CarbonBenchmark

Peer benchmarking of sustainability metrics in VSME reporting

Barrie Wilkinson, James Brownsey

1. Executive Summary

The problem

The reporting of sustainability metrics is becoming more widespread as regulators and other stakeholders increasingly request and mandate that companies make detailed disclosures regarding a company's impact on environmental, social and governance issues.

The new VSME Standard developed and published by EFRAG in December 2025 is designed to encourage adoption of sustainability reporting practices amongst SMEs in the EU.

One of the key issues for smaller companies is understanding what their numbers mean and putting them into some form of context. The large consultancy houses often play a role in collecting and aggregating peer benchmarks for the larger companies but no such services are typically available for smaller companies.

Smaller companies also lack access to the same level of in-house resources and expertise to gain comfort that their methodologies and assumptions are robust.

How peer benchmarking can help

If smaller companies can gain access to comparable disclosures from peer companies from within their industry, they can start to gain some comfort that their approaches are in line with best practices within their industry and also to identify areas where their results are out of line. In this latter case, the companies can conduct further research to understand whether they need to improve their approached or at least understand whether there are good reasons for the deviations.

About Carbon Benchmark Ltd

Carbon Benchmark created the TurboVSME platform that is gaining adoption by accountancies that are looking for a white labelled, cloud-based solution for use with end user reporting companies.

One of the unique features of the TurboVSME platform is that it gives accountancies and end users access to industry-specific peer benchmarks for the various metrics reported under VSME.

The first generation of benchmarks in the tool were developed based on the public disclosures of roughly 10,000 large, publicly-listed companies. This first generation covers GHG emissions, water and energy metrics.

Given that the TurboVSME has a high volume of VSME submission being inputted by qualified accountants, the plan is to enrich these benchmarks overtime to included disclosures from SMEs and for the full suite of VSME metrics.

The approach

Our approach very much builds on previously established practices of peer benchmarking in other areas such as financial reporting. Many of the terms and concepts we use in this paper will be familiar to practitioners who have constructed peer benchmarks in other fields:

- ***Segmentation***: the clustering of the data into segments so that company characteristics are more homogeneous within a segment than they are across segments.

- ***Normalisation***: dividing one metric by another to create ratios which are more easily compared and aggregated across companies.
- ***Treatment of outliers***: to prevent dubious extreme observations from distorting the estimation of population statistics.
- ***Benchmarking***: the comparison of the results of an individual company to another company or to the peer-averages and percentiles.
- ***Estimation***: the application of the benchmarks in estimating the metrics of companies for guidance purposes
- ***Proxying***: the mapping of low-data segments to the nearest proxy segment.
- ***Backtesting***: the comparison of estimated metrics to actual observed metrics.

Limitations of the current approach

The main limitation of the current approach is that we are aiming to serve the SME community with our peer benchmarks but in the short term only have access to reliable data from larger companies to construct the benchmark. While normalisation of the data (e.g. expressing a metric as per FTE or per dollar of revenue) adjusts for the size of a company, it is likely true that smaller companies can not be assumed to be a scaled down version of a larger company from the same industry.

However, we do believe that having access to large company benchmarks is a great deal better than having access to no benchmarking data at all.

As mentioned earlier, the long term solution to this problem is to start collecting disclosure from smaller companies which is already underway via our TurboVSME platform.

Report Conclusion

We believe that the results of our work are promising and could represent a major step forward in bringing much needed context to the sustainability reporting of smaller companies. Through our research we were able to successfully construct peer benchmarks for 126 different industries and have since mapped these benchmarks to the 1,000+ 4-digit NACE codes used in the VSME standard.

The rest of this report does a deeper dive into the estimation of peer benchmarks for the detailed category-wise Scope 3 emissions as defined by the GHG Protocol. We aim to update this whitepaper periodically with results from other VSME metrics.

2. How we created the GHG emissions peer benchmarks

This section describes the process we followed to create our category-wise, normalized GHG emissions peer benchmarks for 126 different industries.

Our approach to sourcing the data

The data used in this analysis was compiled by the data science team at Carbon Benchmark Ltd and curated by our assurance team. We started with a universe of roughly 10,000 publicly listed companies and performed an automated search for their 2022 and 2023 sustainability reports, annual reports, or integrated reports[1].

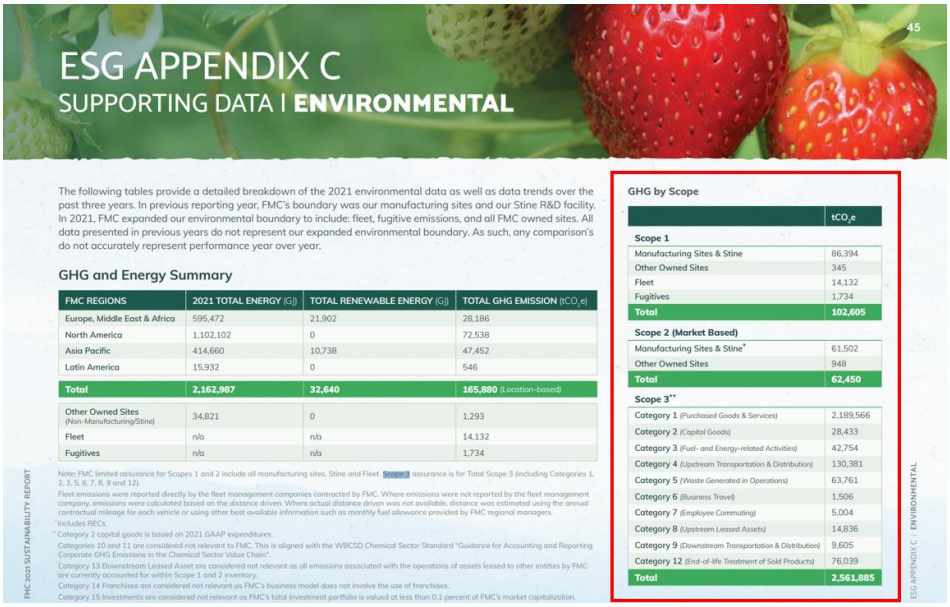
On average, our search yielded roughly 5 reports per company meaning that we reviewed approximately 50,000 sustainability reports as part of our research.

