas have the highest potential (4.14), the cost is generally below 75 € per ton in most of these areas.

In 4.15 it can be seen that the potential for coal has higher concentration points, this is probably due to the higher specific emission an the fact that Lingite is obtained directly from the mining sites instead of being transported, the cost tends to be lower in the highest concentration points an higher in the points away from these areas.

Comparatively, industrial sources 4.16 show a moderate potential in comparison to both fossil sources with significantly lower cost than natural gas. The distribution is even but more sparse, costs are significantly lower as well. These sources have the most desirable pair of cost and distribution.

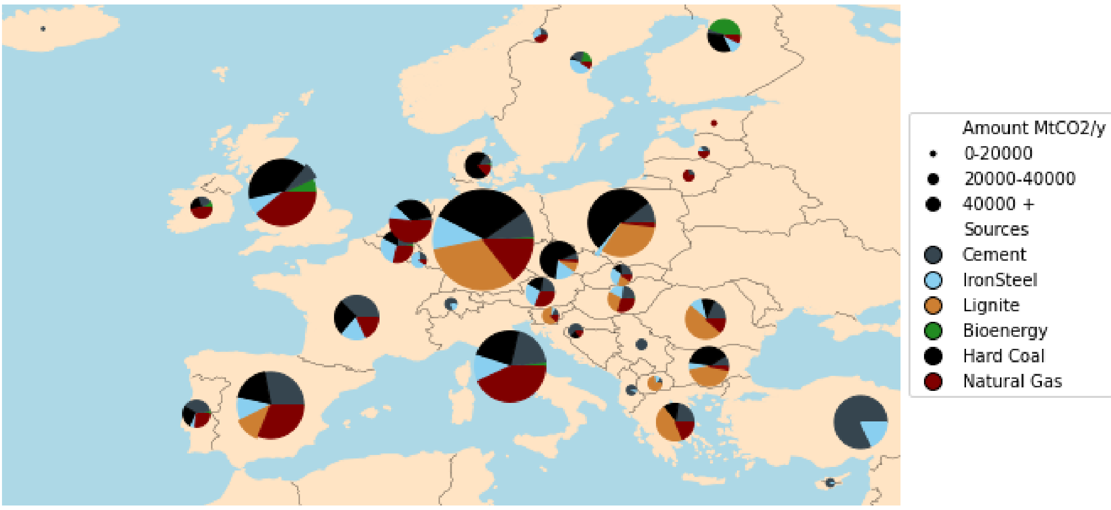


Figure 4.10: General view of the potential and cost of carbon capture.

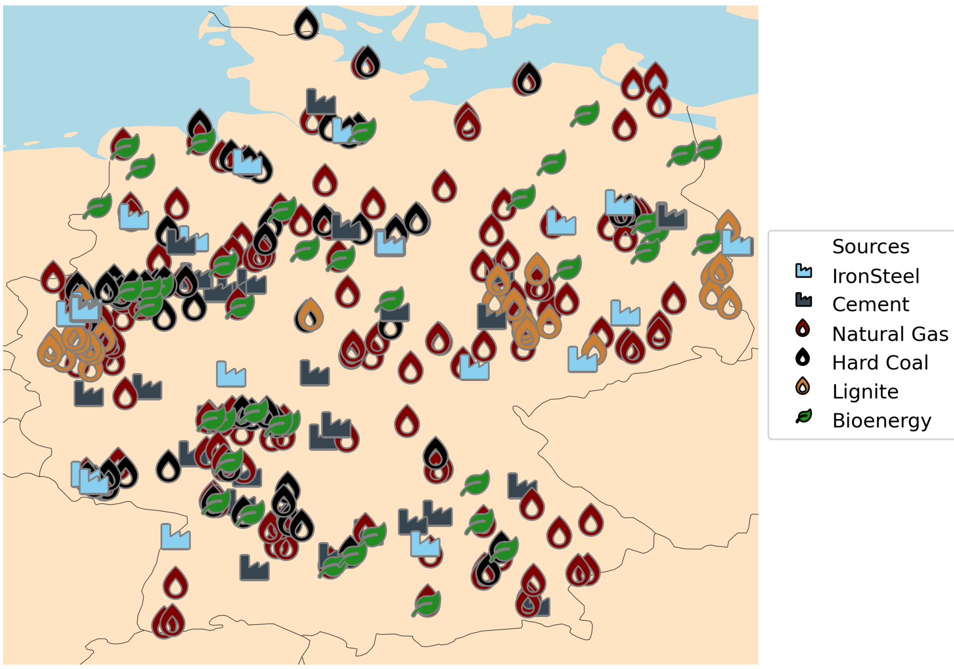


Figure 4.11: Germany view of the potential and cost of carbon capture, the values are €/tCO_2

Bioenergy 4.17 shows the lowest potential but also the lowest prices of all the energy sources, not surprisingly, there is reasons so much work is being done on improving BECCS. Not very strong conclusions can be drawn from the current work as the data sources are very sparse.

Figure 4.12: Italy view of the potential and cost of carbon capture, the values are €/tCO_2

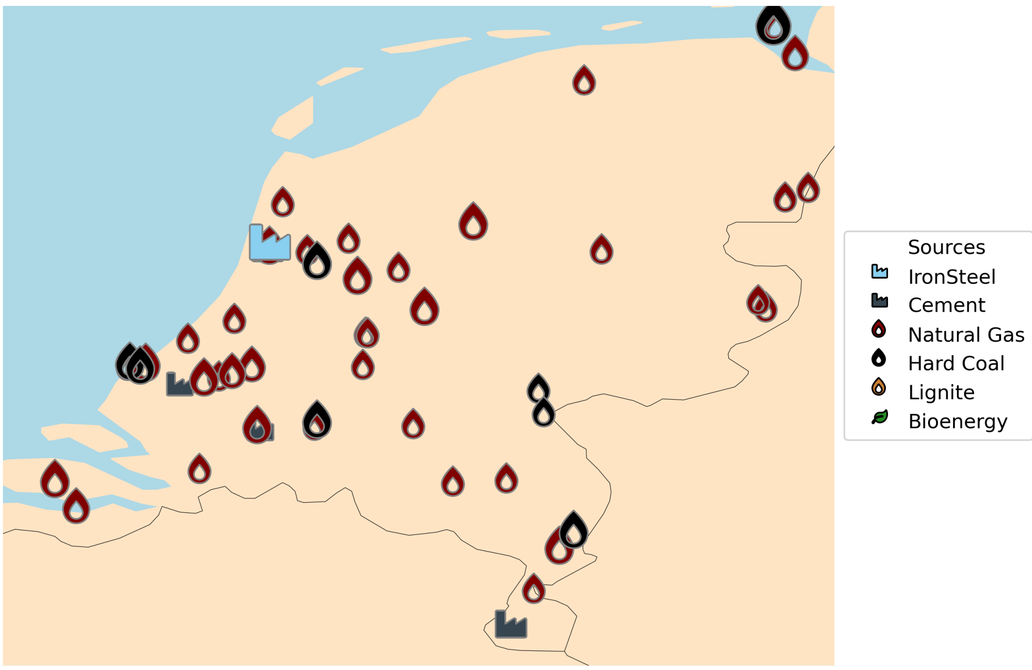


Figure 4.13: Netherlands view of the potential and cost of carbon capture, the values are €/tCO_2

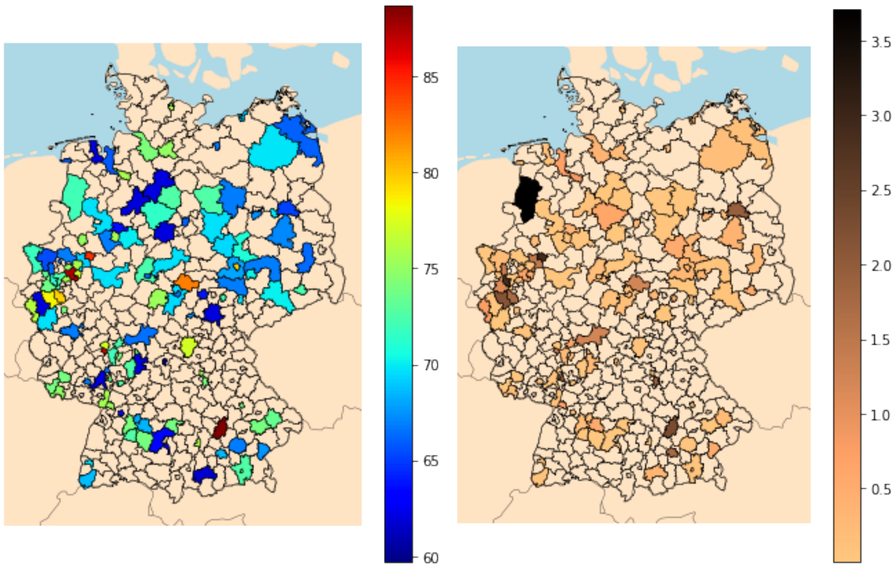


Figure 4.14: Left: German natural gas power plant cost distribution for captured CO2 in 2019€/t. Right: Potential distribution in Mt/y. Here the potential is evenly distributed across the territory

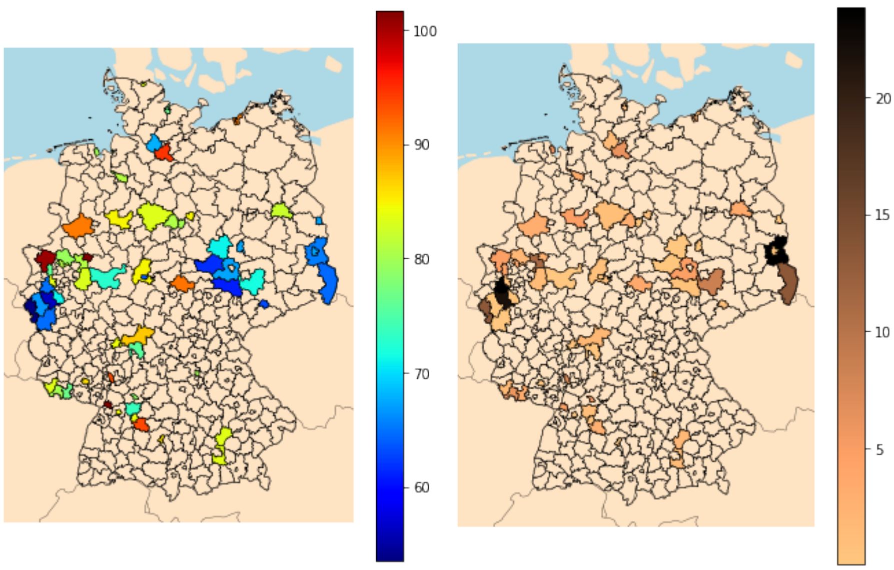


Figure 4.15: Left: German coal power plant cost distribution for captured CO2 in 2019€/t. Right: Potential distribution in Mt/y. In this case, there is a formation of focal points in major mining territories

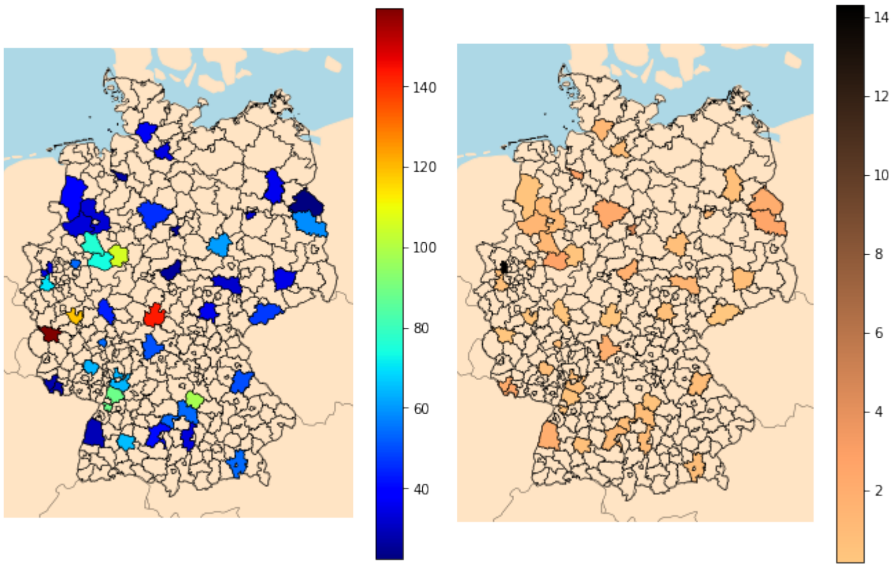


Figure 4.16: Left: German industrial process cost distribution for captured CO2 in 2019€/t. Right: Potential distribution in Mt/y. industry is sparser than natural gas plants yet has lower costs.

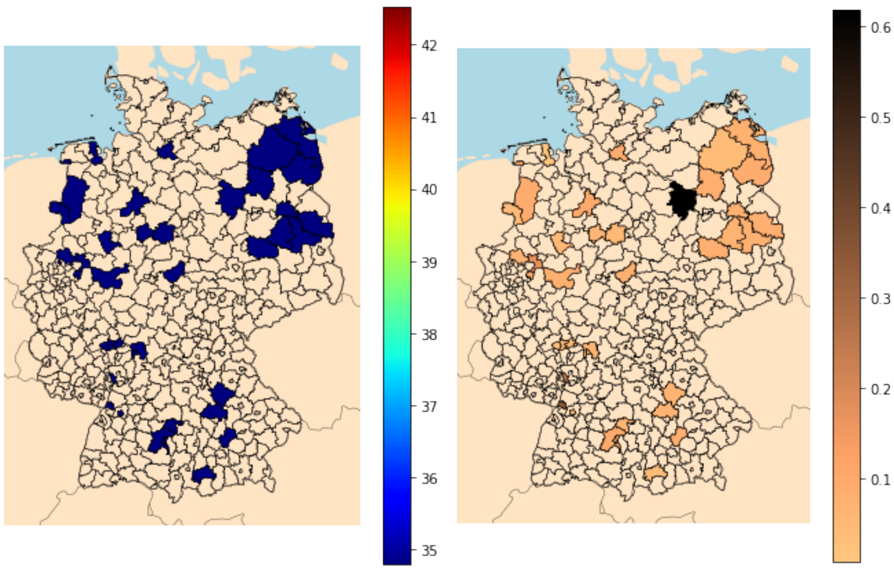


Figure 4.17: Left: German Bioenergy plant cost distribution for captured CO2 in 2019€/t. Right: Potential distribution in Mt/y. The data used for the production of these distributions does not allow strong conclusions

4.4 Cost Potential Curves

To be better able to appreciate the potential and cost distribution of a particular region, the cost curve visualization is more helpful. We can visualize in figure 4.18 the general distribution of potential and cost across Europe. And in figure 4.19 the values for example countries can be seen. It can be appreciated that the values are in the order of hundreds of megatons for the European countries, this with the current park of energy and industrial producers. The cost range spans from 15 € per ton of CO2 corresponding to the lower industrial potential costs up to 120 € per ton.

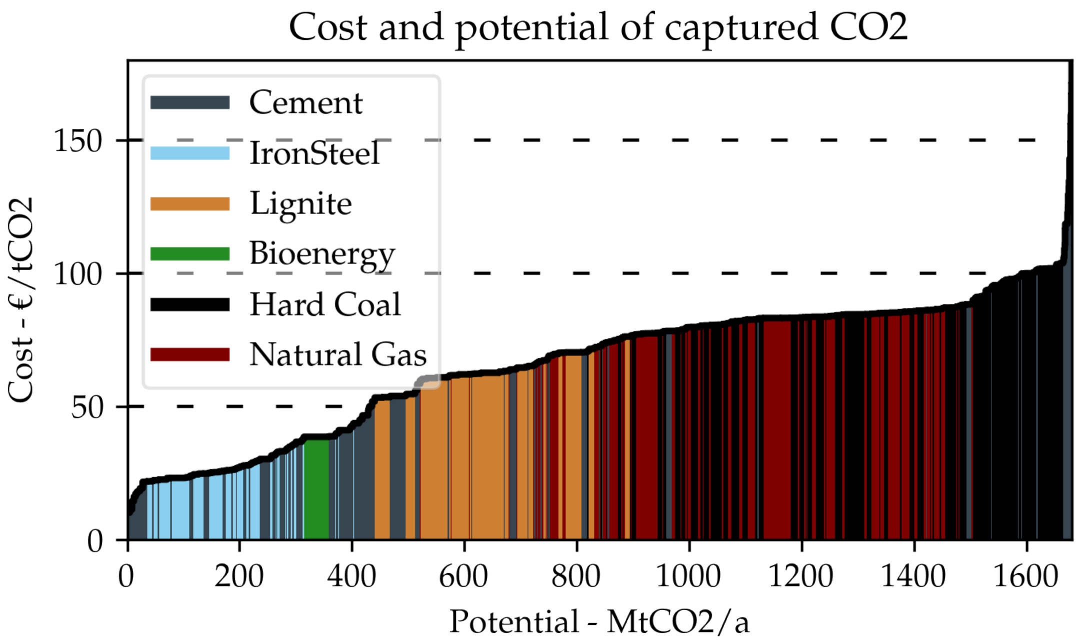


Figure 4.18: Cost potential curve for the countries in EU27, the cost is in €/tCO2 and the potential in Mt/y

Figure 4.20a shows the curve for Germany here a distribution of the different sources can be seen. The highest costs are found to be Natural Gas and Hard Coals, while the lowest costs are the industrial processes and bio energy sources. These last 2 have also the lowest potential, yet it is still significantly high, specially for industrial processes, this can be seen in figure 4.20b.

A summary of the potential and cost can be seen in table 4.4:

It should be noticed that the potential cost for capture of hard coal sources seems to be higher than gas.The reasoning behind this is that the average capital costs obtained from the estimation where reasonably low and the fuel cost relatively low for natural gas plants. The values used are

Source	Sum of potential Mt/y	Mean cost 2019€/t
Bioenergy	44	38.66
Cement	334	43.41
Hard Coal	482	84.29
IronSteel	189	38.89
Lignite	277	62.77
Natural Gas	353	74.20

Source	Sum of potential Mt/y	Mean cost 2019€/t
Bioenergy	3.7	38.66
Cement	30	45.23
Hard Coal	114	84.19
IronSteel	37	32.61
Lignite	113	63.43
Natural Gas	50	70.96

Table 4.4: Summary of cost potential; Left: Europe. Right: Germany.

