Universiteit Utrecht

Carbon Intensity Indicators

A comparison of carbon intensity indicators to benchmark companies within a sector

Master Thesis Energy Science
System Analysis track
Utrecht University in cooperation with Ecofys

Date: September 18th 2012 - March 12th 2013
ECTS: 37.5 ECTS

Author: Esther Eggink - 3151271
e.e.eggink@students.uu.nl

Supervisor Ecofys: Giel Linthorst
Manager unit Supply Chains at Ecofys
Supervisor Utrecht University: Prof. dr. Kornelis Blok
Professor of Sustainable Energy, Utrecht University

Executive summary

Global action is needed to ensure that the global temperature rise is limited to a maximum of 2°C. There is a growing gap between where global emissions are heading and where they need to be (UNEP, 2011). The private sector is responsible - and should take the responsibility - for one wedge of this gap. Benchmarking at company level is a powerful tool to enable target setting and monitor emission reductions. The first part of this study shows the inventory that was created of the current sustainability rankings and the carbon indicators used. From the literature used for this overview, it can be concluded that there are no global company rankings based on quantitative carbon emissions. World-wide, emission intensities across sectors vary in such a way, that a sector specific approach is to be preferred. A target based benchmark, on the basis of which the carbon performance of the largest companies of a sector can be compared, would enable sector wide reduction targets.
The aim of this study was to find the carbon intensity indicator that best measures the technical and operational carbon performance at company level, for the electricity sector, the air transport sector and the steel sector.

In this study, an emission parameter and an activity parameter were combined to form a carbon intensity indicator. The activity parameter was either a monetary or a physical parameter. For each sector, both monetary and physical carbon intensity indicators were selected. Company data for each of these resources were drawn from publicly available sources. This data was used to rate both the data quality, by means of a quantification model, and the aggregation quality. The aggregation quality was determined by the initial aggregation level of the data and the quality of re-aggregation to company level. At a lower level of aggregation, data was more product specific and structural effects were less, but re-aggregation to company level was required. And in case of multiple product indicators, the way to combine them to one indicator at company level should be researched.
In addition to the data and aggregation quality rating, interviews were held with sector stakeholders to obtain feedback on the selected indicators, their benchmark experiences and their opinions.

Sector specific indicators with adequate levels of both data and aggregation quality were found for two of the three sectors. The selected indicator for the steel sector was useful for a first estimation, but, for it to be of more use, would require additional data to further disaggregate to primary and secondary steel. The selected indicators employed for the electricity sector, the air transport sector and the steel sector were respectively 'Scope 1 emissions (with heat allocation) per megawatt hour (MWh) produced electricity', 'Scope 1 emissions per total passenger kilometres' and 'tonne CO_2-e per tonne crude steel'. As regards the air transport indicator, a suitable way was found to combine cargo tonne kilometres and passenger kilometres by means of a weight translation factor. A volume correction factor was also researched, but the data quality of this factor turned out to be too poor to be used. The main drawback of the quality of this last indicator was the comparability and the availability of high quality data.

The stakeholders interviewed in the electricity and the air transport sector approved of the selected physical indicators for their sectors and found the monetary indicator less use full. According to the stakeholder of the steel sector, further disaggregation of the selected physical indicator would be required for it to become functional. The quality of the publicly available data was regarded as the biggest problem by all stakeholders.

Company level benchmarking will be most useful for high emitting sectors in which the majority of emissions can be allocated to only a few products or activities; however, it will always entail a certain level of aggregation and is therein different from product specific benchmarking. This higher level of aggregation makes company level benchmarking an excellent tool for target setting and guiding emission reductions in the sector. These overviews can be used by sector organisations, NGO's and governments.

It is recommended to further improve the quality of emission and performance data in all three sectors, by means of clear, internationally accepted reporting guidelines for each one of them. As thorough operational knowledge are essential, it can best be achieved in co-operation with sector organisations and the CDP, and might also be of interest to the Global Reporting Initiative. The guidelines should include consensus on the reporting methodology as well as on the level of aggregation of the benchmark indicator.

For the steel industry, more in depth research is needed to obtain an adequate level of aggregation quality, and disaggregation to primary and secondary steel is recommended. For further research it is advisable to perform a similar analysis of other high emitting sectors; data disclosure agreements are highly recommended to improve the quality of research.

Acknowledgement

Over the past six months I have worked hard, and with much enthusiasm, to write this master thesis about carbon intensity indicators at Ecofys. This would not have been possible without the help and support of the people around me.

First of all, I would sincerely like to thank my two supervisors Giel Linthorst and Kornelis Blok. Giel has been an amazing coach to me and supported me throughout the challenging process of conducting this research. And Kornelis' great scientific expertise helped me to deal with the content.

Furthermore, I would like to thank my Ecofys colleagues for their valuable advice and for making me feel at ease at Ecofys. Special thanks go to the colleagues in my office who tirelessly answered all my questions, and more importantly, provided me with a stimulating and lively work atmosphere.

Finally, I would like to thank my friends and family for their support and encouragement throughout the process, and especially my aunt, who, because of time-constraints, kindly agreed to perform a last-minute language-check on my work

Universiteit Utrecht

ECOFYS

Contents

1. Introduction

The effect of greenhouse gases (GHG) on global warming has been known about for more than a century (Arrhenius, 1906). At the global Earth Summit in Rio de Janeiro in 1992, nations agreed on a Framework Convention on Climate Change. In 1997, the Kyoto Protocol was adopted which legally binds developed countries to emission reduction targets up to 2012. During recent UNFCCC climate conferences in Copenhagen (2009), Cancun (2010) and Durban (2011) it was internationally agreed that global average temperature rise should be limited to below 2°C to prevent dangerous anthropogenic interference with the climate system.

Nevertheless GHG emissions are still rising in spite of the commitments made by the leaders of the international community at these conferences. There is a growing gap between where global emissions are heading (under current national emission reduction pledges), and where they need to be to ensure global temperature rise is limited to max 2°C (UNEP, 2011). To highlight this trend, the United Nations Environment Programme (UNEP) published a synthesis report called 'Bridging the Emissions Gap' in November 2011 (UNEP, 2011) and an update of this report in 2012 (UNEP, 2012). In these reports it is shown that by 2020, the emission gap between business-as-usual emissions (i.e. if no pledges are implemented) and emissions consistent with the maximum of 2°C temperature rise target scenario would be about twelve gigatonne CO_2-equivalent (Gt CO_2-e). This is nearly as much as current total GHG emissions of the world's energy supply sector. Altogether the pledges made by the national governments add up to a maximum of six Gt CO_2-e. This means that, at the most, only half of the gap is addressed national governments under the UNFCCC (Figure 1.1).

Despite this current lack of commitment by national governments, the report underlines an economically viable emission reduction potential of up to 17±3 Gt CO_2-e by 2020 (UNEP, 2011). As a response, employees of the firm Ecofys published the article 'Wedging the emission Gap' in Nature Climate Change in June 2012. They proposed a bottom-up approach of 21 major global initiatives which together come to produce an emission reduction of ten Gt CO_2-e by 2020. A big part of these reductions would come from company initiatives and umbrella organisations. In this approach, leading organisations within different sectors play an important role to involve the different stakeholders of the wedges. The total emission reduction of this approach goes significantly beyond the pledges towards the UNFCCC (Blok, et al., 2012).

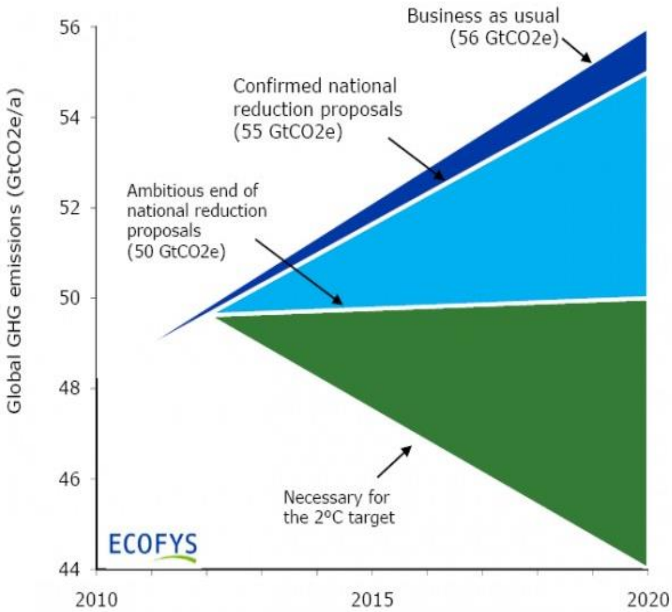

Figure 1.1: Wedging the emission gap (Blok, et al., 2012).

In recent years this growing attention for emission reductions has resulted in an increased amount of sustainability indexes and rankings (ERM, 2010). Both at sector and company level different methods of involvement and emission reduction targets are set. Companies realise that optimising their efficiency in corporate operations reduces their carbon emissions, stimulates innovation and achieves cost savings (CDP, 2009). Other drivers behind these initiatives are increased regulation, increased energy prices, climate change risks and an increased demand for sustainability from investors and consumers (ERM, 2010). But the wide range of targets is not directly comparable and it is difficult to judge the impact. The absence of a standard framework for setting emissions reduction targets has led to a patchwork of company specific targets, which have developed from individual company priorities and market forces (CDP, 2009). Research carried out by the Carbon Disclosure Project among companies concluded that: "harmonisation of targets has advantages, but a 'one-size-fits-all' cross-industry approach is not a favoured option within a voluntary process." A more sector specific approach would reduce the influence of sector specific elements.

Both companies and sector organisations agree that benchmarking is a powerful tool to improve performance and to compare competitors. Many large companies have put a lot of effort and money into their ranking, which in turn increased the amount of commercial rankings with different approaches. Several stakeholders started to act at approximately the same time, which led to the development of multiple approaches. An overview of the different methodologies and approaches can be found in a report from ERM called: "Company GHG Emissions Reporting" (ERM 2010). The threshold to participate is, in most of the voluntary reporting initiatives, relatively low compared to the obligatory initiatives. In most initiatives only general protocols are used to set out the reporting guidelines like setting boundaries, calculating emissions and verification. This is not specific enough to meet the needs of some sectors (ERM, 2010). To ensure that the overall policy goal of maximising economy-wide GHG reductions is met, any new scheme must include strong measures aimed at setting company GHG reduction targets (ERM, 2010).

Environmental Resources Management (ERM) compared different methods and initiatives to report GHG emissions on behalf of the European Commission. They conclude that "Few methods or initiatives provide incentives such as benchmarks, league tables and financial penalties/rewards. And only a few of the major methods and initiatives link GHG reporting to target setting." They identified this as an obvious gap in the way towards significant economy wide emission reductions (ERM, 2010). This report and the increased attention to sustainability rankings by companies, show that companies are willing to improve. But there is a need for clear benchmarking, in order that targets can be set that actually bridge the emission gap and that industries can be easily monitored by governments or non-governmental organisation (NGO) programmes (or both). Companies can also thus be supported and rewarded for any effective carbon emission reducing initiatives.

Currently, most available benchmarks give a false impression of the actual carbon emission performance of companies, because they use many different indicators to compare overall sustainability. The outcome heavily depends on the Key Performance Indicators (KPI) used. In order to bridge the gap, a solely carbon emission indicator is needed to benchmark carbon emissions performance. In addition, many rankings compare the emission intensity of companies across sectors while by nature these can be very different. This study will focus on a sector specific approach. 'Wedging the gap' shows initiatives with a large reduction potential for different sectors (Blok, et al., 2012). To set off these initiatives, specific carbon reduction targets should be set by companies and sectors. A carbon benchmark is needed to compare those companies. In this research, different parameters for carbon benchmarking are compared for different sectors in order to determine the best fitting benchmark indicator per sector and to enable sector specific target setting.

2. Problem definition

First the aim and demarcation of this research will be presented, followed by the research question and sub-questions. At the end of this chapter a general research outline is given for further clarification.

2.1 Aim

In the introduction it is explained why bottom-up emission reduction initiatives are needed. Benchmarking at company level is a powerful tool to support those emission reduction initiatives and establish real emission reduction targets. Different intensity indicators can be used to benchmark carbon performance. The best indicator to use for company level benchmarking depends on the available data quality, the carbon intensity of the products and the homogeneity of the sector. These aspects significantly differ per sector. The aim of this research is to inventories the current indicators used in sustainability benchmarks. And to select the carbon intensity indicator most appropriate to measure the technical and operational carbon efficiency of the companies' activities; the selected indicator can be used for target setting at company level within a certain sector.

2.2 Demarcation

This research has been carried out within the following framework:

- Carbon indicators are researched, but no overall sustainability indicators, in order to ensure real economy-wide emission reductions.
- Only carbon indicators which can be used for target setting and benchmarking at company level are researched.
- Only '*intensity*' indicators are researched, since indicators for absolute emission reduction targets cannot be used to compare different sized companies.
- Carbon indicators comparing the progress of companies' progress over the years are not researched, as this study focuses on the comparison of carbon performance between companies within the same sector.
- Due to time limitations, only the electricity sector, air transport sector and steel sector will be researched. The methodology used might be useable for other sectors, too. The selected sectors are among the most energy intensive ones, and represent different levels of complexity. The sector selection is further explained in paragraph 3.3.
- The emission data is based on the Greenhouse Gas Protocol Corporate Accounting (WRI/WBCSD, 2004). Thus Scope 1 and 2 emissions of the entire companies will be compared. Scope 3 emissions will be excluded since very little data is available, and double counting must be avoided[1].
- Also in order to prevent the comparison between apples and oranges, a scientific methodology is applied to increase the quality of the currently used indicators.
- Comparison is made at company level, not at product level; therefore a certain amount of aggregation will be required.

[1] The direct emissions of a company are Scope 1 emissions. The indirect emissions due to electricity or heat use are Scope 2 emissions. All the emissions in the value chain of a company are Scope 3 emissions. These concepts are further explained in paragraph 3.1 "Important definitions".

2.2 Research question

This study is based on the following research question:

"Which carbon intensity benchmark indicator for target setting purposes, best measures the technical and operational carbon performance at company level, and can be measured with publicly available data in the industries: electricity, air transport and steel?"

2.3 Sub-questions

Sub-questions were formed to help solve the main research question.

- Which quantitative intensity indicators are used in the current available sustainability benchmarks to rank greenhouse gas emissions?
- In literature, which criteria determine the quality of carbon intensity indicators?
- How can the selected criteria be combined to rate and compare indicators?
- What is the opinion of stakeholders of the specific sector on the selected indicators, the created rankings and the used methodology?

2.4 Research outline

In order to find an answer to the research question and the sub-questions, literature was consulted first. An overview of the current experiences with benchmarking has been provided in the research background of Chapter 3. In this chapter the boundaries and definitions of key concepts are described and the selection of sector is explained. Current sustainability benchmarks are analysed and a graphical overview is provided. Finally, it is shown how, literature was used to find criteria on which the comparison between indicators could be based. Chapter 4 explains which methodology was used to rank the indicators. The results of the indicator analyses of the three chosen sectors, i.e. electricity, air transport sector and steel, are discussed in chapter 5. At the end of each sector the pitfalls and sector specific conclusions are described. In Chapter 6 the final conclusion of this research is presented, followed by Chapter 7, in which discussion and recommendations can be found. References and appendices can be found at the end. The first two digits in tables and figures refer to the paragraph in which they are located.

3. Methodological background

The methodological background starts by outlining some important definitions and concepts used in this research, and is followed by an explanation of the sector selection. Next presented is a literature study on the benchmark methodologies currently used, which are subsequently reviewed. A graphical overview of these methodologies completes this section. More detailed information about these methodologies can be found in Appendix A. From literature are selected the criteria on which the indicator comparison will be based.

3.1 Important definitions

In this paragraph the definitions and concepts essential for this research are explained.

Benchmarking

Benchmarking provides a means to compare the carbon emissions within one company or plant to that of similar facilities, or to national or international best practice carbon emission levels. Benchmarking can compare plants, processes or systems (Price, et al., 2008).

Energy efficiency indicators

These indicators can be based on either economic or physical measures of activity. Studies indicate a poor relationship between economic and physical indicators. The latter are recommended, at least for the energy-intensive sectors of the manufacturing industry (Phylipsen, et al., 1997).

Carbon Intensity Indicators (CII)

Companies either use absolute emission reduction targets or intensity reduction targets. An intensity indicator is needed, to compare different sized companies. Carbon intensity indicators are comparable to energy efficiency indicators. Both indicators have a numerator, where the amount of carbon or energy is expressed, and a denominator, where the economic activity or product is expressed. But different energy carriers have different carbon contents, resulting in different carbon emissions per unit of energy. In order to use energy efficiency indicators, fuel-specific disaggregation is required (Phylipsen, et al., 1997). In this research only carbon intensity indicators are analysed. The numerator and denominator of the carbon intensity indicator are referred to as parameters and not as indicators, to clarify which part of the intensity indicator is meant. Economic activity can be measured in monetary terms or in physical terms (Phylipsen, et al., 1997). Therefore three different kinds of parameters are distinguished: an emission parameter, a monetary parameter and a physical parameter.

Emission categories

There are different ways to calculate the emissions of a company. Under the WBCSD/WRI GHG Protocol Corporate Standard there are three reporting sub-categories for emissions:

• **Scope 1** (direct GHG emissions) should always be reported. These emissions occur due to fuel combustion in activities that a company directly controls.

• **Scope 2** (electricity/heat, indirect GHG emissions) should always be reported. These emissions occur off-site at power stations due to electricity/heat usage by the reporting company. They are not strictly ‘direct’ emissions in that, they arise from third party installations but would normally be attributed to a company its operations as the end user of the electricity/heat.

• **Scope 3** (other indirect GHG emissions) is an optional reporting category. These emissions occur due to activities which are not directly controlled by the reporting company but are associated with and influenced by their activities.
In this research Scope 1 and 2 emissions are used as emission parameters. Scope 3 emissions are excluded because of the poor data availability and quality, and also to avoid double counting.

Monetary parameters
Value-based activity parameters can be provided at different levels of aggregation. E.g. GDP at national level, and value added at sectoral level (Phylipsen, et al., 1997). At company level different monetary parameters are used for economic purposes. The most product-related monetary parameters are turnover, revenue and sales. These are all directly related to the product by the formula: price*product. Other, often used economic parameters are profit, EBITDA and added value. These are less dependent on the product but partly also on external economic inputs. Different levels of aggregation can be distinguished i.e. holding turnover, company turnover, division turnover or unit turnover.

Physical parameters
Physical production basically represents the amount of saleable product. A product can either be a good or a service. A widely used physical activity parameter in the manufacturing industry is production volume (in tonnes of product). At a lower level of aggregation different types of product may be distinguished (Phylipsen, et al., 1997).

Aggregation
Many of the experts interviewed and consulted agree that comparing apples and oranges is one of the biggest pitfalls of benchmarking. Ideally, a comparison is made between indicators with the same denominator. But, in practice, there are few companies with only one, identical product. The definition of aggregate is taking all units as a whole, i.e. the total. Therefore, a certain level of aggregation is always needed to be able to compare on company level. At a high level of aggregation, one indicator is used for all products i.e. turnover. In this way all activities and all products are taken as a whole. At a lower level of aggregation; disaggregated, the whole company is separated into parts with the same uniform product. In the case of different products, a way to re-aggregate to one indicator is needed to reduce the structural effects when comparing at company level. This process is easier when emissions for each separate activity are known. Since, in this research, a comparison is made at company level, a trade-off has been made between the data quality and the level of aggregation.

Sector complexity
In this research a variety of sectors are selected. The selected sectors are relatively homogeneous which means they are less complex than heterogeneous sectors. Homogeneous sectors have fewer products which are clearly described products compared to heterogeneous sectors which have many products and include more process-related output. Examples of less complex sectors are the electricity sector, the cement sector and the manufacture of iron ores. These are all sectors with a clearly described bulk product. More complex sectors are the chemical sector, the land transport sector and the health sector.

3.2 Data sources

In this paragraph, the data sources are listed which are used to gather the company specific emission- and performance data.

Annual (sustainability) reports

Most companies provide an annual report with monetary and performance data. This data is often provided for investors and can therefore be considered reliable. In the last couple of years many companies have also provided an annual sustainability report, in which absolute emission data can often be found.

CDP Database

The Carbon Disclosure Project is an independent not-for-profit organisation holding the largest database of primary corporate climate change information in the world. Thousands of organisations, from across the world's major economies, measure and disclose their greenhouse gas emissions and climate change strategies through the CDP. The CDP puts this information at the heart of financial and policy decision-making (CDP, 2012).

PLATTS Power database contains data from all the electricity companies in Europe. Assets per energy supplier and the corresponding capacities per year are provided.

Emission Trading Scheme (ETS) database

This database is used for a cap-and-trade system regulated by the European Union. Participation is obligatory for the companies of certain energy intensive industries. Verified Scope 1 and 2 emissions have to be reported for each controlled plant. For this study the ETS database is combined with the PLATTS database and used for the analysis of the electricity sector.

3.3 Sector selection

In this study, sector specific emission intensity indicators are compared. Due to time limitations not all sectors can be analysed. In this paragraph the sector selection is explained.

Significant emission reductions are more likely to happen in high emitting sectors. First the sectors with many high emitting companies are selected. The 2000 biggest turnover companies are ranked based on company emissions in 2008 (Jong, 2011). 240 companies of this ranking emit more than ten million tonnes of CO_2-e per year. Classifying these 240 companies into the ISIC divisions[2] resulted in a pre-selection of six sectors, which have more than ten companies in the top 240. These sectors are listed in Table 3.3. Further selection is based on the homogeneity of the products of the sectors. A variety of sectors with different complexity levels will provide the best possible choice of suitable benchmark indicators for target setting. It is expected that the gas and petroleum sector and coke and petroleum products sector, have a similar uniform product to the basic metal sector. These two sectors are therefore excluded. The land transport and air transport sectors are also expected to

[2] The International Standard Industrial Classification (ISIC) of all economic activities, Revision 4, published in 2008. This is a revision of the General Industrial Classification of Economic Activities within the European Communities, known by the acronym NACE and was initially published by Eurostat in 1970. This revision organises companies into industrial sections based on similar products and production processes. It starts with economic sections A to Q, where each section is divided into many more divisions, groups and classes.

have a similar homogeneity level. The air transport sector was selected because the boundaries of the activities in this sector are described more accurately. For some sectors it is possible to further specify to more detailed economic activities. Thus only electricity companies from the "Electricity and heat" sector are selected; and only steel companies from the "Basic metals" sector are opted for. Companies do not always fit in one sector; for example some electricity companies might also produce steam or gas. In chapter 5 Results, it is discussed which part of the companies' activities are included.

Researched sectors	No. large companies >10 $MtCO_2$	ISIC rev 4 code		Expected complexity
Electricity and heat	78	D35.1	Electric power generation, transmission and distribution	1
		D35.2	Manufacture of gas; distribution of gaseous fuels through mains	
		D35.3	Steam and air conditioning supply	
Basic metals	47	C24.1	Manufacture of basic iron and steel	2
		C24.2	Manufacture of basic precious and other non-ferrous metals	
		C24.31	Casting of iron and steel	
		C24.32	Casting of non-ferrous metals	
Gas and petroleum	37	B06.1	Extraction of crude petroleum	2
		B06.2	Extraction of natural gas	
Land transport	12	H49.1	Transport via railways	3
		H49.2	Other land transport	
		H49.3	Transport via pipeline	
Air transport	11	H51.1	Passenger air transport	3
		H51.2	Freight air transport	
Coke and petroleum products	11	C19.1	Manufacture of coke oven products	2
		C19.2	Manufacture of refined petroleum products	

Table 3.3: The ISIC codes of the sectors with more than ten companies who emit >10 $MtCO_2$-e (ISIC, 2008). The sector complexity is indicated with a 1 for homogeneous sectors and 3 for more heterogeneous sectors.

3.4 Review company rankings and indicators used

The purpose of this section is to give an overview of the current benchmark methodologies. To put together this overview, a literature research was performed based on extensive research by the European Commission (ERM, 2010), American researchers (Price, et al., 2008), and other sources (Rate the Raters, 2012). The benchmark methodologies and initiatives which were used to create a company ranking are listed in Appendix A, together with the different benchmark methodologies and the used indicators. In order to understand the methodology, the goal of the benchmark, the target group and the scope of the research are listed as well.

Various methodologies and indicators have been used by the benchmark initiatives. Some benchmark methodologies focus on actual carbon performance while others focus on overall sustainability. And some use quantitative indicators while others use qualitative indicators. Figure 3.4 illustrates the focus of the benchmark methodology and the used indicators. The size of the circle indicates the amount of rankings on that particular spot. For example, five rankings use quantitative indicators to benchmark solely GHG emissions.
On analysing these rankings, three main incentives with associated target groups can be distinguished and are indicated with different colours in Figure 3.4.