Somewhere buried within these reports, we would sometimes find the GHG emissions disclosures of the reporting company such as the one shown **in the table** below:

---

[1] "Integrated report" is a relatively new term for reports that cover both financial information (e.g. balance sheet and P&L) and non-financial information (such as sustainability-related disclosures)

Figure 1: 2022 GHG emissions disclosure by FMC, an agricultural commodities company



ESG APPENDIX C
SUPPORTING DATA | **ENVIRONMENTAL**

45

The following tables provide a detailed breakdown of the 2021 environmental data as well as data trends over the past three years. In previous reporting year, FMC's boundary was our manufacturing sites and our Stine R&D facility. In 2021, FMC expanded our environmental boundary to include: fleet, fugitive emissions, and all FMC owned sites. All data presented in previous years do not represent our expanded environmental boundary. As such, any comparison's do not accurately represent performance year over year.

GHG and Energy Summary

FMC REGIONS	2021 TOTAL ENERGY (GJ)	TOTAL RENEWABLE ENERGY (GJ)	TOTAL GHG EMISSION (tCO_2e)
Europe, Middle East & Africa	595,472	21,902	28,186
North America	1,102,102	0	72,538
Asia Pacific	414,660	10,738	47,452
Latin America	15,932	0	546
Total	**2,162,987**	**32,640**	**165,880** (Location-based)
Other Owned Sites (Non-Manufacturing/Stine)	34,821	0	1,293
Fleet	n/a	n/a	14,132
Fugitives	n/a	n/a	1,734

Note: FMC limited assurance for Scopes 1 and 2 include all manufacturing sites, Stine and Fleet. Scope 3 assurance is for Total Scope 3 (including Categories 1, 2, 3, 5, 6, 7, 8, 9 and 12).
Fleet emissions were reported directly by the fleet management companies contracted by FMC. Where emissions were not reported by the fleet management company, emissions were calculated based on the distance drivers. Where actual distance driven was not available, distance was estimated using the annual contractual mileage for each vehicle or using other best available information such as monthly fuel allowance provided by FMC regional managers.
* Includes RECs.
** Category 2 capital goods is based on 2021 GAAP expenditures.
Categories 10 and 11 are considered not relevant to FMC. This is aligned with the WBCSD Chemical Sector Standard "Guidance for Accounting and Reporting Corporate GHG Emissions in the Chemical Sector Value Chain".
Category 13 Downstream Leased Asset are considered not relevant as all emissions associated with the operations of assets leased to other entities by FMC are currently accounted for within Scope 1 and 2 inventory.
Category 14 Franchises are considered not relevant as FMC's business model does not involve the use of franchises.
Category 15 Investments are considered not relevant as FMC's total investment portfolio is valued at less than 0.1 percent of FMC's market capitalization.

GHG by Scope

	tCO_2e
Scope 1	
Manufacturing Sites & Stine	86,394
Other Owned Sites	345
Fleet	14,132
Fugitives	1,734
Total	**102,605**
Scope 2 (Market Based)	
Manufacturing Sites & Stine*	61,502
Other Owned Sites	948
Total	**62,450**
Scope 3**	
Category 1 (Purchased Goods & Services)	2,189,566
Category 2 (Capital Goods)	28,433
Category 3 (Fuel- and Energy-related Activities)	42,754
Category 4 (Upstream Transportation & Distribution)	130,381
Category 5 (Waste Generated in Operations)	63,761
Category 6 (Business Travel)	1,506
Category 7 (Employee Commuting)	5,004
Category 8 (Upstream Leased Assets)	14,836
Category 9 (Downstream Transportation & Distribution)	9,605
Category 12 (End-of-life Treatment of Sold Products)	76,039
Total	**2,561,885**

FMC 2021 SUSTAINABILITY REPORT
ESG APPENDIX C | ENVIRONMENTAL

Our criteria for choosing which companies are included in the peer datasets

Here are the criteria we used for including a company's disclosure in our peer benchmarks:

