

Article

Overallocation and Correction of Carbon Emissions in the Evaluation of Carbon Footprint

Alex Gurvich ^{1,*}  and Germán G. Creamer ^{2,*} 

¹ PFM Asset Management, Philadelphia, PA 19103, USA

² School of Business, Stevens Institute of Technology, Hoboken, NJ 07030, USA

* Correspondence: alex@gurvich.net (A.G.); gcreammer@stevens.edu (G.G.C.)

Abstract: This paper points to several carbon footprint data distortions that overallocate carbon footprints to individual companies, and to several carbon data intricacies that lead to improved data integrity. Data distortion due to the same company being listed in multiple geographical jurisdictions or through different share classes overstates Emissions Scope 1 by 4.6%, Emissions Scope 2 by 5.5%, Emissions Scope 3 by 10.6% and Reserves by 6.0%. Data distortion due to index construction by having different market capitalization in representative indices overallocates Emissions Scope 1 by 33.9%, Emissions Scope 2 by 27.6%, Emissions Scope 3 by 21.3% and Reserves by 57.2%. A significant amount of carbon data is not precise but is estimated by third-party providers through proprietary techniques. The estimated data for Scope 1 Emissions is 46.4% for the companies in the index. In addition, carbon data is stale, resulting in 94.5% of data being two years old or more. Usage of carbon data in a present format may incorrectly remove some companies from portfolios (negative screen, complete removal) or incorrectly reduce some companies' weight in a portfolio (partial screen, fractional removal).



Citation: Gurvich, A.; Creamer, G.G. Overallocation and Correction of Carbon Emissions in the Evaluation of Carbon Footprint. *Sustainability* **2021**, *13*, 13613. <https://doi.org/10.3390/su132413613>

Academic Editors: Tomonobu Senjyu and Alessandro Franco

Received: 21 October 2021

Accepted: 3 December 2021

Published: 9 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: carbon footprint; energy finance; carbon emissions; Socially Responsible investment; Corporate Sustainability; carbon risk factors; emissions scope

1. Introduction

The methodology for calculating carbon footprints and understanding how carbon footprint affects portfolio construction and valuation is challenging on several levels. The tasks include data classification, reporting, sources, timing and attribution. Understanding the intricacies of carbon footprint data is critical because carbon data is not reported in financial statements, nor is it required by financial regulators. When carbon data is provided, it is provided by the companies directly or is estimated by third parties, while in some instances it is not provided at all. In addition, carbon data is non-standardized, unlike financial data under GAAP, and thus the quality of data is mixed. This research aims to address some of these carbon data issues.

The challenge presented by data classification is a result of multiple factors. Some data is classified by emissions and reserves. Within emissions, there are three classifications by scope, and within each classification a carbon footprint comes from different sources such as coal, oil, or gas. Another problem with data reporting is that a portion of data is actual, while the rest is estimated by third parties, even though most analysis treats data as actual. An additional challenge with data sources stems from the fact that some data is reported directly by the company, while the rest is aggregated by third parties. Further, data timing proves challenging because for the same reporting year by a vendor, the actual data is anywhere from one to three years old. However, all analysis treats data as current. A further issue with data attribution is that the carbon footprint for some companies is double or triple-counted because the same company may be listed in different jurisdictions or through different security listings, and most studies attribute a full carbon footprint to

a company for each listing, resulting in a larger than appropriate carbon footprint. Yet another challenge with data attribution is that the full carbon footprint is allocated to a company, whereas the company's market capitalization weight is not fully represented in an index, even though most studies attribute the full carbon footprint to the company in question without regard to market weight of the company in the index.

Carbon emissions and reserve data have a relatively short history, and it is highly structured [1]. The carbon data is self-reported by companies or estimated through proprietary methodologies by a handful of data aggregators (MSCI, CDP, South Pole (S&P), Trucost, etc.), who publish it on their proprietary platforms. These providers aggregate raw company data, structure the available information, and estimate missing data [2]. Because there is no uniform or consistent way of comparing carbon data between data providers, such data is poorly correlated between these various sources [3].

Although carbon data reporting is improving, and every month adds another data point to the carbon data time series, most investors and researchers are still focused on using derivative carbon scores from proprietary data providers instead of raw carbon data (carbon footprint). In cases where raw carbon data is used, it is received directly from data providers in an "as is" format, with no scrutiny for data accuracy or integrity.

This paper analyzes carbon data and introduces several carbon data adjustment steps that (a) attribute a direct carbon footprint to a company's operations, (b) reduce overallocation of carbon footprints by adjusting for multiple listings and (c) reduce overallocation of carbon footprints by adjustment based on market capitalization within the index at the individual company level. This paper also (a) addresses issues with actual and estimated carbon footprint data and (b) discusses the time lag in carbon data reporting.

2. Literature Review

Usage of Environmental and Carbon data can be divided into the following three broad categories: Corporate Sustainability (CS) (or Socially Responsible (SR)); Environmental, Social and Governance (ESG) and Carbon (C). Each of these categories is presented to researchers as scores or rankings (MSCI—CarbonMetrics Methodology and Definitions, 2014). These scores are provided for each company and are based on the data provider's proprietary data aggregation methodologies, with typical rankings ranging from 1 for the worst ranked company to 10 for the best ranked. Several research papers still use derivative carbon scores [4,5], while more recent research papers use raw carbon data [6–13] as provided directly by data providers. Only Gaspar et al. [14] and Gurvich [15] demonstrate how to remove data flaws and correct for data integrity.