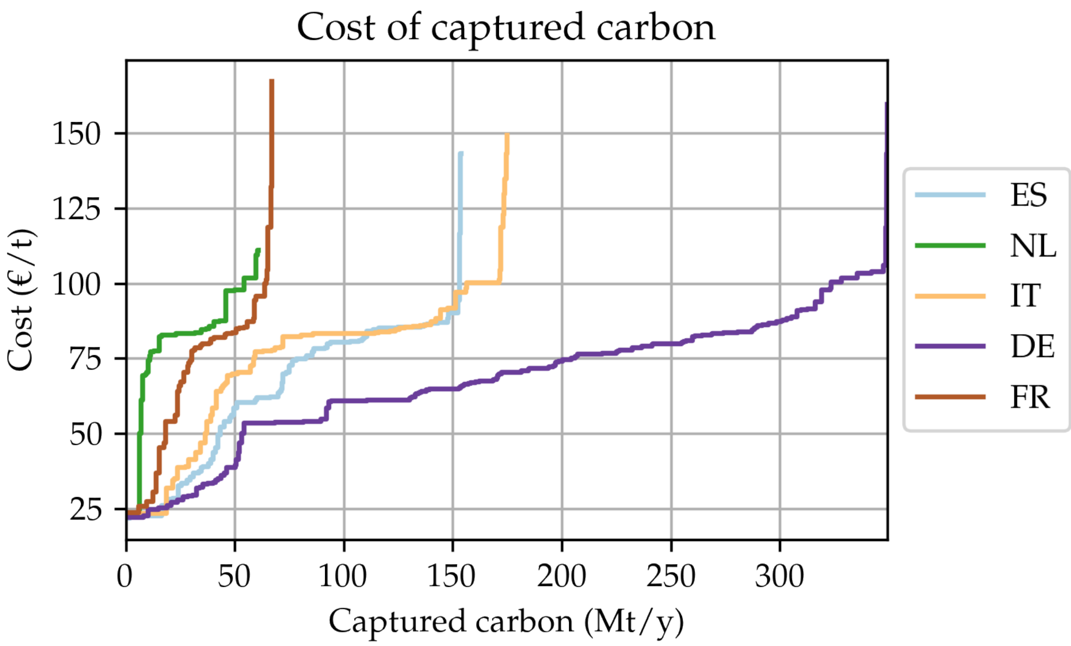


Figure 4.19: Cost potential curve for example countries, the cost is in €/tCO2 and the potential in Mt/y

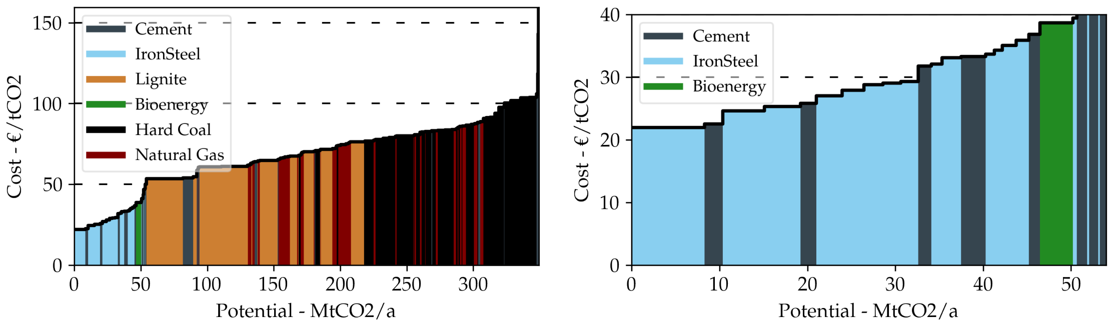


Figure 4.20: (a.) Cost potential curve for Germany for all the different sources. (b.) Cost potential curve for Germany for lower potential sources.

based on the North American Market, these might not hold true for Europe.

4.5 Scenario Development

Until now the visualizations provided are for the existing emission sources.Based on Keramidas K. [2021] an overview of the development of different scenarios is done. Figures 4.21 a and b show the cost potential curves of the 2050 Reference business as usual and 1.5C scenarios respectively. In both scenarios coal has stopped being a carbon source, the amounts from natural gas are around 4 times bigger in the reference scenario.

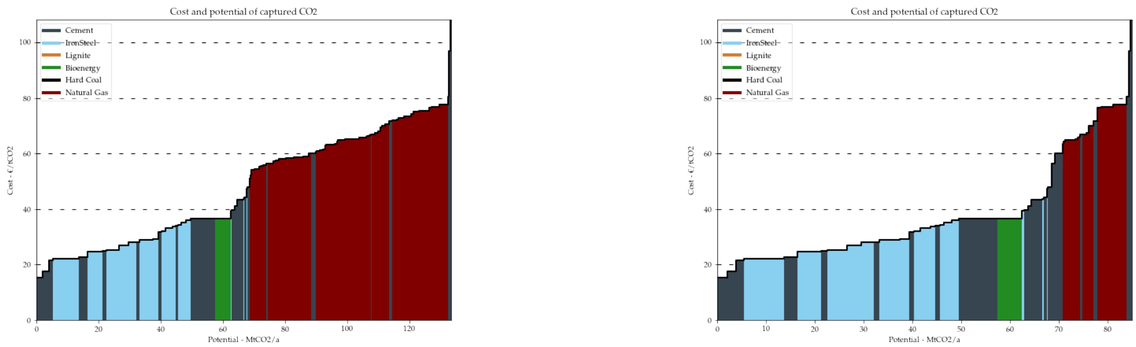


Figure 4.21: (a.)Germany reference scenario at 2050, Curve.(b.) Germany 1.5°C scenario at 2050, Curve

In the geographical scope, seen in figures 4.22 left and right. It can be observed that for Germany the geographical distribution is not strongly afflicted in the more ambitious decarbonization scenarios, the industrial hot spots stay almost the same. The cost distribution is equally unaffected. This can be attributed to the Kohleaustieg ([Bundesministerium der Justiz und für Verbraucherschutz, 2020]) .

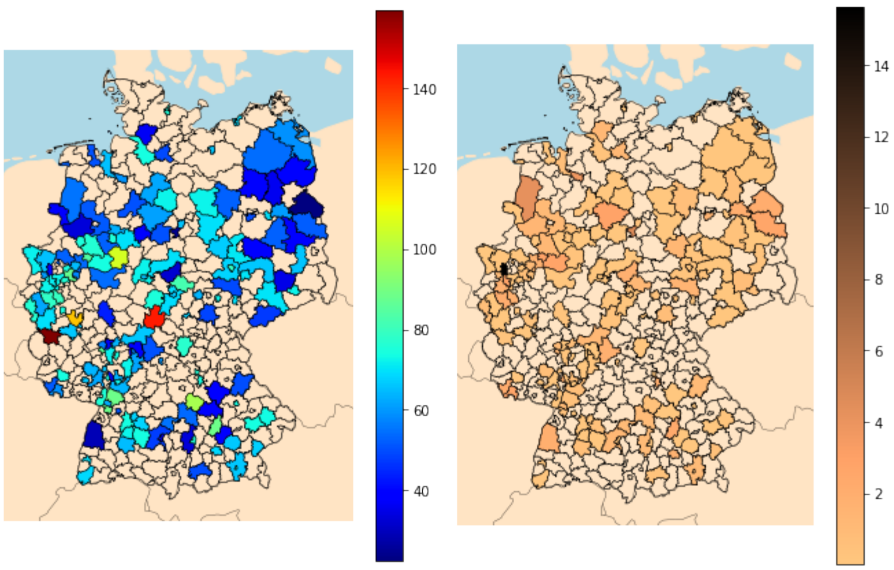


Figure 4.22: (left) Captured CO2 cost distribution of Germany at 2050 in the reference scenario, [2019€/t].(right) Captured CO2 potential distribution of Germany at 250 in the reference scenario, [Mt/y].

In the European context, the pie chart distribution is used to denote the effects on the potential. This can be seen in figures 4.24 and 4.25. In the 1.5C there is less potential and fossil fuels are only used in some cases; Italy and Netherlands have the largest share of Hard coal whereas Spain France and Germany of Natural Gas. In the reference scenario Natural Gas plays stronger role and Lignite is used significantly more in Easter European countries.

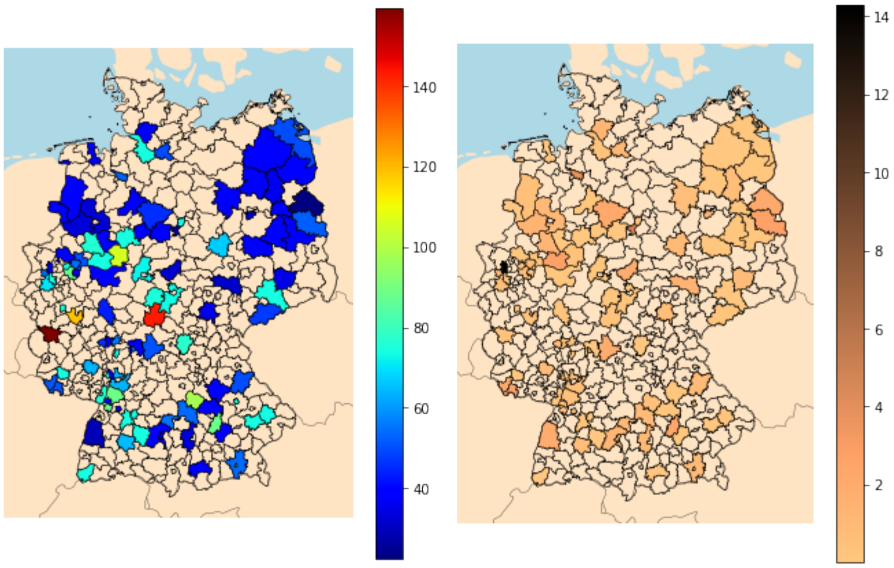


Figure 4.23: (left) Captured CO2 cost distribution of Germany at 2050 in the 1.5C scenario, [2019€/t].(right) Captured CO2 potential distribution of Germany at 250 in the 1.5C scenario, [Mt/y].

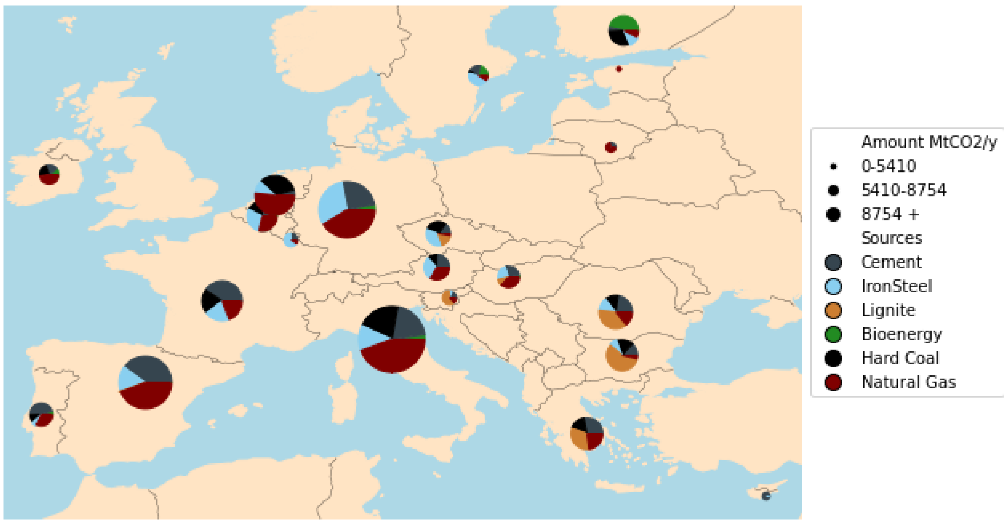


Figure 4.24: European potential reference scenario at 2050.

A more objective quantitative analysis of the scenario development can be done using the curves, figures 4.26a and 4.26a represent the curve development for the 27 countries in the European union and figures 4.27b and 4.27a just for Germany. It can be seen that in both pairs the lower potential component stays constant, this is because only the power plant park development is considered in the scenarios; this bottom part corresponds to the industrial processes. From looking at the European distribution, the constant part of t he curve corresponding to the industrial technologies starts around 300 Mt/y and has a max cost of 40 €/t. If we observe the case of Germany, the industrial sources point corresponds to an amount higher than 50 Mt/y with a max cost of 40 €/t. In the European context, the max amount at the reference scenario is 820 Mt/y while for the 1.5°C is less than 510 Mt/y; a significant difference yet both of the amounts are way higher than the potential demand of 25 Mt/y. Looking at the distribution of the German scenarios inferences on the role of Natural gas can be done; if the demand goes beyond 90 Mt/y there would be need of

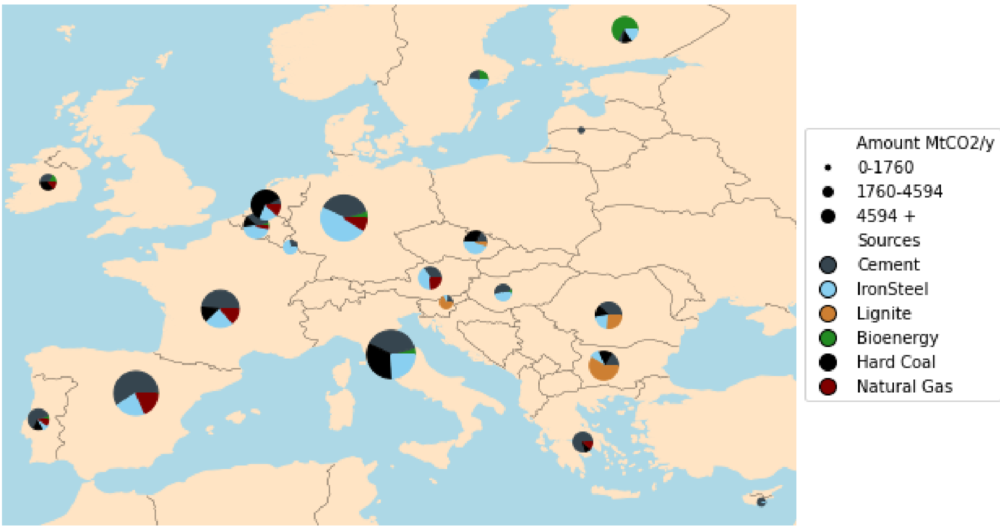
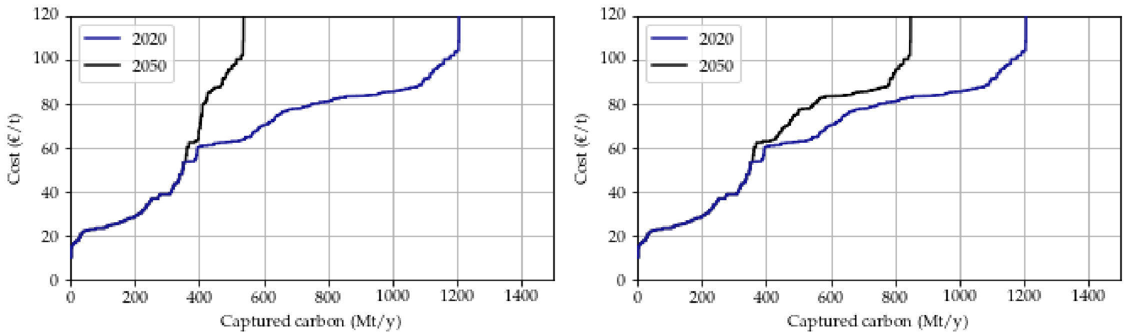


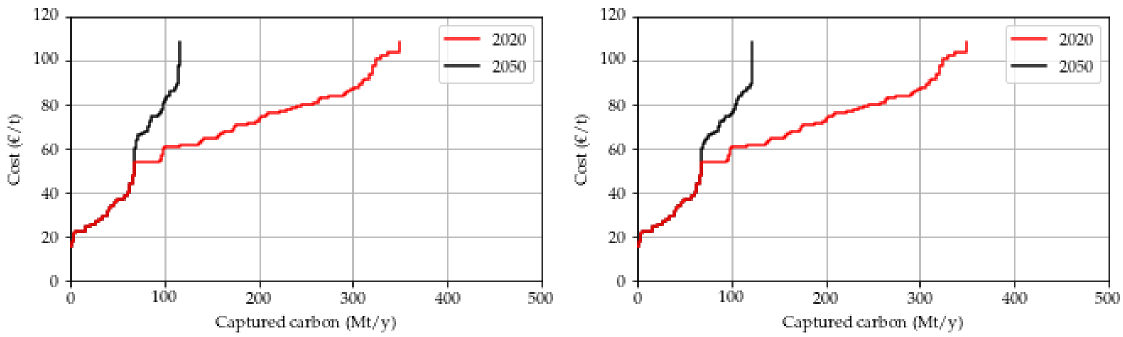
Figure 4.25: European potential 1.5°C scenario at 2050.



(a) 1.5°C scenario. (b) Business as usual scenario

Figure 4.26: Cost potential curve for the countries in EU27 at years 2020 and 2050

alternative sources in the 1.5°C scenario. Figure 4.21 shows this.