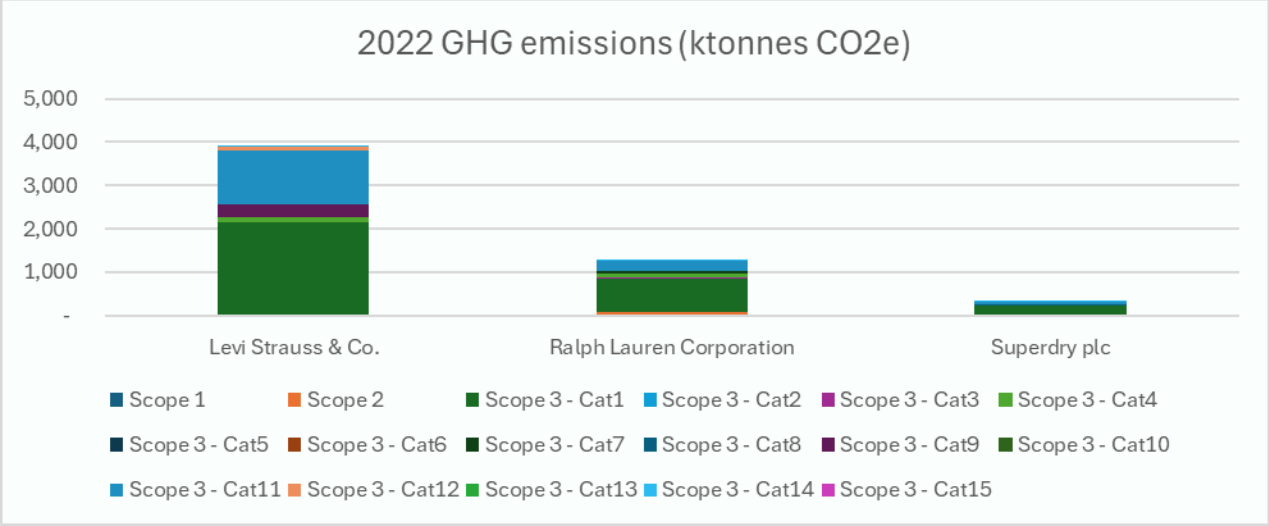
Criteria	Reason
The reporting company needs to be **publicly-listed**	External disclosures made by publicly-listed companies face a lot of scrutiny and most numbers go through some level of assurance. While many private companies have also started to make voluntary disclosures most of these have not yet been through assurance and are therefore of varying quality. We will likely change our position on this in the future as more private companies move to gain assurance for their disclosures as part of their regulatory submissions.
The reporting company needs to have disclosed **Scope 1, Scope 2 and Scope 3 (including the detailed 15 sub-categories of Scope 3**).	Our aim is to create peer benchmarks that are both comprehensive and granular. As a result, we chose to exclude companies that chose not to offer a complete and detailed breakdown of their emissions.

While disclosure practices varied somewhat by industry, roughly 26% of the companies in the sample of 10,000 publicly listed companies met the above criteria meaning that our 2022 dataset contained detailed category-wise disclosures for roughly 2,600 companies spanning 126 industries. This meant that, on average, we had a sample of roughly 20 companies per industry which is generally considered sufficient to estimate the mean and median of a population without excessive measure error. However, for certain sub-categories of Scope 3 and for certain industries we were confronted with low data samples and were forced to fall back on alternative approaches such as proxying and aggregation.

The raw data

The chart below shows the raw "absolute" GHG emissions disclosures for a sub-sample of well-known companies from our Apparel Manufacturing dataset.

Figure 2. A sub-sample of absolute emissions disclosures, Apparel Manufacturing



The problem with attempting to analyze the above dataset is that Superdry plc is a much smaller company than Levi Strauss & Co, so it makes little sense to compare these raw results or aggregate them into statistics.

Let's define the raw absolute emissions of a company in tonnes of CO2e as follows:

$$Emissions\ of\ company\ i\ in\ GHG\ reporting\ category\ c = E_{i,c}$$

Normalization of the data

After some discussions with a group of carbon accountancy vendors that have been using our peer benchmarks in their software, we converged on two different size metrics (revenue and employee headcount). These metrics were very pragmatic choices due to their reliability and widespread availability for all companies:

$$Revenue\ of\ company\ i = R_i$$

$$Headcount\ of\ company\ i = H_i$$

Table 1: Normalization metrics for a sub-sample of Apparel Manufacturers

Company Name	**2022 Revenue USD**	**Employee Headcount**
Levi Strauss & Co.	6,169	19,100
Ralph Lauren Corporation	6,219	14,900
Superdry plc	744	3,350

Source: Yahoo Finance company pages: https://uk.finance.yahoo.com/quote/{ticker}

We welcome feedback on these choices of normalization metrics and are very much open to deploying alternative normalization metrics in the future.

We can now divide the raw emissions by our normalization metrics to create two different versions of the category-wise Carbon Benchmarks (CBR and CBH) at the level of an individual company, i:

Equ. 1a (revenue normalized)

$$CBR_{i,c} = \frac{E_{i,c}}{R_i}$$

Equ. 1b (headcount normalized)

$$CBH_{i,c} = \frac{E_{i,c}}{H_i}$$

If we look at these two normalized metrics for the 3 companies introduced earlier, we are now able to make a meaningful comparison across companies because the data has been adjusted for size.