Carbon Data

Looking at the usage of raw carbon data, Derwal et al. [16], Görgen et al. [4], Tan et al. [5] and Konar et al. [17] used carbon scores to show that companies with lower carbon scores have either better returns or lower risk, while Hao et al. [8], Heaps et al. [18], Andersson et al. [6], Jacob [11], Matsumura et al. [12], Yongqing et al. [5] and Cheema-Fox et al. [7] used raw carbon data to show higher performance or lower risk. Harris [9,10], Gaspar et al. [14] and Gurvich [15] also used raw carbon data but dived deeper into the carbon data structure and separated carbon data analysis based on three carbon emissions Scopes.

Derwal et al. [16] used eco-efficiency scores, which were a precursor to the carbon scores, to show that companies with higher scores also had higher average returns than companies with lower scores. Görgen et al. [4] used carbon-related indicators, such as value chain, public perception, and adaptability, to create a Brown Minus Green risk factor. Tan et al. [5] used carbon scores to show that carbon and climate risks are not fully priced by investors and that carbon intensive industries do not provide adequate risk-adjusted returns. Konar et al. [17] used environmental finance variables, such as toxic releases and environmental lawsuits, to show a relationship with a firm's intangible value.

Hao et al. [8] showed that a carbon efficient portfolio outperforms on an absolute return basis while it underperforms on a risk-adjusted basis. Heaps et al. [18] observed that

companies with a low-carbon footprint and/or low-carbon intensity can outperform their respective benchmarks and provide better risk-adjusted returns. Andersson et al. [6] used raw carbon data to create a lower carbon footprint portfolio without reducing portfolio returns. Jacob [11] proposed low-carbon portfolios that can provide lower portfolio risk relative to the benchmark. Matsumura et al. [12], showed a negative association, based on raw carbon data, between carbon emissions and firm value. Yongqing et al. [13] demonstrated that Australian regulations on carbon emissions negatively affect companies' asset values and cash flows.

Harris [9] stratified raw carbon into different emissions Scopes to show how a low-carbon intensity portfolio produces significant outperformance relative to high carbon intensity portfolios. Harris [10] also built an Efficient Minus Intensive factor to show a positive risk premium that is uncorrelated from other risk factors. Gaspar et al. [14] used raw carbon data, stratified into different emissions Scopes, to build low-carbon multi-asset portfolios that exhibit similar returns and risk-adjusted returns as the benchmark. Cheema-Fox et al. [7] proposed decarbonization factors.

In general, previous papers use secondary sources or proprietary data to evaluate carbon footprints. Some studies use carbon scores, while others use raw carbon data and stratify data for more granular analysis. In addition, some providers utilize word ratings instead of scores, although the meaning is the same. Other studies use carbon scores developed by data providers through proprietary and non-transparent methodologies to link raw carbon data with financial performance. Several studies use raw carbon data as measured in tons of CO₂ and referred to as carbon volume, while others use a carbon intensity indicator which is defined as carbon volume divided by the company's sales. However, none of these papers use raw carbon data to correct carbon allocation errors generated by multiple listings and index allocation.

This paper explicitly uses raw, primary data carbon instead of derivative carbon scores, and proposes a method to adjust the data to reduce carbon allocation errors. This research reallocates the corporations' carbon footprints, correcting the following two types of distortions: multiple listings and index allocation. Multiple listings distortion refers to the number of times the same company is listed as unique in the MSCI ACWI Index. Index allocation distortion refers to the over/under-allocation of carbon footprints based on different companies' market capitalization values in the MSCI ACWI index relative to the company's total market capitalization in the MSCI IMI Index. As a result, the reallocated carbon emissions are separated within Scope 1, Scope 2 and Scope 3, as well as within carbon reserves.

The correction of carbon allocation errors proposed in this paper can be used to develop new or more accurate carbon risk factors that can lead to new portfolio optimization strategies that optimize risk-adjusted return as well as carbon financial and operational efficiency as in Gurvich and Creamer [19].

3. Carbon Data Structure and Usage

3.1. Data Classification—Emissions and Reserves

This research uses raw carbon emissions for data analysis and avoids proprietary Corporate Sustainability (or Socially Responsible) scores, Environmental, Social and Governance scores or Carbon scores/rankings.

Carbon data classification is relevant in terms of how it is calculated and presented. Carbon footprints are generally classified in terms of carbon (CO₂) emissions and carbon (CO₂) reserves. Both are measured in millions of tons of carbon dioxide equivalent. Carbon emissions refer to the carbon dioxide gas emitted during an industrial process, transportation activities or production of heat. The reason emissions are calculated in "equivalents" such as tons of carbon dioxide equivalents ("tCO₂e") is because there are many more gases (e.g., CH₄) produced during industrial processes that also contribute to greenhouse gases, although for the purpose of universality all gases are converted to a carbon dioxide equivalent as presented in Table 1.

Table 1. Carbon Dioxide Equivalent Conversion.

Greenhouse Gases	Tons of Carbon Dioxide Equivalent (tCO ₂ e)
Carbon Dioxide (CO ₂)	1
Methane (CH ₄)	28–36
Nitrous Oxide (N ₂ O)	265–298
Chlorofluorocarbons (CFCs)	7390–22,800

Source: Environmental Protection Agency.