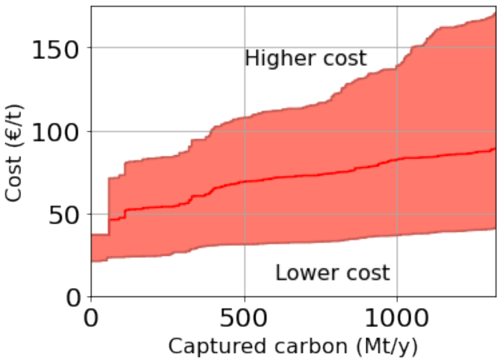


(a) 1.5°C scenario. (b) Business as usual scenario

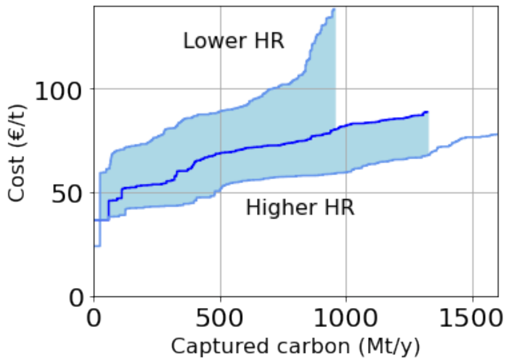
Figure 4.27: Cost potential curve for the countries in Germany at years 2020 and 2050

4.6 Sensitivity Analysis

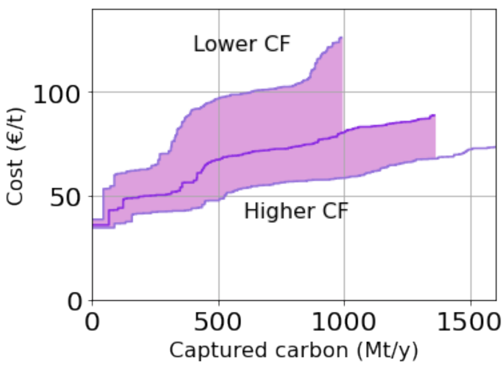
An analysis of the effect of the input values given to the curve production pipeline was done by executing the framework varying the input parameters independently. In the equation 3.15 all the values play a very specific role, and we divide in three types for the matter of analyzing the sensitivity of the framework. The first of them contains all the values that are kept constant based on scope definition; these are the fixed cost factor and the hours in a year. The second group consist on fuel and plant specific parameters which are bound to the input source they come from, these are the fuel cost, the emission factor and the heat rate; all of these values vary inherently in the estimation process. The last group and the one where this analysis is done to is the one associated to the assumptions obtained from doing the literature review, the variation of these values can be seen at the beginning of this section (4) and they consist of the investment costs and the efficiency loses; additionally despite being consistently more or less 90% across studies the effect of capture efficiency was addressed in different levels.



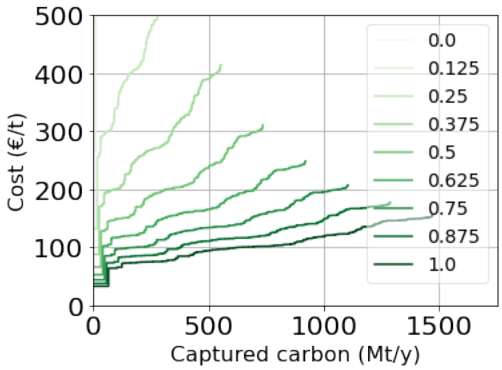
(a) Effect of investment cost. Gas: 431 - 1504 €/MW, Coals 1940 - 3430 €/MW.



(b) Effect of heat rate. 675 - 1421 for gas power plants. For coals, the errors from 4.2 were used.



(a) Effect of capacity factor. Lower values are 30% for all sources besides Bio which is 10%. Upper are 70% and 50%.



(b) Effect of capture efficiency. All the levels are technically feasible but at different investment costs, the sources used have a value of 90% as reference.

Figure 4.29: Effect of variables in the cost potential curves, values are in table 4.1

The way these numbers were varied based on the reference values their ranges can be found int table 4.1. For capacity factors, a grid based on the values of table 3.8 plus minus 0.2 is considered to account for possible operational changes. Capture efficiency will be explored as a grid from 100% capture rate to a value close enough to 0 to see the magnitude of the effect. All of this can be observed in figure 4.29.

The effect of investment has the strongest effect on the cost uncertainty putting the values below 500 Mt/y already above the 100 € mark for the upper limit, the spread increases with the captured potential.

Having a more favorable energy penalty causes an increment of cost, this can be easily explained by looking at the base equation as, although the energy penalty has some effect on fuel cost it has a greater effect on the produced amount, making the value of the denominator increment at a higher rate than the former; this effect on the amount can be seen in the lower part of the graph. This should not mean by any means that having worse efficiency is more economically viable, this is just a caveat of looking at CO_2 strictly as a commodity. If for example, a source of energy that is not the own power plant is used this graph could easily have an inverted shape, addressing this is beyond the scope of this work.

Capacity factor has a very similar effect as of the heat rate penalty, this is because it affects the emissions and energy production in relation to a fixed project investment.

Having a worse capture efficiency drastically increases the cost at the lowest efficiencies. The effects are small at the highest ones. In actual projects this may allow a small window for tuning the quality of the project with its cost.

4.6.1 Data input sensitivity

To demonstrate the robustness of the process and validate it. The utilization of an alternative data source is used to build the cost potential curves. This data is an internal DLR collection. It has a very fine granularity of the properties of the power plants, it also has its own method to approximate efficiency, this can be used to validate the proxy method used in this work.

Figure shows the potentials based on the power plant matching data set at the left and the internal one at the right.

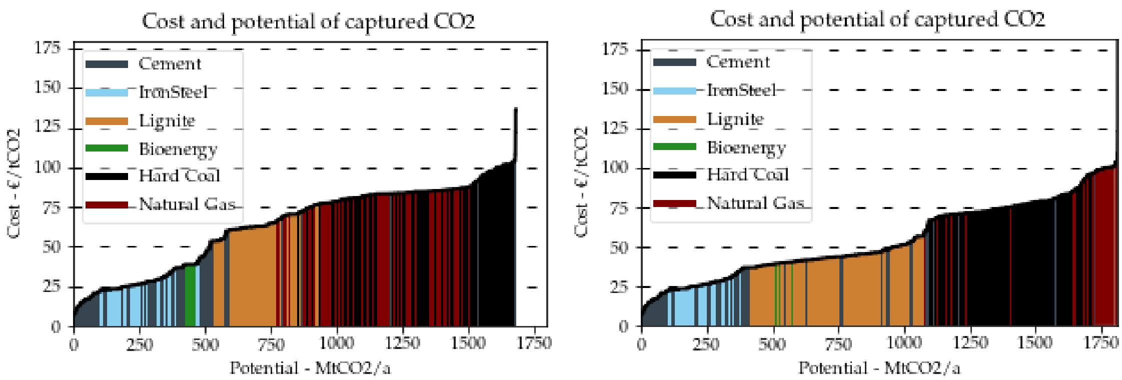


Figure 4.30: (left) PPM European cost potential Curve. (right) Internal dataset cost potential curve

In the power plant matching data there is less lignite plants than in the internal one, plants are in general less energy efficient for the internal data, this can be seen from the lower costs, this difference is specially relevant when looking at the effect on hard coal power plants where the cost shifts to higher relative values. The total amounts are higher for the power plant matching data. It was found out that the ground of this difference was the usage of German data as the only source of training data for the regression. This can be better appreciated by looking at figures **??** and **??** where it can be noted that the difference is less pronounced as in the European curve.

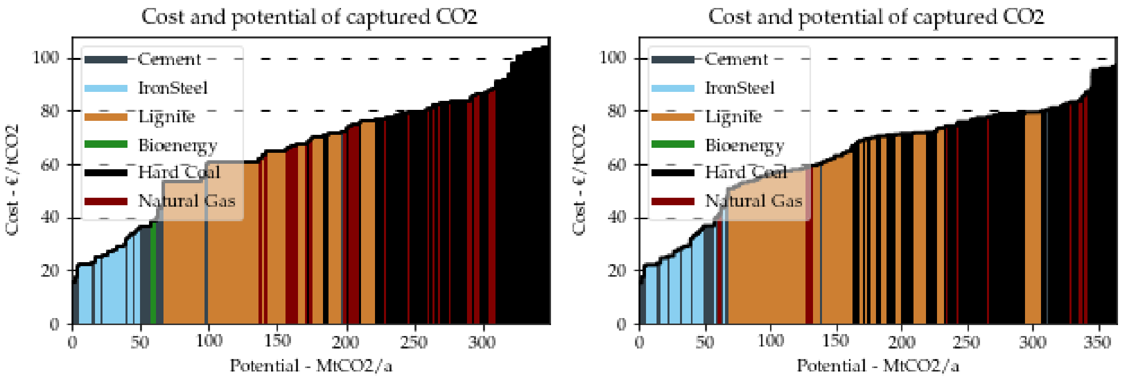


Figure 4.31: (left) PPM German cost potential Curve. (right) Internal dataset cost potential curve

	Germany		Europe	
Source	PPM	Internal	PPM	Internal
Hard Coal	0.41	0.39	0.41	0.40
Lignite	0.38	0.37	0.38	0.37
Natural Gas	0.46	0.50	0.48	0.51

Table 4.5: Average efficiencies of the different datasets

4.7 Emission validation

A frame of reference is needed to give the results a better grounded sense. For doing so a comparison with other data sources that refer to essentially the same entity is done. There is various projects that collect, standardize and report emission data obtained from annual reports. An example of such is the EDGAR [Crippa and Janssens-Maenhout, 2020]. The project PRIMAP from the Postdam Institute for climate impact research already does the work of compiling, cross-validating and reporting the data of this and many other projects [Gütschow et al., 2019]. PRIMAP reports their values using the IPCC emission control notation. The interes lays on the energy sector, the fields associated with it are the following:

- IPC1 Energy
- IPC1A Fuel Combustion Activities (We would have to do a substraction of the non desired fields)
- IPC1B Fugitive Emissions from Fuels (I think we an dismiss these as we are concerned only by the main emissions)
- IPC1B1 Solid Fuels
- IPC1B2 Oil and Natural Gas
- IPC1B3 Other Emissions from Energy Production
- IPC1C Carbon Dioxide Transport and Storage (currently no data available)

The industrial emission codes, represent the emissions of more industries than the ones that are considered in the scope of this thesis.

From the energy codes, the only related to our research are the Fuel combustion activities, because the capture would always happen at the Fluegas bus. The IPC1B codes represent Fugitive emissions, these cannot be captured so they are kept out of the scope. For the remaining values the dataset reports two scenarios:

- HISTCR: In this scenario country-reported data (CRF, BUR, UNFCCC) is prioritized over thirdparty data (CDIAC, FAO, Andrew, EDGAR, BP).
- HISTTP: In this scenario third-party data (CDIAC, FAO, Andrew, EDGAR, BP) is prioritized over country-reported data (CRF, BUR, UNFCCC)

We can visualize them in figure 4.32

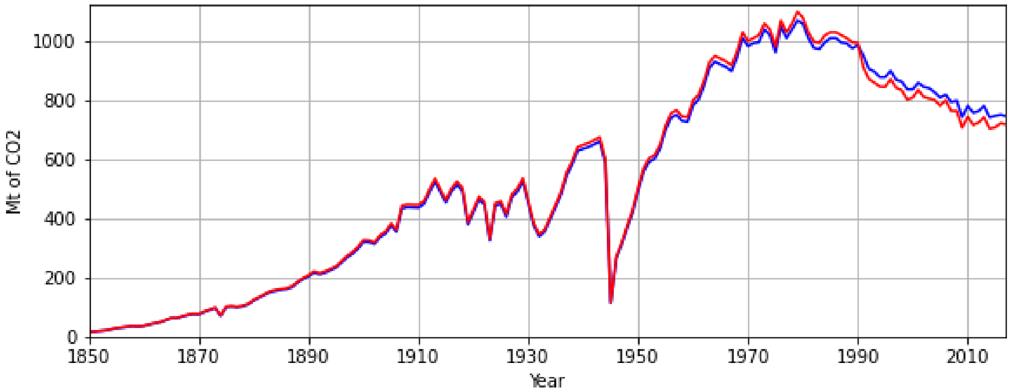


Figure 4.32: CO2 from combustion activities across the last century. HISTCR is red and HISTTP is blue.

The scenarios are pretty consistent with each other but their spans are to wide, for comparison it is smart only to take the most recent years as reference. We cut down the data to include only the years after 2010, this is seen in figure 4.33

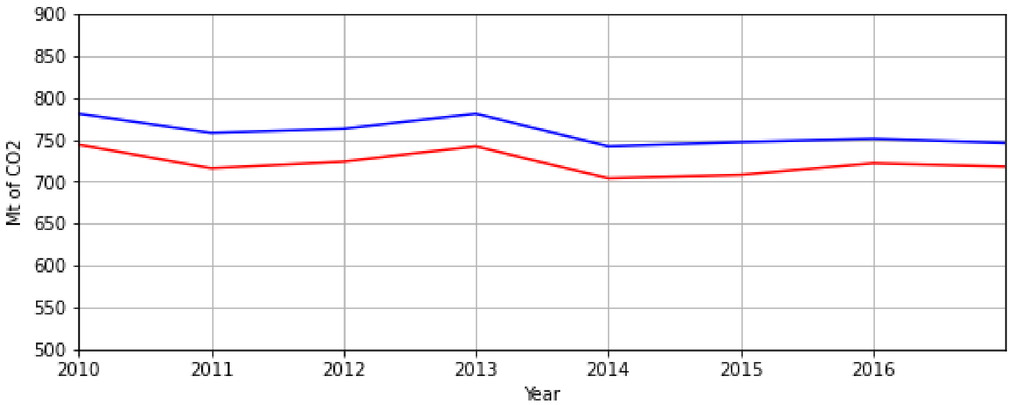


Figure 4.33: CO2 from combustion activities for years after 2010. HISTCR is red and HISTTP is blue.

The values here are more consistent wth each other, by averaging them we roughly measure the working value for CO2 emissions is 740 MtCO2 / y. Our measured captured amounts for Germany fossil fuel mix is 474 Mt/y. The difference is high but still within the same order of magnitude, this hints that they at least hold some connection. The differences come from various factors in the associated to the assumptions and calculations. First of all, the reported value in the distributions is captured potential, not actual emissions, so by default given the fact that we are considering at maximum 90% efficiency of the capture process, the amounts will be lower. Additionally, the capacity factors assumed for the calculations are probably very different from the actual activity of the power plants through the years in which these values were reported. There is also the effect of the emission factor of the fuel that was used in the power plant, every fuel is different in their carbon content and in this study the value is very generalized for each type of fuel. Some of the fuels reported in Power plant matching were excluded due to scope, these are all kind of oils and cokes, in the PRIMAP emission report they could be included in the total sum. There is also the factor of reliability of the input data, there is no assurance that all the power plants that were considered in PRIMAP were considered in the PPM dataset, or any other data source; this is a factor that can't be completely measured because the input source that we use is based only in openly available data whle the energy sector has a diverse range of private and public administrations whose data may not be always available, yet they may report to the GHG monitoring institutions on which the emission values are based on.

One last thing that can be a very important factor in adding uncertainty to these amounts is the carbon capture module configuration. In this case we are assuming one in which the energy for the process is obtained from the same power plant, this plays a huge role because the required energy is "Added" to the retrofitted plant, which causes higher "specific emissions" before capture. The different values of carbon dioxide as well as the representation of potential sources of irregularity can be seen in figure 4.34

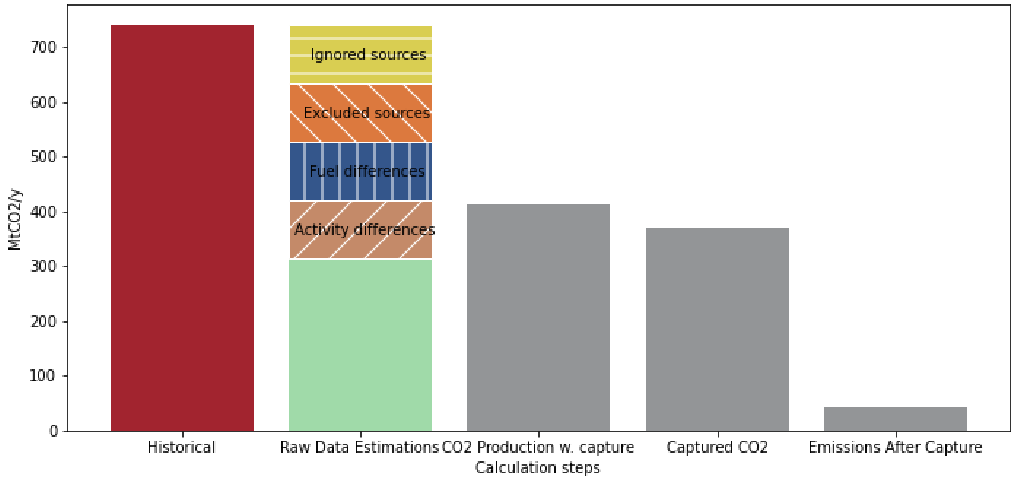


Figure 4.34: CO2 amounts based on the different points of the calculation reference and the values reported by PRIMAP, the stripped squares do not represent actual values.

Discussion

With the curves produced in the last step the questions set at the beginning can be answered, for doing so we will keep our scope on Germany. To compare the different sources of carbon a common ground has to be set. This will be done by listing characteristics that are preferred on a CO_2 source. First spatial availability is listed, ERE profiles are out of the scope of this study, with that as an incognita this criteria is valued on the basis on how well distributed is the data across the territory,three levels are established: concentrated for sources that tend to bundle in small geographical locations, distributed for sources that span across a wide territory and sparse for sources that are distributed across large territories with small quantities. The second criteria is the temporal availability, this will be assessed in a binary way on whenever the presence at 2050 in the reference and 1.5°C scenarios is still above 25 Mt/y. The third criteria is cost, which will take three levels shown in the map legends; the range of 0 to 20 will be labeled as best, 21 to 40 will be good and 41-60 as acceptable. Table 5.1 shows the tabulation of these criteria for each source

Source	**Criteria**	**Level**
Natural Gas	Spatial	Distributed
	Temporal	Yes
	Price	Acceptable
Hard Coal	Spatial	Concentrated
	Temporal	No
	Price	Acceptable
Lignite	Spatial	Concentrated
	Temporal	No
	Price	Acceptable
Bioenergy	Spatial	Distributed
	Temporal	Yes
	Price	Good
Cement	Spatial	Distributed
	Temporal	Yes
	Price	Best
Iron and Steel	Spatial	Sparse
	Temporal	Yes
	Price	Best

Table 5.1: Assessment criteria tabulation

Criteria are given values to perform an objective evaluation.Spatial availability, the most desirable distribution is "Distributed" so it is assigned the highest value, a 2, the second one is Concentrated which is assigned a value of 1 and the least desirable is sparse, assigned a 0. Temporal availability in this case is binary so it is assigned a 2 when it 2050 is reached significantly and a 0 when not. As for the price, there are 3 levels, and will be assigned the number 0 to 2 accordingly. Performing a summation of the assigned values yields the scores seen in figure 5.2.