Figure 3a. revenue-normalized, kgCO2e/USD revenue, Apparel Manufacturing

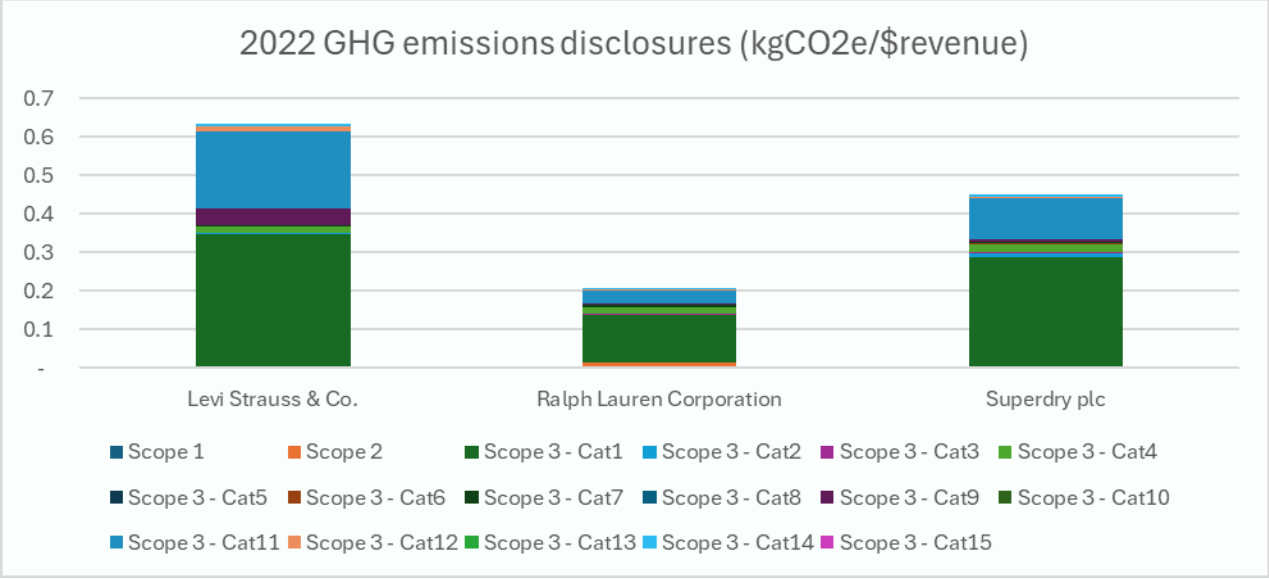
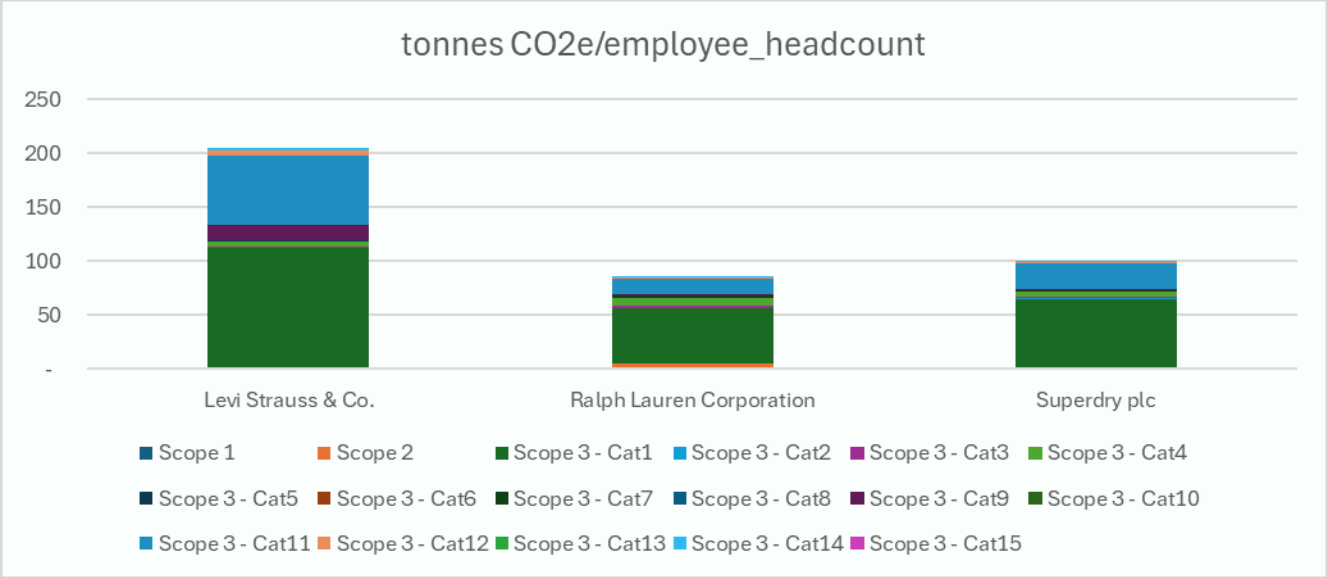


Figure 3b headcount-normalized, tonnes CO2e/employee headcount, Apparel Manufacturing



To be clear, **the above metrics are not the same thing as "spend-based" metrics** which have a bit of a bad reputation with the carbon accounting profession. **The companies in our sample have used a range of approaches including market-based, location-based, activity-based, and supplier-specific estimates**. Our use of revenue and headcount is merely a way to adjust for this size of the company and preserves the intelligence that went into the calculations.

Treatment of outliers

In the case of the 3 companies above, the normalized results have a degree of consistency and give some comfort that we have formed a homogenous grouping from which we can start estimating population statistics. However, in some cases the results after normalization showed some company estimates to be an order of magnitude or more too high or too low relative to their industry peers and in many cases our team were

unable to find good justification for why this might be the case. In such cases, we chose to remove these companies from the sample to avoid polluting our benchmarks with data that appeared prone to measurement error.

We experimented with several different methods for dealing with outliers including Winsorization (the automatic replacement of extreme values beyond a certain percentile) and Trimming (the automatic removal of extreme values beyond a certain percentile) but concluded that each outlier was best dealt with on a case-by-case basis and ultimately left the inclusion/exclusion decision down to our expert data team.

As an example of one of these judgment decisions, we found that Apple Inc had much lower emissions when normalized for revenue than the other Consumer Electronics companies but decided to keep Apple Inc in the peer dataset since we concluded there were plausible reasons why their emissions might be lower per unit of revenue (e.g. Apple Inc charges more for a mobile phone than most competitors. Their products also tend to have lower energy consumption for category 11 compared to some other consumer electronic products in the same price range e.g. TVs).