Carbon emissions are subdivided into three Scopes. Scope 1 refers to direct emissions, such as a company's office heating or emissions released during manufacturing or production processes. Scope 2 refers to indirect emissions due to a company's purchased electricity. For example, during an industrial process, the company in question buys electricity from a utility company, and emissions produced by the electricity provider are attributed to the emissions of the company that uses that electricity. Scope 3 refers to indirect emissions (primarily) downstream of the value chain and emissions produced by the company's sold products. For example, when individuals who own and drive cars produce carbon emissions, these emissions are counted as Scope 3 emissions for the car manufacturer. All three Scopes are attributed to one company, raising the issue of double or triple counting, as Scope 2 or Scope 3 emissions of one company are Scope 1 emissions of another company. Clearly, an investor needs to decide what Scopes are considered when evaluating a company's emission.

Carbon reserves are fossil fuels that are still in the ground and have not been extracted yet. The most relevant fossil fuels, in order of decreasing magnitude of carbon dioxide emissions, are coal, oil and gas. It is important to note that if these fossil fuels are still in the ground or extracted but not yet used (i.e., burned), they emit no carbon emissions. Since reserves are potential emissions, an investor needs to assess the likelihood of reserves being extractable from the ground and the likelihood that they will be extracted. In addition, reserves are aggregated in millions of tons of carbon dioxide equivalents, or MtCO₂e, which equates to the carbon effect from coal, oil and gas. However, coal, oil and gas have different CO₂ equivalents according to their potential carbon footprint (Table 2).

Table 2. Fossil Fuels CO₂ Equivalents.

Fossil Fuel	Relative CO ₂ Production
Coal	1.00
Oil	0.75
Gas	0.50

Source: US Energy Information Administration.

3.2. Data Reporting—Actual vs. Estimated

Carbon footprint data is typically aggregated by data providers, such as MSCI, CDP, South Pole (S&P) and Trucost. Data is classified as actual, reported, or estimated. Actual data is directly calculated for a specific activity, reported data is reported by the company, while estimated data is compiled by a data provider. Unlike financial data, which typically follows Generally Accepted Accounting Principles (GAAP), carbon data does not follow universal reporting schemes and is not required to be reported. In addition, companies may report only one out of three Scopes of emissions data (Table 3).

Typically, carbon data providers cover around three thousand companies for carbon footprint data; however, the amount of estimated data is significant as Table 4 shows for MSCI ACWI.

Table 3. Companies with Available Carbon Data.

Category	Companies	Percent
Total MSCI ACWI (June 2015)	2483	100%
Carbon Emissions Scope 1	2435	98%
Carbon Emissions Scope 2	2424	98%
Carbon Emissions Scope 3	901	36%
Carbon Reserves	136	5%

Source: MSCI.

Table 4. Actual vs. Estimated Scope 1 Carbon Emissions Data.

Category	Companies	Percent
Total MSCI ACWI (July 2016)	2481	100.0%
Reported Carbon Data	1240	50.0%
Estimated Carbon Data	1150	46.4%
No Carbon Data	91	3.6%

Sources: MSCI.

Each data provider has its proprietary estimation models, which are continually updated. Thus, changes in the models lead to the re-evaluation of previously estimated carbon data. In addition, each provider has several estimation models, depending on the sector or any other attribute. Carbon data from different providers varies significantly, as shown by Abele (2016), where the highest correlation between four vendors was 0.37 and the lowest was -0.14 . Clearly, if carbon data is to be used for assessing value, the use of estimated data will pose a challenge for accuracy.

3.3. Data Timing Lag

Carbon data is reported with a significant lag. Data providers need to collect and input various data sources, but there is a natural delay in obtaining carbon data. In addition to the actual data, data providers must calculate estimated data, which also adds to the data reporting lag. Typical lag within a data set is between one and three years as can be observed for MSCI ACWI in Table 5.

Table 5. Time Lag of Carbon Data Reporting.

Reporting Year	Companies	Percent
Data from 2013	9	0.4%
Data from 2014	2334	94.1%
Data from 2015	109	4.4%
Data from 2016	0	0.0%
No data	29	1.2%
Total MSCI ACWI (June 2016)	2481	100%

Source: MSCI June 2016.

Table 5 represents an example of data lags at one point in time in June 2016; however, the entire research analyses data from January 2010 to March 2017.

4. Methodology

4.1. Data Attribution for Multiple Listings

This paper proposes that a corporation's carbon footprint should be assessed based on combining total market capitalization from various geographical listings and/or share classes in the MSCI ACWI index and allocating carbon footprint proportionally. This methodology eliminates double or triple counting and overallocation of a company's

carbon footprint, reallocating carbon emissions and reserves as presented in Equations (1) and (2) according to total market capitalization (Equation (3)).

$$ACE_i = \delta_i TCE \quad (1)$$

$$ACR_i = \delta_i TCR \quad (2)$$

$$TMC = \sum MC_i \quad (3)$$

ACE_i allocated carbon emissions for one company in jurisdiction i (or share class i)

δ_i market capitalization in ACWI index for one company in jurisdiction i (or share class i) divided by market capitalization in ACWI IMI index in one company in jurisdiction i (or share class i)

TCE total carbon emissions for the company

ACR_i allocated carbon reserves for one company in jurisdiction i (or share class i)

TCR the total carbon reserves for the company

TMC total market capitalization for one company

MC_i market capitalization for one company in jurisdiction i (or share class i)

4.2. Data Attribution for Index Allocation

Another attribute of the MSCI carbon data is that it allocates a company's total reported carbon footprint (emissions and reserves) to the company within the MSCI ACWI index, even though the company's total market capitalization is not fully included in the MSCI ACWI index. Since investors invest in the MSCI ACWI index and not the MSCI ACWI IMI index, their holdings do not represent the full market capitalization of each company in their portfolio. This distorts and typically overestimates the carbon footprint for each company.