Using this simplified criteria outputs a pattern that was found earlier in Naims [2016]. Cement sits as the best option, Bioenergy as a very viable one, finding consistent with what was found in (Möllersten et al. [2003]), but given the sparse data that was used in this work, these are only initial indicators of the potential. Natural gas power appears as a third close contender, its viability is mostly related to the fact that it won't be dropped as easily from the energy market as the coals.Iron and steel are relatively strong, one problem is that there is decarbonisation pathways for

Source	**Points**
Natural Gas	4
Hard Coal	1
Lignite	1
Bioenergy	5
Cement	6
Iron and Steel	4

Table 5.2: Assessment scores, Bioenergy values are indicative.

the steel process which could make them even less viable as carbon sources [Toktarova et al., 2020] The coals have weaknesses, such as being constrained to small territories and being expensive to capture.

Regardless of the scenario, at least for Germany, the production of 25 MtCo2/y can be achieved from a technical potential perspective and it does not represent a bottleneck in the bigger scheme of a low relative GHG emissions CO_2 supply chain. The only differences among scenarios is the role of fossil fuels at higher demands; if these demands are going to exist or not is beyond the scope of this work, Naims [2016] estimates a global demand of up to 10000 Mt/y at a long term but the share of German demand can't be estimated from this number.

The question of the economical viability of fossil fuels, can be answered using the information generated in this study. Coal based energy production, despite having high usages in the current year 2021, is going to be dropped from the energy mix, so the only way in which CCS retrofitting can be marginally viable is by building the retrofits now so the projects have at least 17 years to payoff, each subsequent year will make these projects less viable. Natural gas power has better potential, but it is among all the other fuels one of the most expensive. However, cost estimates for retrofitted coal power plants lie in the range of 70-100 €/t, almost double the amount of the other alternatives in this study. Since the cost of consecutive end-user products such as synthetic fuels currently has to compete in a free market with fossil alternatives, this would require stark increases in the CO2 price that seem currently implausible.

From the sensitivity analysis it was found out that the parameters related to the activity of the processes have a strong influence in the potential and the cost that can shift the viability of implementing carbon capture in them. This uncertainty can hardly be addressed in a generalized manner as each project has different particularities related to their location, age and necessities of their surrounding industrial environment. Another important limitation in this work is the scope of industrial processes, there is opportunity to address the role of industries like Ammonia and Ethanol production [Naims, 2016]. Some other scope limitations is also the exclusion of the role of the power market to address the profitability of a retrofit project in a higher economic scope and although this work tries to step into the economic modeling field by employing indexes and basic costing functions there is a big window of opportunity on considering factors like grid stability and reserve capacities

There is a wide selection of paths on which this study can be expanded upon, if the interest for doing so arises. The role of Combined Heat and Power provision can be of importance given the strong dependency of fossil fuels they have in the European context. There is also a rich variety of potential business models that can stir the cost in a significant way, configurations with secondary more modern energy sources for example.

On the development side, the code is highly modular, quality that can be used for continuous improvement. This would allow the inclusion of higher resolution in the cost composition like separating the fixed and variable operational costs or adding including new technologies like the ones mentioned earlier. The architecture and efficiency have however a lot of room of improvement but it is flexible enough to have further means of cross validation by including new data sources.

Conclusions

For the production of this work a wide variety of activities had to be performed. An exhaustive literature review of carbon capture technologies, usage of captured carbon and methodologies for the measurement of the cost of carbon capture. A wide exploration for possible data sets and index development sources was done, this was hand to hand with the need of data cleaning and analysis.In order to make the results reproducible, the methods applied in the thesis are transparently published online in an open source repository.

The identified strengths of this framework is flexibility as, although it is not directly used to draw conclusions, more than one data source is used to produce the outputs with very good success, the modularity allows to modify the steps if one has a different scope or point of reference and the plotting is done in a straight forward and comprehensible way.

There is a lot of opportunity of improvement of the framework. The possibility of integrating Industrial source development scenarios was explored and ruled viable, it was kept out of the scope of this work because of time constraints. Also the inclusion of more industrial sources and a better granularity of the bioenergy sector can be achieved with available information. As a note, the bioenergy sector complexity may motivate a separate framework, or a variation of the one used in this work dedicated to BECCS.

Despite the qualitative nature of the criteria to answer the questions stated at the beginning, the answers are in line with those found for other scopes in the literature reviewed such as Fröhlich [2019] and Naims [2016]. One of the major conclusions is that, fossil fuels are far from a viable option for being a captured CO2 source given the fact that the demand levels stated in this work for Germany can be met by other sources. This is not the end of the story for the usage of coal as a source of energy, the Integrated Gasification Combined Cycle was kept out of the scope because it is a relatively new technology, there has to be analysis in that regard to completely rule out the clean potential of coals.

The bottom line from the author is that the conclusions drawn from this work are based on static data, but the energy system is evolving relatively fast in the current times (2021). This means that the viability can't be completely assure from the data here alone, it has to be analyzed in a more complex system of dimensions where factors like market, environmental impact, energy demand are considered. This was within the objectives from the start of this work, the output here can be used along other tools to draw more concise conclusions and it is what is expected from whoever uses the information and data here provided.

The software framework produced during the writing of this thesis is openly published in the DLR VE repository.

Appendices

Appendix A: Input Data

Table 1: Raw tabulated values of literature

Study	Territory	Technology	CaptureTech	Fuel	fuel_general	P_GROSS	P_NET	P_NET_CC	P_AUX	P_MAX
Bergerson Lave a 2007	Canada	SCPC	MEA	illinois_6	coal	nan	nan	nan	nan	480.0
Bergerson Lave a 2007	Canada	IGCC	SELEXOL	illinois_6	coal	nan	nan	nan	nan	480.0
Bergerson Lave a 2007	Canada	SCPC	MEA	illinois_6	coal	nan	nan	nan	nan	480.0
Bergerson Lave a 2007	Canada	IGCC	SELEXOL	illinois_6	coal	nan	nan	nan	nan	480.0
Bergerson Lave b 2007	Canada	SCPC	MEA	Lignite	coal	nan	nan	nan	nan	nan
Bergerson Lave b 2007	Canada	IGCC	SELEXOL	Bituminous	coal	nan	nan	nan	nan	nan
Bergerson Lave b 2007	Canada	NGCC	MEA	natural_gas	natural_gas	nan	nan	nan	nan	nan
NREL 2012	Europe	NGCC	MEA	natural_gas	natural_gas	nan	615.0	nan	nan	nan
NREL 2012	Europe	SCPC	MEA	Coal	coal	606.0	606.0	nan	nan	nan
NREL 2012	Europe	IGCC	SELEXOL	Coal	coal	590.0	590.0	nan	nan	nan
NREL 2012	Europe	SCPC	MEA	Coal	coal	606.0	606.0	nan	nan	nan
CAESAR 2011	Europe	SCPC	MEA	Bituminous	coal	819.0	754.0	nan	65.0	nan
CAESAR 2011	Europe	IGCC	SELEXOL	Bituminous	coal	441.0	391.0	nan	nan	nan
CAESAR 2011	Europe	NGCC	MEA	natural_gas	natural_gas	nan	829.0	nan	nan	nan
GCCS 2011	Australia	SCPC	MEA	illinois_6	coal	580.0	550.0	nan	30.0	nan
GCCS 2011	Australia	SCPC	MEA	illinois_6	coal	580.0	550.0	nan	30.0	nan
GCCS 2011	Australia	UCPC	MEA	illinois_6	coal	576.0	550.0	nan	26.6	nan
GCCS 2011	Australia	IGCC	SELEXOL	illinois_6	coal	748.0	366.0	nan	112.0	nan
GCCS 2011	Australia	NGCC	MEA	natural_gas	natural_gas	570.0	560.0	nan	10.0	nan
IEAGHG 2014	Europe	SCPC	MEA	Bituminous	coal	nan	1030.0	nan	nan	nan
IEAGHG 2014	Europe	IGCC	SELEXOL	Bituminous	coal	nan	804.0	nan	nan	nan
Rubin 2007	USA	SCPC	MEA	illinois_6	coal	575.0	528.0	nan	nan	nan
Rubin 2007	USA	IGCC	SELEXOL	illinois_6	coal	615.0	538.0	nan	nan	nan
Rubin 2007	USA	NGCC	MEA	natural_gas	natural_gas	517.0	507.0	nan	nan	nan
NETL 2015	USA	SCPC	MEA	illinois_6	coal	581.0	550.0	nan	31.0	nan
NETL 2015	USA	SUBC	MEA	illinois_6	coal	580.0	550.0	nan	30.0	nan
NETL 2015	USA	NGCC	MEA	natural_gas	natural_gas	641.0	630.0	nan	11.0	nan
Davison 2007	Europe	SCPC	MEA	Bituminous	coal	nan	758.0	666.0	nan	nan
Davison 2007	Europe	SCPC	MHI	Bituminous	coal	nan	758.0	676.0	nan	nan
Davison 2007	Europe	SCPC	OXYFUEL	Bituminous	coal	nan	758.0	532.0	nan	nan
Davison 2007	Europe	IGCC	SELEXOL	Bituminous	coal	nan	776.0	676.0	nan	nan
Davison 2007	Europe	NGCC	MEA	natural_gas	natural_gas	nan	776.0	662.0	nan	nan
Davison 2007	Europe	NGCC	MHI	natural_gas	natural_gas	nan	776.0	692.0	nan	nan
Davison 2007	Europe	NGCC	OXYFUEL	natural_gas	natural_gas	nan	776.0	440.0	nan	nan
Dave from Hu 2011	China	SCPC	MEA	China	coal	600.0	570.0	412.0	nan	nan
Li 2011	China	NGCC	MEA	natural_gas	natural_gas	nan	nan	nan	nan	nan
Li 2011	China	SCPC	MEA	Bituminous	coal	nan	nan	nan	nan	nan
Li 2011	China	SUBC	MEA	Bituminous	coal	nan	nan	nan	nan	nan
Wu 2011	China	SCPC	MEA	China	coal	nan	600.0	nan	nan	nan
Wu 2011	China	IGCC	SELEXOL	China	coal	nan	600.0	nan	nan	nan
Renner 2014	Europe	SCPC	MEA	illinois_6	coal	nan	nan	nan	nan	400.0
Renner 2014	Europe	NGCC	MEA	natural_gas	natural_gas	nan	nan	nan	nan	400.0
Renner 2014	China	SCPC	MEA	illinois_6	coal	nan	nan	nan	nan	400.0
Renner 2014	China	NGCC	MEA	natural_gas	natural_gas	nan	nan	nan	nan	400.0
CO2CRC 2017	Australia	SCPC	MEA	victorian_brown	coal	nan	500.0	nan	nan	360.0
CO2CRC 2017	Australia	SCPC	MEA	victorian_brown	coal	nan	500.0	nan	nan	360.0
CO2CRC 2017	Australia	SCPC	MEA	victorian_brown	coal	nan	500.0	nan	nan	360.0
CO2CRC 2017	Australia	SCPC	MEA	black_coal	coal	nan	450.0	nan	nan	333.0
CO2CRC 2017	Australia	SCPC	MEA	black_coal	coal	nan	450.0	nan	nan	333.0
CO2CRC 2017	Australia	SCPC	MEA	black_coal	coal	nan	450.0	nan	nan	333.0
Dillon 2013	USA	CFB	MEA	Petcoke	coal	nan	129.0	nan	nan	nan
Dillon 2013	USA	SUBC	MEA	Bituminous	coal	nan	616.0	nan	nan	nan
Dillon 2013	USA	SUBC	MEA	subbituminous	coal	nan	1500.0	nan	nan	nan
Dillon 2013	USA	SUBC	MEA	Lignite	coal	nan	1010.0	nan	nan	nan
Dillon 2013	USA	SUBC	MEA	Bituminous	coal	nan	1800.0	nan	nan	nan

Table 2: Raw tabulated values of literature, continuation

Study	P_MIN	Repower	Retrofit	CF	ELEC_EFF	HR	HR_UNITS	EF_FUEL	EF_FUEL_UNITS	EF_PLANT
Bergerson Lave a 2007	450.0	False	False	0.75	0.401	nan	nan	nan	nan	1.8
Bergerson Lave a 2007	450.0	False	False	0.75	0.344	nan	nan	nan	nan	1.9
Bergerson Lave a 2007	450.0	False	True	0.75	0.401	nan	nan	nan	nan	1.8
Bergerson Lave a 2007	450.0	False	True	0.75	0.344	nan	nan	nan	nan	1.9
Bergerson Lave b 2007	nan	False	False	0.75	0.42	nan	nan	nan	nan	0.825
Bergerson Lave b 2007	nan	False	False	0.75	0.34	nan	nan	nan	nan	0.9
Bergerson Lave b 2007	nan	False	False	0.75	0.5	nan	nan	nan	nan	0.438
NREL 2012	nan	False	False	0.85	nan	6705.0	BTU_KWH	117.0	lb_MMBTU	nan
NREL 2012	nan	False	False	0.85	nan	9370.0	BTU_KWH	215.0	lb_MMBTU	nan
NREL 2012	nan	False	False	0.85	nan	9030.0	BTU_KWH	215.0	lb_MMBTU	nan
NREL 2012	nan	False	True	0.85	nan	9370.0	BTU_KWH	215.0	lb_MMBTU	nan
CAESAR 2011	nan	False	False	0.85	0.455	nan	nan	nan	nan	763.0
CAESAR 2011	nan	False	False	0.85	0.469	nan	nan	nan	nan	734.0
CAESAR 2011	nan	False	False	0.85	0.583	nan	nan	nan	nan	351.8
GCCS 2011	nan	True	False	1.0	0.39	9.2	GJ_MWh	nan	nan	804.0
GCCS 2011	nan	True	False	1.0	0.39	9.14	GJ_MWh	nan	nan	800.0
GCCS 2011	nan	True	False	1.0	0.44	8.07	GJ_MWh	nan	nan	707.0
GCCS 2011	nan	True	False	1.0	0.41	8.76	GJ_MWh	nan	nan	753.0
GCCS 2011	nan	False	False	1.0	0.508	7.09	GJ_MWh	nan	nan	362.0
IEAGHG 2014	nan	False	False	0.9	0.441	nan	nan	nan	nan	746.0
IEAGHG 2014	nan	False	False	0.85	0.441	nan	nan	nan	nan	836.0
Rubin 2007	nan	True	False	0.75	0.393	nan	nan	nan	nan	811.0
Rubin 2007	nan	True	False	0.75	0.372	nan	nan	nan	nan	822.0
Rubin 2007	nan	True	False	0.75	0.502	nan	nan	nan	nan	367.0
NETL 2015	nan	True	False	0.85	0.39	8740.0	BTU_KWh	204.0	lb_MMBTU	nan
NETL 2015	nan	True	False	0.85	0.407	8379.0	BTU_KWh	204.0	lb_MMBTU	nan
NETL 2015	nan	True	False	0.85	0.515	6629.0	BTU_KWh	119.0	lb_MMBTU	nan
Davison 2007	nan	False	False	0.85	0.44	nan	nan	nan	nan	743.0
Davison 2007	nan	False	False	0.85	0.44	nan	nan	nan	nan	743.0
Davison 2007	nan	False	False	0.85	0.44	nan	nan	nan	nan	743.0
Davison 2007	nan	False	False	0.85	0.431	nan	nan	nan	nan	763.0
Davison 2007	nan	False	False	0.85	0.556	nan	nan	nan	nan	379.0
Davison 2007	nan	False	False	0.85	0.556	nan	nan	nan	nan	379.0
Davison 2007	nan	False	False	0.85	0.556	nan	nan	nan	nan	379.0
Dave from Hu 2011	nan	False	False	0.9	0.414	nan	nan	nan	nan	806.0
Li 2011	nan	False	True	0.45	0.49	nan	nan	nan	nan	419.0
Li 2011	nan	False	True	0.75	0.44	nan	nan	nan	nan	743.0
Li 2011	nan	False	True	0.68	0.365	nan	nan	nan	nan	895.0
Wu 2011	nan	False	False	0.685	nan	8441.0	BTU_KWH	nan	nan	669.0
Wu 2011	nan	False	False	0.685	nan	8361.0	BTU_KWH	nan	nan	663.0
Renner 2014	800.0	False	False	0.85	0.45	nan	nan	90.945	kg_GJ	nan
Renner 2014	800.0	False	False	0.85	0.6	nan	nan	90.945	kg_GJ	nan
Renner 2014	800.0	False	False	0.85	0.45	nan	nan	90.945	kg_GJ	nan
Renner 2014	800.0	False	False	0.85	0.6	nan	nan	90.945	kg_GJ	nan
CO2CRC 2017	2100.0	False	False	0.9	0.27	nan	nan	110.0	kg_GJ	nan
CO2CRC 2017	2100.0	False	True	0.9	0.27	nan	nan	110.0	kg_GJ	nan
CO2CRC 2017	2100.0	False	True	0.9	0.27	nan	nan	110.0	kg_GJ	nan
CO2CRC 2017	450.0	False	False	0.9	0.38	nan	nan	90.0	kg_GJ	nan
CO2CRC 2017	450.0	False	True	0.9	0.38	nan	nan	90.0	kg_GJ	nan
CO2CRC 2017	450.0	False	True	0.9	0.38	nan	nan	90.0	kg_GJ	nan
Dillon 2013	nan	False	True	0.9	0.25	nan	nan	95.0	kg_GJ	nan
Dillon 2013	nan	False	True	0.9	0.31	nan	nan	93.0	kg_GJ	nan
Dillon 2013	nan	False	True	0.9	0.31	nan	nan	97.0	kg_GJ	nan
Dillon 2013	nan	False	True	0.9	0.32	nan	nan	110.0	kg_GJ	nan
Dillon 2013	nan	False	True	0.9	0.35	nan	nan	97.0	kg_GJ	nan