Our choice of sectors and industries

We decided to adopt the sector and industry hierarchy used by Yahoo Finance as a starting point for our segmentation to make it easier to link the emissions disclosures to the financial information of the publicly listed companies. This proved to be very useful in the normalization step described above. Note - if required we can easily re-cut the segmentation into alternative industry hierarchies such as ISIC or NAICS.

The **full list of 126 industries** we adopted is listed in **Appendix A**.

Aggregation

We then aggregated the data into industry-specific, category-wise benchmarks by taking the simple average of the normalized benchmarks of individual companies within each industry as follows:

Equ. 2a (revenue normalized)

$$CBR_{Industry,c} = \frac{\sum_{i=1}^{n} CBR_{i,c}}{n}$$

Equ. 2b (headcount normalized)

$$CBH_{Industry,c} = \frac{\sum_{i=1}^{n} CBH_{i,c}}{n}$$

Where n is the number of companies for which we have data available within that industry. For each of the 126 industries, we have a total of 18 category benchmarks (Scope 1, 2, and 3 plus the 15 sub-categories of Scope 3).

We then have 2 versions of the benchmarks: 1) normalized by revenue and 2) normalized by headcount, meaning that there are a total of 4,536 benchmarks in total (126 x 18 x 2) in our final dataset.

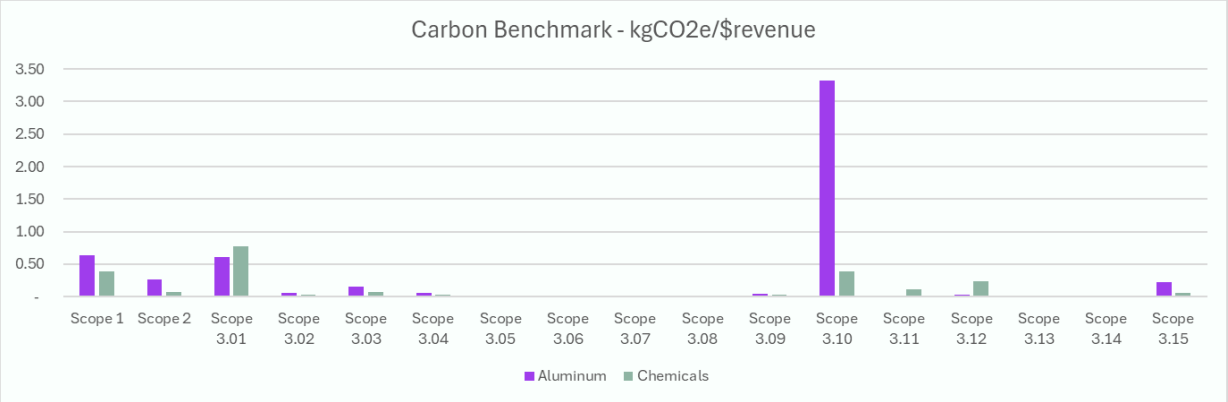
Proxy mapping for low-data industries

For 92 of the 126 industries, we have calculated the benchmarks directly by averaging the normalized results of the companies within that industry. For 34 of the industries, there was insufficient data available, so we have proxied those industries to an industry that we believed to represent a good match in terms of them having a similar profile of GHG emissions-relevant activities

3. Application of the peer benchmarks to various use cases

Below is an example of our revenue-based, category-wise peer benchmarks for 2 industries.

Fig 4. Carbon Benchmark (revenue-based) for the Aluminium and Chemicals industry



Let's now look at 2 well-known companies from these 2 industries:

Company Name	Industry	2022 Revenue $MM	Employee Headcount
Alcoa	Aluminium	12,451	13,600
BASF SE	Chemicals	91,807	111,855

Estimation using our peer benchmarks

The power of our peer benchmarks can now be illustrated by the fact that we can estimate the detailed category-wise emissions of these two companies, solely based on the revenue and headcount information in the above table.

Revenue-based estimation

$$Emissions_{Alcoa,c} = CBR_{Aluminium,c}.Revenue_{Alcoa}$$

$$Emissions_{BASF,c} = CBR_{Chemicals,c}.Revenue_{BASF}$$

Headcount-based estimation

$$Emissions_{Alcoa,c} = CBH_{Aluminium,c}.Headcount_{Alcoa}$$

$$Emissions_{BASF,c} = CBH_{Chemicals,c}.Headcount_{BASF}$$

Backtesting of results

For these two companies, we have the luxury of knowing their actual 2022 GHG emissions disclosures which allows us to backtest the benchmark-based estimates against their actual emissions disclosures:

Figure 5a. Backtesting of peer benchmarking estimate vs actual for Alcoa

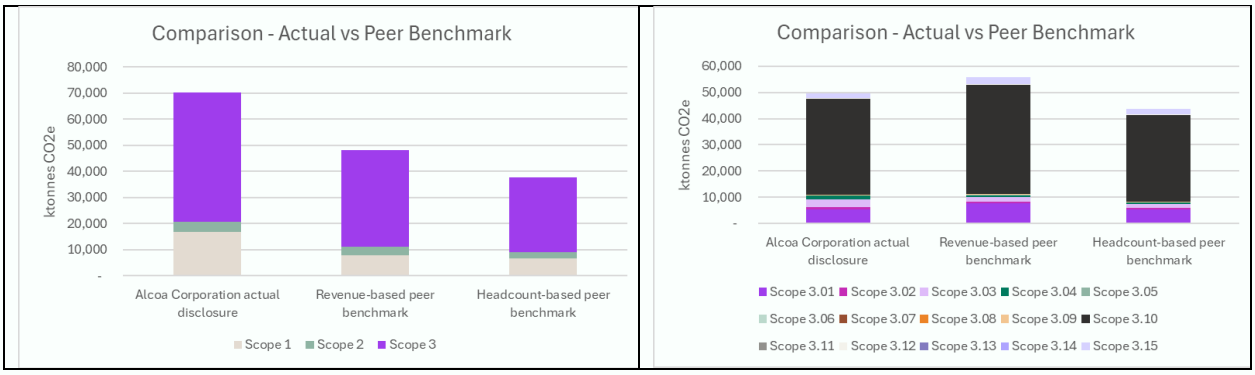
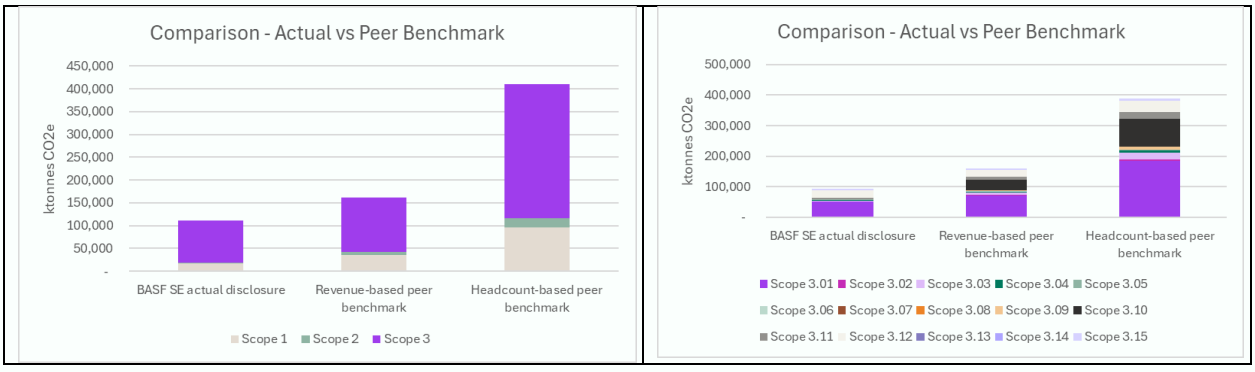
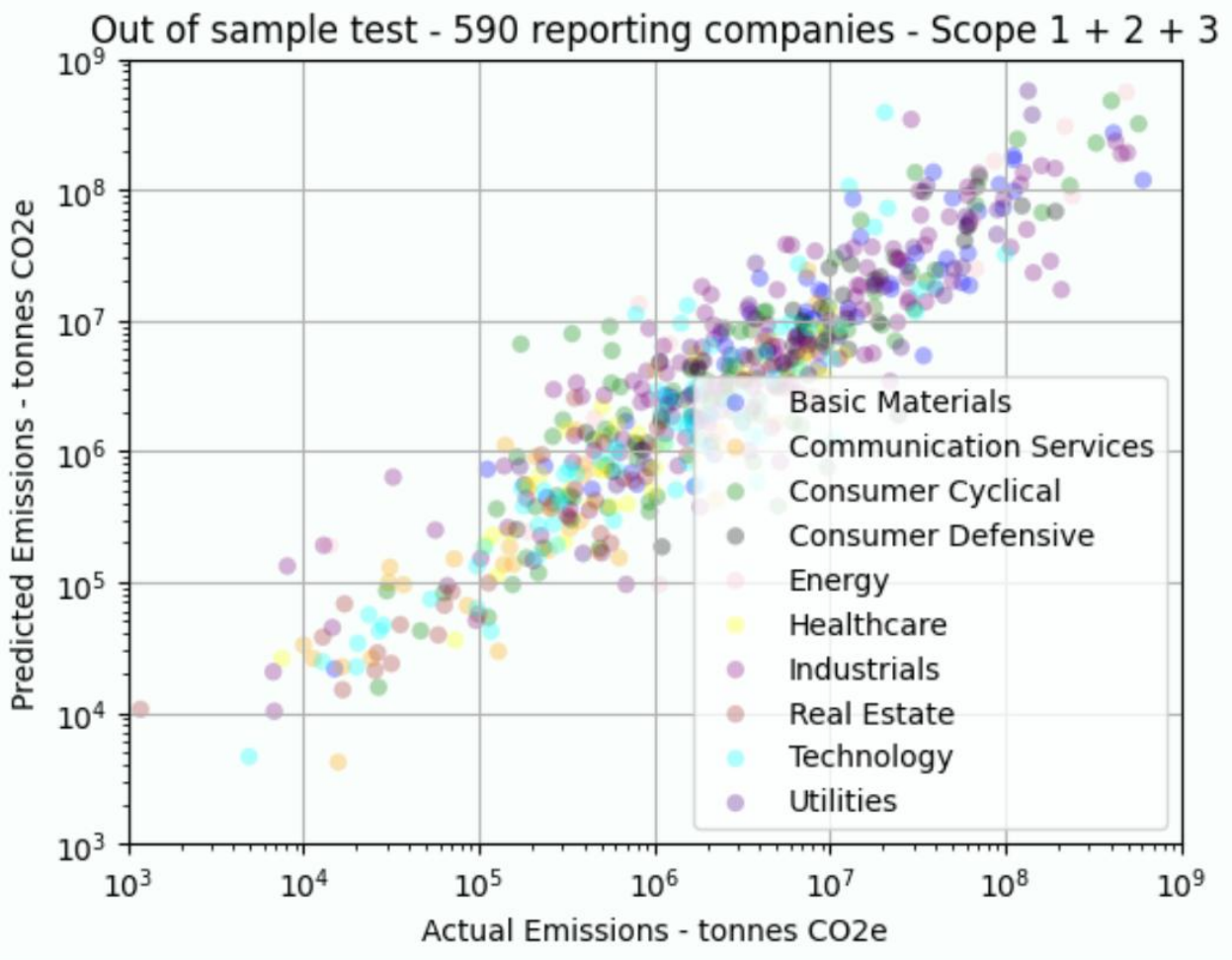