This distortion in carbon footprint allocation arises due to MSCI's index construction methodology (MSCI 2020). MSCI developed an MSCI ACWI IMI (Investable Market Index) index [20], based on the MSCI Global Investable Market Indices (GIMI) Methodology. This methodology aims to provide exhaustive coverage of the relevant investment opportunity set with an emphasis on index liquidity, replicability and investment potential. This index covers all investable large-, mid- and small-cap securities across the developed and emerging markets, targeting approximately 99% of each market's free-float adjusted market capitalization.

However, the MSCI ACWI IMI index is used primarily as a reference index rather than as an investable index. The MSCI ACWI IMI index is not practical for investing due to many reasons, some of which include a very large number of securities, variable liquidity, number of shares available for trading, country regulations etc., to name a few. For institutional investors, it is the MSCI ACWI Index (as opposed to MSCI ACWI IMI index) that is the most used and referenced global index for investing and benchmarking purposes.

In the construction of the MSCI ACWI index, MSCI starts with the reference MSCI ACWI IMI index and makes several transformations, as part of which companies are included with their respective market capitalization. The first transformation is to apply an 85% free float-adjusted market capitalization. This first transformation is used across the board for all large- and mid-capitalization stocks that are included in the MSCI ACWI index (small-capitalization stocks are not included in the MSCI ACWI index). The second transformation applies a proprietary variable adjustment to free float market capitalization. This second transformation is determined by defining a size reference for the MSCI ACWI IMI index and specifying a range of 0.5 times to 1.15 times (MSCI 2020).

The result of these transformations indicates that when a company is included in the MSCI ACWI index, its market capitalization can vary from 42.50% to 97.75% of its full market capitalization listed in the MSCI ACWI IMI index.

This MSCI index construction technique leads to the distortion of actual carbon volume allocated to a company, and carbon volume must therefore be adjusted accordingly. Hence, in this paper, carbon emissions and reserves are allocated based on the proportion of the company's market capitalization in the MSCI ACWI index (Equations (4) and (5)). This methodology reduces the overestimation of the company's carbon footprint.

$$ACE = \gamma TCE \quad (4)$$

$$ACR = \gamma TCR \quad (5)$$

ACE allocated carbon emissions for one company
 γ proportion of a company's market capitalization in the ACWI index to market capitalization in the ACWI IMI index
TCE total carbon emissions for the company
ACR allocated carbon emissions for one company
TCR total carbon reserves for the company

5. Results

Carbon data aggregators go through the process of collecting, cleaning and attributing carbon to companies. As data is reported, each provider allocates carbon in a different way for each company. The data is thus dependent on the data provider's process. In the case of MSCI's data, several data attributes affect data integrity.

5.1. Multiple Listings

The correction of carbon allocation due to multiple listing according to equations 1 and 2 is reported in Table 6. For the period from January 2010 to March 2017, the number of double- or triple-listed companies ranged from 67 to 171 listings. Over the analyzed period, on average 1.6% of companies were double or triple listed; however, the Emissions Scope 1, 2 and 3 and Reserves were overestimated by 4.6%, 5.5%, 10.6% and 6% respectively.

Table 6. Average Overstatement of Emissions and Reserves: 2010–2017.

	Companies in Index and Multiple Listings	Emissions Scope 1 (CO ₂ Tons M)	Emissions Scope 2 (CO ₂ Tons M)	Emissions Scope 3 (CO ₂ Tons M)	Reserves (CO ₂ Tons M)
Original Volume	2450 companies	8252	1525	13,682	336,333
Adjusted Volume	40 multiple-listed	7889	1445	12,374	317,439
Overstated Change	1.6%	4.6% ***	5.5% ***	10.6% ***	6.0% ***

Source: MSCI. Note: *** *t*-test mean difference significance level at 99%.

Descriptive statistics for emission and reserves are in Tables 7 and 8 below.

Table 7. Descriptive Statistics for Emissions and Reserves for MSCI ACWI Index.

MSCI ACWI Index	Emissions Scope 1 (CO ₂ Tons M)	Emissions Scope 2 (CO ₂ Tons M)	Emissions Scope 3 (CO ₂ Tons M)	Reserves (CO ₂ Tons M)
Observations	87	87	87	22
Mean	8252	1525	13,682	336,333
St. Deviation	815	211	21,674	15,875
Maximum	9079	1897	106,321	367,754
Minimum	6597	1135	571	320,093

Source: MSCI.

Tables 7 and 8 highlight the differences between the original carbon footprint and corrected carbon footprint due to correction for multiple listings. It is not a surprise that the mean for corrected data is lower than in the original index, as we remove and

reduce the total amount of emissions and reserves. What is surprising is that data variation, as measured by standard deviation, maximum and minimum, has not changed significantly. We may attribute this to initially high data dispersion, especially for Scopes 2 and 3, as well as Reserves.

Table 8. Descriptive Statistics for Emissions and Reserves for MSCI ACWI Index Corrected for Multiple Listings.

MSCI ACWI Index Corrected	Emissions Scope 1 (CO ₂ Tons M)	Emissions Scope 2 (CO ₂ Tons M)	Emissions Scope 3 (CO ₂ Tons M)	Reserves (CO ₂ Tons M)
Observations	87	87	87	22
Mean	7889	1445	12,374	317,439
St. Deviation	794	209	21,054	13,955
Maximum	8768	1819	103,562	346,654
Minimum	6331	1054	570	303,391

Source: MSCI.