The three main incentives are:

- A group of commercial rankings made for Socially Responsible Investors (SRI) interested in Environmental, Social and Governance issues can be distinguished. These kinds of rankings are indicated with red. Mostly methodologies focus on overall sustainability and qualitative indicators are used.
- Another group can be distinguished which actually focus on quantitative emission reductions. These are either obligatory or voluntary initiatives for high emitting companies to increase energy efficiency and support knowledge sharing. Most of them are based on a top-down approach, set up nationally by governments or government supported institutions to reach (inter)nationally agreed targets. Most of these rankings do not create a company ranking. They are indicated with green. Three of these initiatives create a ranking and are not initiated by the government. These can be considered bottom-up initiatives. They are indicated with dark green. Of these three, CDP is the only global initiative and Canadian Industry Program for Energy Conservation (IPEC) is the only one using solely quantitative indicators.
- The last group of rankings is made with publicity and media attention as main driver. The target group of the rankings are consumers or readers of the specific magazine. These benchmarks are indicated with orange.

From this analysis it can be concluded that most of the benchmarks focussed on quantitative emissions, are focussed on obligatory emission reductions or created for socially responsible investors. In order to use full reduction potential of companies, it is preferable that the benchmark is designed by branch organisation or companies themselves. There are no initiatives which rank global companies based on their quantitative carbon performance. These ranking initiatives are important to enable target setting and help bridge the emission gap (Blok, et al., 2012).

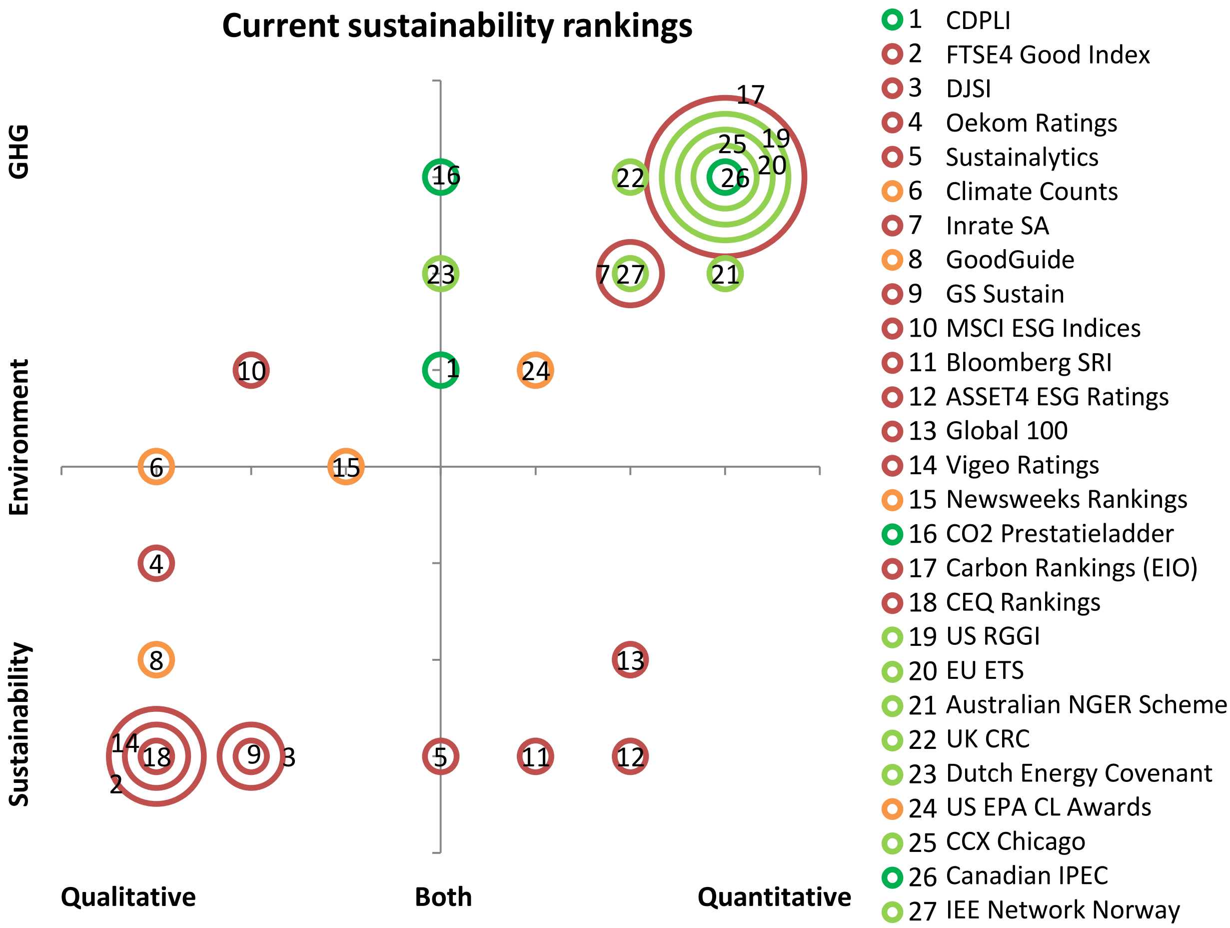

Figure 3.4: The sustainability benchmarks of the literature research plotted based on their choice of indicators and focus.

3.5 Rating criteria

The function of benchmarking is to enable comparisons to be established. This can be done in many ways. There is no preferred indicator; it depends on the purpose of the benchmark. In this research, the purpose is target setting at company level. The intensity indicators should therefore reflect the carbon performance of the whole company as well as possible. A good benchmark is ideally based on similar processes but each company is unique. Therefore, benchmarking at company level will always entail a certain level of aggregation and certain limitations. Experts indicate that the main limitations of the quality of indicators are the homogeneity of the sector, the uniformity of the products and the quality of the available data. In this research, the relatively homogeneous sectors were chosen as a starting point. Therefore, it is expected that the quality of the available data will be the main limitation. To combine theory and practice, a theoretical framework about the level of aggregation is combined with the quality of the available data and interviews with sector stakeholders.

3.5.1 Aggregation triangle

In 1997 Utrecht University performed a study in order to compare the energy efficiency in the manufacturing industry (Phylipsen, et al., 1997). This study explained that the energy consumption within a country or sector is determined by three main factors:

- Level of activities (volume)
- The mix of different kinds of activities (structure)
- The energy consumed per activity level (intensity)

Similar factors can be used for the carbon performance of companies. The *volume* effect (the influence of changes in the activity level on total carbon emission) can be taken into account by using an intensity indicator (dividing the carbon emissions by the activity level). Such indicators are always approximations of the carbon performance because they may contain a higher or lower number of structural effects (Phylipsen, et al., 1997). Structural effects caused by the mix of different kinds of activities (*structure*) are fewer in relatively homogeneous sectors, like the electricity or cement sector, than in sectors with many different products, like the chemical sector. The main problem is how to allocate the emissions to the various activities. The *intensity* effect is described by the emission intensity per separate activity. This becomes a problem when different activities are combined in one indicator. The possible differences in intensity should be accounted for.

Each of the three factors: volume, structure and intensity, can be determined for various aggregation levels. Disaggregate indicators are more detailed and have therefore fewer structural effects. For a carbon intensity indicator you need an emission parameter and an activity parameter. At the highest levels of aggregation, only one indicator is used for the carbon intensity, for example company emissions in combination with company turnover. The indicators at this level include many different products (the structural effect) and different intensities per product (intensity effect). Further disaggregation towards more product specific parameters will provide the possibility of sorting out the structural effects more and more. The amount of date needed increases with the level of disaggregation. At a more disaggregate level, a parameter for each separate identified product is used, for example the emissions of an electricity plant in combination with the produced megawatt hours (MWh) of that plant. Depending on the goal of the analysis, a lower level of aggregation may be necessary; distinguishing between different types of product and/or different specific plants. In this research, *companies* are benchmarked and therefore the comparison is done at company level. Either one aggregate indicator at company level may be used or disaggregate data is used and re-aggregated to company level. This implicates that a trade-off should be made between using disaggregate data to limit the structural effects and the quality of re-aggregation towards company level Figure 3.5.1. Usually, only aggregate monetary performance and disaggregate physical performance data are available (Figure 3.5.2).

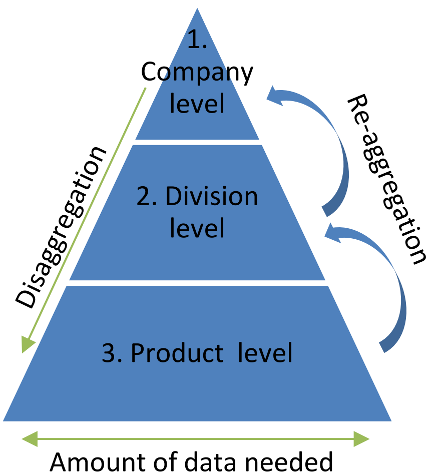

Figure 3.5.1: Aggregation triangle based on the Energy Efficiency Indicator pyramid (Phylipsen, et al., 1997).

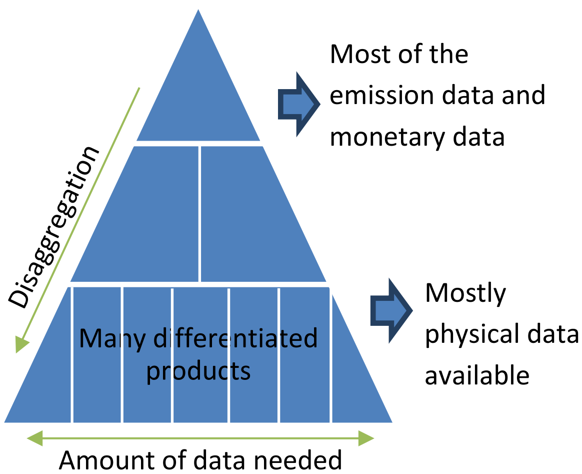

Figure 3.5.2: Structural effects in the aggregation triangle.

From a scientific point of view, the carbon intensity is determined by the technical and operational efficiency. Although improvement of the energy efficiency in many cases is also efficient in economic terms, this does not necessarily have to be the case (Phylipsen, et al., 1997). Economic improvement also depends on other monetary inputs. Therefore monetary parameters are less related to carbon efficiency, compared to the physical product of the activities, independent of the level of aggregation. Literature shows that for manufacturing industries, physical parameters are more appropriate than monetary parameters due to the homogeneity of products and the direct relation with products of the activity. These sectors include: iron & steel, cement, glass, sugar, alkalis & chloride, ammonia, pulp & paper, aluminium, refined oil products, other metals (Cu, Pb, Zn, Ni, etc.), other chemicals (ethylene, etc.) and mining and extraction (Martin, et al., 1994). In other sectors, an economic parameter may be more appropriate as the mix of products might be very diverse and the possibilities for re-aggregation are very low. Another reason to use monetary parameters is if the sector is more service-oriented than production.
The last indicator selection criterion is whether the indicator is useful for the sector stakeholders and whether it offers the possibility of reflecting improvement and differences between companies.

Based on the consulted literature and the opinion of experts of Ecofys, the following criteria are used to select the best carbon intensity indicator per sector:

- Quality of aggregation: The disaggregation level of the separate parameters and the quality of re-aggregation to a company level carbon intensity indicator.
- Quality of data: The quality of aggregation is limited by the quality of the available data. A trade-off between the quality of data and the level of aggregation should be made.
- Current experience with benchmarking in the sector. This criterion is not quantified but qualitatively included in the final recommended indicator per sector.

3.5.2 Carbon intensity parameters

The following carbon intensity indicators and parameters could be distinguished from literature and analysed sustainability rankings.

Physical carbon intensity indicators

Generally the Physical Carbon Intensity Indicators can be seen as the ratio of emissions over physical activity (tCO_2-e/P) where both variables can be aggregate quantities. The aggregation of all GHG emissions is more trivial than in the case of primary energy but different GHG reporting standards provide guidelines for doing this (WRI/WBCSD, 2004). Aggregation of the different physical activities is more complex (Blok, 2007). In the book "Introduction to Energy analysis", Blok describes the Energy Efficiency Intensity and the Physical Production Index of the activities of a company in a specific sector in a specific year:

$$EEI = \frac{E}{PPI} \quad \text{and} \quad PPI = \sum_{x=1}^{n} w_x \cdot P_x$$

Where:

EEI = Energy Efficiency Index (or physical energy intensity)

PPI = the Physical Production Index (or activity index)

E = Energy use

P_x = the amount of production of product x

w_x= the weighting factor for product x

n = the number of products

The EEI can be rewritten for carbon emissions instead of energy use

$$CII = \frac{C}{PPI}$$

Where:

CII = Carbon Intensity Indicator

C = the amount of GHG emissions in tonne CO_2*-e.*

The activity index describes how much of a product is produced. A common parameter to describe this amount of product is weight, like tonnes of produced steel. Not all products can be measured in weight. Other parameters to describe the product activities, reported by the researched sustainability benchmarks, are listed below:

- Used space → *Emissions per* m^2 *(also per cubic area; per cubic area per degree day)*
- Electricity → *Emissions per MWh*
- Employers → *Emissions per full-time employee equivalent (also per hours worked; per operating hour; per guest night; per capita; per patient days)*
- Thermal comfort → *Emissions per* m^3C *(degree-volume air)*
- Sustenance (food) → *Emissions per J (food)*
- Structural materials → *Emissions per* $MPa^2/_3*m^3$ *(tensile strength x volume)*
- Freight transport → *Emissions per ton-km (also ton per nautical mile)*
- Passenger transport → *Emissions per passenger-km (also employee km; vehicle-km)*
- Hygiene → *Emissions per* m^3C *(temperature degree-volume of hot water) and Nm (work)*
- Communication → *Emissions per bytes*
- Illumination → *Emissions per lumen-seconds*

Monetary Carbon Intensity Parameters

The products manufactured by the activities are generally sold for money. The monetary value for the manufacture of these products can also be used as an indicator. The following monetary indicators are used by one or more of the analysed sustainability rankings:

- Turnover → *Emissions per US$ turnover*
- Profit → *Emissions per US$ profit*
- Earnings Before Interest, Taxes, Depreciation and Amortisation → *Emissions per US$ EBITDA*
- Value added → *Emissions per US$ value added*

4. Methodology

The previous section dealt with the methodological background of this research and provided an overview of the current benchmark methodologies, as well as giving a theoretical model for the level of aggregation. This section elaborates on how this theoretical framework is used to perform an analysis of the different quantitative emission intensity parameters.

Step 1: Literature summary

The websites of sector organisations were consulted to get an indication of the current share of emissions of the sector and to get an indication of the current importance of emission benchmarking in the sector. To obtain some information about some large companies in the sector, the different products they offer and their corresponding activities were used as search criteria to get an overview of the possible performance indicators and high emitting activities. If needed experts from Ecofys and additional sector specific studies were consulted to get an overview of the current emissions of the sector, activities, available databases and interesting companies.

Step 2: Indication data availability

Since the quality of the available data is likely to be the biggest limitation in this analysis, a quick scan of the available data of some of the largest companies in the specific sector was performed. This scan focussed mainly on the availability of emission data and what part of the companies activities are included in the available performance data. This is done by looking at publicly available data like the CDP database and data from annual (sustainability) reports which are found at the company websites. Other known databases suggested by Ecofys experts like ETS, PLATTS and databases from other rankings are considered depending on the sector.

Step 3: Carbon Intensity Indicator Comparison

Step 3.1 Parameter selection

The list of carbon intensity indicators found in the literature presented in section 3.5.2 are considered for each sector. Scope 1 and 2 emissions are preferably used as parameters for companies' emissions because guidelines for this way of reporting are commonly accepted and supported by WRI-WBCSD. At least one monetary parameter and one physical parameter are selected for each sector, based on their availability in literature, as well as through logical reasoning and expert opinions.

Step 3.2 Data collection

To analyse the indicators, data was obtained from at least ten companies with the highest absolute emissions within the sector. The companies were selected based on the list of the 2000 largest companies based on turnover, ranked on Scope 1 and 2 emissions, reported in a previous piece of research performed on behalf of Ecofys (Jong, 2011). When no emission or performance data for the particular company was found, the next, sector specific company in the list was investigated. Data for the selected indicators was collected for a period of about four years depending on the available data. A period of four years was used to get a better average company performance and differences were analysed over the years. Problems with the selected indicators were noted down during the data collection. In some cases, the selected parameters were changed or new parameters added depending on further research of the available data. The level of aggregation of the parameters might also have required some modification.

Step 3.3 Indicator positioning

The indicators are rated on data quality and aggregation quality. The data quality is rated on a quantification model in Figure 4.1. Insufficient data is attributed the lowest quality value; zero, and high quality data, which is verified, is attributed the highest quality value; 5. Data which is reported but not verified scores lower than verified data. Incomplete or incomparable data which needed to be significantly modified is scored based on the used modification factor. If the factor(s) is product or company specific it is scored with a 3 and if the factor(s) is industry specific it is scored a 2.
In case the factor is based on not enough data or an overall average is used, it is rated with a 1. In all these cases, the way the values are included, depends on the influence of the used factor(s) on the end score. For each parameter a motivation is given for the indicated value.

5. Verified data
4. Reported data
3. Estimated data based on product specific values
2. Estimated data based on industry specific values
1. Estimated data based on rough estimates
0. No sufficient data

Figure 4.1: Quantification of the data quality based on a similar model created by Ecofys (Ecofys, 2013).

The aggregation quality is determined based on level of disaggregation of the separate parameters, and the quality of re-aggregation, when combining the parameters to one intensity indicator. The formula is presented in Figure 4.2.

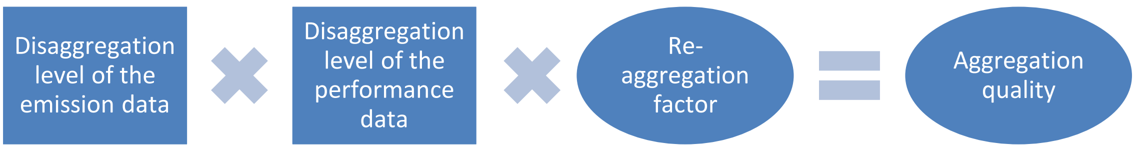

Figure 4.2: The formula used for the aggregation quality. The square form indicates a value based on one of the parameters and the circle indicates a value based on the combined intensity indicator.

The disaggregation level of the parameters is rated based on the Aggregation Triangle in Figure 3.5.1. The selected parameters are visualised in a sector specific aggregation triangle. Value 1 for company level and value 3 for product level. In some cases, comments to support the appointed figure have been provided.
The re-aggregation factor is given when the parameters are combined with an intensity indicator. It is based on the quantification model in Figure 4.3. The value 1 represents high quality re-aggregation and the value 0 represents low quality re-aggregation. This re-aggregated indicator should be as homogeneous as possible, and include as many emitting activities of the company as possible. The possibility of creating one indicator really depends on whether a way can be found to combine the disaggregate parameters. Consequently, reasons for supporting each re-aggregation value given to the indicators are provided.

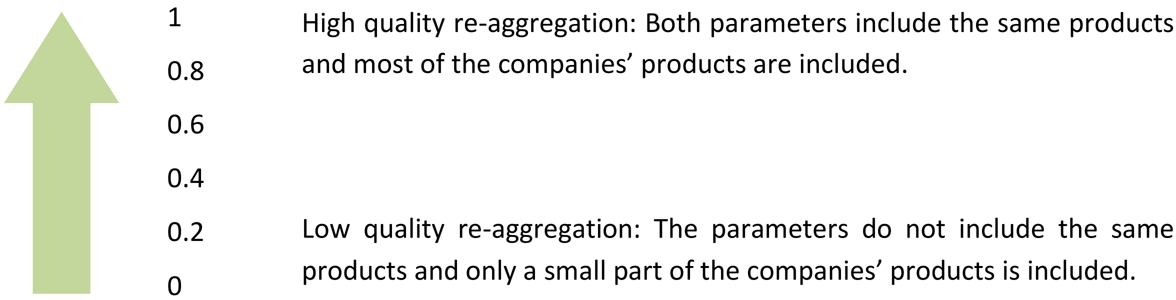

Figure 4.3: Quantification of the quality of re-aggregation into a re-aggregation factor.

To compare the quality of the different indicators, they are plotted in a graph with the data quality on the Y-axis and the aggregation quality on the X-axis.

Step 4: Company rankings

In addition to the analysis mentioned above, the collected data is used to create a company ranking for each of the indicators. These rankings are made to provide an overview of the collected data, and to enable further analysis, such as the influence of the different indicators on the rankings. It might also provide an indication of the influence of the choice for certain technologies or activities by the company on the carbon efficiency. The rankings are also used to provide information on the data quality and the different reporting methodologies used by the companies through the analysis and verification of the numbers or graphs which stand out.

Step 5: Verification method

Step 5.1 Comparison

A first verification step is to check if a similar benchmark results in the same ranking. A search for a similar ranking for each sector was undertaken. Ecofys experts, familiar with the sector, were consulted. Where this information was available, the created ranking, intensities and methodology was compared.

Step 5.2 Interviews

The second verification step was feedback from sector stakeholders. The results of the indicator analysis and the created rankings were presented to different stakeholders (experts, branch organisations and companies). The results of the indicator analysis of Step 3 and 4 were presented to the sector stakeholders. Subsequently, interview questions were used to get feedback about the indicator selection and methodology. Additional question were asked to gain information about the current experience with benchmarking in the sector, the usability of the different indicators and data disclosure problems.

The feedback provided by the stakeholders is included in the report. A list of the main questions and the summary of the interviews can be found in Appendix B.

5. Results

In this chapter the results of the three sector analyses are presented. For each sector, data is collected and analysed, and sector opinions are collected through interviews. At the end of each sector analysis a short conclusion is presented. First the electricity sector is treated, followed by the Air transport sector, and then the Steel sector.

5.1 Electricity sector

The sector analysis consists of seven paragraphs. In the first two paragraphs a short overview of the relevant activities in the electricity sector is presented as well as an indication of the available data. In the third paragraph the data collection and the rating of the indicators are explained. This paragraph concludes with a chart which shows the different indicators plotted, based on data and aggregation quality. For further analysis of the indicators, the collected data is presented as a company ranking for each indicator. The fifth paragraph provides a summary of the interview with Eurelectric, followed by the main pitfalls, experienced during this sector analysis. A short sector specific conclusion is presented in paragraph seven.

5.1.1 Literature summary of relevant processes within sector

The electricity sector has the largest share in the global anthropogenic carbon emissions. The majority of carbon is released, when primary energy carriers in different forms of fossil fuels are burned. Superheated steam is generated in a boiler and expanded through a turbine that is connected to an electricity generator. In modern facilities, a number of recycle loops are used, in which the steam is reheated and subsequently further expanded, to optimise the efficiency (Blok, 2007). However, the carbon intensity of the process depends mainly on the used fuel.

An increased proportion of electricity is produced by plants for combined generation of heat and power (CHP). In such a case the plant is dimensioned on a certain industrial heat demand. While producing heat, electricity is generated as well. In this case the separate efficiencies of heat and electricity production are lower, but the combined efficiency is very high. The emissions of the burned fuels could be allocated to the produced heat and to the electricity. There are other ways to produce electricity with significantly less carbon emissions i.e. renewable energy sources like solar, wind and water power. Nuclear power, mainly fuelled by uranium, is also a low carbon option for electricity generation. Nuclear power poses many other problems and risks than carbon emissions. Since the indicators for this research are solely focused on carbon, these other risks will not be taken into account.

There is no cheap and efficient way to store electricity. Therefore production is adapted to the fluctuating electricity demand. The majority of the electricity production is generated by large companies, due to the high investment costs of technology. Each of these companies uses a variety of primary energy sources. Different studies indicated that carbon intensity improvement is mainly caused by fuel switching and an increased share of renewables rather than efficiency improvement of the used plants.

The produced electricity is often traded between electricity companies many times before it is sold to the end consumer. The electricity sold to the end consumer often consists of electricity generated by different power plants owned by different companies.

5.1.2 Indication data availability

Initially, ten electricity companies, with the highest absolute emissions, were selected for this research. But, when checking the data availability on company websites, annual (sustainability) reports and the CDP database, very little emission data was found of the selected companies, especially of the companies outside of Europe. In some cases, mainly in Asia, not even monetary or physical performance data is disclosed. Only four out of the ten, reported their emission and performance data. A quick scan of the fifteen companies, subsequently on the list of highest emitters made by Jong, resulted in emission data of another three companies.

In 2011 Ecofys performed a benchmark study for European electricity companies with the use of the EU ETS emission database in combination with PLATTS (Ecofys, 2011). Due to the mandatory reporting guidelines of EU ETS this data is of high quality and consistent. Some electricity companies did report their emissions to the CDP or disclose them in an annual sustainability report, but the comparability of this data is not as good as the data of the EU ETS, because both sources are voluntarily and have no strict reporting guidelines. Correcting for inconsistencies and missing data points would require several assumptions. Therefore, the emission and production data of the study performed by Ecofys is used in this analysis. Ten companies are researched in this study, selected based on market share and focus on renewables. Five of these companies were also included in the initial global company selection. Of the other five, three do belong in the top 50 of highest emitting global electricity companies (Iberdrola, EDP and SSE). And two are relatively small companies (Dong and Statkraft). In this research, data quality is more important than the size of the company, therefore the ten companies, from the study done by Ecofys, are selected. In addition, one of the largest emitting global electricity companies is added, Tepco, to indicate possible geographical differences. The emission and production data of Tepco is found in their annual sustainability reports. The monetary data of all eleven electricity companies is found in their annual reports, accessed via their company websites.