Figure 5b. Backtesting of peer benchmarking estimate vs actual for BASF



Below are the results of out-of-sample backtesting on a sample of 590 companies with the following results:



Out-of-sample testing is a form of “blind test” where the model predicts the result without access to the actual observed result. We can then reveal the actual emissions of the company and compare it to our prediction. As can be seen, our benchmarks are clustered around the 45-degree diagonal indicating that we offer an unbiased estimate of emissions for companies that haven’t yet disclosed their emissions.

The correlation between the predicted and observed is roughly 70%. We have identified 4 main reasons why the actual emissions of a company might deviate from our industry-average predictions:

1) Problems with the accuracy of our model due to data limitations
2) Problems with mapping of companies to individual industries when they in fact span multiple industries
3) Modelling errors in the company's own estimates
4) Actual differences in the company's activities vs their peers (e.g. if the company is decarbonizing faster than its peers in certain areas)

We aim to reduce the problems of 1 and 2 through time through the collection of more data and the use of "composite" benchmarks for individual companies. The onus is then on the reporting company to understand and explain whether any remaining deviations are caused by point 3 vs point 4.

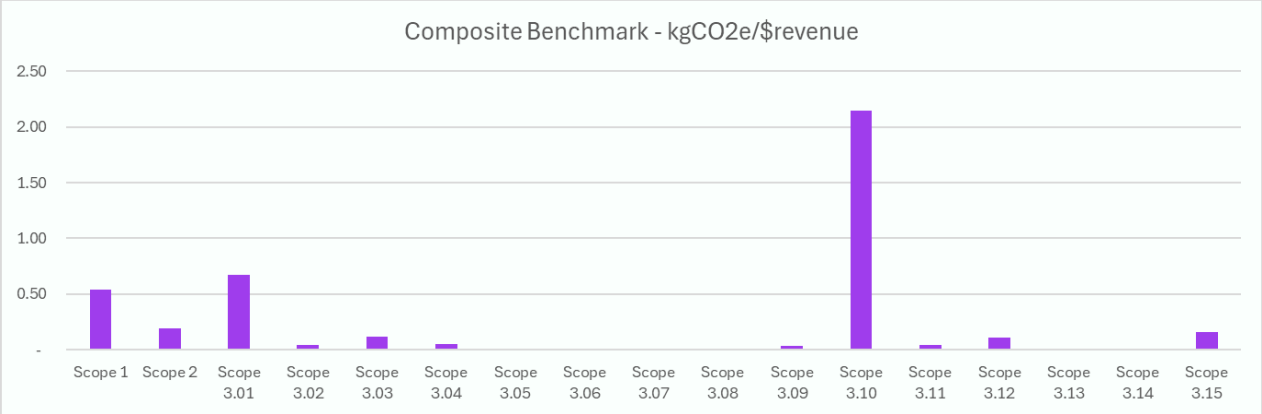
Composite benchmarks

Many companies span multiple industries and our industry-specific peer benchmarks can be easily adapted to cater to this. For example, if a company had revenue weighted 60% to 40% to the Aluminium and Chemicals industries, we can simply construct the composite benchmark by weighting the individual industry benchmarks by the relevant revenue weights:

$$Composite\ Benchmark_c = 0.6\ CBR_{Aluminium,c} + \ 0.4\ CBR_{Chemicals,c}$$

This leads to the following composite profile in the above case.

Fig 6. Composite peer benchmark (60% Aluminium 40% Chemicals)



Financed Emissions – for banks, insurance companies, and asset managers

Financial Institutions play a unique role in financing the activities and future sustainability transition of companies and are under pressure to document how their lending, bond, and equity portfolios are decarbonizing through time.

The 2022 guidance on Finance Emissions from PCAF states the following:

"Limited data is often the main challenge in calculating financed emissions. However, data limitations should not deter financial institutions from starting their GHG accounting journeys. **Beginning with estimated or proxy data can help financial institutions identify emission-intensive hotspots in lending and investment portfolios**"

There is currently no better "proxy" available for estimating a reporting company's emissions than the industry-specific, category-wise, normalized peer benchmarks that we have introduced in this report.

We have already proxied the emissions of more than 10,000 publicly listed companies using our revenue-based and headcount-based benchmarks and we are currently

working with banks and asset managers to apply our rich set of industry-specific benchmarks into SME lending and Private Equity markets.

Additional use cases

Once we open the potential to estimate accurate category-wise emissions of a company based on very limited information (industry, revenue, and headcount) then this opens a whole world of use cases for several different stakeholders:

For individual reporting companies

- Peer benchmarks can help guide a company's materiality assessment.
- Peer benchmarks can help a company prioritize its data collection efforts.
- Peer benchmarks can be used to sanity-check results.
- Peer benchmarks can be used to populate missing categories while internal models are under development.