With certain variations, the overestimations of Emissions Scope 1 and 2 converged to a central value in the period 2014–2017 (Figure 1). Similarly, the number of names in the index does not vary extensively.

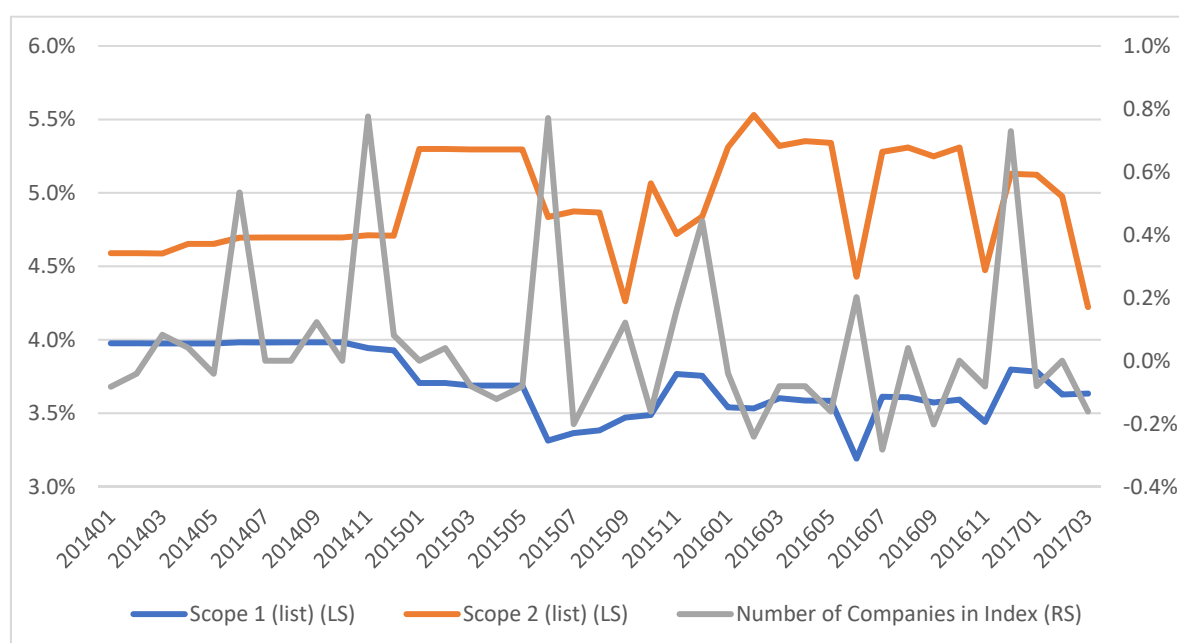


Figure 1. Overstatement of Emissions Scope 1 and 2 due to multiple listings. Number of companies in the index is on the right axis (2014–2017).

The relative stability of Scope 1 and Scope 2 overstatements may be due to the stability of the ACWI index. The number of names added or removed in any given month varies between -0.3% and 0.8% .

As an example, MSCI provides identical carbon data for the same company, within an equity index, which is listed in different jurisdictions and/or through different share classes. This causes double and triple counting of carbon data for multiple listed companies as observed in the carbon emissions and carbon reserve columns for BHP Billiton in Table 9. Using its market capitalization for UK and Australia index listings in Equations (1) and (2), the carbon emissions and carbon reserve weighted columns show an important adjustment.

Table 9. Appropriate way to allocate carbon footprint for BHP Billiton.

Asset Name	Issuance	Share Class	Carbon Emissions (Mt)	Carbon Reserves (Mt)	Market Capitalization (\$M)	Market Capitalization %	Carbon Emissions (Mt) Weighted	Carbon Reserves (Mt) Weighted
BHP BILLITON PLC	United Kingdom	PLC	38	3907	26,597	37.4%	14	1460
BHP BILLITON LIMITED	Australia	Limited	38	3907	44,600	62.6%	24	2447
Total			76	7814	71,197	100.0%	38	3907

Source: MSCI June 2016.

Table 10 shows that in June 2016, 40 companies had multiple listings, of which 3 companies were triple-listed and 37 companies were double-listed. These 40 companies with multiple listings represent 1.6% of the 2481 companies listed in the MSCI ACWI Index in June 2016. However, the overstatement of carbon footprint is significantly larger: Emissions Scope 1, 2 and 3 as well as Reserves were overstated by 3.2%, 4.4%, 14.3% and 4.6%, respectively (Table 11). A complete list of 40 double- and triple-listed companies for June 2016 is listed in Appendix A.

Table 10. Number of Companies with Multiple Listings.

	Companies
Double-listed companies	37
Triple-listed companies	3
Total companies with multiple listings	40
Companies in MSCI ACWI Index	2481
Percentage of companies with multiple listings	1.6%

Source: MSCI June 2016.

Table 11. Overstatement of Emissions and Reserves, typical month.

	Companies with Multiple Listings	Emissions Scope 1 (CO ₂ Tons M)	Emissions Scope 2 (CO ₂ Tons M)	Emissions Scope 3 (CO ₂ Tons M)	Reserves (CO ₂ Tons M)
Original Volume	2481 companies	8702	1454	17,235	324,815
Adjusted Volume	40	8433	1392	15,085	310,458
Overstated Change	1.6%	3.2%	4.4%	14.3%	4.6%

Source: MSCI June 2016.