5.1.3 Carbon Intensity Indicator Comparison

In this paragraph the data collection of the selected parameters is explained, followed by the rating of the separate parameters and the combined intensity indicators, as described in the methodology. In the last section of this paragraph, the indicators are plotted in a graph based on their data- and aggregation quality.

5.1.3.1 Parameter selection

From the list of parameters shown in paragraph 3.5.2 the parameters listed in Table 5.1.1 are chosen.

Emission parameters	Physical parameter	Monetary parameter
Scope 1 emissions without heat allocation in tCO_2-e Scope 1 emissions with heat allocation in tCO_2-e	MWh produced	Turnover in US$

Table 5.1.1: The selected parameters used to form the intensity indicators for the electricity industry.

5.1.3.2 Data collection

In the research performed by Ecofys in 2011 the emission and production data is calculated as follows: the ETS database provides verified carbon emissions and produced MWh of every fossil power plant in Europe. The emissions of the different plants of one company are added up based on operational control[3]. These form the Scope 1 emissions without heat allocation. For the parameter with heat allocation, the effect of heat production in CHP plants is included. The heat ratio is estimated per technology on the base of average heat ratios of all fossil power plants in Europe. Depending on the used technology the substitution or the power loss method is used for allocating the emissions to heat. The used heat allocation methods can be found in Appendix C.

The produced electricity reported in ETS data is only based on fossil power plants. The share of renewables is not included yet. Therefore the verified production data of the fossil plants from ETS is combined with the PLATTS database where an overview of the installed capacities of all electricity plants in Europe is provided. Only the nuclear and renewable capacities in PLATTS are used, and combined with a load factor to get the production data. The load factors are gathered from the annual reports of the owners of the plants. In some cases, no load factor was given. In that case, literature is consulted for the average load factor of the specific technology in Europe. The produced electricity of the fossil and renewable plants of one company is added up based on operational control. In this way relatively high quality, disaggregate emission and activity data is provided. The annual sustainability reports of Tepco are used for their emission and production data.

Turnover data of all companies are found in their annual reports. Almost no product specific monetary data is found. Only the aggregate company turnover including all the products; produced electricity, traded electricity, sold heat and possibly sold gas. Further disaggregation based on the available data is not possible.

5.1.3.3 Indicator comparison

The selected carbon performance intensity indicators consist of one of the two emission parameters of Table 5.1.1 combined with one activity parameter; the physical parameter produced MWh or the monetary parameter turnover. The electricity sector is a quite homogeneous sector, with a relatively uniform activity. Electricity is produced, in some cases combined with heat production. Overall only two other activities can be distinguished: buying and re-selling electricity produced by other companies and gas production.

The emission data available by the ETS database is plant specific, the level of disaggregation of this data is not at company level. The different plants can be seen as divisions. These emissions are the direct emissions from electricity and heat production installations. For one of the emission indicators an allocation methodology is used for heat to account for the efficiency loss of electricity production when cogenerating heat. This emission indicator is disaggregated to product level. This results in an emission parameter at division level and one at product level. The emissions of traded electricity are very hard to measure because it is traded many times. Estimating these emissions would pollute the high quality data of the ETS database and are therefore excluded from the emission parameter. The emissions of a company due to gas production are not taken into account in the emission parameter because gas production is not directly part of the electricity industry and mainly electricity companies are compared.

The turnover data is only available at company level, no further disaggregation is possible. It includes the four different products: traded electricity, produced electricity, heat and gas, while the emission data only includes produced electricity and heat. The quality of the aggregation level of the indicator

[3] The WBCSD/WRI GHG Protocol Corporate Standard offers two distinct approaches for corporate reporting: Operational control, where a company accounts for 100% of the GHG emissions from operations over which it has control. It does not account for GHG emissions from operations in which it owns an interest but has no control. Or the equity share approach, where a company accounts for GHG emissions from operations according to its share of equity in the operation.

would be better if more gas and traded electricity specific turnover data would be available. Then these two products could be excluded from the monetary parameter and the two parameters would include the same products, but up till now this is not the case. This causes a discrepancy between the included products of the emission data and the monetary data.
The emission parameters are also combined with the disaggregated product parameter: produced electricity. The separate plant data of both parameters can be added up to re-aggregate to company level. The quality of re-aggregation is on the one hand better for the emission indicator with heat allocation because in that case the numerator entails the same uniform activity as the denominator. But on the other hand is heat as product excluded from the benchmark. Based on Aggregation Triangle presented in Figure 3.5.1, a specific aggregation triangle of the electricity industry is shown in Figure 5.1.1. It is a graphical representation of the aggregation level of the different products.

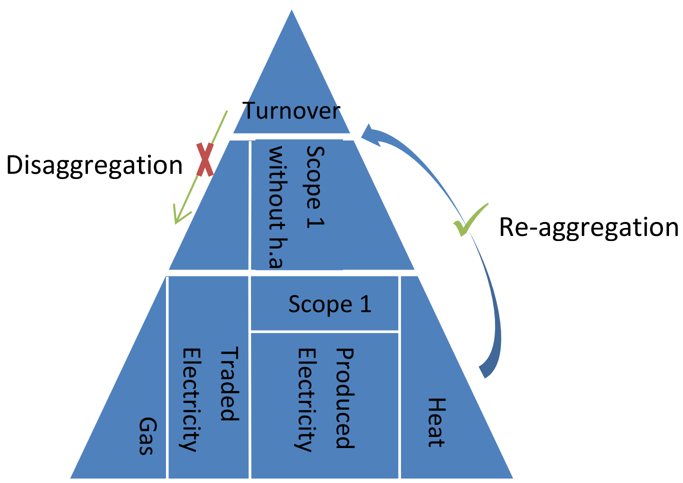

Figure 5.1.1: The aggregation pyramid of the electricity sector. h.a. stands for heat allocation.

To compare the different indicators, first the level of disaggregation and the quality of the used data of the separate parameters are rated. The level of disaggregation based on Figure 5.1.1 and the level of data quality based on Figure 4.1. This is presented in Table 5.1.2, extra comments on the given value are shown on the right.

Parameters	Disaggregation level		Data quality	
Scope 1 emissions without heat allocation	2	Division level, plant specific emissions from ETS	4	Verified data
Scope 1 emissions with heat allocation (h.a.)	3	Product level, electricity specific emissions	3.5	Verified data with use of plant specific average heat production.
Turnover	1	Company level	4	Verified data
MWh produced	3	Product level	3	Verified data with use of average load factors for renewable energy production per plant.

Table 5.1.2: The disaggregation level and the data quality of the separate parameters in the electricity industry.

The two parameter values needed for one indicator are combined to create the indicator value. When the different parameters are combined, re-aggregation to company level is needed. And ideally the same products are included in the denominator and the numerator of the indicator. To rate this process of combining the different levels of aggregation, a quality of re-aggregation factor is included in the product of the different disaggregation levels. This re-aggregation factor of the different indicators is presented in Table 5.1.3. Because the methodology of the collected data is clear, no extra modification on the data quality is needed when the parameters are combined.

Indicators	Re-aggregation factor	
Scope 1 with h.a./Turnover	0.4	Parameters include different products.
Scope 1 with h.a./MWh	0.6	Parameters include same products but heat, resold electricity and gas are excluded.
Scope 1 without h.a./MWh	0.6	Parameters include almost the same products (except for heat) but resold electricity and gas are excluded.

Table 5.1.3: The quality of re-aggregation factor of the combined parameters into indicators.

This results in the following values for the aggregation quality and the data quality of the different indicators, presented in Table 5.1.4.

Indicators	Aggregation quality	Data quality
Scope 1 with h.a./Turnover	1.2	14
Scope 1 with h.a./MWh	5.4	11
Scope 1 without h.a./MWh	3.6	12

Table 5.1.4: Indicator positioning based on aggregation and data quality.

To provide an overview of the quality of the indicators they are plotted in a graph Figure 5.1.2, with on the Y-axis the data quality and on the X-axis the aggregation quality.

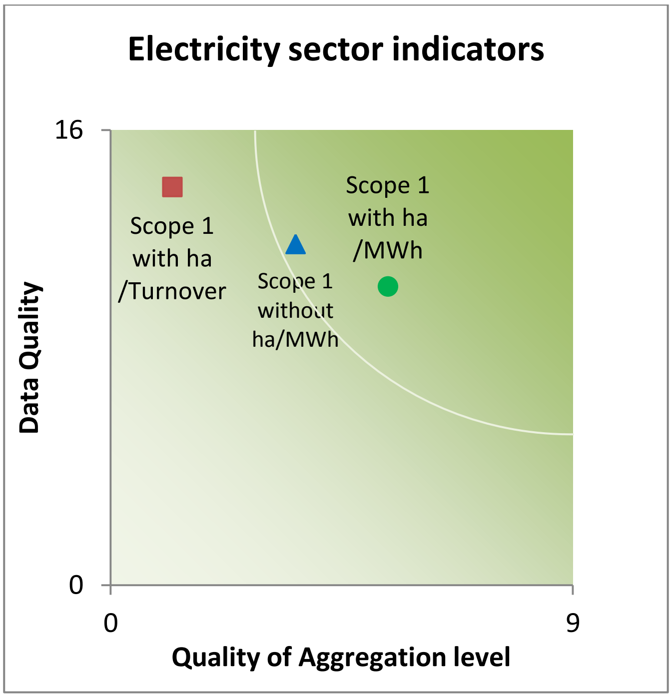

Figure 5.1.2: The different Carbon Intensity Indicators of the electricity sector plotted based on aggregation and data quality. The quality decreases with the green colour intensity. The bow encloses an area which gives an indication of indicators with a certain quality level.

From Figure 5.1.2 it can be concluded that the physical indicators have a much better aggregation quality than the monetary indicator. This can be explained by the initial high level of dis-aggregation and the homogeneity of the product electricity which enables a high quality re-aggregation. Because the monetary parameter has a different level of aggregation, the excluded products in the emission parameter could not be separated in this monetary parameter. The difference in included products causes the low re-aggregation factor. Further separation of the revenue of the different products like traded electricity and gas is needed to increase the quality of this indicator.

Physical indicators, which are already used by CDP and ETS amongst others, are better carbon intensity indicators than the monetary parameter when looking at the data and aggregation quality. Allocating for heat has a higher level of disaggregation, but it also includes the chosen for a heat allocation methodology. The emission allocation to heat is done at plant level and the used methodology is well accepted in literature. Based on this, it is advised to use the physical indicator with heat allocation. To further analyse the different indicators and the used data, the company data is put into a ranking based on each indicator.

5.1.4 Company rankings

The following graphs provide an overview of the company specific data. The indicators are further analysed, compared and verified.

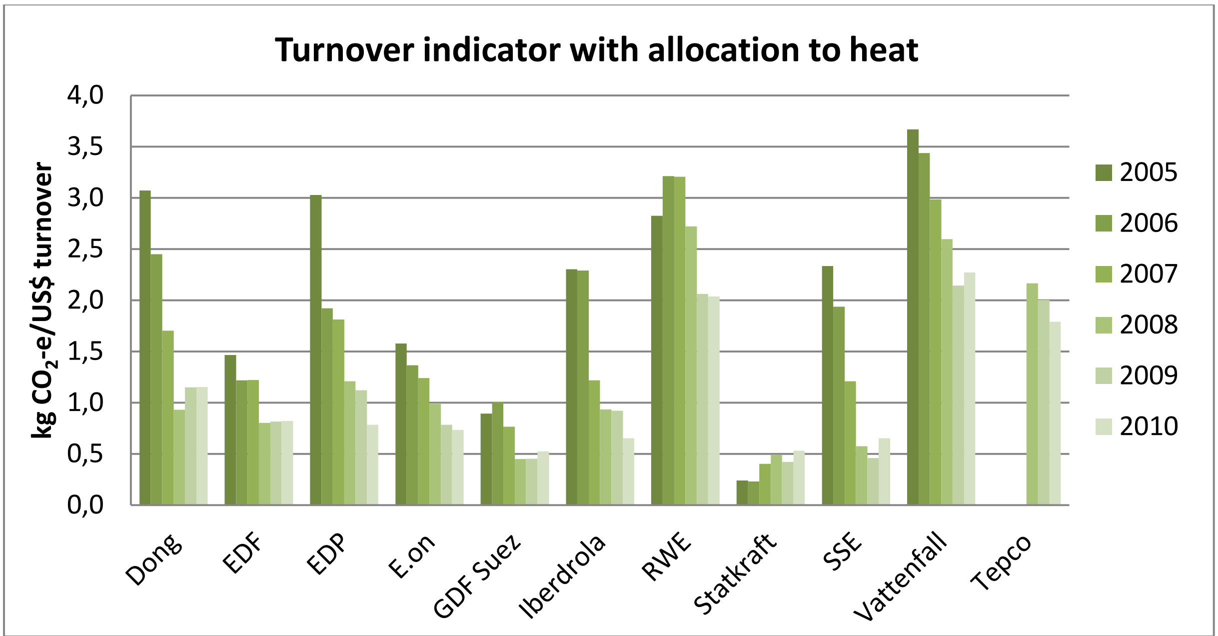

Figure 5.1.3: A ranking of eleven electricity companies based on a combination of an emission- and a monetary parameter. The emission parameter is corrected for heat production.

The use of the monetary indicator in Figure 5.1.3 shows a clear reduction in carbon intensity for almost all companies over the last six years. In the last two years this reduction changed into an increasing carbon intensity for more than half of the companies. This can be explained by the economic crises which hit in 2009. Vattenfall described in their annual report that in 2011 their turnover was smaller due to a smaller electricity demand, but that this was partly offset by an increased gas demand. This offset is not included in the physical indicators shown in Figure 5.1.4 and 5.1.5.

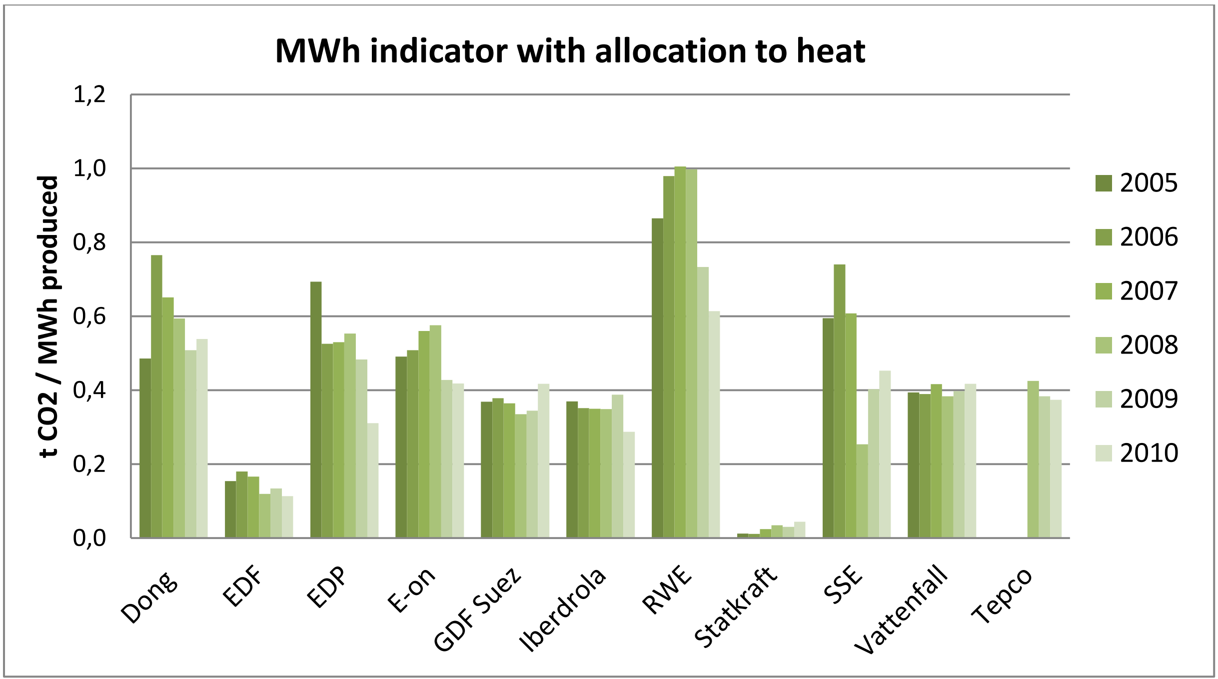

Figure 5.1.4: A ranking of eleven electricity companies based on a combination of an emission- and a physical parameter. The emission parameter is corrected for heat production.

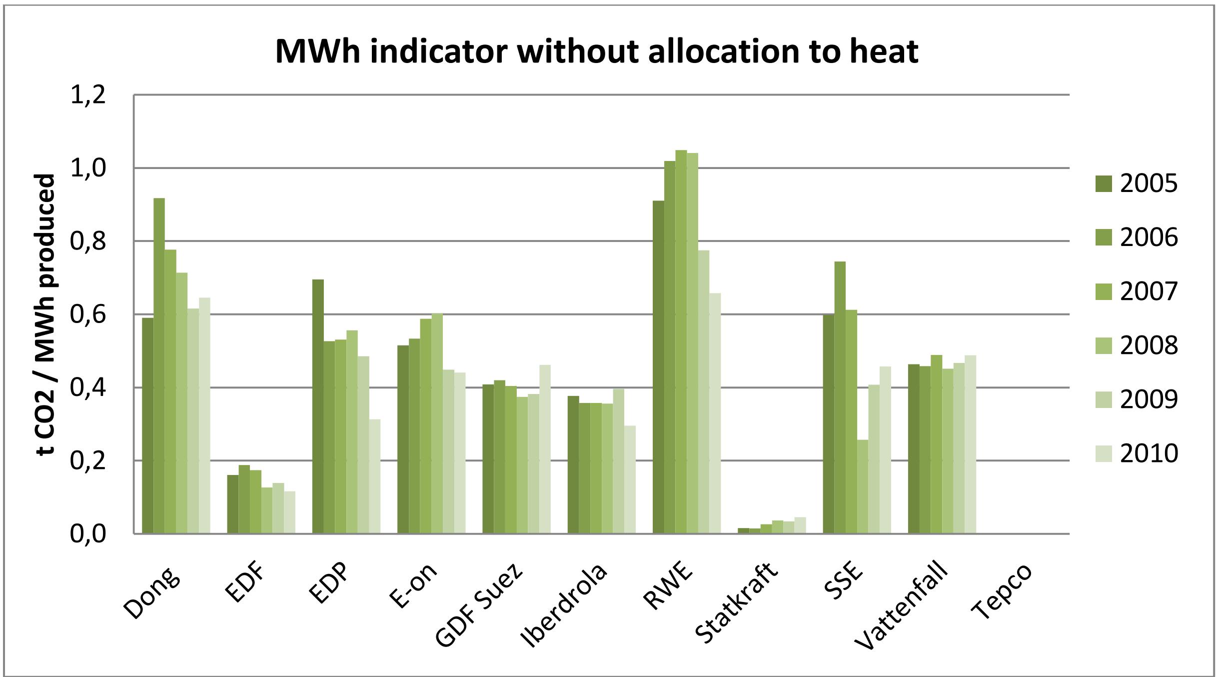

Figure 5.1.5: A ranking of eleven electricity companies based on a combination of an emission- and a physical parameter. The emission parameter is not corrected for heat production.

The physical carbon intensity is more constant over the years compared to the monetary indicator. In the last two years an increasing carbon intensity is seen at five companies in both Figure 5.1.4 and 5.1.5. This can be explained by a smaller electricity demand due to the economic crisis. Installations might not have been used on their optimal capacity and cheap fuels for the base load like coal might have taken a relatively larger share of the emissions than the variable load of natural gas.

Statkraft has a small physical intensity indicator due to large shares of hydropower and EDF has a small intensity indicator due to large shares of nuclear power. This small intensity value is less visible in the monetary indicator of Figure 5.1.3. The monetary indicator shows a very different ranking than the physical indicator. This shows that they are not interchangeable. One of them provides a better reproduction of the technical and operational efficiency than the other.

The intensities of Figure 5.1.5 are slightly higher than the intensities of Figure 5.1.4, because no emissions are allocated to heat production, but they result in similar rankings. It can be concluded that for these companies, allocation to heat has a small effect. This is partly explained by the small amount of heat produced, compared to the amount of electricity. All researched companies are located in Europe except for Tepco, when comparing companies from over the whole world, larger disparities in heat production can be expected, for example in Russia. But for these companies the two physical indicators can be considered interchangeable, because they create a similar ranking.

5.1.5 Verification method

Underneath two verification steps are done. The first is a comparison with another sector specific ranking with a similar goal. For the electricity sector the Top 100 ranking of the Energy Intelligence is used. They have ranked the carbon intensity based on two indicators: tCO_2-e/MWh and the share of renewables. The other is feedback from stakeholders on the methodology and outcome of this research.

Comparison with Energy Intelligence (EI) ranking

All utility companies included in this research are also present in the top 100 Green utilities of EI (Roos, 2012). EI ranked hundred companies, which together represent about 55% of the world's total power generating capacity[4]. The ranking EI created based on solely quantitative carbon intensity, in tCO_2-e/MWh, has the same three companies in the top 3 as the ranking of the physical indicators in this research. Statkraft has the lowest carbon intensity, followed by EDF and then Iberdrola. This confirms that the quality of the used data is good and that the physical indicator is useful.

Feedback from Interviews

Jesse Scott, head of the Environment & Sustainable Development Policy unit of Eurelectric, recognises the above presented ranking based on the physical indicators and the intensity development along the graph. Eurelectric agrees tCO_2-e/MWh is the best indicator to benchmark electricity companies. Eurelectric explained that the largest differences between companies are related to culture difference focussing on short or long term and the current assets. Large changes over the years within a company are also mostly due to changes in their assets and the used fuel composition. She did not know any additional indicators which could have been researched for this purpose, although for Eurelectric, it would be interesting to look at the impact of the introduction of new technologies on the carbon intensities of companies.

[4] Each company was awarded up to 300 points — up to 100 based on carbon dioxide emissions intensity and up to 200 based on renewable generating capacity, in absolute and relative terms. The main ranking lists all 100 companies with their points, rank and total capacity as an indication of size. Three other tables are created to show the top 20 performers under each criterion: their CO_2 emissions per MWh, proportion of renewable energy to total capacity, and size of renewable energy capacity in gigawatts. The emissions table includes all carbon-free generating capacity; nuclear, hydro and non-hydro renewables — while the other two tables include only non-hydro renewables.

5.1.6 Methodological or data pitfalls

During the indicator selection and the data collection the following sector specific problems occurred.

Data availability and quality

There is almost no transparent emission data found of the largest electricity companies outside of Europe. European companies are ahead in that field. In the US and Asia some companies provide emission data from certain demonstration plants, or emission indicators without any disclosure about the used methodology or scope. In most cases the data quality is too low to be useful for comparison. When no clear guidelines for reporting are used, many inconsistencies can occur; most of them described by ERM in "Principles of company GHG reporting" (ERM, 2010). Most of the electricity companies in Europe provide some emission data, but the methodology of the reported data was often not transparent and included some uncertainties. Jesse Scott from Eurelectric confirms that electricity companies are very conservative with data disclosure. Therefore the ETS database and the PLATTS database were used. The ETS data can be considered of high quality but is only available for European companies. The data of the PLATTS database is consistent, but has a lower quality than the ETS. To look for possible differences with companies outside of Europe, Tepco from Japan was included in the benchmark, but only three years of emission data was available and no heat production. Scott emphasised that benchmarking, and therewith voluntary data disclosure, is more of added value to companies with multiple small generation units than for large companies, which always have old inefficient and newer plants.

Heat allocation

When heat is produced for commercial use, it can be considered a useful by-product, rather than waste. Co-generating heat lowers the electricity production efficiency, but increases the overall efficiency compared to separate production of both products. When heat is considered as waste, co-generation would be penalised while it is very efficient when used properly. There are multiple ways for heat allocation to account for this loss in efficiency. Which allocation method must be used can be debated. In this research the power loss method and the substitution method are used (Appendix C). When looking at the rankings, it is shown that heat allocation has no effect on the ranking. When using an allocation method, certain efficiency assumptions are needed for the electricity and heat ratio. This ratio is location specific; sometimes country specific average factors or world average factors are used. The impact of the choice for a factor is expected to be low based on the small differences in carbon intensity for using an allocation method or not.

Traded Electricity

It could be debated if only the produced electricity should be benchmarked or the traded electricity as well. Since electricity is traded many times, produced electricity is a more robust parameter and therefore used in this research. Since the carbon performance is partly dependent on the share of renewable capacity, it is assumed that all companies have access to renewable sources.

5.1.7 Sector specific conclusions and recommendations

In the electricity sector the common used indicator tCO_2-e/MWh with heat allocation has the highest level of disaggregation. High quality European data is available and stakeholders from this sector consider this indicator the most fair.