For assurance professionals

- Assurance professionals can use peer benchmarks to identify a potential error or misstatement in a company's reporting GHG emissions.

For regulators

- Use of peer benchmarks to challenge a company's materiality assessment.
- Use of peer benchmarks to challenge a company's methods and assumptions.
- Use of peer benchmarks to calibrate a "standardized floor" for certain categories.
- Use of peer benchmarks to compare the decarbonization progress of companies within their industry.

Appendix A: The full list of sectors and industries in our model

Sector	Industry
Basic Materials	Agricultural Inputs
Basic Materials	Aluminum
Basic Materials	Cement
Basic Materials	Chemicals
Basic Materials	Copper
Basic Materials	Gold
Basic Materials	Lumber & Wood Production
Basic Materials	Other Industrial Metals & Mining
Basic Materials	Other Precious Metals & Mining
Basic Materials	Paper & Paper Products
Basic Materials	Silver
Basic Materials	Specialty Chemicals
Basic Materials	Steel
Communication Services	Advertising Agencies
Communication Services	Broadcasting
Communication Services	Electronic Gaming & Multimedia
Communication Services	Entertainment
Communication Services	Internet Content & Information
Communication Services	Publishing
Communication Services	Telecom Services
Consumer Cyclical	Apparel Manufacturing
Consumer Cyclical	Apparel Retail
Consumer Cyclical	Auto & Truck Dealerships
Consumer Cyclical	Auto Manufacturers
Consumer Cyclical	Auto Parts
Consumer Cyclical	Department Stores
Consumer Cyclical	Footwear & Accessories
Consumer Cyclical	Furnishings, Fixtures & Appliances
Consumer Cyclical	Gambling
Consumer Cyclical	Home Improvement Retail
Consumer Cyclical	Internet Retail
Consumer Cyclical	Leisure
Consumer Cyclical	Lodging
Consumer Cyclical	Luxury Goods
Consumer Cyclical	Packaging & Containers
Consumer Cyclical	Personal Services
Consumer Cyclical	Recreational Vehicles
Consumer Cyclical	Residential Construction

Consumer Cyclical	Resorts & Casinos
Consumer Cyclical	Restaurants
Consumer Cyclical	Specialty Retail
Consumer Cyclical	Textile Manufacturing
Consumer Cyclical	Travel Services
Consumer Defensive	Beverages - Brewers
Consumer Defensive	Beverages - Non-Alcoholic
Consumer Defensive	Beverages - Wineries & Distilleries
Consumer Defensive	Confectioners
Consumer Defensive	Discount Stores
Consumer Defensive	Education & Training Services
Consumer Defensive	Farm Products
Consumer Defensive	Food Distribution
Consumer Defensive	Grocery Stores
Consumer Defensive	Household & Personal Products
Consumer Defensive	Packaged Foods
Consumer Defensive	Tobacco
Energy	Oil & Gas Drilling
Energy	Oil & Gas E&P
Energy	Oil & Gas Equipment & Services
Energy	Oil & Gas Integrated
Energy	Oil & Gas Midstream
Energy	Oil & Gas Refining & Marketing
Healthcare	Biotechnology
Healthcare	Diagnostics & Research
Healthcare	Drug Manufacturers - General
Healthcare	Drug Manufacturers - Specialty & Generic
Healthcare	Health Information Services
Healthcare	Healthcare Plans
Healthcare	Medical Care Facilities
Healthcare	Medical Devices
Healthcare	Medical Distribution
Healthcare	Medical Instruments & Supplies
Healthcare	Pharmaceutical Retailers
Industrials	Aerospace & Defense
Industrials	Airlines
Industrials	Airports & Air Services
Industrials	Building Products & Equipment
Industrials	Business Equipment & Supplies
Industrials	Conglomerates
Industrials	Consulting Services
Industrials	Electrical Equipment & Parts
Industrials	Engineering & Construction

Industrials	Farm & Heavy Construction Machinery
Industrials	Industrial Distribution
Industrials	Infrastructure Operations
Industrials	Integrated Freight & Logistics
Industrials	Marine Shipping
Industrials	Metal Fabrication
Industrials	Railroads
Industrials	Rental & Leasing Services
Industrials	Security & Protection Services
Industrials	Specialty Business Services
Industrials	Specialty Industrial Machinery
Industrials	Staffing & Employment Services
Industrials	Tools & Accessories
Industrials	Trucking
Industrials	Waste Management
Real Estate	Real Estate - Development
Real Estate	Real Estate - Diversified
Real Estate	Real Estate Services
Real Estate	REIT - Diversified
Real Estate	REIT - Healthcare Facilities
Real Estate	REIT - Hotel & Motel
Real Estate	REIT - Industrial
Real Estate	REIT - Mortgage
Real Estate	REIT - Office
Real Estate	REIT - Residential
Real Estate	REIT - Retail
Real Estate	REIT - Specialty
Technology	Communication Equipment
Technology	Computer Hardware
Technology	Consumer Electronics
Technology	Electronic Components
Technology	Electronics & Computer Distribution
Technology	Information Technology Services
Technology	Scientific & Technical Instruments
Technology	Semiconductor Equipment & Materials
Technology	Semiconductors
Technology	Software - Application
Technology	Software - Infrastructure
Technology	Solar
Utilities	Utilities - Diversified
Utilities	Utilities - Independent Power Producers
Utilities	Utilities - Regulated Electric
Utilities	Utilities - Regulated Gas

Utilities	Utilities - Regulated Water
Utilities	Utilities - Renewable