5.2. Index Allocation

Adjusting a corporation's carbon footprint based on the proportion of each company's market capitalization in the index according to Equations (4) and (5) reduces the overallocation of the carbon footprint as reported in Table 12. The MSCI ACWI Investable Market Index is larger than the MSCI ACWI Index by 24.8% during the period January 2010 to March 2017; however, the overstatement of carbon footprint is larger: Emissions Scope 1, 2 and 3 as well as Reserves were overstated by 33.9%, 27.6%, 21.3% and 57.2%, respectively, in the same period.

Table 12. Average Carbon Volume Index Allocation Overestimation: 2010–2017.

Index Allocation Overestimation	Market Cap (\$M)	Emissions Scope 1 (CO ₂ Tons M)	Emissions Scope 2 (CO ₂ Tons M)	Emissions Scope 3 (CO ₂ Tons M)	Reserves (CO ₂ Tons M)
MSCI ACWI IMI Index	40,102,176	8252	1525	13,682	336,333
MSCI ACWI Index	32,126,859	6164	1195	11,275	213,970
Overestimation	24.8% ***	33.9% ***	27.6% ***	21.3% ***	57.2% ***

Source: MSCI. Note: *** *t*-test mean difference significance level at 99%.

Descriptive statistics for MSCI ACWI IMI and MSCI ACWI indices are in Tables 13 and 14 below.

Table 13. Descriptive Statistics for MSCI ACWI IMI Index.

MSCI ACWI IMI Index	Market Cap (\$M)	Emissions Scope 1 (CO ₂ Tons M)	Emissions Scope 2 (CO ₂ Tons M)	Emissions Scope 3 (CO ₂ Tons M)	Reserves (CO ₂ Tons M)
Observations	87	87	87	87	22
Mean	40,102,176	8252	1525	13,682	336,333
St. Deviation	5,533,205	815	211	21,674	15,875
Maximum	49,101,066	9079	1897	106,321	367,754
Minimum	28,900,125	6597	1135	571	320,093

Source: MSCI.

Table 14. Descriptive Statistics for MSCI ACWI Index Corrected for Overestimation.

MSCI ACWI Index	Market Cap (\$M)	Emissions Scope 1 (CO ₂ Tons M)	Emissions Scope 2 (CO ₂ Tons M)	Emissions Scope 3 (CO ₂ Tons M)	Reserves (CO ₂ Tons M)
Observations	87	87	87	87	22
Mean	32,126,859	6164	1195	11,275	213,970
St. Deviation	4,902,105	596	147	19,024	15,537
Maximum	39,962,384	6805	1450	93,924	237,744
Minimum	22,563,915	4982	925	431	196,226

Source: MSCI.

Tables 13 and 14 highlight the differences between original carbon footprints and corrected carbon footprints due to reassessment for overestimation. The means and variation data have been significantly reduced in the MSCI ACWI Index, after correction, relative to the MSCI ACWI IMI index. This makes sense, as the difference in market capitalization between the two indices is high (24.8%) and the corrections for overestimation were significant (33.9%, 27.6%, 21.3%, 57.2%) for all three Scopes and Reserves.

As an example, Adaro Energy shows a reduction of 45% of its stated carbon allocation when it is adjusted by its participation in the index (Table 15) for June 2016.

Table 15. Appropriate methodology to allocate carbon footprint for Adaro Energy.

Asset Name	Market Capitalization MSCI ACWI IMI (\$M)	Market Capitalization MSCI ACWI (\$M)	Carbon Emissions (Mt)	Carbon Reserves (Mt)	Portion of Market Capitalization in Index %	Emissions (Mt) Weighted	Carbon Reserves (Mt) Weighted
ADARO ENERGY TBK PT	\$2057.8	\$926.0	1.3	2197.3	45.0%	0.6	988.8

Source: MSCI June 2016.

Figure 2 shows that in the last 4 years of analysis (2014–2017), the overestimation in MSCI ACWI of Emissions Scope 2 converged to central values, while the overestimation of Emissions Scope 1 increased from 33% to about 39%. Conversely, the ACWI index stayed relatively constant relative to the ACWI IMI at around 81%.

The more recent upward change in overestimation of Scope 1 may be attributable to both the higher number of companies reporting their carbon volume and to a better quality of reporting. As there is increased pressure on companies to disclose their carbon footprint, more are compelled to report their Scope 1 emissions. Similarly, as carbon data becomes more standardized, companies have a better ability to provide a more accurate carbon footprint. However, for Scope 2, carbon footprint data is more stable. This may be because Scope 2 carbon footprint is an indirect measure of carbon footprints derived from the company's usage of electricity by an external energy provider. Therefore, this data is

never direct and always estimated, and the estimation techniques may not change in their accuracy.