EU ETS in combination with PLATTS form a suitable European data source, although the load factors of the renewable sources entail a certain level of uncertainty. Disclosure of the produced electricity by renewable sources will be of increased importance and increase the data quality. A suitable way to allocate emissions to heat is available, but the impact of heat allocation in this mainly European ranking is small and since it is plant specific it complicates the calculations. Globally this might be different. Global comparable or transparent data is hardly available. Experts do not expect voluntary disclosure of this data by companies in the future. Clear guidelines, minimum administration or obligatory global policy could be a solution.

Nuclear energy has a low carbon intensity and will therefore score well in the carbon intensity ranking. It should be noted that nuclear power has many other risks and disadvantages which are not included in this benchmark. This benchmark is only focused on quantitative greenhouse gas reductions.

5.2 Air transport sector

This sector analysis consists of the same seven paragraphs, as the sector analysis of the electricity sector. In the first two, a short overview of the relevant activities in the air transport sector is presented and an indication of the available data. In the third paragraph the data collection and the rating of the indicators is explained. This paragraph concludes with a graph, where the different indicators are plotted, based on data- and aggregation quality. For further analysis of the indicators, the collected data is presented in a company ranking for each indicator. In the fifth paragraph a summary of the interviews with Air France-KLM and ICAO is presented, followed by the main pitfalls, experienced during this sector analysis. A short sector specific conclusion is presented in the last paragraph.

5.2.1 Literature summary of relevant processes within sector

The air transport sector accounts for approximately 2% of the world's carbon emissions (WRI, 2000). Emissions from air transport are not included in country specific emissions and country specific targets. Therefore it is of extra value to benchmark this sector and stimulate global target setting initiatives to reduce emissions. The carbon performance of an airline is not only dependent on the available technologies, but also depends on the product the airplane is carrying (cargo or passengers), the speed, the distance and the type of airplane (Kroesen, 2013). Studies indicate that fuel efficiency of aviation can be improved by 40–50% by 2050 through a variety of means, including technology, operation, and management of air traffic (GEA, 2012). Most benchmarks in the air transport sector only focus on fuel efficiency because this is crucial for the costs. These benchmarks are quite aggregate (Kroesen, 2013). It is noted by the ICAO that the possession of a modern fleet of aircrafts is not sufficient as an efficiency indicator; factors as operational management, air traffic management, and airport services, etc. must also be taken into consideration (ICAO, 2011). When benchmarking the fuel efficiency, the carbon emissions due to the use of different (bio) fuels might not be included. In this research not only the flight emissions are included but also the Scope 1 emission on the ground at the gate. Data of fourteen airlines is searched, selected based on the highest absolute emissions in 2008. Two of those companies were merged or reorganised which resulted in incomparable data and one of those companies did not provide any emission. Eleven companies are further researched and ranked. Usually the reported emission data is based on the GHG protocol but this is not obliged and some companies might report their emissions in a favourable way for their company.

5.2.2 Indication data availability

Almost all of the large airlines provide a publicly accessible annual report with financial data like the annual turnover and multiple traffic performance parameters. Passenger-kilometre is the most straight forward performance parameter and very easily accessible. It represents the amount of kilometres times the amount of passengers transported over this distance. In 2010 and 2011 most of the eleven companies provided an annual CSR report with different emission parameters and some of them reported their emissions to the CDP. About half of the researched airlines reported both their Scope 1 and 2 emissions in accordance with the guidelines of the Greenhouse Gas protocol of the WRI and WBCSD.

Some airlines provided a separation of the p-km's made for short flights and for long-haul flights. Emission intensities for short flights are higher; they are carried out at a lower altitude and the emissions due to take-off and landing are relatively big. Companies with many short-haul flights could therefore be advantaged. Correcting for this would indirectly stimulate short-haul flights.

Short-haul flights are more emission intensive and could be substituted by train or other ways of transport. Aside from that few guidelines for reporting, the separation in flight distance are set. For example some companies use the separation short-haul, long-haul. Others use domestic and international flights and again others separate based on geographical region. Therefore this level of disaggregation will not be included.

Disaggregation towards cargo or passengers kilometres is possible. Overall the data availability of emissions in this sector is better than other sectors. This is confirmed by expert Sam Brand from ICAO.

5.2.3 Carbon Intensity Indicator Comparison

In this paragraph the data collection of the selected parameters is explained, followed by the rating of the separate parameters and the combined intensity indicators, as described in the methodology. In the last section of this paragraph, the indicators are plotted in a graph based on their data- and aggregation quality.

5.2.3.1 Parameter selection

From the list of parameters shown in paragraph 3.5.2 the parameters listed in Table 5.2.1 are chosen.

Emission parameters	Physical parameters	Monetary parameters
Scope 1 (ktons CO_2-e) Scope 2 (ktons CO_2-e) Operating emissions (ktons CO_2-e)	Revenue passenger kilometre (p-km) Available seat kilometre (as-km) Total revenue tonne kilometre (Total t-km)	Total turnover (bill. US$)

Table 5.2.1: The selected parameters used to form the intensity indicators for the air transport industry.

All companies reported total company turnover as well as passenger and cargo turnover separately. This is the best monetary parameter available. But the emissions are strongly related to flight kilometres and the monetary value of a flight kilometre differs a lot. It is heavily dependent on the market demand and supply, fuel prices, location and timing. Therefore it is expected that in the air transport sector a monetary parameter does not represent the actual emission performance of an airline. Only one monetary parameter is considered to indicate the difference with a product related parameter. This will be total company turnover.

Reported physical parameters for carbon intensity are revenue passenger kilometre (p-km), available seat kilometre (as-km) and sometimes total revenue tonne kilometre (total t-km).

5.2.3.2 Data collection and further parameter selection

The emission data of seven of the eleven companies is collected from their annual sustainability reports, accessed through the company website. Of four companies the Scope 1 and 2 emissions are collected from the CDP database, in those cases the CDP data was more specified and transparent than the annual reports. In about half of the annual sustainability reports, the defined terms Scope 1 and 2 emissions are not used, but operating emissions, ground emissions and MWh electricity used. Operating emissions are direct emissions due to airplane fuel burn. Most companies use the term ground emissions, for the direct emissions due to fuel burn on the ground by, for example, ground support vehicles or auxiliary power units (FAA, 2005). Scope 1 emissions include both operational- and ground emissions. In case the electricity use in MWh is reported, an average emission factor of the country where the airline is based is used, to calculate the Scope 2 emissions.

In the last three years, the average operating emissions of all airlines consist of 98% of total emissions (Scope 1 and 2). Table 5.2.2 shows, how the emissions of all the airlines, over the last three years, are distributed over the reported categories.

Reported categories	Average	Reported range	
Operating emissions/ Scope 1 in the air	98%	95,8% - 99,7%	Airplane fuel combustion (kerosene and biofuel)
Ground emissions/ Scope 1 on the ground	1%	0,2% - 1,6%	Fuel combustion ground vehicles and heating/cooling
Scope 2	1%	0,3% - 3,9%	Electricity use

Table 5.2.2: Emission distribution of total emissions (Scope 1 and 2) in the air transport industry.

It is considered using the operating emissions as parameter because they account for about 98% of all emissions. But since there is no clear guideline for reporting these emissions, and there is a guideline for the reporting of Scope 1 emissions, Scope 1 emissions will be used for the numerator of the carbon intensity indicator. Three out of the eleven companies, reported only their operating emissions. To determine the Scope 1 emission of these companies, 1% ground emissions are added to the operating emissions. The same three companies did not report their Scope 2 emissions or their MWh electricity used. And, companies which did report their Scope 2 emissions used different calculation methods at different accuracy levels. Most companies used the average electricity use of two or three sites to calculate the total kWh. Since Scope 2 emissions account for approximately 1% of total emissions in this sector, these are excluded from the analysis. A check is performed to see if using both Scope 1 and 2 emissions or only Scope 1 emissions would cause a different ranking, but since all companies shift approximately 1%, this is not the case. Both Scope 1 and 2 emissions as well as only Scope 1 emission can be used as emission metric (consistency in use is needed). Because not all companies in this research reported their Scope 2 emissions, only Scope 1 emissions are used as emission parameters in the air transport indicators.

The company turnover and physical parameters are found in the annual reports, accessed through the company websites. All companies reported the amount of p-km and the capacity in available seat km (as-km). As-km is a parameter used to show the flight capacity. Using this parameter, the load factor of the flights would not be included. The load factor is part of the operational efficiency and therefore this parameter is not further researched.

Most airlines report both their cargo-tonne kilometres (cargo t-km) and passenger kilometres (p-km) flown. The relative amount of cargo transport varies per company; for the researched companies this varies between 1% and 30%, based on the cargo turnover found in the annual reports. Looking at passenger transport alone might undervalue companies with a large share of cargo transport. Since cargo transport and passenger transport are more often combined by airlines, a parameter including both cargo and passengers is researched as well: total t-km. This indicator is reported by seven of the eleven companies. An average translation factor is used: 10.8 passengers/tonne, provided by the ICAO (ICAO, 2007). A similar translation factor is reported by some of the researched companies in their annual sustainability report; this ranges from 7.2 – 11.5 passengers/tonne. The reported total t-km, by some of the companies, depends heavily on this translation factor; the amount of passenger km is a lot higher than the amount of cargo km. Since this translation factor is an industry average, it is better to use it for the small amount of cargo km than for the large amount of passenger km. Therefore a new indicator is added: total p-km. The same average translation factor of ICAO is used, but with this indicator the amount of cargo tonnes is translated into passengers

One tonne passengers take up a lot more space than one tonne cargo; this causes different emission factors per tonne. The emission factors for passenger tonnes and cargo tonnes are found in the annual sustainability reports of four companies; Table 5.2.3. The average emissions of one passenger tonne kilometre are twice as high, as the emissions of an average cargo tonne kilometre.

	Lufthansa	Emirates	Air France/KLM	Singapore Airlines	Average
gCO_2-e/passenger t-km	1888	747	1082	940	
gCO_2-e/cargo t-km	672	555	441	540	
Pass. t-km/cargo t-km	**2,8**	**1,3**	**2,5**	**1,7**	**2,1**

Table 5.2.3: The average volume correction factor based on the passenger and cargo intensities reported by 4 companies.

If this difference is not accounted for, large cargo carriers are advantaged. Because in the combined passenger and cargo indicator, their share of 'low emitting cargo tonnes' is higher. A volume correction factor, could account for this difference. The last physical indicator researched, combines both passengers and cargo to total p-km and uses the volume correction factor for the translated cargo tonnes (total p-km+).

5.2.3.3 Indicator positioning

One emission parameter and five activity parameters can be distinguished from this data analysis listed in Table 5.2.4. Both the emissions and the turnover data are gathered at the highest level of aggregation. Data for the four physical parameters are based at the product level of aggregation. Based on the model presented in paragraph 3.5.1, the level of aggregation of the different parameters is graphically shown in Figure 5.2.1.

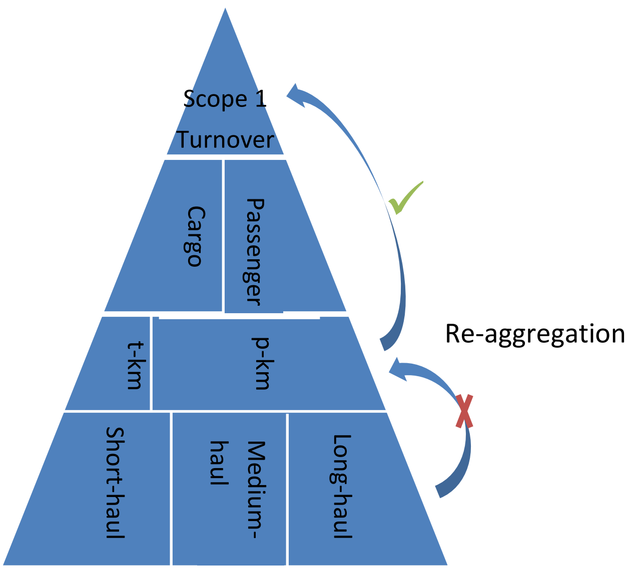

Figure 5.2.1: The aggregation pyramid of the air transport sector, which consists of 1 company level, 1 division level and 2 product levels.

To compare the different indicators, the level of disaggregation and the quality of the used data of the separate parameters are rated. The level of disaggregation based on Figure 5.2.1 and the level of data quality based on Figure 5.1 in the methodology chapter. This is presented in Table 5.2.4, extra comments on the given value for data quality are shown on the right.

Parameters	Disaggregation level		Data quality	
Scope 1 emissions	1	Company level	3.5	Mostly verified Scope 1 emissions but some different reporting methodologies which are partly corrected with company specific factors.
Turnover	1	Company level	4.0	Verified data.
p-km	3	Product level	3.5	Verified data but methodology is unclear.
Total t-km	3	Product level	2.0	Industry average weight translation factor used for all passenger km.
Total p-km	3	Product level	2.5	Industry average weight translation factor used for all cargo t-km.
Volume corrected total p-km	3	Product level	1.5	Industry average weight translation factor used combined with a roughly estimated volume correction factor for all cargo t-km.

Table 5.2.4: The disaggregation level and the data quality of the separate parameters in the air transport industry.

The four physical parameters have a different data quality. The amount of passenger kilometres is the most robust physical performance data; all airplane companies directly provide this data. The total t-km and total p-km parameters combine these passenger kilometres with cargo tonne kilometres by using an average weight translation factor. The total t-km parameter depends more on this weight translation factor than the total p-km parameter, because the amount of passengers is larger than the amount of tonnes cargo. The last parameter also translates the cargo tonnes into passengers with the weight translation factor, but, additionally, uses the volume correction factor, which is based on only four companies.

Next to the level of disaggregation per parameter, the quality of re-aggregation is rated when the parameters are combined to intensity indicators based on Figure 5.3. This is shown in Table 5.2.5.

Indicators	Re-aggregation factor	
Scope 1/turnover	0.2	No re-aggregation is needed, but turnover is too much dependent on economic factors and too little related to the emissions.
Scope 1/p-km	0.4	Parameters include different products. P-km includes no cargo km.
Scope 1/total t-km	0.6	Re-aggregating cargo and passengers correcting for the weight difference, but very dependent on the used translation factor for p-km.
Scope 1/total p-km	0.8	Re-aggregating cargo and passengers correcting for the weight difference, less dependent on the used translation factor for all t-km.
Volume corrected total p-km	1	Re-aggregating cargo and passengers correcting for the weight and volume difference.

Table 5.2.5: The quality of re-aggregation factor of the combined parameters into indicators.

The aggregation quality is based on the product of the separate disaggregation parameters (Table 5.2.4) times the re-aggregation factor of the indicators (Table 5.2.5). The data quality value is the product of the separate data quality parameters (Table 5.2.4). This results in the following values for the aggregation quality and the data quality of the different indicators, presented in Table 5.2.6.

Indicators	Aggregation quality	Data quality
Scope 1/turnover	0.2	14
Scope 1/p-km	1.2	12
Scope 1/total t-km	1.8	7.0
Scope 1/total p-km	2.4	8.8
Scope 1/volume corrected total p-km	3.0	5.3

Table 5.2.6: Indicator positioning based on aggregation and data quality.

To provide an overview of the quality of the indicators they are plotted in a graph Figure 5.2.2, with on the Y-axis the data quality and on the X-axis the aggregation quality.

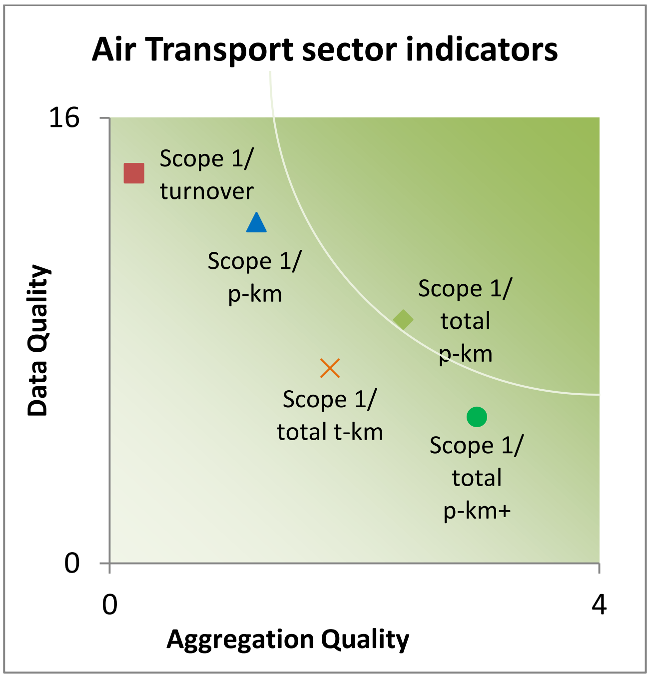

Figure 5.2.2: The different Carbon Intensity Indicators of the air transport sector plotted based on aggregation and data quality. The quality decreases with the green colour intensity. The bow encloses an area which gives an indication of indicators with a certain quality level.

The overview shows the negative influence of the translation and correction factor needed for re-aggregation on the data quality of most of the physical indicators. And in return the positive influence these factors have on the aggregation quality. Improving the data quality of these factors could significantly improve the data quality of the total p-km and total p-km+ indicators. Based on this, it is advised to use the total p-km indicator. To further analyse the different indicators and the used data, the company data is put into a ranking based on each indicator.

5.2.4 Company rankings

Collected data is used to identify and analyse differences and similarities in the formed rankings per indicator. Comparing the monetary indicator in Figure 5.2.3 with the physical indicators in the Figures 5.2.4 to 5.2.7, it is clear that it creates a total different ranking. This can easily explained by the weak relation between the monetary revenue and the used energy per km. It is minimally related to the functional products and therefore creates a wrong perception of the carbon performance of the company. Since there is enough data available for more product related indicators, the monetary indicator should not be used for the benchmarking of airplane companies.

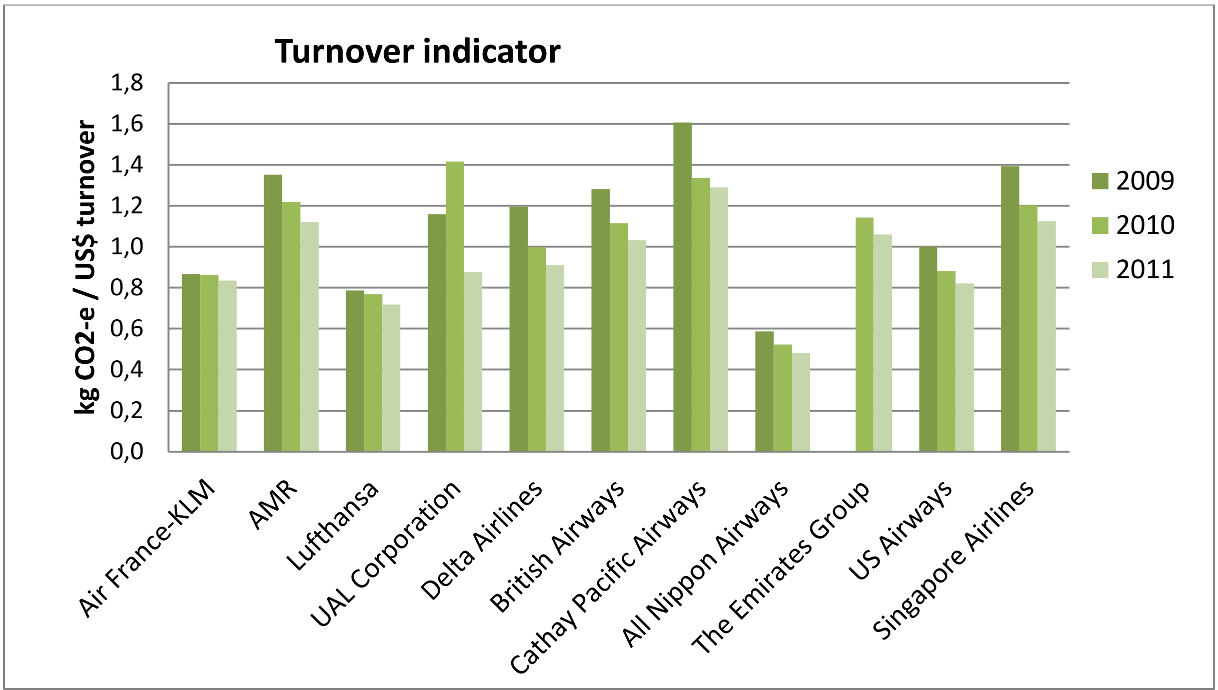

Figure 5.2.3: The eleven air transport companies compared on base of turnover, a monetary carbon intensity indicator.

The p-km indicator is frequently used by airlines to indicate their level of activity. For companies with a low share of cargo transport, this is a sufficient parameter to indicate their level of activity. This indicator could also be used to benchmark the passenger transport of airlines with a larger share of cargo. But for company level benchmarking this indicator is not sufficient; it excludes all cargo transport, while airlines combine cargo and passenger flight more often. In this research, all Scope 1 emissions are included. Therefore, companies with a relatively large share of cargo transport are disadvantaged in Figure 5.2.3. The other researched physical indicators have a way of including the cargo km. Weight is an important parameter of the amount of emissions per kilometre and it makes sense to combine the parameters on base of amount of tonnes. This results in the ranking of total t-km in Figure 5.2.5 and the ranking of total p-km in Figure 5.2.6.

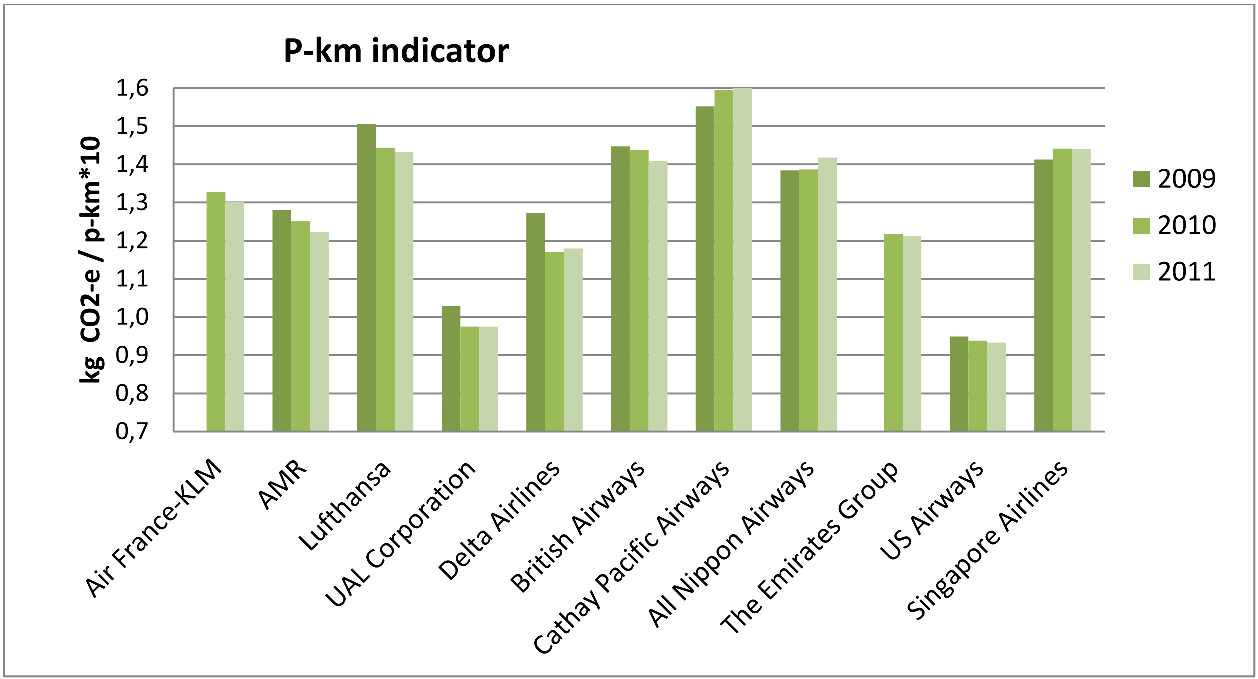

Figure 5.2.4: The eleven air transport companies compared on base of total Scope 1 emissions/revenue passenger kilometres (p-km), a physical carbon intensity indicator. Cargo kilometres are not included in the activity parameter.