Table 3: Raw tabulated values of literature, continuation

Study	EF_PLANT_UNITS	YEAR_EMM	YEAR_EMM_UNITS	CAPITAL_C_REF	CAPITAL_C_UNITS	CAP_LCOE	CAP_LCOE_UNITS	FCF	DISC_RATE	FOM
Bergerson Lave a 2007	lb_KWh	nan	nan	nan	nan	0.027	CURR_KWH	nan	10.0	0.0
Bergerson Lave a 2007	lb_KWh	nan	nan	nan	nan	0.035	CURR_KWH	nan	10.0	0.0
Bergerson Lave a 2007	lb_KWh	nan	nan	nan	nan	0.021	CURR_KWH	nan	10.0	0.0
Bergerson Lave a 2007	lb_KWh	nan	nan	nan	nan	0.031	CURR_KWH	nan	10.0	0.0
Bergerson Lave b 2007	kg_kwh	6.5	billton_year	1600.0	CURR_KW	nan	nan	nan	8.0	0.0
Bergerson Lave b 2007	kg_kwh	7.2	billton_year	1800.0	CURR_KW	nan	nan	nan	8.0	0.0
Bergerson Lave b 2007	kg_kwh	3.5	billton_year	nan	nan	0.009	CURR_KWH	nan	8.0	0.0
NREL 2012	nan	nan	nan	1230.0	CURR_KW	nan	nan	0.11	nan	6.31
NREL 2012	nan	nan	nan	2890.0	CURR_KW	nan	nan	0.106	nan	23.0
NREL 2012	nan	nan	nan	4010.0	CURR_KW	nan	nan	0.106	nan	31.0
NREL 2012	nan	nan	nan	2890.0	CURR_KW	nan	nan	0.106	nan	23.0
CAESAR 2011	kg_MWh	nan	nan	1930.0	CURR_KW	nan	nan	0.094	nan	36000.0
CAESAR 2011	kg_MWh	nan	nan	1950.0	CURR_KW	nan	nan	0.097	nan	50.0
CAESAR 2011	kg_MWh	nan	nan	630.0	CURR_KW	nan	nan	0.097	nan	28.7
GCCS 2011	kg_MWh	nan	nan	1919.0	CURR_KW	nan	nan	0.152	nan	19378000.0
GCCS 2011	kg_MWh	nan	nan	1921.0	CURR_KW	nan	nan	0.152	nan	19922000.0
GCCS 2011	kg_MWh	nan	nan	2013.0	CURR_KW	nan	nan	0.152	nan	19431000.0
GCCS 2011	kg_MWh	nan	nan	2618.0	CURR_KW	nan	nan	0.152	nan	38292000.0
GCCS 2011	kg_MWh	nan	nan	711.0	CURR_KW	nan	nan	0.152	nan	7823890.0
IEAGHG 2014	kg_MWh	nan	nan	1447.0	CURR_KW	nan	nan	nan	8.0	47350000.0
IEAGHG 2014	kg_MWh	nan	nan	1447.0	CURR_KW	nan	nan	nan	8.0	106806600.0
Rubin 2007	kg_MWh	nan	nan	1442.0	CURR_KW	nan	nan	0.148	nan	nan
Rubin 2007	kg_MWh	nan	nan	1567.0	CURR_KW	nan	nan	0.148	nan	nan
Rubin 2007	kg_MWh	nan	nan	671.0	CURR_KW	nan	nan	0.148	nan	nan
NETL 2015	nan	nan	nan	1960.0	CURR_KW	nan	nan	0.124	nan	nan
NETL 2015	nan	nan	nan	2026.0	CURR_KW	nan	nan	0.124	nan	nan
NETL 2015	nan	nan	nan	685.0	CURR_KW	nan	nan	0.124	nan	nan
Davison 2007	kg_MWh	nan	nan	1408.0	CURR_KW	nan	nan	nan	10.0	56.0
Davison 2007	kg_MWh	nan	nan	1408.0	CURR_KW	nan	nan	nan	10.0	56.0
Davison 2007	kg_MWh	nan	nan	1408.0	CURR_KW	nan	nan	nan	10.0	56.0
Davison 2007	kg_MWh	nan	nan	1613.0	CURR_KW	nan	nan	nan	10.0	64.0
Davison 2007	kg_MWh	nan	nan	499.0	CURR_KW	nan	nan	nan	10.0	20.0
Davison 2007	kg_MWh	nan	nan	499.0	CURR_KW	nan	nan	nan	10.0	20.0
Davison 2007	kg_MWh	nan	nan	499.0	CURR_KW	nan	nan	nan	10.0	20.0
Dave from Hu 2011	kg_MWh	nan	nan	688.0	CURR_KW	nan	nan	nan	7.0	33.9
Li 2011	kg_MWh	nan	nan	562.0	CURR_KW	nan	nan	nan	8.0	34.84
Li 2011	kg_MWh	nan	nan	664.0	CURR_KW	nan	nan	nan	8.0	38.906
Li 2011	kg_MWh	nan	nan	675.0	CURR_KW	nan	nan	nan	8.0	42.5
Wu 2011	kg_MWh	nan	nan	1106.0	CURR_KW	nan	nan	nan	6.0	16600000.0
Wu 2011	kg_MWh	nan	nan	2548.0	CURR_KW	nan	nan	nan	6.0	53500000.0
Renner 2014	nan	nan	nan	1765.0	CURR_KW	nan	nan	nan	8.0	nan
Renner 2014	nan	nan	nan	647.0	CURR_KW	nan	nan	nan	8.0	nan
Renner 2014	nan	nan	nan	851.0	CURR_KW	nan	nan	nan	8.0	nan
Renner 2014	nan	nan	nan	407.0	CURR_KW	nan	nan	nan	8.0	nan
CO2CRC 2017	nan	nan	nan	2945.25	CURR_KW	nan	nan	nan	8.0	42.075
CO2CRC 2017	nan	nan	nan	0.0	CURR_KW	nan	nan	nan	8.0	42.075
CO2CRC 2017	nan	nan	nan	0.0	CURR_KW	nan	nan	nan	8.0	42.075
CO2CRC 2017	nan	nan	nan	2371.5	CURR_KW	nan	nan	nan	8.0	34.425
CO2CRC 2017	nan	nan	nan	0.0	CURR_KW	nan	nan	nan	8.0	34.425
CO2CRC 2017	nan	nan	nan	0.0	CURR_KW	nan	nan	nan	8.0	34.425
Dillon 2013	nan	nan	nan	0.0	CURR_KW	nan	nan	nan	12.5	nan
Dillon 2013	nan	nan	nan	0.0	CURR_KW	nan	nan	nan	12.5	nan
Dillon 2013	nan	nan	nan	0.0	CURR_KW	nan	nan	nan	12.5	nan
Dillon 2013	nan	nan	nan	0.0	CURR_KW	nan	nan	nan	12.5	nan
Dillon 2013	nan	nan	nan	0.0	CURR_KW	nan	nan	nan	12.5	nan

Table 4: Raw tabulated values of literature, continuation

Study	FOM_UNITS	VOM	VOM_UNITS	FC	FC_UNITS	LIFE	CAPITAL_CC	CAP_LCOE_CC	FOM_CC	VOM_CC
Bergerson Lave a 2007	CURR_KW	0.013	CURR_KWh	27.0	CURR_TON	30.0	nan	0.036	0.0	0.034
Bergerson Lave a 2007	CURR_KW	0.011	CURR_KWh	27.0	CURR_TON	30.0	nan	0.035	0.0	0.022
Bergerson Lave a 2007	CURR_KW	0.007	CURR_KWh	27.0	CURR_TON	30.0	nan	0.036	0.0	0.034
Bergerson Lave a 2007	CURR_KW	0.01	CURR_KWh	27.0	CURR_TON	30.0	nan	0.035	0.0	0.022
Bergerson Lave b 2007	CURR_KW	0.012	CURR_KWh	11.0	CURR_TON	30.0	2550.0	nan	0.0	0.034
Bergerson Lave b 2007	CURR_KW	0.001	CURR_KWh	49.0	CURR_TON	30.0	2800.0	nan	0.0	0.013
Bergerson Lave b 2007	CURR_KW	0.042	CURR_KWh	7.0	CURR_MMBTU	30.0	nan	0.019	0.0	0.05
NREL 2012	CURR_KWY	0.004	CURR_KWh	6.12	CURR_GJ	40.0	3750.0	nan	18.4	0.01
NREL 2012	CURR_KWY	0.004	CURR_KWh	2.64	CURR_GJ	40.0	6560.0	nan	35.2	0.006
NREL 2012	CURR_KWY	0.007	CURR_KWh	2.64	CURR_GJ	40.0	6600.0	nan	44.4	0.011
NREL 2012	CURR_KWY	0.004	CURR_KWh	2.64	CURR_GJ	40.0	6920.0	nan	58.4	0.01
CAESAR 2011	CURR_Y	0.0	CURR_KWh	4.5	CURR_GJ	30.0	3011.0	nan	65.0	0.002
CAESAR 2011	CURR_KWY	0.002	CURR_KWh	4.5	CURR_GJ	30.0	2643.0	nan	56.0	0.002
CAESAR 2011	CURR_KWY	0.001	CURR_KWh	6.5	CURR_GJ	30.0	969.0	nan	43.2	0.002
GCCS 2011	CURR	0.005	CURR_KWh	2.61	CURR_GJ	30.0	3464.0	nan	29701000.0	0.009
GCCS 2011	CURR	0.005	CURR_KWh	2.61	CURR_GJ	30.0	3440.0	nan	28499000.0	0.009
GCCS 2011	CURR	0.004	CURR_KWh	2.61	CURR_GJ	30.0	3404.0	nan	27701000.0	0.008
GCCS 2011	CURR	0.005	CURR_KWh	2.61	CURR_GJ	30.0	3413.0	nan	38469000.0	0.006
GCCS 2011	CURR	0.001	CURR_KWh	7.4	CURR_GJ	30.0	1447.0	nan	11620570.0	0.002
IEAGHG 2014	CURR_Y	0.001	CURR_KWh	2.5	CURR_GJ	25.0	2771.0	nan	69275200.0	0.004
IEAGHG 2014	CURR_Y	0.002	CURR_KWh	2.5	CURR_GJ	25.0	3157.0	nan	112520000.0	0.002
Rubin 2007	nan	0.013	CURR_KWh	1.2	CURR_GJ	nan	2345.0	nan	nan	0.03
Rubin 2007	nan	0.013	CURR_KWh	1.2	CURR_GJ	nan	1567.0	nan	nan	0.018
Rubin 2007	nan	0.015	CURR_KWh	6.0	CURR_GJ	nan	1091.0	nan	nan	0.021
NETL 2015	nan	0.018	CURR_KWh	2.78	CURR_GJ	nan	3467.0	nan	nan	0.03
NETL 2015	nan	0.019	CURR_KWh	2.78	CURR_GJ	nan	3524.0	nan	nan	0.03
NETL 2015	nan	0.005	CURR_KWh	5.81	CURR_GJ	nan	1481.0	nan	nan	0.01
Davison 2007	CURR_KWY	0.0	CURR_KWh	2.2	CURR_GJ	25.0	1979.0	nan	79.0	0.0
Davison 2007	CURR_KWY	0.0	CURR_KWh	2.2	CURR_GJ	25.0	2043.0	nan	81.0	0.0
Davison 2007	CURR_KWY	0.0	CURR_KWh	2.2	CURR_GJ	25.0	2205.0	nan	88.0	0.0
Davison 2007	CURR_KWY	0.0	CURR_KWh	2.2	CURR_GJ	25.0	2204.0	nan	88.0	0.0
Davison 2007	CURR_KWY	0.0	CURR_KWh	7.8	CURR_GJ	25.0	869.0	nan	34.0	0.0
Davison 2007	CURR_KWY	0.0	CURR_KWh	7.8	CURR_GJ	25.0	887.0	nan	35.0	0.0
Davison 2007	CURR_KWY	0.0	CURR_KWh	7.8	CURR_GJ	25.0	1532.0	nan	61.0	0.0
Dave from Hu 2011	CURR_KWY	0.0	CURR_KWh	3.6	CURR_GJ	20.0	1257.0	nan	62.85	0.0
Li 2011	CURR_KWY	0.0	CURR_KWh	13.2	CURR_GJ	25.0	1045.32	nan	59.576	0.0
Li 2011	CURR_KWY	0.0	CURR_KWh	5.2	CURR_GJ	25.0	1108.88	nan	62.25	0.0
Li 2011	CURR_KWY	0.0	CURR_KWh	5.2	CURR_GJ	25.0	1221.75	nan	76.925	0.0
Wu 2011	CURR	0.0	CURR_KWh	5.2	CURR_MMBTU	40.0	1780.0	nan	42200000.0	0.0
Wu 2011	CURR	0.0	CURR_KWh	5.2	CURR_MMBTU	40.0	3822.0	nan	96600000.0	0.0
Renner 2014	nan	0.012	CURR_KWh	4.34	CURR_GJ	40.0	3225.0	nan	nan	0.022
Renner 2014	nan	0.006	CURR_KWh	11.6	CURR_MMBTU	25.0	1589.0	nan	nan	0.013
Renner 2014	nan	0.003	CURR_KWh	3.8	CURR_GJ	40.0	1370.0	nan	nan	0.017
Renner 2014	nan	0.001	CURR_KWh	10.25	CURR_MMBTU	25.0	733.0	nan	nan	0.002
CO2CRC 2017	CURR_KWY	0.002	CURR_KWh	1.8	CURR_MMBTU	30.0	6311.25	nan	49.725	0.009
CO2CRC 2017	CURR_KWY	0.002	CURR_KWh	1.8	CURR_MMBTU	30.0	3748.5	nan	53.55	0.011
CO2CRC 2017	CURR_KWY	0.002	CURR_KWh	1.8	CURR_MMBTU	30.0	2983.5	nan	45.9	0.009
CO2CRC 2017	CURR_KWY	0.002	CURR_KWh	2.0	CURR_MMBTU	30.0	5355.0	nan	42.075	0.007
CO2CRC 2017	CURR_KWY	0.002	CURR_KWh	2.0	CURR_MMBTU	30.0	3136.5	nan	57.375	0.008
CO2CRC 2017	CURR_KWY	0.002	CURR_KWh	2.0	CURR_MMBTU	30.0	2187.9	nan	53.55	0.008
Dillon 2013	nan	0.0	CURR_KWh	30.0	CURR_TON	30.0	4000.0	nan	nan	0.0
Dillon 2013	nan	0.0	CURR_KWh	30.0	CURR_TON	30.0	2900.0	nan	nan	0.0
Dillon 2013	nan	0.0	CURR_KWh	30.0	CURR_TON	30.0	2600.0	nan	nan	0.0
Dillon 2013	nan	0.0	CURR_KWh	30.0	CURR_TON	30.0	1400.0	nan	nan	0.0
Dillon 2013	nan	0.0	CURR_KWh	30.0	CURR_TON	30.0	900.0	nan	nan	0.0

Table 5: Raw tabulated values of literature, continuation

Study	HR_CC	ELEC_EFF_CC	CAP_EFF	P_AUX_CC	EFF_BASIS	BASIS
Bergerson Lave a 2007	nan	0.292	0.9	nan	HHV	2002USD
Bergerson Lave a 2007	nan	0.293	0.9	nan	HHV	2002USD
Bergerson Lave a 2007	nan	0.292	0.9	nan	HHV	2002USD
Bergerson Lave a 2007	nan	0.293	0.9	nan	HHV	2002USD
Bergerson Lave b 2007	nan	0.33	0.9	nan	HHV	2005USD
Bergerson Lave b 2007	nan	0.28	0.9	nan	HHV	2005USD
Bergerson Lave b 2007	nan	0.43	0.9	nan	HHV	2005USD
NREL 2012	10800.0	nan	0.85	nan	nan	2010USD
NREL 2012	12600.0	nan	0.85	nan	nan	2010USD
NREL 2012	11800.0	nan	0.85	nan	nan	2010USD
NREL 2012	12600.0	nan	0.85	nan	nan	2010USD
CAESAR 2011	nan	0.334	0.9	135.0	LHV	2008EUR
CAESAR 2011	nan	0.364	0.9	0.0	LHV	2008EUR
CAESAR 2011	nan	0.499	0.905	nan	LHV	2008EUR
GCCS 2011	13.22	0.272	0.9	77.0	HHV	2009USD
GCCS 2011	12.73	0.283	0.9	81.0	HHV	2009USD
GCCS 2011	10.83	0.332	0.9	68.0	HHV	2009USD
GCCS 2011	12.61	0.32	0.9	64.0	HHV	2009USD
GCCS 2011	8.24	0.437	0.9	28.0	HHV	2009USD
IEAGHG 2014	nan	0.352	0.9	nan	LHV	2013EUR
IEAGHG 2014	nan	0.355	0.9	nan	LHV	2013EUR
Rubin 2007	nan	0.299	0.9	nan	HHV	2005USD
Rubin 2007	nan	0.372	0.9	nan	HHV	2005USD
Rubin 2007	nan	0.428	0.9	nan	HHV	2005USD
NETL 2015	10953.0	0.312	0.9	63.0	HHV	2011USD
NETL 2015	10508.0	0.325	0.9	61.0	HHV	2011USD
NETL 2015	7466.0	0.457	0.9	31.0	HHV	2011USD
Davison 2007	nan	0.348	0.875	nan	LHV	2005EUR
Davison 2007	nan	0.353	0.9	nan	LHV	2005EUR
Davison 2007	nan	0.354	0.908	nan	LHV	2005EUR
Davison 2007	nan	0.345	0.85	nan	LHV	2005EUR
Davison 2007	nan	0.474	0.85	nan	LHV	2005EUR
Davison 2007	nan	0.496	0.85	nan	LHV	2005EUR
Davison 2007	nan	0.447	0.972	nan	LHV	2005EUR
Dave from Hu 2011	nan	0.299	0.9	nan	HHV	2010USD
Li 2011	nan	0.41	0.85	nan	HHV	2010USD
Li 2011	nan	0.38	0.85	nan	HHV	2010USD
Li 2011	nan	0.29	0.85	nan	HHV	2010USD
Wu 2011	11724.0	nan	0.9	nan	LHV	2010USD
Wu 2011	10074.0	nan	0.9	nan	LHV	2010USD
Renner 2014	nan	0.36	0.9	nan	LHV	2011EUR
Renner 2014	nan	0.52	0.9	nan	LHV	2011EUR
Renner 2014	nan	0.36	0.9	nan	LHV	2011EUR
Renner 2014	nan	0.52	0.9	nan	LHV	2011EUR
CO2CRC 2017	nan	0.195	0.9	nan	LHV	2017USD
CO2CRC 2017	nan	0.155	0.9	nan	LHV	2017USD
CO2CRC 2017	nan	0.195	0.9	nan	LHV	2017USD
CO2CRC 2017	nan	0.267	0.9	nan	LHV	2017USD
CO2CRC 2017	nan	0.267	0.9	nan	LHV	2017USD
CO2CRC 2017	nan	0.267	0.9	nan	LHV	2017USD
Dillon 2013	nan	0.11	0.9	nan	HHV	2009USD
Dillon 2013	nan	0.21	0.9	nan	HHV	2009USD
Dillon 2013	nan	0.2	0.9	nan	HHV	2009USD
Dillon 2013	nan	0.2	0.9	nan	HHV	2009USD
Dillon 2013	nan	0.25	0.9	nan	HHV	2009USD