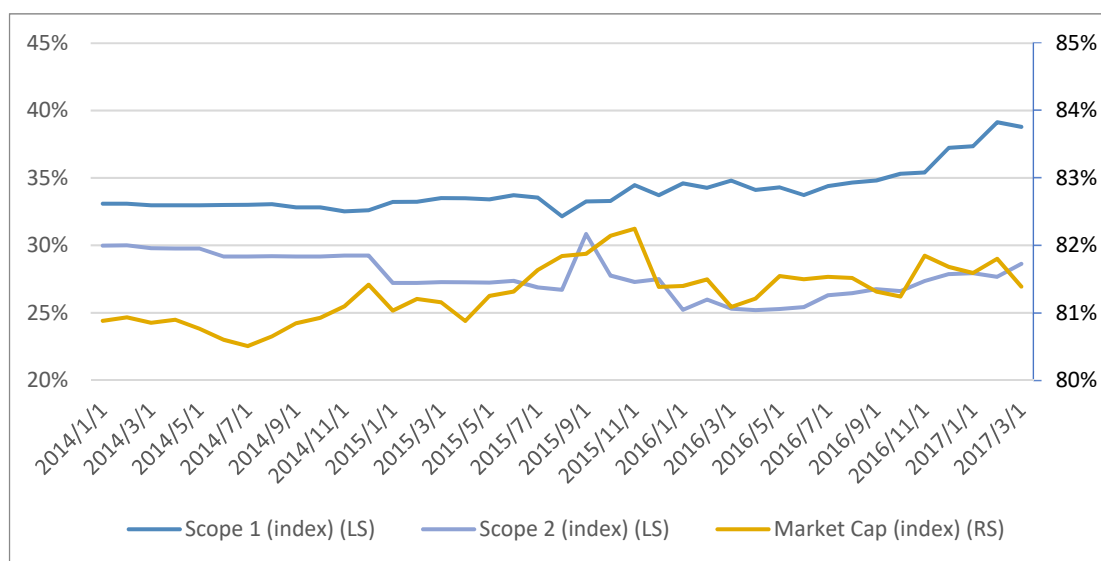


Figure 2. Overestimation of Emissions Scope 1 and 2 in the MSCI ACWI due to index allocation on left side axis. Market cap index is on the right side axis (2014–2017).

5.3. Correlations

The original, as well as the adjusted series of Emissions Scope 1 and 2 are highly correlated (0.89–0.9), while both series with Emissions Scope 3 and Reserves are smaller (Table 16). The lowest correlation of all the series with Emissions Scope 3 can be explained because of the difficulty in calculating this indicator, as it includes indirect emissions downstream of the value chain and those produced by the company's sold products.

Table 16. Correlations between Emission Scope 1 (S1), Scope 2 (S2), Scope 3 (S3) and Reserves (R).

	S1 Index	S2 Index	S3 Index	R Index	S1 List	S2 List	S3 List	R List	S1	S2	S3	R
S1 index	1.00											
S2 index	0.89	1.00										
S3 index	0.31	0.15	1.00									
R index	0.49	0.72	0.16	1.00								
S1 list	0.99	0.86	0.36	0.31	1.00							
S2 list	0.90	1.00	0.14	0.66	0.87	1.00						
S3 list	0.31	0.15	1.00	0.16	0.35	0.14	1.00					
R list	0.57	0.79	−0.08	0.93	0.45	0.74	−0.08	1.00				
S1	1.00	0.87	0.32	0.30	0.99	0.88	0.32	0.44	1.00			
S2	0.89	1.00	0.13	0.67	0.86	1.00	0.13	0.75	0.87	1.00		
S3	0.34	0.18	1.00	0.16	0.38	0.16	1.00	−0.08	0.35	0.15	1.00	
R	0.54	0.77	0.00	0.96	0.40	0.72	0.01	0.99	0.40	0.73	0.01	1.00

Note: “index” and “list” refer to adjustments due to index allocation and multiple listings, respectively.

MSCI's index construction poses a challenging issue to investors who need to assess a company's carbon volume output and link it to the company's valuation, as carbon volume is overallocated to companies and distorts data analysis. This paper recognizes this issue and appropriately allocates carbon volume to each company.

6. Discussion

Actual data reporting is an important consideration in understanding carbon data structure. Companies either report their carbon footprint or they do not, and there is no requirement for reporting. For example, for Scope 1, in a typical reporting month, only 50% of data is reported by companies, whereas estimated data is at 46% and the rest is neither reported nor estimated. The unreported data is estimated by the provider. This is significant because estimation techniques are specific to each data provider and the estimation methodologies are opaque.

Timing of data reporting is important, as there is a significant lag between a company's data report and distribution by the provider. This is because the provider requires data from the individual company or else scrape data from an external data source. It then needs to aggregate available data and estimate the data from the non-reporting companies. In a typical reporting month, the carbon data is three years old for 0.4% of the companies in the index, two years old for 94.1% of companies and one year old for 4.4% of the companies, with no data for the remaining 1.1% of companies. This creates another distortion in analyzing and attributing carbon footprints since the information is stale. The one redeeming feature is that such data is stale for most of the companies in the index.

A significant issue with carbon data attribution is a mismatch of a company's actual market capitalization relative to the capitalization included in the MSCI ACWI Index. The full market capitalization is included in the MSCI ACWI IMI Index, whereas only partial market capitalization is included in the MSCI ACWI Index. The MSCI ACWI Index is the most common investor benchmark; however, it includes only partial free float market capitalization of companies. This MSCI index construction technique overestimates carbon allocation at the company level and carbon volume must be adjusted accordingly as presented in the previous section. Since many investors hold or benchmark themselves to the MSCI ACWI Index, they overallocate the amount of carbon their investment portfolio holds.

7. Conclusions and Policy Implications

This research points to several carbon data distortions that may significantly overallocate carbon footprints to individual companies and proposes a methodology for corporate carbon allocation. These distortions and discrepancies may skew carbon analysis for either portfolio decarbonization or portfolio valuation purposes. Data distortion that stems from having the same company listed in multiple geographical jurisdictions or through different share classes overstates Emissions Scope 1, 2 and 3 and Reserves by 4.6%, 5.5%, 10.6% and 6.0%, respectively. While data distortion due to index construction by having different market capitalization in representative indices excessively allocates Emissions Scope 1, 2 and 3 and Reserves by 33.9%, 27.6%, 21.3% and 57.2%, respectively.