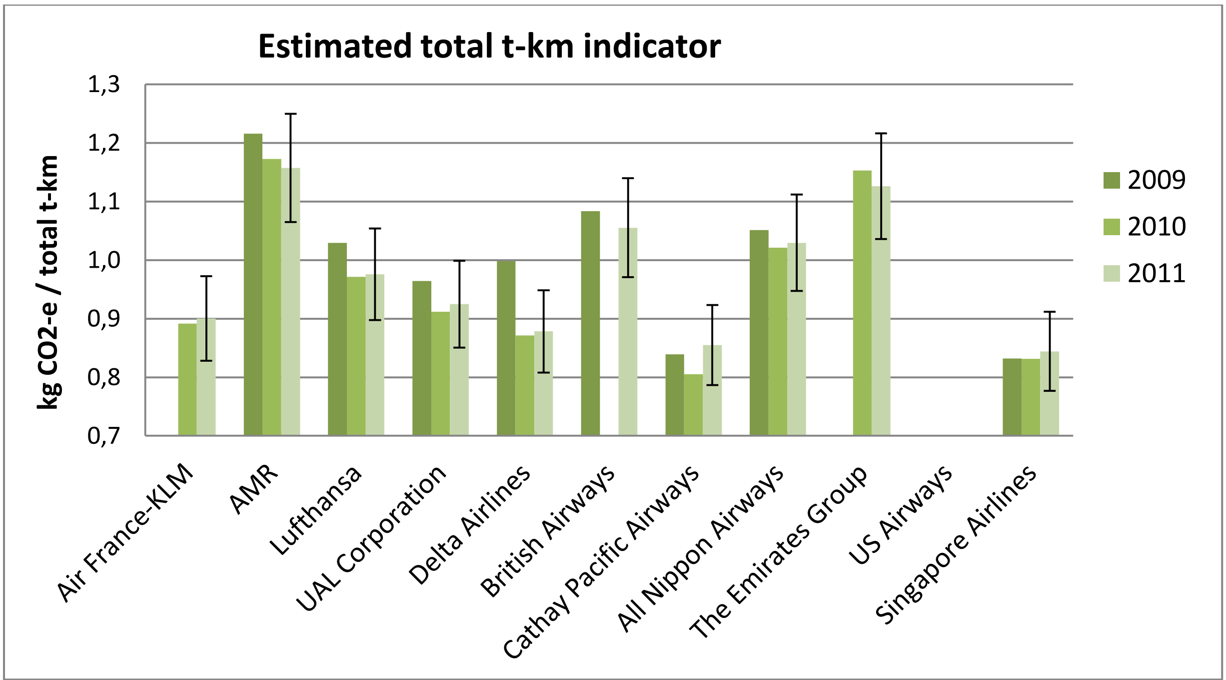

Figure 5.2.5: The eleven air transport companies compared based on Scope 1 emissions/total tonne-kilometres (t-km), a physical carbon intensity indicator which includes both cargo and passenger tonnes. The average weight translation factor, published by ICAO, of 10.8 p-km/t-km is used to translate the passengers into tonnes. The error-bars are 8%.

Comparing the total t-km ranking in Figure 5.2.5 with the p-km ranking in Figure 5.2.4 shows that companies with a significant share cargo t-km of the total t-km, like Cathay Pacific (50%) and Singapore Airlines (45%), score much better in the total t-km ranking, because their cargo km are included. A disadvantage of total t-km is that it is very dependent on the factor which translates the amount of passengers to tonnes. Because the amount of t-km is often smaller than the amount of p-km, the total p-km intensity in Figure 5.2.6 is less dependent on this factor. The impact of the translation factor on the indicators is shown in a sensitivity analysis presented in Table 5.2.7.

Translation factor	Singapore Airlines (45% cargo)				AMR (12% cargo)			
p-km/t-km	Sc.1/total t-km		Sc.1/total p-km		Sc.1/total t-km		Sc.1/total p-km	
7.8	0.70	-17%	0.90	+15%	0.86	-25%	1.11	+4%
8.8	0.75	-11%	0.85	+9%	0.96	-17%	1.10	+2%
9.8	0.80	-5%	0.82	+4%	1.06	-8%	1.08	+1%
10.8	**0.84**	**0%**	**0.78**	**0%**	**1.16**	**0%**	**1.07**	**0%**
11.8	0.88	+5%	0.75	-4%	1.25	+8%	1.06	-1%

Table 5.2.7: Sensitivity analysis on the translation factor for weight on the total t-km intensity and the total p-km intensity. 10.8 p-km/t-km is the average factor (ICAO, 2007). Singapore airlines is a large cargo carrier and AMR relatively small.

The reported range of the translation factor by six of the researched companies is 7.2-11.5. Only the Emirates Group reported a translation factor lower than 10 p-km/t-km, excluding this value the reported range is 10-11.5. If the translation factor of airlines would vary between 9.8 and 11.8, then the intensity of the total t-km indicator would differ up to 8% and the intensity of the total p-km up till 4%. It can be concluded that of these two indicators, total t-km is less robust but total t-km has the advantage that it is already reported by some companies and it is easier to visualise passengers in weight than to visualise cargo in passengers.

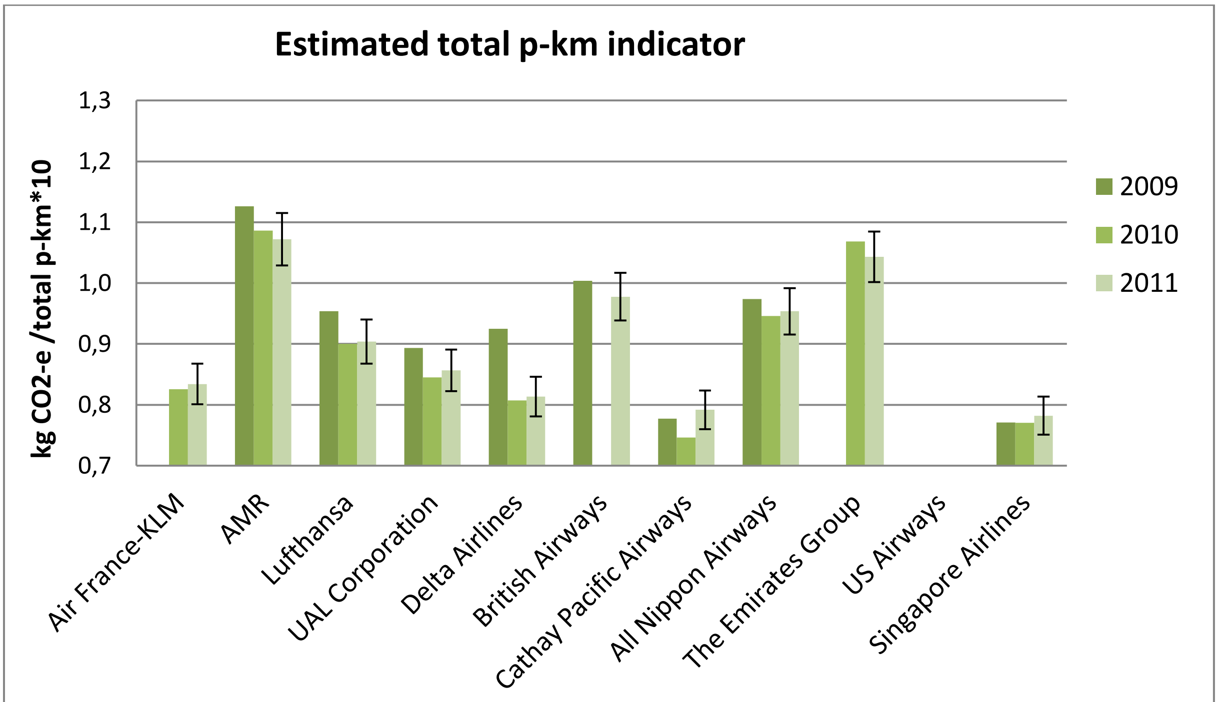

Figure 5.2.6: The eleven air transport companies compared based on Scope 1 emissions/total passenger-kilometres (p-km), a physical carbon intensity indicator which includes both passenger and cargo kilometres. An average weight translation factor, published by ICAO, of 10.8 p-km/t-km is used to translate the weight of the cargo into an equally heavy amount of passengers. The error bars are 4%.

The above mentioned weight translation assumes that average emissions of one tonne passenger-km is equal to average emissions of one tonne cargo-kilometre, but the amount of tonnes passengers fitting in one airplane is smaller than the amount of tonnes cargo. The relative weight of the airplane causes a difference in emission intensity per tonne load. Table 5.2.3 shows the separate emission intensity of cargo tonne kilometres (cargo t-km) and passenger tonne kilometres (passenger t-km) of four companies. On average this results in an emission ratio of 2.1 passenger t-km/cargo t-km. This means that on average carrying one tonne cargo has similar emissions as about two tonnes of passengers. The last ranking is the total p-km indicator corrected for the volume difference in Figure 5.2.7. All cargo t-km are not only translated to cargo p-km by multiplying with the average translation factor but are also divided by the average emission ratio to become equivalent to the emissions due to one passenger kilometre. The impact of the translation factor on the indicators is shown in a sensitivity analysis presented in Table 5.2.8.

Translation factor	**Singapore Airlines (45% cargo)**		**AMR (12% cargo)**	
p-km/t-km	Sc.1/total p-km		Sc.1/total p-km	
1.3	0.88	-15%	1.10	-4%
1.7	0.97	-6%	1.13	-1%
2.1	**1.03**	**0%**	**1.15**	**0%**
2.5	1.08	+5%	1.16	+1%
2.9	1.12	+8%	1.17	+2%

Table 5.2.8: Sensitivity analysis on the volume correction factor. 2.1 is an estimation for the average explained in Table 5.2.3.

The reported range of the volume correction factor by four of the researched companies is 1.3-2.8. The Emirates Group also reported the translation factor of 1.3. When excluding this value, the reported range is 1.7-2.8. If the translation factor of airlines would vary between 1.6 and 2.6, the intensity of the total p-km indicator would differ up to 8%.

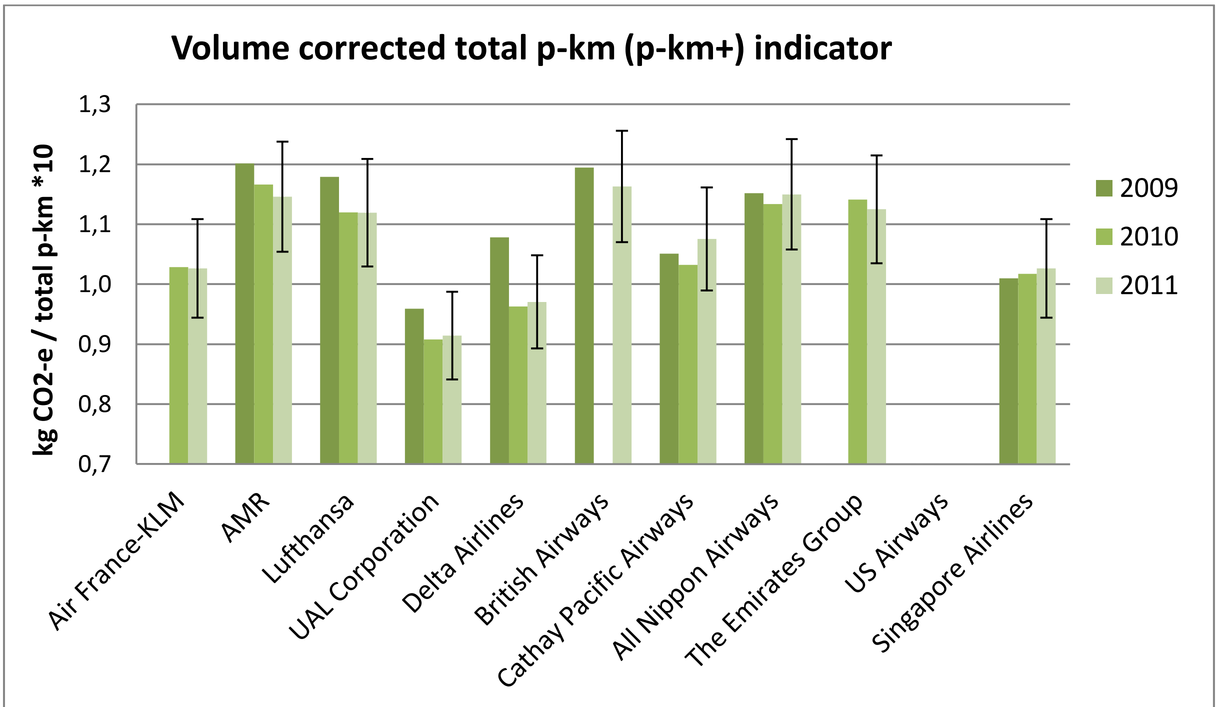

Figure 5.2.7: The eleven air transport companies compared based on Scope 1 emissions/total passenger-kilometres (p-km), a physical carbon intensity indicator which includes both passenger and cargo kilometres. An average weight translation factor, published by ICAO, of 10.8 p-km/t-km is used and a volume correction factor of 2.1. An 8% error bar indicates the sensitivity of the volume correction factor.

The use of only the cargo translation factor would stimulate cargo transport, since it would bring the overall emission intensity down. The use of the volume correction factor would filter out this advantage for cargo carriers. This volume correction might change over time and is based on very little data. Further data collection and research is needed to refine this factor. To reduce the uncertainty range reporting guidelines on how to measure this factor are needed.

5.2.5 Verification method

No publicly available ranking provides any sustainability scores of the air transport sector. Therefore only the interview with Fokko Kroesen, Environmental Manager at KLM and the interview with Sam Brand and Neil Dickson, environmental officers at the International Civil Aviation Organisation (ICAO) are used for verification of the used methodology, the decisions made and the results.

Kroesen agrees that monetary indicators are not good, because they are too little related to the emitting activity. He thinks the researched physical indicators are interesting, disaggregating to cargo and passengers is necessary. He approves of the volume correction factor for the emission intensity difference between transporting one tonne cargo and one tonne passengers. But, according to Kroesen, further disaggregation is needed. Roughly three different business models can be distinguished within the air industry: short-haul/business flights, cargo and medium-/long-haul passenger flights. He would ideally also disaggregate towards these three business models.

Dickson from ICAO doubts some of the used data. In the total p-km+ ranking, including the volume correction, has Delta airlines, a lower carbon intensity than the Emirates Group, while Delta airlines has older aircrafts. This might be due to factors besides technological efficiency, like electricity use at the airports or more luxurious flights, with more space per person. Kroesen has also doubts about publicly disclosed data. According to him, some clear reporting guidelines are missing. For example, the way passenger kilometres or cargo tonne kilometres are reported. This can either be the Great Circle Distance flown or the flight plan[5]. Reporting the flight plan kilometres is always advantageous and might give a tilted view of the operational efficiency; detours should not be rewarded. Another methodological issue is the reported weight of passengers and if extra weight for passenger chairs, food and personnel is included. This would influence the translation factor and the volume correction factor between cargo and passengers.

5.2.6 Methodological or data pitfalls

During the indicator selection and the data collection the following sector specific problems occurred

Data availability and quality

Overall emission reporting is done by most airlines, but reporting guidelines are missing. A lot of emission and activity data is publicly available and based on the GHG protocol. This protocol is not sector specific, though and still many uncertainties about the calculation methodologies exist. ICAO is working on a sector specific guiding document. Some important guidelines which are needed and recognised by ICAO are: reporting of the Great Circle Distance between airports instead of the flight plan and a standard average weight for passengers to be used in fuel efficiency metrics. Other guidelines for data reporting faced during the gathering of this research are consistent use of certain emission parameters. Scope 1 and 2 emissions are favoured, because guidelines for these emissions exist. Use of other emission related parameters like kWh electricity use, litre fuel, kilograms fuel, etc.

[5] Great Circle Distance is the shortest distance between the departure airport and the destination airport of the passenger or piece of cargo. The flight plan is the actual flown distance; this may include detours.

are less suitable for emission intensity, because related emissions may differ per company or situation.

If the air transport industry wants to disaggregate to different flight distances, agreements on the way of reporting are needed. Due to the lack of reporting guidelines, cheap airlines are able to find a way to be ranked first in certain benchmarks. This decreases the trust in the sector and should be avoided.

5.2.7 Sector specific conclusions and recommendations

The total p-km indicator is the carbon intensity indicator with a relatively high quality of data and a high quality of aggregation. The use of the total p-km+ indicator, including the volume correction, increases the aggregation quality but the data quality of this correction factor is still quite low. Further research on this volume correction factor is needed. As ICAO already recognises guidelines on an international standard passenger weight is needed as well. From this research it is recommended to research if this factor is company or region specific or a global average factor may be sufficient.

5.3 Steel sector

This sector analysis consists of the same seven paragraphs, as the sector analysis of the electricity sector and the air transport sector. In the first two, a short overview of the relevant activities in the air transport sector is presented and an indication of the available data. In the third paragraph the data collection and the rating of the indicators is explained. This paragraph concludes with a graph, where the different indicators are plotted, based on data- and aggregation quality. For further analysis of the indicators, the collected data is presented in a company ranking for each indicator. In the fifth paragraph a summary of the interviews with ArcelorMittal and Eurofer is presented, followed by the main pitfalls, experienced during this sector analysis. A short sector specific conclusion is presented in the last paragraph.

5.3.1 Literature summary of relevant processes within sector

The iron and steel industry is one of the largest emitting sectors of the world. According to the International Energy Agency, the iron and steel industry accounts for approximately 4-5% of total world carbon emissions (Worldsteel, 2009). This research focuses on steel production which is separately benchmarked from iron production, because there are many large steel producers which produce only steel. Separating these products leads to a lower level of aggregation of the benchmark.

Steel is produced from pig iron or scrap. Three main routes can be discerned in steelmaking (IPCC, 2007) and (Worldsteel, 2009):

1) The primary route is the Blast Furnace (BF) route which produces steel from primary raw materials. This route reduces iron ore to iron in blast furnaces using mostly coke or coal. The iron is then processed into steel (IPCC, 2007).

2) The Direct Reduction (DR) route, which uses natural gas to produce direct reduced iron which is mainly used as an alternative iron input in electric arc furnaces.

3) The third route melts scrap steel in Electric Arc Furnaces (EAF) to produce crude steel that is further processed (IPCC, 2007). This results in secondary steel which has a lower quality than primary steel. Figure 5.3.1 shows the main production routes used to make steel and their share in the world. An indication of the different average carbon dioxide intensity of the routes is given by a study done for Ultra Low CO_2 Steel resp. 1.97, 1.10 and 0.45 tCO_2-e/tonne crude steel (Birat, 2009). The average carbon dioxide intensity of the steel sector in 2009 was 1.9 tCO_2-e/tonne crude steel (Birat, 2009) (Worldsteel, 2009). The carbon intensity values include direct and indirect emissions from coke making, sintering, iron making, casting and rolling. Mining is not included.

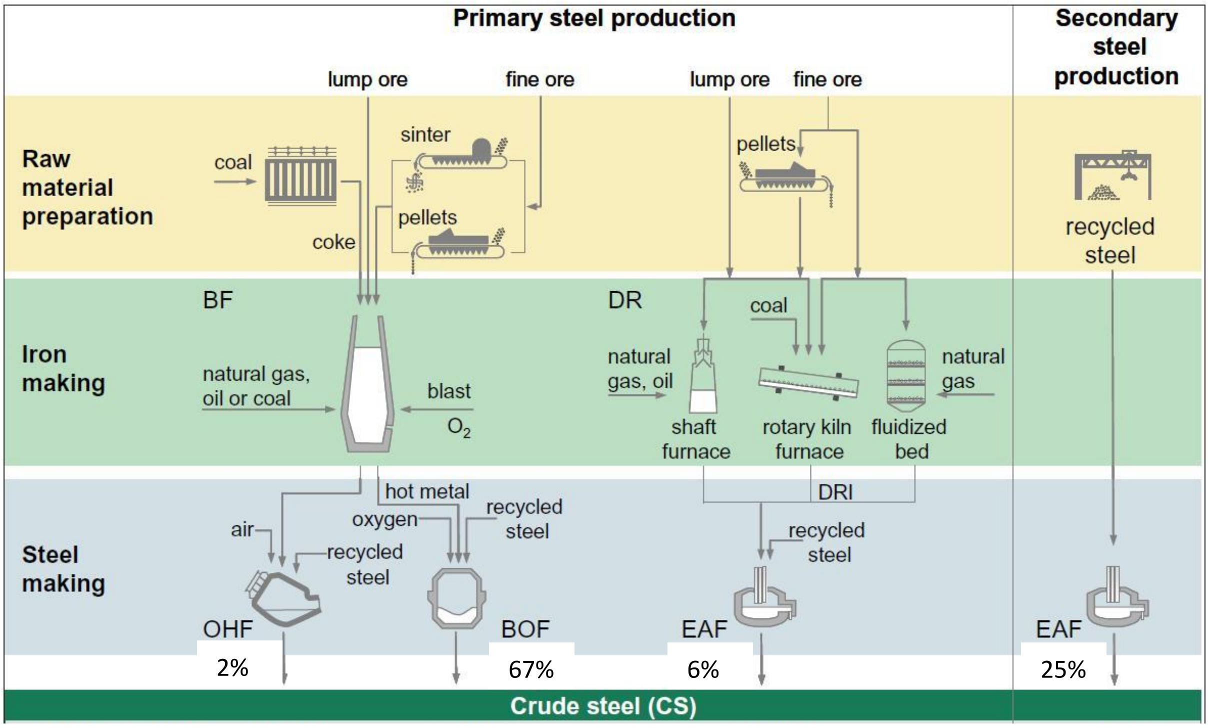

Figure 5.3.1: Steel production routes and the production shares in the world. BF: Blast Furnace; OHF: Open Hearth Furnace; BOF: Basic Oxygen Furnace; DR: Direct Reduction; EAF: Electric Arc Furnace. This figure is for illustrative purposes only, as the steelmaking process can vary from one facility to another (Worldsteel, 2009).

5.3.2 Indication data availability

In annual sustainability reports and the CDP database, carbon emissions are reported by about half of the large steel companies. Most reported their Scope 1 and 2 emissions, but, for example Tata steel reported its Scope 1 and 2 emissions, including a part of Scope 3. Other companies reported their emissions without any further explanation on the used calculation method. Some very large companies, like ArcelorMittal, which also have mines, often report their Scope 1 and 2 emissions excluding mining.

Most of the companies which reported their emissions also reported produced tonne crude steel as activity parameter. The exact way of reporting is different per company; some companies reported tonnes of sold steel or tonnes of shipped steel. But overall tonnes of crude steel production are an accepted way of reporting. If other metals were produced by the same company, the production in tonnes was always separately reported. Some companies also provided an overview with the tonnes of specific processed-product like tonnes hot rolled steel or tonnes steel sheets.

Most companies also indicated their own calculated physical carbon intensity in their annual sustainability reports or at the CDP. The used parameters for this intensity indicator were often based on their own reporting methodology of the separate parameters.

In annual reports company turnover data is found. Some very large companies with different products, like Sumitomo, provide separate turnover data for the steel division.

5.3.3 Carbon Intensity Indicator Comparison

In this paragraph the data collection of the selected parameters is explained, followed by the rating of the separate parameters and the combined intensity indicators, as described in the methodology. In the last section of this paragraph, the indicators are plotted in a graph based on their data- and aggregation quality.

5.3.3.1 Parameter selection

From the list of parameters shown in paragraph 3.5.2 and based on the available data the parameters listed in Table 5.3.1 are selected.

Emission parameter	Physical parameter	Monetary parameter	Reported physical indicator
Scope 1 and 2 (ktons CO_2-e)	Produced crude steel (tonnes CS)	Company turnover (bill. US$)	tCO_2-e/tonne crude steel (tCS)

Table 5.3.1: The selected parameters used to form the intensity indicators for the steel industry.

Next to these separate parameters, the reported physical carbon intensity is also collected. The final products of the steel industry can be very diverse. One of the categorisations ArcelorMittal uses are flat products, long products and pipes and tubes. Other categorisations are sheets, wires, stainless steel etc. Often steel companies claim to have a unique kind of a certain category (Eurofer, 2013). Most of the emissions take place during the iron making process and the crude steel making process. Therefore tonne of crude steel production is used as physical parameter and this is not further specified into final-products. If the company also produces significant amounts of other metals, the steel turnover is used.

5.3.3.2 Data collection

The steel companies with the highest emissions worldwide are compared. For this research, data of at least ten companies is collected, in order to get a reasonable comparison. 25 Companies are sought to assemble emission and performance data. Sixteen of those 25 companies did not provide their absolute Scope 1 or 2 emissions in publicly available data. They did not report to the CDP, nor did they have an annual (sustainability) report nor were their carbon emissions reported on their websites. In two of those cases, relative emissions compared to a baseline were provided, but since the used methodology was unknown, those were not useable. It can be concluded that there is only little transparent emission data publicly available. To come to a total amount of ten steel companies, one European steel company, further down the list of highest emitting companies, was added. This was done to save time on further research of companies with a low chance of finding emission data. This company is Rautaruukki, which reported its absolute emission to the CDP.

Of the ten researched companies, four did not report their Scope 2 emissions in all years, which is therefore estimated based on the reported electricity and a world average emission factor of 0.504 tCO_2/MWh (IEA, 2012), or a company average of other years. Tata steel included part of Scope 3 emissions, since it is unknown which part is Scope 3 and which part Scope 1 and 2, this is not corrected. Only the emissions due to steel production of ArcelorMittal are taken into account because of the size of the holding and the mining activities. ArcelorMittal South Africa is excluded from the holding annual report and therefore also seen accounted as a separate company in this research. POSCO included the carbon credits it has bought.

Due to interest of investors most companies publicly disclosed an annual report with performance data. Turnover and crude steel production was relatively easy to find. For one company, the annual report data was only available in Chinese and one company was a private company and did not provide an annual report.

5.3.3.3 Indicator positioning

One emission parameter in combination with one physical parameter and one monetary parameter are researched. The company reported physical intensity indicator is also included. Both the emissions and the turnover data are gathered at the highest level of aggregation. Data for the physical parameter is based at the product level of aggregation. Based on the Aggregation Triangle presented in Figure 4.4.1, the level of aggregation of the different parameters is graphically shown in Figure 5.3.2.