Table 6: Codebook of input data

Name	dtype	Range	Description	Values	Value_description	Necessary	Interchangeable with
Study	string		Identifier for the study responsible author or...			Yes	
Publication_Year	int	Integers from 0 to 2020	Year in which the study was published			Yes	
Territory	case	See Values	General region to which the study is assigned,...	China	Some studies are done explicitly considering C...	No	
				USA	Mosts studies are reported in USD, but only so...		
				Europe	Some studies are done in the European Context		
				Australia	There is a Handful of studies done in Australi...		
				Canada	Canada has specific studies too		
Technology	case	See Values	General power t echnology to which the study i...	NGCC	Natural Gas Combined Cycle	Yes	
				SCPC	Super Critical Pulverized Coal		
				UCPC	Ultra Super Critical Pulverized Coal		
				SUBC	Subcritical Pulverized Coal		
				CFB	Coal Fluidized Bed		
				IGCC	Integrated Gasification Combined Cycle		
CaptureTech	case	See Values	Capture technology assumed in the study	OXYFUEL		Yes	
				MEA			
				SELEXOL			
				MHI			
Fuel	string		Fuel name identifier, it is the particular nam...			No	
fuel_general	string	See Values		coal	Any type of coal, ex. Bituminous, Lignite	Yes	
				natural_gas	Natural gas in general		
P_GROSS	float	Real positive numbers	Gorss installed plant capacity, including auxi...			No	
P_NET	float	Real positive numbers	Net installed capacity, or nominal capacity, i...			No	
P_NET_CC	float	Real positive numbers	Some reports have values for the net capacity ...			No	
P_AUX	float	Real positive numbers	Auxilary power of the plant			No	
P_MAX	float	Real positive numbers	If the report considers a range for it's power...			No	
P_MIN	float	Real positive numbers	If the report considers a range for it's power...			No	
Repower	bool	{True, False}	If true, the power plant had additional power ...			No	
Retrofit	bool	{True, False}	If true, the study is done in a retrofit case			No	
CF	float	0-1	Capacity factor of the power plant, also calle...			No	
ELEC_EFF	float	0-1	Electric efficiency of the reference power pla...			Yes	HR
HR	float	Real positive numbers	Heat rate of the reference power plant			Yes	ELEC_EFF
HR_UNITS	string	unit strings *	Units of the Heat rate, Heat per electricity			Yes	
EF_FUEL	float	Real positive numbers	Emission factor of the fuel, if it is reported...			Yes	EF_PLANT, YEAR_EMM
EF_FUEL_UNITS	string	unit strings *	Units of the fuel emissionfactor, mass of CO2 ...			Yes	
EF_PLANT	float	Real positive numbers	Plant specific emission, if reported. It is us...			Yes	EF_FUEL, YEAR_EMM
EF_PLANT_UNITS	string	unit strings *	Units of specific emission of the plant, mass ...			Yes	
YEAR_EMM	float	Real positive numbers	Yearly emissions from the plant, used to calcu...			Yes	EF_FUEL, EF_PLANT
YEAR_EMM_UNITS	string	unit strings *	Units of the yearly emissions, amount of CO2 p...			Yes	
CAPITAL_C_REF	float	Real positive numbers	Capital cost of the reference plant used in th...			Yes	CAP_LCOE
CAPITAL_C_UNITS	string	unit strings *	Units of the capital cost, could be per KW or ...			Yes	
CAP_LCOE	float	Real positive numbers	Calculated component of capital cost of LCOE f...			Yes	CAPITAL_C_REF
CAP_LCOE_UNITS	string	unit strings *	Units of the capital cost, LCOE currency per ...			Yes	
FCF	float	0-1	Fixed cost factor, adjuster value for the inlf...			Yes	DISC_RATE + LIFE
DISC_RATE	float	0-1	Inflation index for the claculation of FCF			Yes	FCF
FOM	float	Real positive numbers	Fixed operational and maintenance costs, if th...			No	FOM
FOM_UNITS	string	unit strings *	Units of the FOM, varies a lot depending on th...			No	
VOM	float	Real positive numbers	Variable operation and maintenance costs , if ...			No	VOM
VOM_UNITS	string	unit strings *	Units of the FOM, varies a lot depending on th...			No	
FC	float	Real positive numbers	Fuel cost reported in the study, if not found,...			Yes	
FC_UNITS	string	unit strings *	Units of fuel cost, could be per energy or amo...			Yes	
LIFE	int	Real positive integers	Life of the proyect assumed for the discountin...			Yes	FCF
CAPITAL_CC	float	Real positive numbers	Capital cost of the capture plant used in the ...			Yes	CAP_LCOE_CC
CAP_LCOE_CC	float	Real positive numbers	Calculated component of capital cost of LCOE f...			Yes	CAPITAL_CC
FOM_CC	float	Real positive numbers	Fixed operational and maintenance costs, if th...			No	
VOM_CC	float	Real positive numbers	Variable operation and maintenance costs , if ...			No	
HR_CC	float	Real positive numbers	Heat rate of the capture power plant, if ADDIT...			Yes	ELEC_EFF_CC
ELEC_EFF_CC	float	0-1	Efficiency of the capture plant, if efficiency...			Yes	HR_CC
CAP_EFF	float	0-1	Capture efficiency of the carbon capture process			Yes	
P_AUX_CC	float	Real positive numbers	Additional power added for the capture process			No	
EFF_BASIS	case	{HHV,LHV}	Heat value basis of the fuel cost and emission...	LHV	Low heat value, or Net heat value. The heat ob...	Yes	
				HHV	High heat value or Gross heat value. The heat ...	Yes	
BASIS	string	basis string ***	Currency basis of the study, it should be cons...			Yes	

Appendix B: Harmonized Values

Table 7: Harmonized values of carbon captue studies

label	region	p_year	power_technology	capture_technology	fuel_name	fuel_type	power_gross	power_net	power_aux	retrofit
Bergerson_Lave_a_SCPC_A	Canada	2007	SCPC	MEA	illinois_6	coal	nan	465.0	nan	False
Bergerson_Lave_a_IGCC_A	Canada	2007	IGCC	SELEXOL	illinois_6	coal	nan	465.0	nan	False
Bergerson_Lave_a_SCPC_B	Canada	2007	SCPC	MEA	illinois_6	coal	nan	465.0	nan	True
Bergerson_Lave_a_IGCC_B	Canada	2007	IGCC	SELEXOL	illinois_6	coal	nan	465.0	nan	True
Bergerson_Lave_b_SCPC_A	Canada	2007	SCPC	MEA	Lignite	coal	nan	nan	nan	False
Bergerson_Lave_b_IGCC_A	Canada	2007	IGCC	SELEXOL	Bituminous	coal	nan	nan	nan	False
Bergerson_Lave_b_NGCC_A	Canada	2007	NGCC	MEA	natural_gas	natural_gas	nan	nan	nan	False
NREL_NGCC_A	Europe	2012	NGCC	MEA	natural_gas	natural_gas	nan	615.0	nan	False
NREL_SCPC_A	Europe	2012	SCPC	MEA	Coal	coal	606.0	606.0	nan	False
NREL_IGCC_A	Europe	2012	IGCC	SELEXOL	Coal	coal	590.0	590.0	nan	False
NREL_SCPC_B	Europe	2012	SCPC	MEA	Coal	coal	606.0	606.0	nan	True
CAESAR_SCPC_A	Europe	2011	SCPC	MEA	Bituminous	coal	819.0	754.0	65.0	False
CAESAR_IGCC_A	Europe	2011	IGCC	SELEXOL	Bituminous	coal	441.0	391.0	nan	False
CAESAR_NGCC_A	Europe	2011	NGCC	MEA	natural_gas	natural_gas	nan	829.0	nan	False
GCCS_SCPC_A	Australia	2011	SCPC	MEA	illinois_6	coal	580.0	550.0	30.0	False
GCCS_SCPC_B	Australia	2011	SCPC	MEA	illinois_6	coal	580.0	550.0	30.0	False
GCCS_UCPC_A	Australia	2011	UCPC	MEA	illinois_6	coal	576.0	550.0	26.6	False
GCCS_IGCC_A	Australia	2011	IGCC	SELEXOL	illinois_6	coal	748.0	366.0	112.0	False
GCCS_NGCC_A	Australia	2011	NGCC	MEA	natural_gas	natural_gas	570.0	560.0	10.0	False
IEAGHG_SCPC_A	Europe	2014	SCPC	MEA	Bituminous	coal	nan	1030.0	nan	False
IEAGHG_IGCC_A	Europe	2014	IGCC	SELEXOL	Bituminous	coal	nan	804.0	nan	False
Rubin_SCPC_A	USA	2007	SCPC	MEA	illinois_6	coal	575.0	528.0	nan	False
Rubin_IGCC_A	USA	2007	IGCC	SELEXOL	illinois_6	coal	615.0	538.0	nan	False
Rubin_NGCC_A	USA	2007	NGCC	MEA	natural_gas	natural_gas	517.0	507.0	nan	False
NETL_SCPC_A	USA	2015	SCPC	MEA	illinois_6	coal	581.0	550.0	31.0	False
NETL_SUBC_A	USA	2015	SUBC	MEA	illinois_6	coal	580.0	550.0	30.0	False
NETL_NGCC_A	USA	2015	NGCC	MEA	natural_gas	natural_gas	641.0	630.0	11.0	False
Davison_SCPC_A	Europe	2007	SCPC	MEA	Bituminous	coal	nan	758.0	nan	False
Davison_SCPC_B	Europe	2007	SCPC	MHI	Bituminous	coal	nan	758.0	nan	False
Davison_SCPC_C	Europe	2007	SCPC	OXYFUEL	Bituminous	coal	nan	758.0	nan	False
Davison_IGCC_A	Europe	2007	IGCC	SELEXOL	Bituminous	coal	nan	776.0	nan	False
Davison_NGCC_A	Europe	2007	NGCC	MEA	natural_gas	natural_gas	nan	776.0	nan	False
Davison_NGCC_B	Europe	2007	NGCC	MHI	natural_gas	natural_gas	nan	776.0	nan	False
Davison_NGCC_C	Europe	2007	NGCC	OXYFUEL	natural_gas	natural_gas	nan	776.0	nan	False
Dave_from_Hu_SCPC_A	China	2011	SCPC	MEA	China	coal	600.0	570.0	nan	False
Li_NGCC_A	China	2011	NGCC	MEA	natural_gas	natural_gas	nan	nan	nan	True
Li_SCPC_A	China	2011	SCPC	MEA	Bituminous	coal	nan	nan	nan	True
Li_SUBC_A	China	2011	SUBC	MEA	Bituminous	coal	nan	nan	nan	True
Wu_SCPC_A	China	2011	SCPC	MEA	China	coal	nan	600.0	nan	False
Wu_IGCC_A	China	2011	IGCC	SELEXOL	China	coal	nan	600.0	nan	False
Renner_SCPC_A	Europe	2014	SCPC	MEA	illinois_6	coal	nan	600.0	nan	False
Renner_NGCC_A	Europe	2014	NGCC	MEA	natural_gas	natural_gas	nan	600.0	nan	False
Renner_SCPC_B	China	2014	SCPC	MEA	illinois_6	coal	nan	600.0	nan	False
Renner_NGCC_B	China	2014	NGCC	MEA	natural_gas	natural_gas	nan	600.0	nan	False
CO2CRC_SCPC_A	Australia	2017	SCPC	MEA	victorian_brown	coal	nan	500.0	nan	False
CO2CRC_SCPC_B	Australia	2017	SCPC	MEA	victorian_brown	coal	nan	500.0	nan	True
CO2CRC_SCPC_C	Australia	2017	SCPC	MEA	victorian_brown	coal	nan	500.0	nan	True
CO2CRC_SCPC_D	Australia	2017	SCPC	MEA	black_coal	coal	nan	450.0	nan	False
CO2CRC_SCPC_F	Australia	2017	SCPC	MEA	black_coal	coal	nan	450.0	nan	True
CO2CRC_SCPC_G	Australia	2017	SCPC	MEA	black_coal	coal	nan	450.0	nan	True
Dillon_CFB_A	USA	2013	CFB	MEA	Petcoke	coal	nan	129.0	nan	True
Dillon_SUBC_A	USA	2013	SUBC	MEA	Bituminous	coal	nan	616.0	nan	True
Dillon_SUBC_B	USA	2013	SUBC	MEA	subbituminous	coal	nan	1500.0	nan	True
Dillon_SUBC_C	USA	2013	SUBC	MEA	Lignite	coal	nan	1010.0	nan	True
Dillon_SUBC_D	USA	2013	SUBC	MEA	Bituminous	coal	nan	1800.0	nan	True

Table 8: Harmonized values of carbon captue studies, continuation

label	repower	capacity_factor	electric_efficiency	electric_efficiency_cc	heat_rate	heat_rate_cc	fuel_emission_factor	plant_emission	capture_efficiency	capital_cost
Bergerson_Lave_a_SCPC_A	False	0.75	0.401	0.292	8977.556	12328.767	0.0	0.816	0.9	nan
Bergerson_Lave_a_IGCC_A	False	0.75	0.344	0.293	10465.116	12286.689	0.0	0.862	0.9	nan
Bergerson_Lave_a_SCPC_B	False	0.75	0.401	0.292	8977.556	12328.767	0.0	0.816	0.9	nan
Bergerson_Lave_a_IGCC_B	False	0.75	0.344	0.293	10465.116	12286.689	0.0	0.862	0.9	nan
Bergerson_Lave_b_SCPC_A	False	0.75	0.42	0.33	8571.429	10909.091	0.0	0.825	0.9	2270.724
Bergerson_Lave_b_IGCC_A	False	0.75	0.34	0.28	10588.235	12857.143	0.0	0.9	0.9	2554.564
Bergerson_Lave_b_NGCC_A	False	0.75	0.5	0.43	7200.0	8372.093	0.0	0.438	0.9	nan
NREL_NGCC_A	False	0.85	0.509	0.316	7074.177	11394.648	0.0	nan	0.9	977.547
NREL_SCPC_A	False	0.85	0.364	0.271	9885.912	13293.756	0.0	nan	0.9	2296.837
NREL_IGCC_A	False	0.85	0.378	0.289	9527.192	12449.708	0.0	nan	0.9	3186.961
NREL_SCPC_B	False	0.85	0.364	0.271	9885.912	13293.756	0.0	nan	0.9	2296.837
CAESAR_SCPC_A	False	0.85	0.455	0.334	7912.088	10778.443	0.0	0.763	0.9	1944.78
CAESAR_IGCC_A	False	0.85	0.469	0.364	7679.181	9890.11	0.0	0.734	0.9	1964.934
CAESAR_NGCC_A	False	0.85	0.583	0.499	6174.957	7214.429	0.0	0.352	0.9	634.825
GCCS_SCPC_A	True	1.0	0.39	0.272	9200.0	13220.0	0.0	0.804	0.9	1525.132
GCCS_SCPC_B	True	1.0	0.39	0.283	9140.0	12730.0	0.0	0.8	0.9	1526.721
GCCS_UCPC_A	True	1.0	0.44	0.332	8070.0	10830.0	0.0	0.707	0.9	1599.838
GCCS_IGCC_A	True	1.0	0.41	0.32	8760.0	12610.0	0.0	0.753	0.9	2080.664
GCCS_NGCC_A	False	1.0	0.508	0.437	7090.0	8240.0	0.0	0.362	0.9	565.07
IEAGHG_SCPC_A	False	0.9	0.441	0.352	8163.265	10227.273	0.0	0.746	0.9	1525.135
IEAGHG_IGCC_A	False	0.85	0.441	0.355	8163.265	10140.845	0.0	0.836	0.9	1525.135
Rubin_SCPC_A	True	0.75	0.393	0.299	9160.305	12040.134	0.0	0.811	0.9	2046.49
Rubin_IGCC_A	True	0.75	0.372	0.372	9677.419	9677.419	0.0	0.822	0.9	2223.89
Rubin_NGCC_A	True	0.75	0.502	0.428	7171.315	8411.215	0.0	0.367	0.9	952.285
NETL_SCPC_A	True	0.85	0.39	0.312	9221.224	11556.072	0.0	nan	0.9	1354.536
NETL_SUBC_A	True	0.85	0.407	0.325	8840.348	11086.57	0.0	nan	0.9	1400.148
NETL_NGCC_A	True	0.85	0.515	0.457	6993.993	7877.078	0.0	nan	0.9	473.397
Davison_SCPC_A	False	0.85	0.44	0.348	8181.818	10344.828	0.0	0.743	0.9	2486.718
Davison_SCPC_B	False	0.85	0.44	0.353	8181.818	10198.3	0.0	0.743	0.9	2486.718
Davison_SCPC_C	False	0.85	0.44	0.354	8181.818	10169.492	0.0	0.743	0.9	2486.718
Davison_IGCC_A	False	0.85	0.431	0.345	8352.668	10434.783	0.0	0.763	0.9	2848.776
Davison_NGCC_A	False	0.85	0.556	0.474	6474.82	7594.937	0.0	0.379	0.9	881.301
Davison_NGCC_B	False	0.85	0.556	0.496	6474.82	7258.065	0.0	0.379	0.9	881.301
Davison_NGCC_C	False	0.85	0.556	0.447	6474.82	8053.691	0.0	0.379	0.9	881.301
Dave_from_Hu_SCPC_A	False	0.9	0.414	0.299	8695.652	12040.134	0.0	0.806	0.9	546.79
Li_NGCC_A	False	0.45	0.49	0.41	7346.939	8780.488	0.0	0.419	0.9	446.651
Li_SCPC_A	False	0.75	0.44	0.38	8181.818	9473.684	0.0	0.743	0.9	527.716
Li_SUBC_A	False	0.68	0.365	0.29	9863.014	12413.793	0.0	0.895	0.9	536.458
Wu_SCPC_A	False	0.685	0.404	0.291	8905.761	12369.523	0.0	0.669	0.9	878.997
Wu_IGCC_A	False	0.685	0.408	0.339	8821.357	10628.674	0.0	0.663	0.9	2025.031
Renner_SCPC_A	False	0.85	0.45	0.36	8000.0	10000.0	0.0	nan	0.9	1620.436
Renner_NGCC_A	False	0.85	0.6	0.52	6000.0	6923.077	0.0	nan	0.9	594.007
Renner_SCPC_B	False	0.85	0.45	0.36	8000.0	10000.0	0.0	nan	0.9	781.298
Renner_NGCC_B	False	0.85	0.6	0.52	6000.0	6923.077	0.0	nan	0.9	373.664
CO2CRC_SCPC_A	False	0.9	0.27	0.195	13333.333	18461.538	0.0	nan	0.9	2753.82
CO2CRC_SCPC_B	False	0.9	0.27	0.155	13333.333	23225.806	0.0	nan	0.9	0.0
CO2CRC_SCPC_C	False	0.9	0.27	0.195	13333.333	18461.538	0.0	nan	0.9	0.0
CO2CRC_SCPC_D	False	0.9	0.38	0.267	9473.684	13483.146	0.0	nan	0.9	2217.362
CO2CRC_SCPC_F	False	0.9	0.38	0.267	9473.684	13483.146	0.0	nan	0.9	0.0
CO2CRC_SCPC_G	False	0.9	0.38	0.267	9473.684	13483.146	0.0	nan	0.9	0.0
Dillon_CFB_A	False	0.9	0.25	0.11	14400.0	32727.273	0.0	nan	0.9	0.0
Dillon_SUBC_A	False	0.9	0.31	0.21	11612.903	17142.857	0.0	nan	0.9	0.0
Dillon_SUBC_B	False	0.9	0.31	0.2	11612.903	18000.0	0.0	nan	0.9	0.0
Dillon_SUBC_C	False	0.9	0.32	0.2	11250.0	18000.0	0.0	nan	0.9	0.0
Dillon_SUBC_D	False	0.9	0.35	0.25	10285.714	14400.0	0.0	nan	0.9	0.0