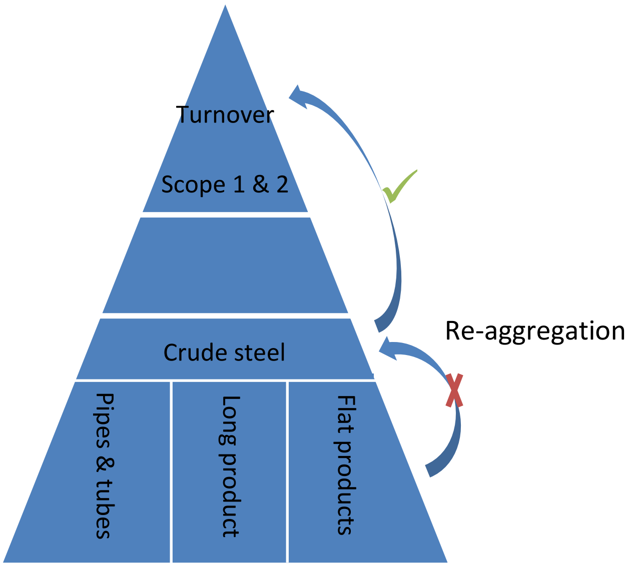

Figure 5.3.2: The aggregation pyramid of the steel sector, which consists of 1 company level, 1 division level and 2 product levels.

To compare the different indicators, the level of disaggregation and the quality of the used data of the separate parameters are rated. The level of disaggregation based on Figure 5.3.2 and the level of data quality based on Figure 4.1 in the methodology chapter. This is presented in Table 5.3.2, extra comments on the given value are shown on the right.

Parameters	Disaggregation level		Data quality	
Scope 1 and 2 emissions	1	Company level, if other metals are produced steel specific emissions.	2.5	Most used emission data is reported but different methodologies are used. Most are corrected based on company specific average data, but for some companies industry average data is used.
Turnover	1	Company level	4.0	Verified data
Produced crude steel	3	Product level	3.5	Verified or reported data. Not always crude steel production is reported, sometimes variations like sold or shipped crude steel.
Reported Emissions/ crude steel	1/3	Combination of company and product level	2.5	Reported data, but methodology is often not transparent which decreases data quality. Some company parts could be excluded.

Table 5.3.2: The disaggregation level and the data quality of the separate parameters in the steel industry.

Next to the level of disaggregation per parameter, the quality of re-aggregation is valued based on Figure 4.3. This is shown in Table 5.3.3.

Indicators	Re-aggregation factor	
tCO_2-e/ Turnover	0.8	In most cases both turnover data as well as emission data are at company level and no re-aggregation is needed. In some cases emission data is specified for the steel production and the parameters include not the same part of the company.
tCO_2-e/ tCS	0.8	Overall the same product is included (crude steel). Emissions due to further modification of the crude steel are allocated to the crude steel production. This causes a small discrepancy between the two parameters, which accounts for only a very small part of the emissions.
Reported tCO_2-e/ tSteel	1.0	The parameters are fitted by the company itself and matching of the two separate parameters is possible. No further re-aggregation is needed.

Table 5.3.3: The quality of re-aggregation of the combined parameters into indicators.

The overall aggregation quality is based on the product of the separate disaggregation parameters (Table 5.3.2) times the re-aggregation factor of the indicator (Table 5.3.3). The data quality value is the product of the separate data quality parameters (Table 5.3.2). This results in the following values for the aggregation quality and the data quality of the different indicators, presented in Table 5.3.4.

Indicators	Aggregation quality	Data quality
Scope 1 and 2/ Turnover	0.8	10
Scope 1 and 2/ tCS	2.4	8,8
Reported tCO2-e/ tSteel	3.0	6,3

Table 5.3.4: Indicator positioning based on aggregation and data quality.

To provide an overview of the quality of the indicators they are plotted in Figure 5.3.3, with on the Y-axis the data quality and on the X-axis the aggregation quality.

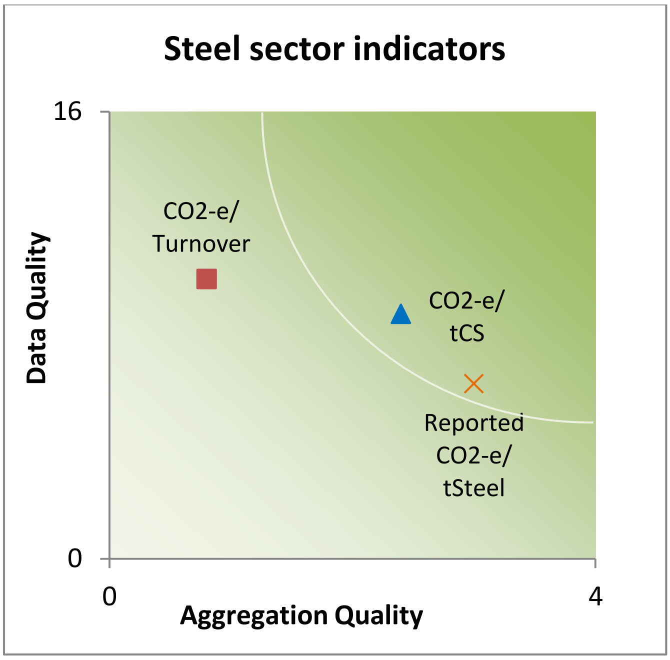

Figure 5.3.3: The different Carbon Intensity Indicators of the steel sector plotted based on aggregation and data quality. The quality decreases with the green colour intensity. The bow encloses an area which gives an indication of indicators with a certain quality level.

This overview shows that, at this moment, the physical indicators are preferable. Separate reporting of parameters has a positive effect on data quality, but reporting of a direct intensity indicator is better for the fit between the emission and the activity parameter which results in a higher aggregation quality. Based on this, it is advised to use separate parameters to create the physical carbon intensity indicator. To further analyse different indicators and used data, company data is put into a ranking based on each indicator.

5.3.4 Company rankings

The collected data is used to identify and analyse differences and similarities in the formed rankings per indicator. In Figure 5.3.4 the ranking based on the turnover is presented.

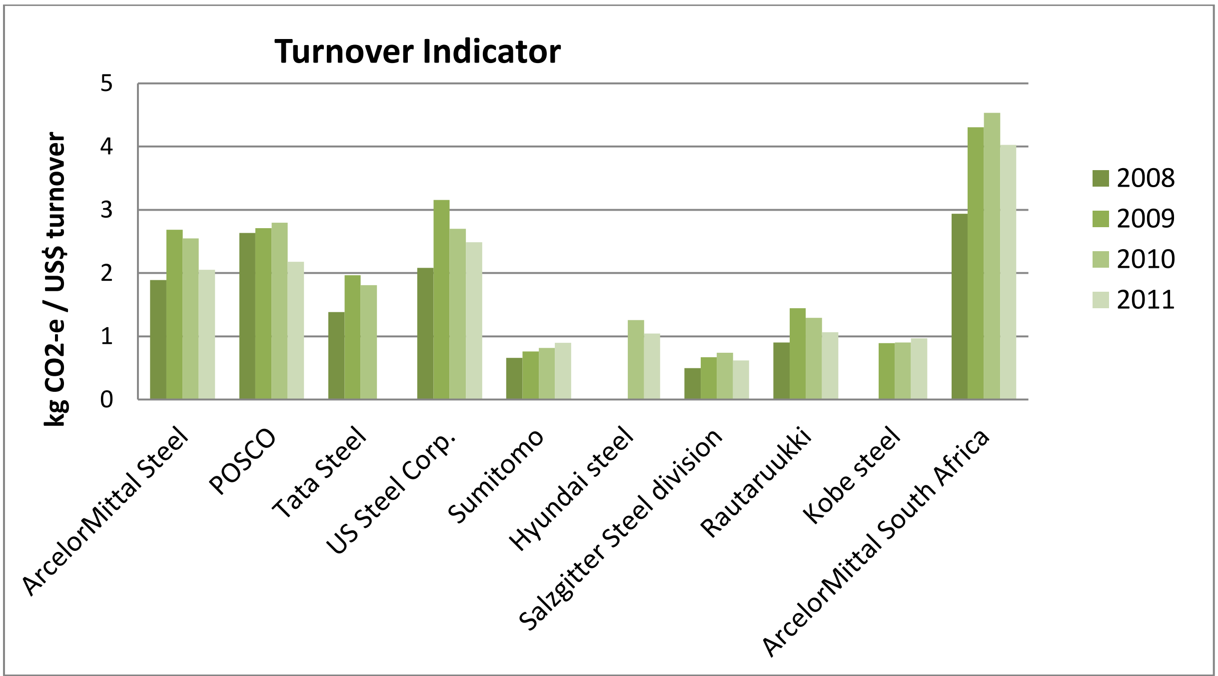

Figure 5.3.4: The ten steel companies compared on base of turnover, a monetary carbon intensity indicator.

Comparing this ranking with the physical indicator ranking in Figure 5.3.6, it is clear that they are significantly different and are not interchangeable. It is notable that in Figure 5.3.4 roughly three groups can be distinguished, respectively groups of 1, 2 and 4 kg CO_2/US$. This could indicate a different calculation methodology, but because the turnover data includes many structural effects, other reasons like size or region of the company can be thought of as well. One of the possibilities is related to the corporate structure of the steel companies. In Figure 5.3.4 total company turnover is used for the company Sumitomo. In Figure 5.3.5 the difference in choice for company turnover in the carbon indicator is presented vs. the carbon indicator with only steel turnover as monetary parameter. Using only steel turnover is at a more disaggregate level, but using this only for Sumitomo, the indicator would on average be five times as high as the other companies. Not all companies report their separate steel turnover and therefore this higher level of disaggregation cannot be used. The difference gives an indication of the structural effects included in this monetary indicator. Further research is needed to investigate the reason behind these differences.

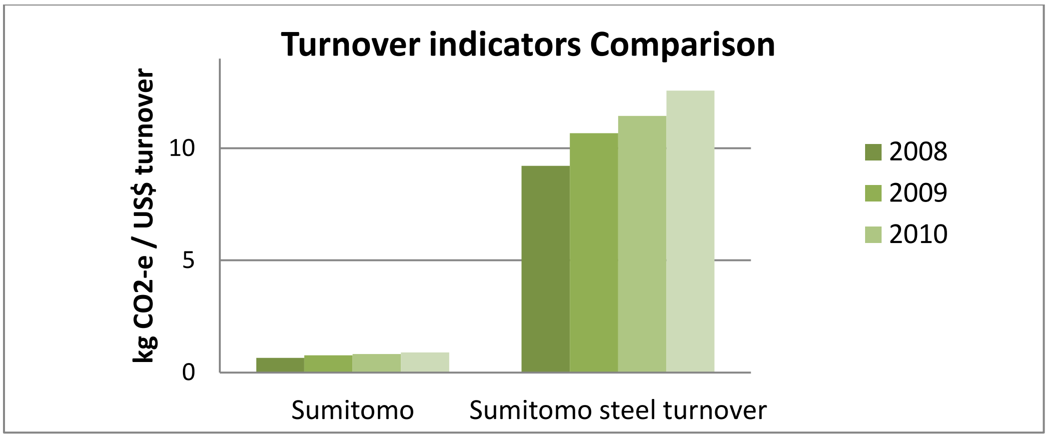

Figure 5.3.5: The difference between the monetary carbon intensity indicators based on company turnover vs. steel division turnover. Sumitomo is used to illustrate this difference.

The ranking based on the crude steel production intensity indicator is presented in Figure 5.3.6.

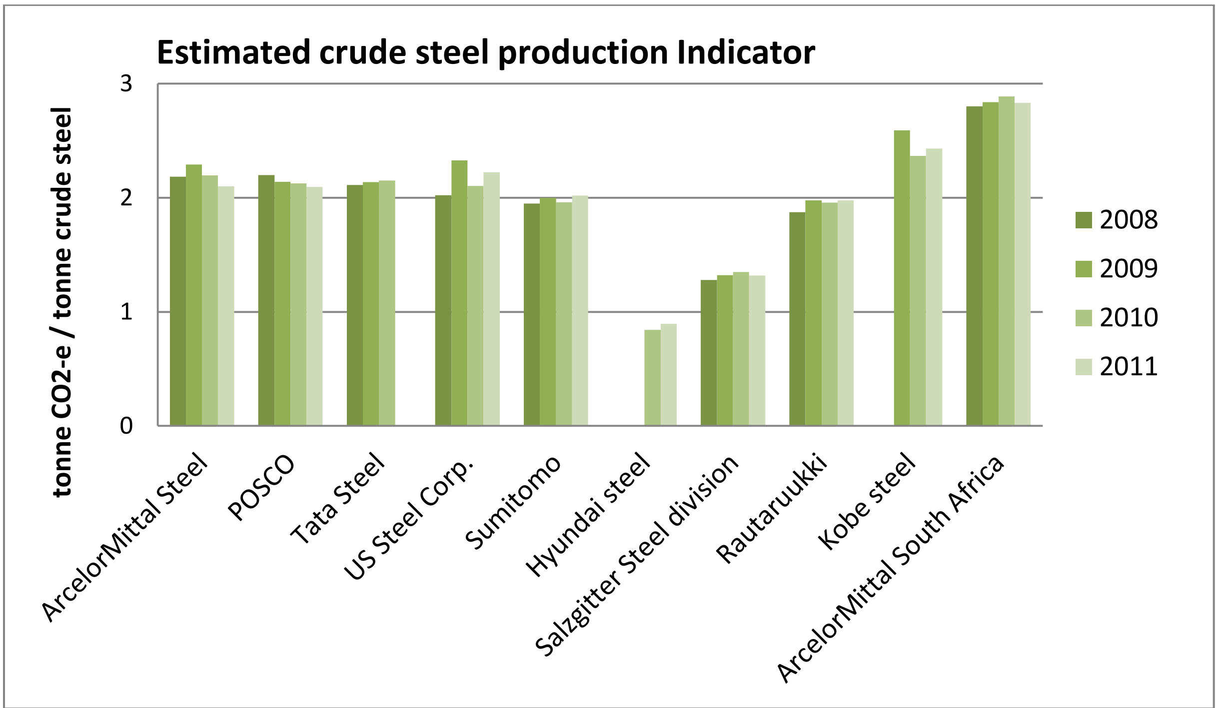

Figure 5.3.6: The ten steel companies compared on base of crude steel production, a physical carbon intensity indicator.

The estimated crude steel production indicator ranking shows quite a stable picture over the years. Two companies have significantly lower emission intensity. The two companies with this lower emission intensity did not directly report their physical carbon intensity as can be seen in Figure 5.3.7. It is likely that this significantly lower carbon intensity is due to a larger share of secondary steel production or the calculation methodology is different. Further disaggregation is needed to separate primary steel from secondary steel production. This separation is made by some of the steel companies, but the way of reporting was very different which resulted in incomparable data. This again shows the importance of reporting guidelines.

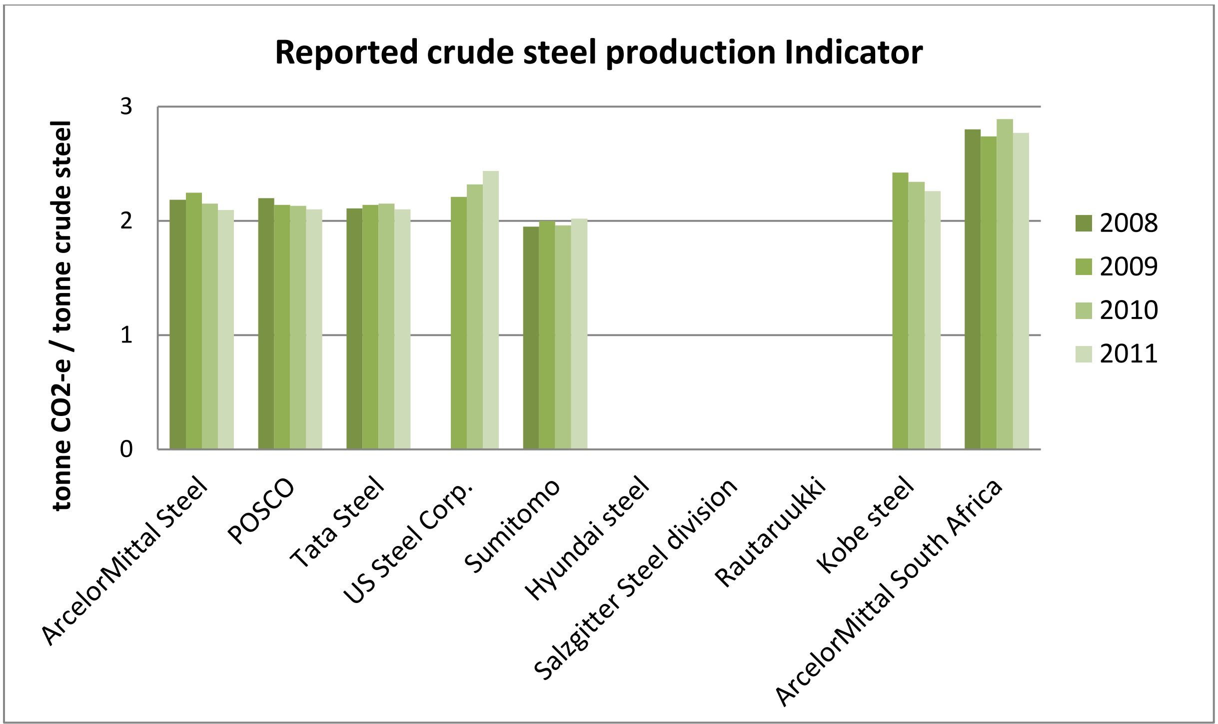

Figure 5.3.7: The ten steel companies compared on base of directly reported intensities. The physical carbon intensity indicator based on crude steel production is used.

5.3.5 Verification method

The CDP database provides scoring of six of the researched steel companies, but since this scoring is based on many other indicators, like carbon management systems, comparison with this benchmark will not be useful for verification. No similar quantitative ranking provides any carbon efficiency ranking or scores of the steel sector. Therefore only interviews will be used as verification method. Interviews are done with Karl Buttiens, Director of Environment & Global CO_2 strategy at ArcelorMittal, Buttiens is well known in the steel sector for his knowledge about benchmarking. Danny Croon, Director Environment at Eurofer and David Valenti, Climate Change Manager at Eurofer. The interviews are used to get feedback on the used methodology, the decisions made and the results.

Buttiens explains that as soon as indicators are defined, distortions are inevitable. Given a free choice of indicator, many steel producers could define one in which they are the best. This is enhanced due to the fact that a long series of carbon intensive processes are benchmarked which are not in general well aligned on site level. Eurofer recognises that compromises and choices are almost implicit in benchmarking; the choices should be limited as much as possible, in order for it not to weaken the methodology. Buttiens emphasises the complexity of the steel production process and the presence of some uncertain cross relational effects between processes. Energy savings in one compartment may adversely affect the product properties which can cause an additional energy input in some downstream process. In his view all those effects should be included to know which company is the most efficient in an overall and unbiased benchmark.

The trends described by the graphs are not directly recognised by Buttiens. This may be due to the lack of a generally accepted methodology of the publicly available figures. Assumptions behind each figure may be unknown. In general, Buttiens believes that the used indicators do not take sufficient care of structural differences between the reporting parties. Further disaggregation is needed, for the benchmark indicator to be useful.

Valenti from Eurofer agrees that the monetary indicator, tCO_2-e/turnover, is not useful; according to him, tCO_2-e/EBITDA would be more interesting, because then "the price of carbon you can survive with" might be expressed. Eurofer agrees a physical indicator is best, but the benchmarking should take place at a lower level of aggregation.

Specific issues which need a reporting guideline are the simultaneous production of several by-products along the main product like coke, gas and electricity. Another important issue is the accounting of scrap which everyone uses to some degree. If this is not taken care of, the scrap input is benchmarked. At the moment there is a discussion going on if scrap is useful or not and if the performance should be neutralised from scrap. Overall both Buttiens and Eurofer have doubts about the usefulness of a company benchmark the way it is presented in this research due to the many limitations of the indicators researched.

5.3.6 Methodological or data pitfalls

During the indicator selection and the data collection the following sector specific problems occurred

Data availability and quality

Data availability and quality is the biggest problem in benchmarking the steel industry. The small amount of consistent environmental data publicly available is remarkable for such a high emitting sector. No absolute emission data was found for 14 of the top 25 high emitting steel companies. Only nine of the top 25 high emitting steel companies reported their emissions in a sustainability report, website or to the CDP. Most steel companies do write something about their green initiatives and targets but immense companies like Nippon steel or Frans Haniel do not provide disclosure of their absolute emissions. For further research data disclosure agreements are needed to be able to benchmark these companies. Performance data was easier to find because this is needed by investors, but different ways of reporting are used. This reduces the quality and therewith the usability of the data. This is confirmed by the sector stakeholders in the interviews.

Another indication of different calculation methodologies is the differences in reported direct intensity and the intensity based on the two parameters shown in Figure 5.3.6 and 5.3.7. And the three steel companies which both reported to the CDP and on their website or annual sustainability report, reported 5% less emission to the CDP. This can be explained by different calculation methodologies in the company or exclusion of some units. To be conservative the figures from the annual reports are used if both reporting ways are used.

Structuring of companies' activities

The steel industry is less homogeneous than first expected with quite some by-products and many different sized companies. In this sector it is important to define boundaries of activities that will be included. For example in this research the by-products are not considered. Neither are the activities iron ore mining and the production of other metals. Therefore it is very important to match the inclusion of these activities between the nominator and the denominator. The monetary indicator is likely to include all these activities, disaggregation is even more important. The physical indicator often has a lower level of aggregation than the emission parameter. This increases the possibility of matching these to parameters. For the physical parameter disaggregation to primary and secondary steel is recommended, because these products have a significantly different intensity. Reporting agreements on the amount of included scrap are needed to be able to do re-aggregate these different products properly.

Reporting issues

A clear reporting issue is related to the amount of scrap used in the production process. It can be debated whether the performance should be neutralised from scrap or is seen as a useful product. The same goes for secondary steel. This steel has a lower quality, but is still used for many applications. A good way to account for this difference should be agreed upon by the involved companies and stakeholders. Another issue is how to account for the production of by-products like coke, gas and electricity. These have value as well and some emissions could be allocated to these products if they are used. More specific data disclosure is needed to account for these products. An issue related to the monetary indicator are the different prices of steel per tonne over the whole world. In China steel is much cheaper than in Europe.

5.3.7 Sector specific conclusions and recommendations

CO_2-e/tonne crude steel is the most accepted benchmark indicator with a relative high data and aggregation quality. In spite the limitations mentioned by the sector stakeholders which are implicit when a benchmark at company level is performed, this physical indicator is recommended to use i.e. for target setting purposes. Reporting guidelines within the industry could even further improve the quality of this indicator.

It is not recommended to directly report the whole intensity indicator because the credibility decreases when the methodology is unknown. When absolute emissions and performance data is separately disclosed, the data can be used for multiple purposes and more easily transformed to a comparable indicator.

Based on the found emission data it is recommended to further investigate the possibilities including Scope 3 emissions. Avoided emissions in this sector could be very interesting. Since these emissions are even more complex and reporting guidelines are less mature, sector involvement is needed including agreements about data disclosure and reporting guidelines.

6. Conclusion and recommendations

Sector specific, target based benchmarking at company level is a powerful tool for establishing and monitoring emission reduction targets. In the first part of this study, an inventory was created of the current sustainability rankings and the carbon indicators used. From this inventory it could be concluded that there are no global company rankings focussed on quantitative emission reductions. Such global rankings can be used for target setting and are expected to be of added value to bridge the emission gap. In the second part of this study different kinds of sector specific indicators were analysed for the electricity sector, the air transport sector and the steel sector. Sector specific indicators with adequate levels of both data and aggregation quality were found for two of the three sectors. The selected indicator for the steel sector still contained several structural effects.

The following sector specific conclusions were drawn from the indicator analyses: the physical indicator ‘tCO_2-e per MWh’ turned out to be the best one available for the electricity sector. The EU ETS database had offered verified emission and performance data at plant level, which could easily be combined. An allocation factor for heat has been used in the case of CHP plants. The main disadvantage of the selected indicator was that heat production, electricity trading and gas production had been excluded.

From the indicator analysis of the air transport sector it could be concluded that the physical indicator ‘kg CO_2-e per total passenger kilometres’ was the best one available for benchmarking at company level. A suitable way for re-aggregating the product level data into a company level indicator was found. This indicator combined the cargo tonne kilometres and passenger kilometres, whereby a weight translation factor was used to translate the cargo tonnes into passengers. An additional volume correction for this same indicator had been researched as well; however the data quality was still too poor to be included. This volume correction factor would have further increased the quality of re-aggregation. Availability of emission data is relatively large in this sector, as most companies provide annual sustainability reports. The main drawback of this indicator was the comparability of the emission data as well as the performance data.

The last indicator analysis showed that the physical indicator ‘tonne CO_2-e per tonne crude steel’ was the best one for the steel sector. But, interviews with stakeholders showed that further disaggregation to primary and secondary steel is still needed. The comparability of performance data, at such a low level of aggregation, should be improved. The selected indicator could, however, be a useful start for further research. Current work on a more product specific benchmark, by ArcelorMittal and Eurofer, could be of use as well.