Table 9: Harmonized values of carbon captue studies, continuation

label	capital_cost_cc	FCF	life	fixed_om	fixed_om_cc	variable_om	variable_om_cc	fuel_cost	lcoe_capex	lcoe_om
Bergerson_Lave_a_SCPC_A	nan	0.11	30.0	0.0	0.0	0.019	0.046	2.853	0.041	0.019
Bergerson_Lave_a_IGCC_A	nan	0.11	30.0	0.0	0.0	0.016	0.03	2.853	0.053	0.016
Bergerson_Lave_a_SCPC_B	nan	0.11	30.0	0.0	0.0	0.01	0.046	2.853	0.032	0.01
Bergerson_Lave_a_IGCC_B	nan	0.11	30.0	0.0	0.0	0.014	0.03	2.853	0.047	0.014
Bergerson_Lave_b_SCPC_A	2906.862	0.11	30.0	0.0	0.0	0.012	0.039	0.741	0.038	0.012
Bergerson_Lave_b_IGCC_A	3191.848	0.11	30.0	0.0	0.0	0.001	0.015	2.332	0.043	0.001
Bergerson_Lave_b_NGCC_A	nan	0.11	30.0	0.0	0.0	0.042	0.057	1.697	0.013	0.042
NREL_NGCC_A	3479.313	0.11	40.0	5.888	17.072	0.003	0.009	1.556	0.014	0.004
NREL_SCPC_A	6086.478	0.11	40.0	21.461	32.659	0.003	0.006	1.421	0.034	0.006
NREL_IGCC_A	6123.59	0.11	40.0	28.925	41.195	0.006	0.01	1.421	0.047	0.01
NREL_SCPC_B	6420.492	0.11	40.0	21.461	54.184	0.003	0.009	1.421	0.034	0.006
CAESAR_SCPC_A	4300.082	0.11	30.0	0.055	0.0	0.0	0.004	5.98	0.029	0.0
CAESAR_IGCC_A	3774.532	0.11	30.0	57.837	79.975	0.002	0.004	5.98	0.029	0.01
CAESAR_NGCC_A	1383.852	0.11	30.0	33.198	61.695	0.001	0.004	5.174	0.009	0.005
GCCS_SCPC_A	3353.948	0.11	30.0	1.031	1.743	0.004	0.009	1.804	0.019	0.004
GCCS_SCPC_B	3330.71	0.11	30.0	1.06	1.672	0.004	0.009	1.804	0.019	0.004
GCCS_UCPC_A	3295.854	0.11	30.0	1.034	1.626	0.004	0.008	1.804	0.02	0.004
GCCS_IGCC_A	3304.568	0.11	30.0	3.063	3.392	0.004	0.006	1.804	0.026	0.004
GCCS_NGCC_A	1401.028	0.11	30.0	0.409	0.67	0.001	0.002	3.801	0.007	0.001
IEAGHG_SCPC_A	3512.898	0.11	25.0	51.658	85.265	0.001	0.004	1.962	0.021	0.008
IEAGHG_IGCC_A	4002.245	0.11	25.0	149.279	177.42	0.002	0.002	1.962	0.023	0.022
Rubin_SCPC_A	2673.173	0.11	nan	nan	nan	0.013	0.034	1.725	0.034	0.013
Rubin_IGCC_A	1786.295	0.11	nan	nan	nan	0.013	0.021	1.725	0.037	0.013
Rubin_NGCC_A	1243.681	0.11	nan	nan	nan	0.015	0.024	1.534	0.016	0.015
NETL_SCPC_A	3262.127	0.11	nan	nan	nan	0.015	0.028	2.248	0.02	0.015
NETL_SUBC_A	3315.758	0.11	nan	nan	nan	0.015	0.028	2.248	0.021	0.015
NETL_NGCC_A	1393.484	0.11	nan	nan	nan	0.004	0.01	3.511	0.007	0.004
Davison_SCPC_A	2807.434	0.11	25.0	69.36	112.07	0.0	0.0	3.935	0.037	0.009
Davison_SCPC_B	2898.225	0.11	25.0	69.36	114.908	0.0	0.0	3.935	0.037	0.009
Davison_SCPC_C	3128.041	0.11	25.0	69.36	124.838	0.0	0.0	3.935	0.037	0.009
Davison_IGCC_A	3126.622	0.11	25.0	79.269	124.838	0.0	0.0	3.935	0.042	0.011
Davison_NGCC_A	1232.774	0.11	25.0	24.771	48.233	0.0	0.0	2.482	0.013	0.003
Davison_NGCC_B	1258.309	0.11	25.0	24.771	49.651	0.0	0.0	2.482	0.013	0.003
Davison_NGCC_C	2173.314	0.11	25.0	24.771	86.535	0.0	0.0	2.482	0.013	0.003
Dave_from_Hu_SCPC_A	1166.266	0.11	20.0	31.631	58.313	0.0	0.0	1.938	0.008	0.004
Li_NGCC_A	969.865	0.11	25.0	32.508	55.276	0.0	0.0	3.356	0.012	0.008
Li_SCPC_A	1028.837	0.11	25.0	36.302	57.757	0.0	0.0	2.799	0.009	0.006
Li_SUBC_A	1133.56	0.11	25.0	39.655	71.372	0.0	0.0	2.799	0.01	0.007
Wu_SCPC_A	1651.514	0.11	40.0	0.645	1.631	0.0	0.0	2.653	0.016	0.0
Wu_IGCC_A	3546.115	0.11	40.0	2.08	3.734	0.0	0.0	2.653	0.037	0.0
Renner_SCPC_A	4031.155	0.11	40.0	nan	nan	0.013	0.027	4.662	0.024	0.013
Renner_NGCC_A	1986.203	0.11	25.0	nan	nan	0.006	0.016	8.827	0.009	0.006
Renner_SCPC_B	1712.46	0.11	40.0	nan	nan	0.003	0.021	4.082	0.012	0.003
Renner_NGCC_B	916.228	0.11	25.0	nan	nan	0.001	0.003	7.8	0.006	0.001
CO2CRC_SCPC_A	5933.071	0.11	30.0	38.069	46.745	0.002	0.009	1.319	0.038	0.007
CO2CRC_SCPC_B	3523.885	0.11	30.0	38.069	50.341	0.002	0.01	1.319	0.0	0.007
CO2CRC_SCPC_C	2804.724	0.11	30.0	38.069	43.15	0.002	0.009	1.319	0.0	0.007
CO2CRC_SCPC_D	5034.121	0.11	30.0	31.147	39.554	0.002	0.006	1.465	0.031	0.006
CO2CRC_SCPC_F	2948.556	0.11	30.0	31.147	53.937	0.002	0.008	1.465	0.0	0.006
CO2CRC_SCPC_G	2056.798	0.11	30.0	31.147	50.341	0.002	0.008	1.465	0.0	0.006
Dillon_CFB_A	3872.919	0.11	30.0	nan	nan	0.0	0.0	0.972	0.0	0.0
Dillon_SUBC_A	2807.866	0.11	30.0	nan	nan	0.0	0.0	0.686	0.0	0.0
Dillon_SUBC_B	2517.397	0.11	30.0	nan	nan	0.0	0.0	0.972	0.0	0.0
Dillon_SUBC_C	1355.522	0.11	30.0	nan	nan	0.0	0.0	0.972	0.0	0.0
Dillon_SUBC_D	871.407	0.11	30.0	nan	nan	0.0	0.0	0.686	0.0	0.0

Table 10: Harmonized values of carbon captue studies, continuation

label	lcoe_fu	lcoe_capex_cc	lcoe_om_cc	lcoe_fu_cc	basis	heat_basis	lcoe	lcoe_cc	captured	cost_of_cc
Bergerson_Lave_a_SCPC_A	0.026	0.048	0.046	0.035	EUR2019	HHV	0.085	0.13	1.009	43.787
Bergerson_Lave_a_IGCC_A	0.03	0.047	0.03	0.035	EUR2019	HHV	0.099	0.112	0.911	14.619
Bergerson_Lave_a_SCPC_B	0.026	0.048	0.046	0.035	EUR2019	HHV	0.067	0.13	1.009	61.664
Bergerson_Lave_a_IGCC_B	0.03	0.047	0.03	0.035	EUR2019	HHV	0.091	0.112	0.911	22.908
Bergerson_Lave_b_SCPC_A	0.006	0.049	0.039	0.008	EUR2019	HHV	0.056	0.096	0.945	42.391
Bergerson_Lave_b_IGCC_A	0.025	0.054	0.015	0.03	EUR2019	HHV	0.068	0.098	0.984	30.596
Bergerson_Lave_b_NGCC_A	0.012	0.022	0.057	0.014	EUR2019	HHV	0.067	0.093	0.458	56.632
NREL_NGCC_A	0.011	0.051	0.012	0.018	EUR2019	nan	0.03	0.081	0.516	99.104
NREL_SCPC_A	0.014	0.09	0.01	0.019	EUR2019	nan	0.054	0.119	1.106	58.385
NREL_IGCC_A	0.014	0.091	0.015	0.018	EUR2019	nan	0.071	0.124	1.036	51.186
NREL_SCPC_B	0.014	0.095	0.016	0.019	EUR2019	nan	0.054	0.13	1.106	68.629
CAESAR_SCPC_A	0.047	0.064	0.004	0.064	EUR2019	LHV	0.076	0.132	0.935	59.383
CAESAR_IGCC_A	0.046	0.056	0.015	0.059	EUR2019	LHV	0.085	0.129	0.851	52.454
CAESAR_NGCC_A	0.032	0.02	0.012	0.037	EUR2019	LHV	0.047	0.07	0.37	62.748
GCCS_SCPC_A	0.017	0.042	0.009	0.024	EUR2019	HHV	0.04	0.075	1.04	33.407
GCCS_SCPC_B	0.016	0.042	0.009	0.023	EUR2019	HHV	0.04	0.074	1.003	33.73
GCCS_UCPC_A	0.015	0.041	0.008	0.02	EUR2019	HHV	0.039	0.069	0.854	35.314
GCCS_IGCC_A	0.016	0.042	0.006	0.023	EUR2019	HHV	0.046	0.07	0.976	24.715
GCCS_NGCC_A	0.027	0.018	0.002	0.031	EUR2019	HHV	0.035	0.051	0.379	42.535
IEAGHG_SCPC_A	0.016	0.049	0.015	0.02	EUR2019	LHV	0.045	0.085	0.841	46.96
IEAGHG_IGCC_A	0.016	0.059	0.026	0.02	EUR2019	LHV	0.061	0.105	0.935	47.563
Rubin_SCPC_A	0.016	0.045	0.034	0.021	EUR2019	HHV	0.063	0.1	0.959	38.187
Rubin_IGCC_A	0.017	0.03	0.021	0.017	EUR2019	HHV	0.067	0.067	0.74	0.194
Rubin_NGCC_A	0.011	0.021	0.024	0.013	EUR2019	HHV	0.042	0.058	0.387	40.523
NETL_SCPC_A	0.021	0.048	0.028	0.026	EUR2019	HHV	0.055	0.103	0.912	51.912
NETL_SUBC_A	0.02	0.049	0.028	0.025	EUR2019	HHV	0.055	0.102	0.875	53.731
NETL_NGCC_A	0.025	0.021	0.01	0.028	EUR2019	HHV	0.036	0.058	0.363	62.014
Davison_SCPC_A	0.032	0.042	0.015	0.041	EUR2019	LHV	0.078	0.098	0.845	22.589
Davison_SCPC_B	0.032	0.043	0.016	0.04	EUR2019	LHV	0.078	0.099	0.834	24.299
Davison_SCPC_C	0.032	0.046	0.017	0.04	EUR2019	LHV	0.078	0.103	0.831	29.957
Davison_IGCC_A	0.033	0.046	0.017	0.041	EUR2019	LHV	0.086	0.104	0.858	21.608
Davison_NGCC_A	0.016	0.018	0.007	0.019	EUR2019	LHV	0.032	0.044	0.4	27.968
Davison_NGCC_B	0.016	0.019	0.007	0.018	EUR2019	LHV	0.032	0.043	0.382	28.574
Davison_NGCC_C	0.016	0.032	0.012	0.02	EUR2019	LHV	0.032	0.064	0.424	74.207
Dave_from_Hu_SCPC_A	0.017	0.016	0.008	0.023	EUR2019	HHV	0.029	0.047	1.004	18.503
Li_NGCC_A	0.025	0.027	0.014	0.029	EUR2019	HHV	0.046	0.071	0.451	56.175
Li_SCPC_A	0.023	0.017	0.009	0.027	EUR2019	HHV	0.037	0.053	0.774	19.818
Li_SUBC_A	0.028	0.021	0.012	0.035	EUR2019	HHV	0.044	0.068	1.014	23.284
Wu_SCPC_A	0.024	0.03	0.0	0.033	EUR2019	LHV	0.04	0.063	0.836	28.149
Wu_IGCC_A	0.023	0.065	0.001	0.028	EUR2019	LHV	0.061	0.094	0.719	45.906
Renner_SCPC_A	0.037	0.06	0.027	0.047	EUR2019	LHV	0.074	0.134	0.819	73.262
Renner_NGCC_A	0.053	0.029	0.016	0.061	EUR2019	LHV	0.068	0.107	0.567	68.354
Renner_SCPC_B	0.033	0.025	0.021	0.041	EUR2019	LHV	0.047	0.087	0.819	49.197
Renner_NGCC_B	0.047	0.014	0.003	0.054	EUR2019	LHV	0.054	0.07	0.567	29.367
CO2CRC_SCPC_A	0.018	0.083	0.015	0.024	EUR2019	LHV	0.063	0.122	1.828	32.206
CO2CRC_SCPC_B	0.018	0.049	0.017	0.031	EUR2019	LHV	0.025	0.096	2.299	31.254
CO2CRC_SCPC_C	0.018	0.039	0.014	0.024	EUR2019	LHV	0.025	0.078	1.828	29.089
CO2CRC_SCPC_D	0.014	0.07	0.012	0.02	EUR2019	LHV	0.051	0.102	1.092	46.756
CO2CRC_SCPC_F	0.014	0.041	0.015	0.02	EUR2019	LHV	0.02	0.076	1.092	51.461
CO2CRC_SCPC_G	0.014	0.029	0.014	0.02	EUR2019	LHV	0.02	0.063	1.092	39.625
Dillon_CFB_A	0.014	0.054	0.0	0.032	EUR2019	HHV	0.014	0.086	2.798	25.706
Dillon_SUBC_A	0.008	0.039	0.0	0.012	EUR2019	HHV	0.008	0.051	1.435	29.99
Dillon_SUBC_B	0.011	0.035	0.0	0.017	EUR2019	HHV	0.011	0.053	1.571	26.336
Dillon_SUBC_C	0.011	0.019	0.0	0.017	EUR2019	HHV	0.011	0.036	1.782	14.311
Dillon_SUBC_D	0.007	0.012	0.0	0.01	EUR2019	HHV	0.007	0.022	1.257	11.933

Table 11: Harmonized values of carbon captue studies, continuation

label	cc_capex	cc_om	cc_fu
Bergerson_Lave_a_SCPC_A	7.558	26.755	9.474
Bergerson_Lave_a_IGCC_A	-6.403	15.316	5.707
Bergerson_Lave_a_SCPC_B	16.558	35.632	9.474
Bergerson_Lave_a_IGCC_B	0.246	16.955	5.707
Bergerson_Lave_b_SCPC_A	11.288	29.27	1.833
Bergerson_Lave_b_IGCC_A	10.865	14.353	5.379
Bergerson_Lave_b_NGCC_A	19.099	33.19	4.343
NREL_NGCC_A	71.756	14.314	13.033
NREL_SCPC_A	50.7	3.306	4.379
NREL_IGCC_A	41.952	5.225	4.01
NREL_SCPC_B	55.169	9.081	4.379
CAESAR_SCPC_A	37.252	3.809	18.322
CAESAR_IGCC_A	31.47	5.445	15.539
CAESAR_NGCC_A	29.959	18.251	14.538
GCCS_SCPC_A	22.12	4.315	6.973
GCCS_SCPC_B	22.624	4.649	6.457
GCCS_UCPC_A	24.978	4.507	5.829
GCCS_IGCC_A	15.778	1.819	7.118
GCCS_NGCC_A	27.765	3.225	11.544
IEAGHG_SCPC_A	33.021	9.125	4.814
IEAGHG_IGCC_A	39.212	4.201	4.151
Rubin_SCPC_A	10.954	22.057	5.177
Rubin_IGCC_A	-9.919	10.112	0.0
Rubin_NGCC_A	12.613	23.0	4.91
NETL_SCPC_A	30.942	15.216	5.754
NETL_SUBC_A	32.388	15.573	5.77
NETL_NGCC_A	37.533	15.931	8.549
Davison_SCPC_A	5.612	6.911	10.066
Davison_SCPC_B	7.305	7.475	9.519
Davison_SCPC_C	11.416	9.131	9.41
Davison_IGCC_A	4.792	7.267	9.55
Davison_NGCC_A	12.997	8.022	6.949
Davison_NGCC_B	14.588	8.901	5.085
Davison_NGCC_C	45.056	19.914	9.237
Dave_from_Hu_SCPC_A	8.618	3.432	6.452
Li_NGCC_A	32.445	13.054	10.676
Li_SCPC_A	10.853	4.296	4.67
Li_SUBC_A	10.893	5.35	7.042
Wu_SCPC_A	16.961	0.2	10.988
Wu_IGCC_A	38.847	0.391	6.669
Renner_SCPC_A	43.577	18.293	11.392
Renner_NGCC_A	36.351	17.624	14.38
Renner_SCPC_B	16.832	22.39	9.975
Renner_NGCC_B	14.166	2.495	12.706
CO2CRC_SCPC_A	24.307	4.199	3.7
CO2CRC_SCPC_B	21.415	4.165	5.674
CO2CRC_SCPC_C	21.444	3.945	3.7
CO2CRC_SCPC_D	36.04	5.336	5.38
CO2CRC_SCPC_F	37.726	8.355	5.38
CO2CRC_SCPC_G	26.316	7.93	5.38
Dillon_CFB_A	19.341	0.0	6.366
Dillon_SUBC_A	27.345	0.0	2.646
Dillon_SUBC_B	22.386	0.0	3.95
Dillon_SUBC_C	10.629	0.0	3.681
Dillon_SUBC_D	9.686	0.0	2.247

Appendix B: Correlations

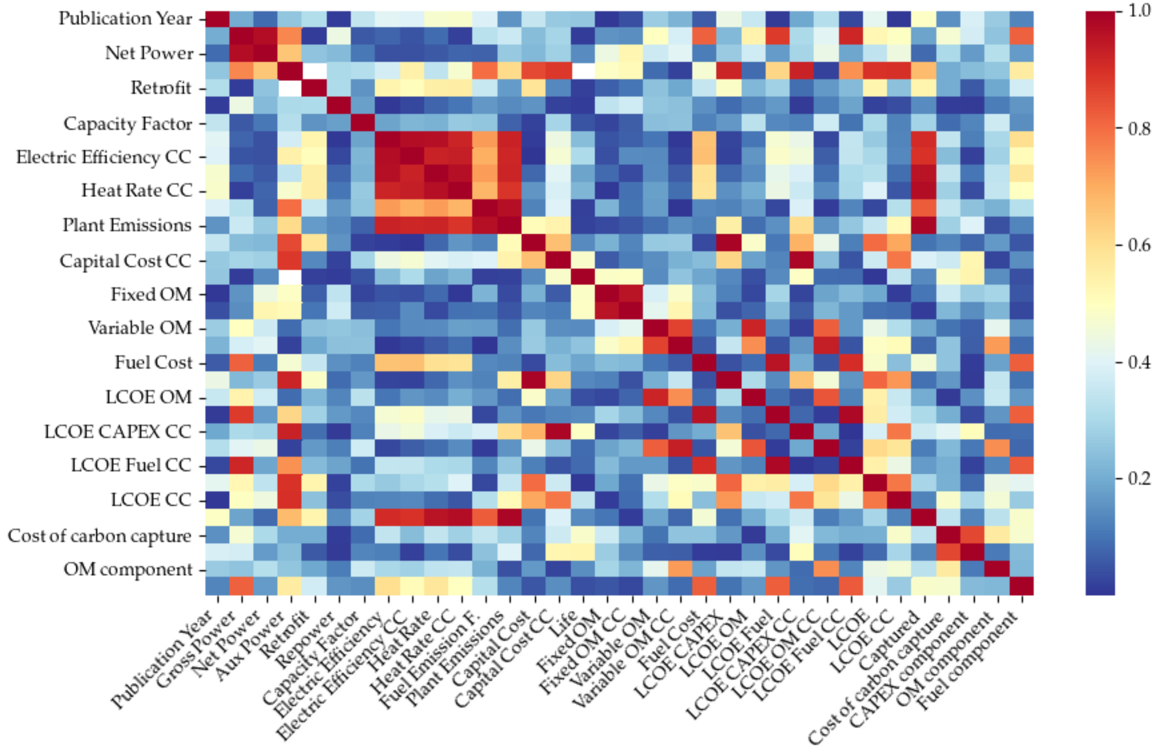


Figure 1: Correlation matrix of all the input and output variables from the harmonization process

Bibliography

Kolmogorov–smirnov test. In *The Concise Encyclopedia of Statistics*, pages 283–287. Springer New York, New York, NY, 2008.

Danish Energy Agency. Technology data generation of electricity and district heating. Technical report, 2016.

Bilogur. Bilogur, (2018). missingno: a missing data visualization suite. journal of open source software, 3(22), 547, https://doi.org/10.21105/joss.00547. *Bilogur, (2018). Missingno: a missing data visualization suite. Journal of Open Source Software.*, 3(22):547, 2018.

Leo Breiman. Random forests. *Machine Learning*, 45:5–32, 2001.

Schlachtberger D. Brown T., Hörsch J. Pypsa: Python for power system analysis. , Date 2018.

S. Budinis, S. Krevor, N. Mac Dowell, N. Brandon, and A. Hawkes. An assessment of ccs costs, barriers and potential. *Energy Strategy Reviews*, 22:61–81, 2018.

Mai Bui, Claire S. Adjiman, André Bardow, Edward J. Anthony, Andy Boston, Solomon Brown, Paul S. Fennell, Sabine Fuss, Amparo Galindo, Leigh A. Hackett, Jason P. Hallett, Howard J. Herzog, George Jackson, Jasmin Kemper, Samuel Krevor, Geoffrey C. Maitland, Michael Matuszewski, Ian S. Metcalfe, Camille Petit, Graeme Puxty, Jeffrey Reimer, David M. Reiner, Edward S. Rubin, Stuart A. Scott, Nilay Shah, Berend Smit, J. P. Martin Trusler, Paul Webley, Jennifer Wilcox, and Niall Mac Dowell. Carbon capture and storage (ccs): the way forward. *Energy Environmental Science*, 11(5):1062–1176, 2018.

Bundesministerium der Justiz und für Verbraucherschutz. Gesetz zur reduzierung und zur beendigung der kohleverstromung, 2020.

ChemicalEngineeringOnline. 2019 chemical engineering plant cost index annual average. `https://www.chemengonline.com/`, Date 2019.

CO2CRC. Retrofitting ccs to coal: Enhancing australia's energy security. Technical report, 2017.

Solazzo E. Huang G. Guizzardi D. Koffi E.-Muntean M. Schieberle C. Friedrich R. Crippa, M. and G. Janssens-Maenhout. High resolution temporal profiles in the emissions database for global atmospheric research. , Date 2020.

Desmond Dillon, John Wheeldon, Robert Chu, Gerald Choi, and Christopher Loy. A summary of epri's engineering and economic studies of post combustion capture retrofit applied at various north american host sites. *Energy Procedia*, 37:2349–2358, 2013.

EEA. Eprtr. `https://prtr.eea.europa.eu/#/faq`, Date 2019.

EIA. Henry hub natural gas spot price. Type, Institution, Date 2020a.

EIA. Assumptions to the annual energy outlook 2020: Coal market module. Technical report, 2020b.

EIA. Coal marktets. `https://www.eia.gov/coal/markets/#tabs-prices-2`, Date 2020c.

EUROSTAT. Nuts administrative units/statistical units. `https://ec.europa.eu/eurostat/de/web/gisco/geodata/reference-data/administrative-units-statistical-units/nuts`, Date 2020.

J. Hörsch F. Gotzens, H. Heinrichs and F. Hofmann. Performing energy modelling exercises in a transparent way - the issue of data quality in power plant databases. *Energy Strategy Reviews*, 23:1–12, 2019.

F20. Achieving the paris climate goals in the covid-19 era. Technical report, 2020.

Thomas Fröhlich. The cost of carbon capture and storage for natural gas combined cycle power plants. Technical report, 2019.

Christoph Gaeldicke. Data and developments concerning german and global energy supply. Technical report, 2019.

Stefania Osk Gardarsdottir, Edoardo De Lena, Matteo Romano, Simon Roussanaly, Mari Voldsund, José-Francisco Pérez-Calvo, David Berstad, Chao Fu, Rahul Anantharaman, Daniel Sutter, Matteo Gazzani, Marco Mazzotti, and Giovanni Cinti. Comparison of technologies for co2 capture from cement production—part 2: Cost analysis. 12(3):542, 2019.

GCCS. Economic assessment of carbon capture and storage technologies. Technical report, 2011.

Aurelien Geron. *Hands-on machine learning with Scikit-Learn and TensorFlow : concepts, tools, and techniques to build intelligent systems.* O'Reilly Media, Beijing ; Boston, first edition. edition, 2017.

The World Bank Group. Commodity markets. Technical report, 2020.

Johannes; Gütschow, Louise; Jeffery, Robert; Gieseke, and Annika Günther. The primap-hist national historical emissions time series (1850-2017). , Date 2019.

Bingyin Hu and Haibo Zhai. The cost of carbon capture and storage for coal-fired power plants in china. *International Journal of Greenhouse Gas Control*, 65:23–31, 2017.

Wenxuan Hu. *Potential analysis of industrial CO2 sources for the provision of synthetic fuels.* PhD thesis, University ofStuttgart, 2019.

IEA. Technology brief e01 coal-fired power. Technical report, 2010.

IEAGHG. Co2 captured at coal based power and hydrogen plants. Technical report, 2014.

IECM. Amine based post combustion co2 capture. Technical report, 2019.

IHS. UOCI. `https://ihsmarkit.com/info/cera/ihsindexes/index.html`, Date 2020a.

IHS. UCCI. `https://ihsmarkit.com/info/cera/ihsindexes/index.html`, Date 2020b.

industryabout. Fossil fuel energy. `https://www.industryabout.com/energy/5575-fossil-fuels-energy`, Date 2019.

IPCC. Carbon capture and storage. Technical report, 2005.

Clemens Gerbaulet Friedrich Kunz Jens Weibezahn, Richard Weinhold. Conventional power plants. Technical report, 2018.

Diaz Vazquez A. Schade B. Tchung-Ming S. Weitzel M. Vandyck T. Wojtowicz K. Keramidas K., Fosse F. Global energy and climate 2020: A new normal beyond covid-19. Technical report, 2021.

Dirk Uwe Sauer Klaus Görner. Konventionelle kraftwerke. Technical report, Nationale Akademie der Wissenschaften Leopoldina acatech – Deutsche Akademie der Technikwissenschaften Union der deutschen Akademien der Wissenschaften, 2016.

Takeshi Kuramochi, Andrea Ramírez, Wim Turkenburg, and André Faaij. Comparative assessment of co2 capture technologies for carbon-intensive industrial processes. *Progress in Energy and Combustion Science*, 38(1):87–112, 2012.

Ung Lee, Alexander Mitsos, and Chonghun Han. Optimal retrofit of a co2 capture pilot plant using superstructure and rate-based models. *International Journal of Greenhouse Gas Control*, 50: 57–69, 2016.

Macrotrends. Euro dollar exchange rate (eur usd) - historical chart. `https://www.macrotrends.net/2548/euro-dollar-exchange-rate-historical-chart`, Date 2020.

Hari C. Mantripragada, Haibo Zhai, and Edward S. Rubin. Boundary dam or petra nova – which is a better model for ccs energy supply? *International Journal of Greenhouse Gas Control*, 82: 59–68, 2019.

Markus Millinger, Tafarte Philip, Jordan Matthias, Hahn Alena, Meisel Kathleen, and Thrän Daniela. *Electrofuels from Excess Renewable Electricity at High Variable Renewable Shares: Cost, Greenhouse Gas Abatement, Carbon Use and Competition.* 2020.

Kenneth Möllersten, Jinyue Yan, and Jose R. Moreira. Potential market niches for biomass energy with co2 capture and storage—opportunities for energy supply with negative co2 emissions. *Biomass and Bioenergy*, 25(3):273–285, 2003.

Jorge Morales Pedraza. Chapter 4 - current status and perspective in the use of coal for electricity generation in the north america region. In Jorge Morales Pedraza, editor, *Conventional Energy in North America*, pages 211 – 257. Elsevier, 2019. ISBN 978-0-12-814889-1. doi: https://doi.org/10.1016/B978-0-12-814889-1.00004-8. URL `http://www.sciencedirect.com/science/article/pii/B9780128148891000048`.

Henriette Naims. Economics of carbon dioxide capture and utilization—a supply and demand perspective. *Environmental Science and Pollution Research*, 23(22):22226–22241, 2016.

openmod. openmod initiative wiki. `https://wiki.openmod-initiative.org/wiki/Main_Page`, Date 2020.

Simon Roussanaly, Amy L. Brunsvold, Erik S. Hognes, Jana P. Jakobsen, and Xianping Zhang. Integrated techno-economic and environmental assessment of an amine-based capture. *Energy Procedia*, 37:2453–2461, 2013.

Edward S. Rubin, Christopher Short, George Booras, John Davison, Clas Ekstrom, Michael Matuszewski, and Sean McCoy. A proposed methodology for co2 capture and storage cost estimates. *International Journal of Greenhouse Gas Control*, 17:488–503, 2013.

Edward S. Rubin, John E. Davison, and Howard J. Herzog. The cost of co2 capture and storage. *International Journal of Greenhouse Gas Control*, 40:378–400, 2015.

Dipak K. Sarkar. Chapter 2 - steam generators. In Dipak K. Sarkar, editor, *Thermal Power Plant*, pages 39 – 89. Elsevier, 2015. ISBN 978-0-12-801575-9. doi: https://doi.org/10.1016/B978-0-12-801575-9.00002-0. URL `http://www.sciencedirect.com/science/article/pii/B9780128015759000020`.

Yvonne Scholz. *Renewable energy based electricity supply at low costs.* PhD thesis, Universität Stuttgart, 2012.

Michael Sterner. Bioenergy and renewable power methane in integrated 100energy systems. Technical report, Universität Kassel, 2009.

Alla Toktarova, Ida Karlsson, Johan Rootzén, Lisa Göransson, Mikael Odenberger, and Filip Johnsson. Pathways for low-carbon transition of the steel industry—a swedish case study. *Energies*, 13(15):3840, Jul 2020. ISSN 1996-1073. doi: 10.3390/en13153840. URL `http://dx.doi.org/10.3390/en13153840`.

Richard Turton. *Analysis, synthesis, and design of chemical processes.* Prentice Hall international series in the physical and chemical engineerin sciences. Prentice Hall, Upper Saddle River, NJ, 5 edition, 2018.

Black Veatch. Cost and performance data for power generation technologies. Technical report, 2012.

Niklas Von der Assen. Selecting co2 sources for co2 utilization by environmental-merit-order curves. 2016.

Berthold Kaufmann Passivhaus Institut Darmstadt Witta Ebel, Wolfgang Feist. Economy and financing of efficiency: new buildings, renovation and step by step retrofit. Technical report, Passivhaus Institut, 2019.

Haibo Zhai. Advanced membranes and learning scale required for cost-effective post-combustion carbon capture. *iScience*, 13:440–451, 2019.

Di Zhang, Mai Bui, Mathilde Fajardy, Piera Patrizio, Florian Kraxner, and Niall Mac Dowell. Unlocking the potential of beccs with indigenous sources of biomass at a national scale. *Sustainable Energy Fuels*, 4(1):226–253, 2020.

Shan Zhou, Shujuan Wang, Chenchen Sun, and Changhe Chen. So2 effect on degradation of mea and some other amines. *Energy Procedia*, 37:896–904, 2013.