In the three sectors researched only aggregate data of adequate data quality was available for the monetary indicators; this resulted in several structural effects. As far as the air transport is concerned, for example, both cargo and passenger transport are generally included in the turnover, though these products have different emission intensities. The relative share of each product is not visible in turnover parameters.

Product specific data was used for physical parameters. A way to re-aggregate the product specific data of each sector to company level was found. This resulted in fewer structural effects in the physical indicators, as compared to the monetary indicators. The stakeholders interviewed agreed that the monetary indicator was not a useful one for benchmarking technical and operational efficiency in their sectors.

This study has clearly revealed the pitfalls of publicly disclosed data in these sectors. The main problem is that publicly available data is often provided on a voluntary basis, and allows companies to choose their own methodology and way of reporting. The result was a variety of reported units, indicators and processes. Some companies used a 'flattering' methodology, setting their own boundaries, or a completely opaque one. It was therefore very hard to estimate the quality of this data and in some cases prevented comparison with other companies. In the stakeholder review it was confirmed that voluntary high quality data disclosure is an issue in all three sectors and is not likely to change without agreements. In their experience, most companies indicated that they are reticent about transparent data disclosure, because they are afraid of misinterpretation and misuse by competitors or activists. Bias and inaccuracy in the re-aggregation methodology hereby pose a major problem.

The main recommendation evolving from this research is that data quality should be further improved. This remains the biggest restriction towards the initial level of disaggregation of the data, and the possibility of high quality re-aggregation to company level. Therefore, improving the data quality and availability will directly affect the indicator's precision.

The quality of emission data could be highly improved through the increase of internationally accepted reporting guidance per sector. Guidance is needed to increase the transparency of calculation methodologies, for both emission and performance parameters. Agreement is needed on which activities of the sector are benchmarked, and accordingly, on what part of a company should be included in the particular sector benchmark. Some sector organisations, like the ICAO, have already initiated this, and are working on reporting guidelines with various countries.

It is recommended that sector organisations initiate sector specific reporting guidance, and potentially decide on benchmark guidelines in cooperation with the companies involved, so as to create a benchmark that is not driven by investment, but by technical efficiency. Preferably, company level benchmarks should be initiated in cooperation with the companies concerned, as data disclosure and in depth operational knowledge is needed. If necessary, data of their members could be disclosed confidentially.

Though the sectors researched are, in general, relatively homogenous, even for these sectors re-aggregation of the different emitting activities was needed. Company level benchmarking would be most useful for high emitting sectors in which the major share of emissions can be allocated to only a few products or activities. It is therefore recommended to perform an indicator analysis for each high emitting sector, and focus on physical indicators. A separate analysis per sector is required so that key products and processes can be identified. In heterogeneous sectors such as the chemical industry, it might be necessary to specify more detailed economic activities, for example, through the use of a subsector benchmark for companies with relatively similar activities.

The data quality of publicly disclosed data constitutes the single largest problem. Although acceptance of the benchmark indicator and reporting guidance would significantly increase the data quality, they might still be insufficient to provide a useful company benchmark.

In order to avoid resistance and increase cooperation with companies, it is recommended to communicate target based benchmark studies at company level in another way. A different terminology could perhaps be used to replace the term 'benchmark'. The latter suggests exactness, whereas benchmarking the efficiency of companies is merely an indication at a higher level of aggregation. The focus on target setting may be an opportunity to communicate the comparisons between companies in a positive way.

The sector interviews showed that carbon performance benchmarking is a "hot topic" among companies and sector organisations. Large disclosure initiatives like CDP and GRI are moving towards a more sector specific approach. This study provides a profound base on which further company level indicator research can be based. The final recommendation is to start target based benchmarking in the high emitting sectors and to first use an indicative indicator. It will increase the discussion on reducing carbon emissions and ways of reporting between companies in these global sectors, and lead to greater co-operation between stakeholders to further improve the quality of this indicator. Even though a certain level of aggregation is inevitable, the higher level of aggregation makes company level benchmarking an excellent tool for target setting and guiding emission reductions in the sector. An overview of the progress of the efficiency improvements can be used by sector organisations, NGO's and governments. It provides companies of each specific sector with a tool to collectively take responsibility for their part of the emission gap.

7. Discussion

Benchmarking in itself is a complex and delicate matter, as it can never be carried out objectively. The purpose of a benchmark always determines which indicator is the best one to use. Thus, in order to choose between indicators, the purpose of the accompanying benchmarks should be defined first. The aim of this study was not to create a specific benchmark, but to investigate the possible indicators from a scientific point of view. So as to come to the best indicator per sector, several important decisions were made. The most important decision in this study was the one determining criteria on which the indicators should be rated, and in this case the criteria were based on aggregation quality, data quality and sector opinion. The main decision was made following the study of various research reports and practical experience by Ecofys, but, as extensively explained in this research, this decision could be debated. Theoretically, the level of aggregation of an indicator should always be intensified. If data availability and quality were unlimited, creating a most disaggregate benchmark would be possible with an increasing amount of effort and time. But this hypothetical situation was not the case in practice, and certainly not when publicly available data is used. A more practical point of view, through the criteria of data quality and sector opinion, was therefore also included.

Another important decision related to the methodology used to quantify the quality of the indicators. This was established through three scoring models (Figure 3.5.1, 4.1 and 4.3). The first one involved data quality quantification based on a model often used by Ecofys, to indicate the data quality (Ecofys, 2013). This scoring was more indicative than fixed, however, and it could be debated whether reported data established through an unknown methodology is of a higher quality than that which has been modified by well-accepted industry average factors. The level of influence of a modification on the final value was also important. This feature was taken into account in a qualitative manner by way of logical reasoning. The level of data quality could be improved by rating all the separate data points; in this study the quality of data was estimated per company, based on the data source used and the influence of modification factors. The average rating of that quality was quantified, and rounded off by half points. The second scoring model, the so-called Aggregation Triangle, was based on a large benchmark study by Utrecht University. In this model the quantification of the aggregation levels provided a mere indication of the differences between indicators, as the amount of aggregation levels and the corresponding numbers were not fixed. The third and final scoring model dealt with the quantification of an estimated re-aggregation quality. This estimation was based on two main requirements, homogeneity and inclusion of the whole company. Both requirements were based on logical reasoning, rather than on literature. All in all, though the evaluation of indicators revealed a certain level of subjectivity, every decision was scientifically argued.

Sector specific decisions and limitations also had some influence on the final conclusion. In the electricity sector only one non-European company was benchmarked, as hardly any global emission or performance data had been disclosed. This may have affected the estimated importance of heat allocation and thereby the selection of the best indicator. Other decisions related to the choice of heat allocation factors employed and the decision to exclude certain activities.

In the air transport sector, comparable data availability was the biggest limitation. An important example of this was found in the way in which passenger kilometres were reported. Although generally it is the most basic performance indicator, companies either use the Great Circle Distance between the destination airports for their reports, or else rely on flight plans. This may have had a large impact on the established rankings, but it was less important for the choice of the best indicator in this study. The decision to use the average weight translation factor, reported by ICAO, rather than the weight translation factor reported by some of the companies, had a significant influence on the carbon intensities. The Emirates Group, in particular, reported a very low weight factor as compared to the average weight factor. As this greatly influenced the ranking, the average weight factor reported by ICAO was used. For further research it may be recommended to investigate whether or not the weight of an average person is significantly different for each airline. The volume correction factor was only based on four companies, and contained many uncertainties. This factor is likely to differ per airplane and per company. For further research it is also recommended to investigate the company acceptance of the total passenger-km indicator, and to compare the volume correction factor of various companies, as these indicators have not been used. Overall reporting guidelines to improve the data quality of the reported indicators are essential in this sector. They are to prevent low cost airlines from finding a methodology to be rated first in some rankings, which would decrease the credibility of benchmarking in this sector.

In the steel sector the main choice concerned the structuring of the activities involved. Some of these were excluded or simplified to increase the quality of re-aggregation. The production of steel is a complex process, and in this study, only the activities up to the fabrication of crude steel were taken into account. The process produces by far the largest share of emissions. Production of other metals and mining of iron ores were excluded. The best selected indicator did not distinguish between primary and secondary steel, which would very likely have highly influenced the ranking. The way of reporting the production of different kinds of steel differs per company, making diversification complex. It is recommended to investigate the possibilities for further disaggregation to this level, by either using average factors and thereby decreasing the data quality, or by finding a way to agree on reporting guidelines.

Smaller general data inconsistencies included different financial years for emission reporting and performance data. Another one related to geographical differences; technology was assumed to be globally accessible, but this is not the case for all natural resources. It could be debated whether the accessibility to renewable energy should be taken into account. Another point of discussion concerned the comparison between larger and smaller companies. In the electricity sector this was the case and it may be unfair. Smaller companies tend to have newer plants and installations. This might provide an unfair advantage for smaller, newer companies, but from an emission point of view it may be desirable.

Nevertheless, the publicly available data turned out to be sufficient for performing a first indicative study, like this one, and can be used as a foundation for further development of benchmarking and improvement of data quality.

References

Arrhenius S. Die vermutliche Ursache der Klimaschwankungen [Journal] // Meddelanden från K. Vetenskapsakademiens Nobelinstitut. - Stockholm : [s.n.], 1906. - 2 : Vol. 1. - pp. 1–10.

Birat J.P. Steel and CO2 – the ULCOS Program, CCS and Mineral Carbonation using Steelmaking Slag [Online] // ULCOS. - ArcelorMittal, 2009. - January 2013. - http://www.ulcos.org/en/docs/Ref01__Birat_slag_finaal.pdf.

Blok K. [et al.] Bridging the greenhouse-gas emissions gap [Journal] // Nature Climate Change. - 2012. - Vol. 2. - pp. 471–474.

Blok K. Introduction to Energy Analysis [Book]. - Utrecht : Techne Press, 2007.

CDP Carbon Disclosure Project Leadership Index [Online]. - 2012. - https://www.cdproject.net/Documents/Guidance/CDP2012ReportingGuidance.pdf.

CDP The Carbon Chasm [Report]. - London United Kingdom : Carbon Disclosure Project, 2009.

Ecofys Benchmark of Dong Energy's position compared to its European competitors on Green Growth Leadership [Report]. - Utrecht : Confidential, 2011.

Ecofys Methodology for the free allocation of emission allowances in the EU ETS post 2012 [Report]. - [s.l.] : European Commission, 2009.

EIO ET Europe 300 Carbon Rankings 2011 Report [Report]. - UK : Environmental Investment Organisation, 2011.

ERM Company GHG Emissions Reporting - – a Study on Methods and Initiatives [Report]. - Manchester : European Commission Environmental Resources Management, 2010.

GEA Global Energy Assessment - Toward a Sustainable Future [Report]. - Cambridge, UK; New York, USA; Laxenburg, Austria : Cambridge University Press, 2012.

Graus W. and Worrell E. Methods for calculating CO2 intensity of power generation and consumption: A global perspective [Journal]. - [s.l.] : Energy Policy, 2010.

Haakman J. Benchmarking the CO2 intensity and analysing the CO2 reduction strategies of utilities in Europe [Report]. - Utrecht : Utrecht University and Ecofys, 2011.

ICAO ICAO Environmental Report [Report]. - Montreal : International Civil Aviation Organization (ICAO), 2007.

ICAO Resolution A37-19: Consolidated statement of continuing ICAO policies and practices related to environmental protection - Climate change. [Conference] // 37th Session of the ICAO Assembly. - [s.l.] : ICAO, 2011.

IEA Energy balances [Report]. - Paris : International Energy Agency (IEA), 2010.

IPCC Industry and Mitigation. [Report] : Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. - Cambridge, United Kingdom and New York, NY, USA : Cambridge University Press, 2007.

ISIC International Standard Industrial Classification of All Economic Activities, Rev.4 [Report] / Economic and Socal Affairs ; United Nations Secretariat. - New York : United Nations Publications, 2008.

Jong Berry de Carbon impact large companies [Report]. - [s.l.] : Ecofys, 2011.

Martin N. [et al.] International comparisons of Energy Efficiency [Conference] // Workshop proceedings. - Berkeley : [s.n.], 1994.

Phylipsen G. J.M., Blok K. and Worrel E. Handbook on International Comparisons of Energy Efficiency in the Manufacturing Industry [Book]. - Utrecht : Utrecht University, 1997.

Price Lynn, Christina Galitsky and Kramer Klaas Jan International Experience with Key Program Elements of Industrial Energy Efficiency or Greenhouse Gas Emissions Reduction Target-Setting Programs [Report]. - Berkeley : Lawrence Berkeley National Laboratory, 2008.

Rate the Raters Rate the Raters, Phase five: Polling the experts [Online] // Sustainability. - September 11, 2012. - October 25, 2012. - http://www.sustainability.com/library/rate-the-raters-phase-five-polling-the-experts-2012#.UIj-4W9mIqw.

Roos P. EI New Energy Top 100 Green Utilities [Journal]. - New York : Energy Intelligence, 2012. - 26 : Vol. 1.

UNEP Bridging the Emissions Gap [Report]. - [s.l.] : United Nations Environment Programme, 2011.

UNFCCC Copenhagen Accord [Report]. - Copenhagen : United Nations Framework Convention on Climate Change, 2009.

UNIDO Global Industrial Energy Efficiency Benchmarking, An Energy Policy Tool working paper [Report]. - Vienna : United Nations Industrial Development Organization, 2010.

Verificatiebureau Benchmarking Rapportage monitoring resultaten van het Convenant Benchmarking [Report] : Monitoringrapport 1999 - 2007. - Utrecht : Verificatiebureau Benchmarking, 2008.

Worldsteel Steel's contribution to a low carbon future [Online] // Worldsteel association. - Worldsteel association, 2009. - January 22, 2013. - http://www.worldsteel.org/dms/internetDocumentList/downloads/publications/Climate-change-position-paper-A4/document/Climate%20change%20position%20paper%20A4.pdf.

WRI/WBCSD GHG Protocol Corporate Accounting and Reporting Standard [Report] : Accounting and Reporting Standard. - Geneve : World Resources Institute and World Business Council of Sustainable Development, 2004.

Experts

Experts	Company	Position	Specialism/supporting role
Kornelis Blok	UU and Ecofys	Professor of Sustainable Energy; Director of science	Wedging the gap; sector approach; scientific reporting
Giel Linthorst	Ecofys	Unit manager Supply chains	Emission benchmarking
Maarten Neelis	Ecofys	Managing consultant industrial processes	Emission trading
Paul Blinde	Ecofys	Managing consultant market based mechanisms	Emission trading
Jeroen Scheepmaker	Ecofys	Senior consultant carbon management	Emission benchmarking
Ewout van der Beek	Ecofys	Supply chain consultant	Electricity sector; emission benchmarking
Rolph Spaas	Ecofys	Customer segment manager	Electricity sector
Rob Winkel	Ecofys	Senior consultant sustainable transport	Air transport sector
Joop Oude Lohuis	Ecofys	Client director	Steel sector

Interviewed stakeholders

Interviewed sector experts	Company	Position
Pedro Faria and Michelle O'Keeffe	CDP	Technical Director and Director of Technical Reporting
Danny Croon and David Valenti	Eurofer	Director Environment and Climate Change Manager
Karl Buttiens	Arcelor Mittal	General manager
Jesse Scott	Eurelectric	Head of the Environment & Sustainable Development Policy Unit
Sam Brand and Neil Dickson	ICAO	Environmental Officers of the Committee on Aviation Environmental Protection
Fokko Kroesen	Air France - KLM	Environmental Manager

Appendices

In the first appendix the analysed company benchmarks have been listed, stating the goal of the benchmark, the scope and methodology. This is followed by the interview questions and the minutes of the stakeholder interviews. In the last appendix the heat allocation methods used, are explained.

Appendix A: Current company rankings

	Name	Goal	Scope/ target group	Methodology and Indicators
1.	Carbon Disclosure Project Leadership Index (CDPLI) www.cdproject.net	Working to drive GHG emissions reduction and sustainable water use by business and cities. A Disclosure and Performance Leadership Index is combined.	The world's largest investors, businesses and governments participate. Thousands of companies, 655 institutional investors. Very extensive analysis. Non-for-profit organisation, participation is voluntary.	Disclosure index scores the specificity and completeness of the data provided. Performance index is a measure of the positive actions that the company has demonstrated through their CDP response. These include actions to promote climate change mitigation, adaptation and transparency. This includes a very extensive qualitative questionnaire, future plans and some quantitative indicators. Some of the KPI's CDP rates: -Emission intensity business metric and absolute emissions. -Target presence, strength/ambition and progress. -Driving emission reduction investment: Payback Period, Marginal Abatement Cost curves of projects, methods and practice examples.
2.	FTSE4 Good Index Series www.ftse.com/Indices/FTSE4Good_Index_Series	The index is designed for the creation of index tracking funds, derivatives and as a performance benchmark. Can be used for Investment, Research, Reference and Benchmarking.	Benchmark and tradable indices for Environmental, Social and Governance (ESG) investors.	KPI's including: Environmental Management, Climate change mitigation and adaptation. Company should report total operational energy or Carbon emissions. For certain sectors the energy efficiency should be reported (Oil & Gas & Coal: end user emissions. Automobiles & Aerospace fuel efficiency improvements above average subsector). The criteria have been designed to help investors minimise ESG risks. Performance indicators: -5% reduction in carbon intensity over the last two years. -Company is in top quartile of companies of subsector when assessed on carbon efficiency metrics in previous two years. -A transformational Initiative to reduce emissions.
3.	Dow Jones Sustainability Index	Global sustainability benchmark. Tracks the stock	Benchmark for Investors with sustainability portfolio	Annual Corporate Sustainability Assessment (CSA) in order to identify companies that are better equipped to recognise and respond to

	(DJSI) www.sustainability-indexes.com	performance of the world's leading companies. Exploiting sustainability insights to generate attractive long-term investment returns.	and companies who want to adopt sustainable best practices. The biggest 2500 companies are approached for voluntary reporting.	emerging sustainability opportunities and challenges presented by global and industry trends. Economic, Environmental and Social indicators with at least 50% of the questionnaire covering industry-specific risks and opportunities. With the only general Environmental criteria: Environmental reporting and Environmental management/policy system.
4.	Oekom Corporate Ratings www.oekom-research.com	As partner of institutional investors and financial service providers, Oekom develops investment strategies that combine sustainability research with a high rate of return.	Benchmark for investors. Companies are selected from investment indices and known sustainable leaders.	Qualitative interview assessment based on Environmental and Social criteria. Classified leaders in their industry or not (from A+ to D- or no info available). Environmental Rating criteria: Environmental Management, Products and Services, Eco-Efficiency. Absolute values compared to previous years for product development.
5.	Sustainalytics Company Ratings www.sustainalytics.com	Investment research firm specialised in environmental, social and governance (ESG) research and analysis.	Rating for investors. Use STOXX Global 1800 Index as source for companies	Co-create STOXX® Global ESG Environmental Leaders, combined indices for investors. Separate Environmental, Social and Governmental indices as well. Sector-specific weighting matrices and sustainability indicators scores. Environmental indices based on many indicators, both quantitative as qualitative. Include CDP reporting. Co-create Newsweek Green Rankings (see further in this table).
6.	Climate Counts www.climatecounts.org	Scores companies on their practices to reduce global warming. 2 goals: Offer consumers an easy-to-use method for making informed purchasing decisions; Provide companies an environmental benchmark with which to identify their standing in relation to their peers.	Rate site for consumers. About 136 of the most well-known companies across 16 major consumer sectors were chosen to rate. Publicly available data and databases like CDP are used.	Qualitative indicators are used. Focus lies on indicators which show inventory, strategy, action and disclosure. No quantitative indicators are used. The effort of target setting is rewarded, not the ambition level. Target needs defined baseline, reduction amount/percentage per year and timeframe.
7.	Inrate Sustainability Assessment www.inrate.com/Site/Services/Sustainabili	Be an integral part of the global financial infrastructure by providing sustainability intelligence that allows capital	Investors in Europe. Over 2500 equities across all major markets, over 100 bond issuers. Customised	Research modules focus on the sustainability of management and operations, products and services and on certain controversial business practices. Depending on the client, a climate change assessment is added, which is a quantitative methodology based on extended input-

	ty-assessments.aspx	markets to redirect investment flows toward a more sustainable economy.	rankings for Investors.	output data, Life Cycle Inventories (LCI) and Life Cycle Assessments (LCA).
8.	GoodGuide www.goodguide.com	Helps consumers quickly evaluate and compare products. Core product categories: personal care, household chemical and food products, among others.	Rate products that comprise the top 80% of current sales in a category, plus innovative products marketed as having health, environmental or social benefits.	Rates products and companies on their health, environmental and social performance. Environment score characterises the potential environmental impacts associated with the manufacture, sale, use and disposal of a product. KPI's for environmental score are: environmental management, transparency and qualitative environmental impact.
9.	Goldman Sachs Sustain (GS Sustain) www.goldmansachs.com/our-thinking/topics/gs-sustain/index.html	Identify companies that can remain successful in the face of a rapidly changing world. Incorporates ESG into picking long-run winners and looking for emergent industries	Investment focus list for investors. Researched more than 750 large companies, in 22 industries. Focus list consist now of 44 companies in 8 industries.	Incorporates corporate governance, social issues with regard to leadership, employees and wider stakeholders and environmental management. Members have to score well on a combination of ESG score and industry positioning and this must translate into improving financial performance. It combines ESG analysis with industry themes and quantitative valuation techniques and highlights emerging industries. Environmental indicators: level of greenhouse gas emissions relative to asset base and other sector specific indicators Management of water, waste & recycling, Suppliers and sourcing, Biodiversity and land use.
10.	MSCI ESG Indices www.msci.com/products/indices/esg/environmental	The Global Environment Indices are benchmarks for investors seeking exposure to companies whose primary source of revenues increase the efficient use of scarce natural resources or mitigate the impact of environmental degradation. Participation is voluntary.	Benchmark for investors.	The ESG Indices consists of 2 main indices. 1. MSCI Global Environment Index is an aggregation of indices based on 5 key themes: Alternative Energy, Clean Technology, Sustainable Water, Green Building and the Pollution Prevention. 2. MSCI Global Climate Index. 100 company leaders in Renewable Energy, Future Fuels, and Clean Technology & Efficiency are selected from the Developed Market equity universe. Criteria include market share, strategic commitment, investment in research and development, intellectual property and reputation.
11.	Bloomberg SRI www.bloomberg.co	Sustainability combines corporate citizenship, risk	Benchmark for Socially Responsible Investors. 5200	Environmental KPI's: Quantitative indicators: Emission/revenue (including acquisitions), Amount of renewable energy credits purchased

	m/bsustainable	management and strategic opportunity – driving our operating costs down, our revenues up, and influencing wider adoption of sustainable practices across the business community.	companies are reporting their data.	by the company as a percentage of total electricity consumed, Travel emissions/employee, Water/employee, Landfill waste/employee, Investment in sustainability projects/revenue, Qualitative indicators: Third party verification of environmental policies, Green building policy, Environmental supply chain management policy.
12.	ASSET4 ESG Ratings www.csrhub.com/datasource/asset4	Provides a new generation of investment research information, which can be integrated with traditional financial analysis. Enables investors to lower their risk exposure.	4256 Companies Covered. Customised rankings for Investors.	250+ key performance indicators (KPIs) based on Environmental, Social and Governance indicators. 135 industries are identified. Environmental indicators: Energy used, Water recycled, CO2 emissions, Waste recycled, Spills and pollution controversies, etc.
13.	Global 100 Most Sustainable Cooperation's www.global100.org	Reinforce, raise awareness and highlight global firms most willing and able to deal with the key social and environmental factors they face in their everyday operations. Made by Corporate Knights (CK) a media, research and investment research company.	Rating for investors. All large companies are considered. Through criteria, a selection is made to research.	11 Social, Safety and Environmental KPI's are used. 4 Environmental KPI's: Total revenue per period/criteria (Energy; GHG emissions; Water; Waste) Progress of the criteria is scored compared to previous years and similar companies.
14.	Vigeo Ratings www.vigeo.com	Give a better insight into the value of intangible assets (reputation, human capital, innovation etc.) and the potential for sustainable value creation.	ESG research on more than 2,500 issuers in both developed and developing countries, across both equity and fixed income asset classes.	Environmental criteria questions based on visibility, content, resources adequacy, control, reporting and strictness of processes. End is a scale of 4 differential scores.
15.	Newsweeks Green	Reveal the planets biggest	As an online magazine it is	Based on three scores:

	Rankings www.thedailybeast.com/newsweek/features/2012/newsweek-green-rankings.html	protectors and polluters. Participation is voluntary.	made for a broad, mostly American, public. Rated the biggest 500 U.S. companies and 500 Global companies	1. Quantitative Environmental Impact Score: Total cost of all environmental impacts of a corporation's global operations (GHG emissions, Water use and Waste disposal) normalised against a companies' annual revenues. 2. Green policies score based on a qualitative assessment of a companies' environmental policies and performance. 3. Reputation score based on a survey among professionals.
16.	CO2 Prestatieladder www.skao.nl	It is a concrete instrument to reward sustainability when granting assignments. It stimulates companies to identify and reduce their own CO2 emissions and permanently look for further reduction possibilities also together with suppliers.	Benchmark for building companies in the Netherlands. Could be extended to other sectors. Participation as a company is voluntary.	5 different levels are defined. Qualitative data is used for scoring to determine level. The main criteria are: Identification of current emissions, continuous mechanism to look for more reduction potentials, installing these identified reduction potentials, sharing of gained knowledge and searching for more reductions opportunities with other parties. The indicator tCO2/euro is used as intensity indicator to compare emissions with the average within the sector. A minimal percentage emission reduction compared to previous years is required.
17.	2011 Carbon rankings by Environmental Investment Organisation (EIO) www.eio.org.uk	An independent non-profit organisation which researches, promotes and implements investment systems designed to incentivise global corporate emissions reduction.	Benchmark made for investors. All large publicly traded companies are rated.	CO2 emission per turnover is used as indicator for the ranking. Emission intensive sectors are partially compensated. Companies are compared within sector.
18.	Covalence Ethical Quote (CEQ) Ranking www.ethicalquote.com	Provides reputation management tools and ESG ratings & data to corporates and investors	Benchmark for investors. Rated 2800 of world's largest companies	50 Qualitative criteria for overall sustainability from the GRI. Environmental criteria include material- and energy use, water- and waste management, pollution, biodiversity, emissions, recycling and compliance with environmental laws.
19.	US Regional Greenhouse Gas Initiative (RGGI) www.rggi.org	Market-based regulatory program to reduce greenhouse gas emissions.	Cap-system for electric power plants in eight states in the US. Participation for plants of >25MW is mandatory.	Installation level direct emissions are measured. Allocation of CO2 permits is done on the base of output, in this case MWh.

20.	EU Emissions Trading Scheme (EU ETS) ec.europa.eu/clima/policies/ets/index_en.htm	Market-based regulatory program to reduce greenhouse gas emissions.	Energy intensive companies in the EU. Targets: NAP process defines allocations by installation. Participation is mandatory.	Installation-level boundary based on meeting combustion capacity or production thresholds. Allocation of CO2 permits is done on the base of defined output.
21.	Australian National Greenhouse and Energy Reporting (NGER) Scheme www.cleanenergyregulator.gov.au	Reducing emissions cost effectively by putting a price on carbon.	Energy intensive companies in Australia. Companies emitting 25 ktCO_2/year participation is mandatory (approx. 500 companies).	Companies have to report Scope 1 and 2 emissions, energy use and waste.
22.	UK Carbon Reduction Commitment (CRC) www.decc.gov.uk/en/content/cms/emissions/crc_efficiency/guidance/guidance.aspx	Reducing emissions of the UK cost effectively to meet internationally set targets.	Around 5,000 organisations in the UK use >6 GWh/year for which participation is mandatory.	One league table with revenue recycling according to performance will be used for all participants. The standard provides a league table and reports annual progress against targets allowing meaningful comparison between companies. Three metrics are used to calculate the score for the league table: early action metric, absolute metric: progress compared to rolling average previous years and growth metric: progress in emission intensity based on tCO_2/revenue.
23.	Dutch Energy Covenant www.benchmarking-energie.nl	Pledge with energy intensive sectors on energy efficiency in exchange for no regulation, to meet internationally set Dutch reduction targets.	Energy-intensive industry >0.5 PJ/yr. 84% of energy intensive enterprises have signed the covenant	Companies must develop an energy efficiency plan and the covenant contains criteria governing the rate of investment. Benchmarked against top 10% of global operators in their sector. Individual plants are benchmarked in terms of production, revenues and specific energy consumption (SEC). Benchmark for top 10% can be determined through three different methods. Region, decile method or best practice. More information on these methods can be found at the reference link.
24.	US EPA Climate Leadership Awards (CLA) www.epa.gov/climateleadership/awards/index.html	National awards program that recognises and incentivises exemplary corporate, organisational, and individual leadership in response to climate change.	One individual and 20 organisations in the private sector who are leading the way in the management and reduction of GHG emissions - both in internal	Partner companies commit to reducing their impact on the global environment by completing a corporate-wide inventory of their GHG emissions based on a quality management system, setting aggressive reduction goals, and annually reporting their progress to EPA. The standard provides a leadership index and reports annual progress against targets. Criteria are: GHG Scope 1 and 2 inventory and

			operations and throughout the supply chain are rewarded. Participation is voluntary.	qualitative GHG reducing management, innovation and initiatives. Four data sources are used to calculate fuel consumption and economic output values for common commercial and industrial sectors. These results are used in combination with emission factors to estimate CO_2 intensity by sector.
25.	Chicago Climate Exchange (CCX) www.theice.com/ccx.jhtml	Comprehensive cap and trade program with an offsets component.	North America's largest GHG reduction program consisting of 92 members. Participation is voluntary.	Verification based trade system for tCO2-e compared to baseline emissions. 6% reduction target by 2010 is mandatory. Baseline is members average emissions during the 1998 – 2001 period (Phase II Baseline may be the single year 2000)
26.	Canadian Industry Program for Energy Conservation (IPEC) oee.nrcan.gc.ca/industrial/technical-info/benchmarking/4377#energy-performance	Aids the adoption of an energy management standard, and accelerates energy-saving investments and the exchange of best-practices information within Canada's industrial sector.	Companies in the manufacturing industries. Participation is voluntary.	Energy performance benchmarking: comparative analysis of energy use (MJ) per unit of physical production (tonne). The plant energy-intensity indicator is based on the energy-intensity indicator and the material input factor for each process in the operational stream. The processes required for a plant are linked by the material input factors to determine how much each process must produce for the plant to make one tonne of product. That number is then multiplied by the process energy-intensity indicator to establish the contribution that the process makes to the overall plant energy-intensity indicator. E.a. Energy-intensity indicator for EcoTech EAF Bar Plant = Bar Plant energy-intensity indicator + (tonnes of cast steel/tonne of bar × Continuous Casting energy-intensity indicator) + (tonnes of liquid steel/tonne of bar × EAF energy-intensity indicator). If not possible benchmark with best-in-class: select best-in-class by looking at companies with lowest costs and highest energy efficiency for all functions.
27.	Industrial Energy Efficiency Network (IEE Network) www.motiva.fi/files/1937/CR_Norway.pdf	Energy efficiency programme with governmental funding to stimulate energy efficiency measures in industry.	780 member companies from 13 energy intensive industry sectors in Norway. Participation is voluntary.	Audit methodology: Phase 1: Walkthrough audit establishing energy management and rough mapping of the main water and energy flows in the company. Benchmark company-data with best practice companies in the same sector. Phase 2: Detailed energy and mass balance analysis of the company, resulting in a comprehensive plan for energy efficiency measures to be undertaken within a given period of time (voluntary).

Appendix B: Interviews with stakeholders

Interview questions

1. Do you benchmark the carbon performance of your members at company level?
2. What are the biggest problems you face?
3. Do you recognise the trends?
4. What characteristics make an intensity indicator practical for you?
5. Can you think of other relevant indicators than the researched ones?
6. What disclosure problems do you face in your industry that may hamper the use of certain indicators for benchmarking?
7. Would the selected indicator be useful for branch organisations and/or policymakers to follow the progress and reductions in the whole sector?

B.1 Interview Carbon Disclosure Project

Pedro Faria – Technical Director
Michelle O'Keeffe - Director of Technical Reporting

CDP recognises the trend of ranking and valuating companies. It is done for quite some years already and still a large amount of work needs to be done. It requires lots of data.
One of the biggest problems is that companies consider themselves special. And because they are different they want to be ranked based on different indicators.

CDP is moving towards more sector specific approaches. They prefer to quantify their methodology further including other indicators than carbon. It still requires a lot of work to be done especially on the quality of their collected data.

CDP agrees with literature that product related indicators are better than turnover related indicators. But in practice it only works well in sectors with homogeneous products. They also use tCO_2-e/MWh as intensity indicator in the electricity sector.

They ask companies to disclose absolute emissions, performance data as well as emission intensity indicators because regularly companies use different boundaries for their intensity figures, which result in different numbers. They think the separate parameter figures overall have a higher quality because the reporting of these numbers are more standardised. Up till now there is less guidance for emission intensity indicators which result in all different kinds of indicators.

They confirm they are not only rating quantitative carbon performance but also include environmental management performance and performance improvement over time which explain the differences in the rankings of the research and the CDP rankings.

Faria emphasised the fact that with benchmarking your point of view really matters because you are trying to find one figure for many things. Investors often prefer the indicator based on turnover and engineers often prefer the indicator based on the produced output.

B.2 Interview Eurelectric

Jesse Scott - Head of the Environment & Sustainable Development Policy unit

Eurelectric is a European branch organisation for electricity generators and distributors (not transmission). Their members are the national associations, but in practice the companies play a role themselves. Scott focuses on climate change and the politics of ETS. Their goal is to be Climate neutral in 2050. They have set an absolute target.
Eurelectric is not benchmarking their members because their main priority/task is to lobby towards the EU. Informing the members (i.e. benchmarking them) becomes second. They had started a voluntary benchmark project some years ago, but only some companies delivered input and no thorough data was collected. She explained that electricity companies are very conservative with data disclosure. Each year Eurelectric creates a power statistics report. Statistics are a big challenge for them.

Eurelectric agrees tCO_2-e/MWh is the best indicator for benchmarking electricity companies and does not know any other indicators which are not researched. She recognises the presented tCO_2-e/MWh ranking and the intensity development of companies over the years. For the carbon performance of companies, their culture is just as important as their current asset base. For example the company RWE currently has a Dutch CEO and he is steering the company away from coal/lignite. The company is re-thinking itself from the inside. Scott estimated however that the expected impact of new technologies will not go far enough to reach the emission reductions that are needed.
Related research, interesting for Eurelectric, would be looking at the impact of new technologies and/or new management on the carbon intensity of a company.

For Eurelectric a benchmark is useful as an instrument for communication with the EU about new technologies and renewables. From an electricity company perspective minimum administration is the most important, she believes. If you run multiple small generation units benchmarking is of more added value than for large companies since those large companies always have a combination of old inefficient and newer plants. The large old plants are often idiosyncratic (unique, large and complex to some degree). The engineers working at (old) large complex plants are often engineers that perfectly know how to operate the plant, but who do not necessarily have the knowledge to collect data. These plants are often operated with minimum amount of staff (e.g. 20 FTE), as a result of which there is not a lot of extra capacity for data collection.

The most important problem for bottom-up benchmarking in this sector is the lack of added value in the experience of the electricity companies. They believe that due to ETS the advantages of process optimisation are already covered. Their general instinct is not to share performance data with others so as to avoid competition. Therefore data availability and cooperation form a problem as well.
Scott experiences the electricity sector as rather conservative, with few drivers, except for legislation, to reduce emissions and combat climate change. But in the past few years their position towards ETS has changed. They prefer now to have a steady EU ETS plan (2030 targets) as soon as possible instead of further delays or even worse: changeable country specific regulation. The sector prefers “any plan” instead of “no plan at all”. The electricity sector would not like the discussion about 2030 ETS targets to be postponed to the new EC in 2015, with perhaps 2030 targets only by 2017. Electricity sector plans are made 10-15 years in advance. For the electricity sector the ETS is seen as the best

decarbonisation instrument (the best single driver). The carbon price should stimulate RE and EE. At the moment the main problem ETS concerns is the level of ambition, rather than reporting guidance.
Overall Scott expects no interest from large European electricity companies in a bottom-up benchmark. She believes the Japanese might be more interested; Japanese companies consider the electricity sector still largely as a sector with voluntary emission reduction commitments instead of formal obligations.

B.3 Interview Air France - KLM

Fokko Kroesen – Environmental manager KLM

Fokko Kroesen is expert on bio fuels and carbon issues within KLM. He explains that most benchmarks in the air transport sector focus on fuel efficiency because it is crucial for the costs. These current available benchmarks are quite aggregate because a lot of the data is not (publicly) disclosed.

The carbon performance is not only dependent on the available technologies but depend on the product the airplane is carrying (cargo or passengers), the speed, the distance and the type of airplane. Therefore disclosure of specific data could tell a lot about the business model, performance and strategy of the company.

According to Kroesen benchmarking with publicly available data is very hard, especially when you want to do it at a lower level of aggregation. At the moment only a branch organisation might have access to more detailed data under strict agreements where they are able to process the data but only publish the aggregate version of the benchmark.
KLM participated in a fuel efficiency benchmark made by IATA. KLM came out as one of the best and therefore they stopped. It was mainly interesting for the ones lagging behind because they were expected to share their good case practices. They are still interested in their position towards other companies though.

The most “secret” data are the specific load factors of the different flights. KLM has created a tool for their website where customers can calculate the specific emission of a certain flight. The calculations behind this tool are complicated and confidential.
Kroesen agrees that the monetary indicator is not good because it is too little related to the product. He thinks the researched physical indicators are interesting, disaggregating to cargo and passengers is necessary. He approves of the volume correction factor for the emission intensity difference between transporting one tonne cargo and one tonne passengers. You can roughly distinguish 3 different business models within the air industry: Short-haul/business flights, cargo and medium-/long-haul passengers. Since these flights are all very different. He would also disaggregate towards the three different products of these business models.

Looking at the different rankings he doubts some of the used data. For benchmarking with publicly disclosed data some clear reporting guidelines are missing. For example the way the p-km or t-km is reported. You can either report the great circle distance or the flight plan. The flight plan is the actual flown distance. Reporting the flight plan is advantageous in this case and might give a tilted view of the efficiency, detours should not be rewarded. Another thing is if the APU is included or not, this is the air conditioning during the flight run by a generator mostly fuelled by kerosene. And if the extra weight in a passenger flight is accounted for. Some companies add some weight to account for passenger chairs, food and personnel. This would influence the p-km/t-km factor and the volume correction factor between cargo and passengers. He says because of these factors many benchmarks publish the wrong rankings with the wrong conclusions. If a ranking is published these factors should be included.

With the increased fuel prices and increased scarcity all companies will improve. But benchmarking should be done in a very transparent way, with sources and uncertainty margins if certain specifics are not disclosed.
The industry is sometimes does not even agree with ETS because for example cargo is heavily afflicted. It is an enormous challenge to assemble all the data, include all these factors and come to an accepted method. He wonders if it is possible and worth the effort.

B.4 Interview International Civil Aviation Organisation (ICAO)

Sam Brand - Environmental officer
Neil Dickson - Environmental officer

The ICAO is as the United Nations agency responsible for international aviation. They have members of all nations and are run by a council with a president. At the moment they work with the Group on International Aviation and Climate Change (GIACC) on a global fuel efficiency benchmark for the aviation industry. In 2012 they formulated a target of an annual fuel efficiency improvement of 2% until 2020 calculated on the volume of fuel used per revenue tonne kilometre (t-km) performed (ICAO, 2011). They capture and produce data and benchmark their members on base of fuel efficiency. “The intent of the carbon metric system is to equitably reward advances in aircraft technologies (i.e. structural, propulsion and aerodynamic) which contribute to reductions in aircraft carbon emissions, and differentiate between aircraft with different generations of these technologies (ICAO, 2011).”

Dickson is expert on benchmarking the technological level and the efficiency level but has little experience with benchmarking the emission data and the used methodology. It is new for them. In the graphs Delta scores better than Emirates, while Delta has older aircrafts than Emirates. This might indeed be due to other factors than technological efficiency, like electricity use at the airports or more luxury flights with more space per person. They are not familiar with this kind of benchmarking, but would find it interesting to know these other factors.

They agree with the chosen level of aggregation and a separation between cargo and passengers. The biggest problem concerning disclosure is the load factor data; this is due to extreme international competitiveness. They also see a separation between business aviation/small jets, short-haul and long-haul. But they do not know yet if this level of aggregation should be considered. They find the research indicators very useful.
Concerning data availability Brand emphasises the example set by the Canadian IPEC. They did an extensive research and publicly disclosed a large amount of high quality data from their members.

B.5 Working group notes of the Committee on Aviation Environmental Protection (CAEP) of ICAO

The Committee on Aviation Environmental Protection (CAEP) of ICAO is working on a CO2 Standard. The most important conclusions from this working paper are listed below (ICAO, 2011).
ICAO should establish rigorous annual reporting, on fuel consumption and fuel efficiency. They should publish the results, in order to provide transparency on performance of the aviation sector. China, as a part of the working group, noted that the possession of a modern fleet of aircraft is not sufficient as an efficiency indicator; such factors as operational management, air traffic management, and airport services, etc. must also be taken into consideration. CAEP endorsed two options for fuel efficiency metrics: Litres of fuel consumed/Revenue Tonne Kilometres and Fuel Mass consumed/ payload x distance. The second metric is being developed by CAEP. In future it will be important to develop appropriate conversion factors for different fuels, including the lower carbon footprint of alternative fuels.

It was noted that different countries use different average weights for passengers for operational purposes. To convert passengers to tonnes, North America appears to use an average of 100 kg per passenger, whereas different averages are used in Japan (75kg, 92.5kg and 102.5kg, depending on class and route) and in China (75kg). CAEP recommends that ICAO establish a standard average weight for passengers to be used in fuel efficiency metrics.
For both metrics, the distance measured should be the Great Circle Distance between airports, rather than the actual distance flown. Use of the constant factor of Great Circle Distance ensures that all types of efficiency improvements (including improvements in air traffic management) are reflected in the metric. In addition, it may be necessary to clarify reporting requirements and guidance between the UNFCCC and ICAO in order to ensure standard and consistent reporting and minimise burden to those reporting.

CAEP discussed the technical issue of a metric for carbon neutral growth that would be expressed in terms of fuel efficiency. Carbon-neutral growth will be achieved when the rate of fuel efficiency improvement is equal to the rate of increase in Revenue Tonne Kilometres. CAEP was unable to reach consensus on further consideration of this issue. China disagrees with the concept of ICAO adopting a goal of carbon neutral growth at this time. France is of the opinion that the metric should also reflect the situation where aviation is part of an emissions trading scheme: carbon neutral growth would be achieved when all emissions from air operators above a fixed cap are compensated by reductions of emissions by other operators. The United States expressed the view that carbon-neutral growth can be achieved by 2025 through technological and operational improvements, in combination with use of alternative aviation fuels. CAEP recommends that ICAO adopt an aspirational fuel efficiency goal to be achieved by 2012: that is, by 2012, the average fuel efficiency of international aviation will not exceed *X* Litres per 100 total t-km.

A best approach to calculation of fuel efficiency would use actual data on fuel consumed and actual data on total t-km performed. CAEP therefore recommends that ICAO put in place comprehensive data collection on aviation emission no later than 2012.
Based on the carbon metric system, the CO2 Standard will aim to reduce aircraft carbon emissions by encouraging the integration of fuel efficient technologies into aircraft design and development. It has been developed such that effective improvements observed through the CO2 Standard will correlate with reductions of carbon emissions by aircraft during day-to-day operations.

B.6 Interview ArcelorMittal and Eurofer the European steel Association

Karl Buttiens – Director of Environment & Global CO_2 strategy
Danny Croon – Director Environment at Eurofer
David Valenti - Climate Change Manager at Eurofer

Since 1988 Karl Buttiens works on benchmarking in the steel industry. At the moment ArcelorMittal and Eurofer work together on a standardised performance benchmark methodology for the metal industry i.e. aluminium, lime, iron and steel. Their methodology is much disaggregated towards intermediate product level and works with scenario roadmaps. According to Buttiens limits should be avoided. The fewer boundaries a benchmarking methodology has the better it is. It will have less chance on artefacts or bias. The entire world should be the boundary and the principle should be scenario comparison.

Ideally the whole world should be taken into account. Eurofer recognises that compromises are almost implicit in benchmarking but they should be limited as much as possible because they will weaken the methodology. Both parties agree that a clear objective is very important with benchmarking. ArcelorMittal uses different benchmark tools in function of purpose. Buttiens explains four main problems: The first is the clarity on what exactly we want to benchmark. Too often it is unclear or one is trying to benchmark several aspects at the same time. You cannot benchmark carbon and energy or energy and cost etc. A choice should be made; either costs, energy or carbon. Because steel production has a long production chain benchmarking is more difficult. Only one clear objective at the time can be benchmarked with a dedicated methodology. Second is the definition of the indicator: as soon as you define an indicator you introduce distortions. Given a free choice of indicator many steel producers could define one in which they are the best. This is enhanced due to the fact that you need to benchmark a long series of carbon intensive processes which are not in general well aligned on site level. Third, there are some uncertain cross relational effects between processes. Energy savings in one compartment may adversely affect the product properties which can cause an additional energy input in some downstream process. These relations are in general quite complex and very hard to quantify so only an overall and unbiased benchmark can give answers. Finally there are some specific issues for which you need to find an acceptable and general methodology such as the simultaneous production of several by-products along the main product like coke, gas and electricity; the accounting of scrap which everyone uses to some degree (if not taken care of, you benchmark the scrap input). At the moment there is a discussion going on if scrap is useful or not and if the performance should be neutralised from scrap.

A good benchmark method should not contain artefacts and should be able to determine the best performer as the winner. If the benchmark does not provide that it will be noticed by the operators and lose credibility. However if it is enforced people will act according to the benchmark and if wrong signals are given the people will go in the wrong direction.

The trends described by the graphs are not recognised by Buttiens. Due to the lack of a generally accepted methodology publicly available figures are rather worthless. A lot of assumptions are behind every figure and in general the assumptions are different or they do not take sufficient care of structural differences between the reporting parties.

Not only data availability but also data quality is an issue. In steel activities you try to identify small deltas on differences between large inputs and products. Very often it is difficult to measure or there is no measurement because production sees no interest (gas flows, wet/ dry weights, LCVs…).
It depends on the purpose of the benchmark what characteristics would make an indicator practical. If you want to improve performance the indicator should identify as directly and precisely as possible the contribution of each operator involved in the final outcome. It should be clear to these operators what their impact is and how they can contribute to improve the final result. Other indicators can be useful (footprint) to estimate global trends impact of macro evolutions etc.

The biggest concern related to disclosure of company data in this industry is disclosure of data which is poorly understood or where there is room for interpretation. People tend to simplify. When you disclose data they might pick only the data they like and take it out of context which may end in the wrong conclusions. It may lead to undue public discussions or pressures to move into an undesirable direction. There is a real concern that the outside world will start to dictate how steel should be made (it is happening cfr. Taranto) and this will most likely not be in the right direction (otherwise the steel makers would not be professionals).

CO_2-e/turnover is not a good indicator according to Valenti from Eurofer because it does not mean anything. Tonne CO_2-e/EBITDA would be more interesting because then you get information on the price of carbon you can survive with. Eurofer agrees a physical indicator is best but the benchmarking should take place at a lower level of aggregation.

According to Buttiens benchmarking on direct emissions is not helpful for the steel industry. The Cap & Trade system will not be good enough in the coming 30 years. Eurofer and ArcelorMittal are determined to define a methodology to compare the industry in a fair way. A physical indicator where they compare every process along the production line to a reference level and in the end combine all those indicators to account for interference between processes.

Appendix C: Heat allocation methods used for the electricity sector

The different heat allocation methods are described in the thesis of Haakman called: *"Benchmarking the carbon intensity and analysing the carbon reduction strategies of utilities in Europe"* (Haakman, 2011).

C.1 Substitution allocation method

The substitution method calculates to what extent carbon emissions are avoided because the heat is not separately produced. The carbon emissions are allocated according to the following equation (Graus, et al., 2010).

Carbon emissions of $P_i = P_i\ ((C_i\ (I_i - H_i/r\ )/(P_i))$ (substitution method)

- $_i$ is the fuel source 1,...., n
- C_i the carbon emission factor per fuel source (tonne CO_2-e/TJ)
- P_i the power production per fuel source (GWh)
- H_i the heat output from CHP plants (GWh)
- r the reference efficiency for heat generation
- I_i is the fuel input power of CHP plants

With this method, the assumption is made that one CHP plant is used instead of two separate plants producing heat and electricity. This method is selected because it takes into account the external benefits: the reduction in fuel use as compared to separate heat generation.

C.2 Power loss allocation method

The power loss method is based on the fact that the electricity output of a CHP plant is lower than a power plant where no useful heat is generated (Blok, 2007). A correction factor is used to calculate the amount of electricity which would have been produced if no heat was generated, as shown in the following equation (Graus, et al., 2010).

Carbon emissions of $P_i = P_i\ ((C_i I_i)/(P_i + sH_i)\ )$ (power loss method)

- $_i$ is the fuel source 1,...., n
- C_i the carbon emission factor per fuel source (tonne CO_2-e/TJ)
- P_i the power production per fuel source (GWh)
- H_i the heat output from CHP plants (GWh)
- s is the correction factor
- I_i is the fuel input power of CHP plants

This method compensates for the actual efficiency drop caused by the production of heat, which makes the carbon emissions of CHP plants comparable with power plants without heat production. For that reason, this methodology has also been selected for the benchmark of the carbon intensity.
For some installations the power loss method will be used, for others the substitution method will be used. The specific method used for every technology can be found in the thesis of Haakman (Haakman, 2011).