Portfolio managers who use quantitative screens to evaluate companies based on their carbon footprints to reduce carbon exposure in their portfolios should use screens based on raw carbon data, as opposed to carbon scores. Carbon scores are opaque, proprietary, and subject to estimation. In addition, using scores from different providers will produce different portfolio results, as scores are poorly correlated among different score providers.

Raw carbon data presents its challenges due to its structure. This paper illustrates several important issues with raw carbon data that distort and overallocate carbon footprints. Usage of carbon data in its present format may incorrectly remove some companies from portfolios (negative screen, complete removal) or incorrectly reduce some companies' weight in a portfolio (partial screen, fractional removal).

In conclusion, this research introduces novel methods of carbon allocation and a straightforward method to correct carbon footprint overallocation that can help researchers, regulators, and investors to identify the proper level of corporate carbon footprints.

Author Contributions: Conceptualization, Data curation, Investigation, Visualization, Writing—original draft: A.G.; Formal analysis, Methodology, Project administration, Validation: A.G. and G.G.C.; Writing—review & editing: G.G.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Restrictions apply to the availability of these data. Data was obtained from MSCI as a paid subscription.

Acknowledgments: The authors thank Suman Banerjee, Jonathan Kaufman, Emmanuel Hatzakis, and Dibyendu Sarkar for their comments in an early version of this paper, two anonymous referees, for their comments and suggestions, and Patrick Jardine for proof-reading the paper. The opinions presented are the exclusive responsibility of the authors.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Companies with Multiple Index Listings.

Asset Name	Listings
AMOREPACIFIC CORP PR	3
HYUNDAI MOTOR CO 2ND PRF	3
LIBERTY GLOBAL PLC	3
21ST CENTURY FOX	2
ALPHABET INC A	2
ALTICE A	2
AP MOLLER-MAERSK A	2
ATLAS COPCO AB-A FR	2
BANCO BRADESCO SA	2
BHP BILLITON LIMITED	2
BMW STA	2
CHOCOLADEFABRIKEN LINDT & SPRUENGLI AG	2
DISCOVERY COMMUNICATNS NEW	2
FORTRESS INCOME	2
GPO DE INVERSIONES SURAMERICANA SA	2
HENKEL STA	2
HEWLETT PACKARD ENTERPRISE CO	2
INTESA SANPAOLO N/C SA	2
INVESTEC	2
KASIKORNBANK (ALIEN)	2
LG CHEM LTD	2
LG HOUSE&HEALTHCARE	2
LIBERTY SIRIUS XM	2
LOJAS AMERICANAS ON	2
MONDI LTD	2
O2 CZECH REPUBLIC A.S	2
PETROBRAS ON	2
RELX	2
ROYAL DUTCH SHELL A	2
SAMSUNG ELECTRONICS	2
SCHIBSTED AKSJESELASKAP	2

Table A1. Cont.

Asset Name	Listings
SCHINDLER HOLDING PC	2
SIAM CEMENT CO THE	2
SURGUTNEFTGAZ (RUB)	2
SWATCH GROUP B	2
TATA MOTORS A	2
TELECOM ITALIA NEW	2
UNDER ARMOUR	2
VALE DO RIO DOCE PNA	2
VOLKSWAGEN STA	2

Source: MSCI June 2016.

References

1. MSCI CarbonMetrics Methodology and Definitions. 2014. Available online: <https://www.msci.com/index-carbon-footprint-metrics> (accessed on 20 October 2021).
2. MSCI Carbon Estimation Methodology. 2014. Available online: <https://www.msci.com/index-carbon-footprint-metrics> (accessed on 20 October 2021).
3. Abele, L. ESG as an Alpha Source. *Auriel Equity Investig.* **2016**. Available online: https://asiakas.kotisivukone.com/files/finsif.kotisivukone.com/helsinki_esg_integration_as_an_alpha_source_auriel_2015_05_06.pdf (accessed on 20 October 2021).
4. Görgen, M.; Jacob, A.; Nerlinger, M.; Riordan, R.; Rohleder, M.; Wilkens, M. Carbon Risk. 2019. Available online: <https://doi.org/10.2139/ssrn.2930897> (accessed on 20 October 2021).
5. Tan, K.S.; Wirjanto, T.; Fang, M. *Managing Climate and Carbon Risk in Investment Portfolios*; Society of Actuaries, 2018. Available online: <https://www.soa.org/globalassets/assets/Files/resources/research-report/2018/managing-climate-carbon-risk.pdf> (accessed on 20 October 2021).
6. Andersson, M.; Bolton, P.; Samama, F. Hedging Climate Risk. *Financial Anal. J.* **2016**, *72*, 13–32. [CrossRef]
7. Cheema-Fox, A.; LaPerla, B.; Serafeim, G.; Turkington, D.; Wang, H. *Decarbonization Factors*; Harvard Business School: Boston, MA, USA, 2019; pp. 20–37.
8. Hao, B.; Soe, A.M.; Tang, K. *Carbon Risk Integration in Factor Portfolios*; S&P Dow Jones Indices: New York, NY, USA, 2018; Volume 9, pp. 53–65.
9. Harris, J. The Emerging Importance of Carbon Emission-Intensities and Scope 3 (Supply Chain) Emissions in Equity Returns". *ET Index Spec. Rep.* **2015**. [CrossRef]
10. Harris, J. The Carbon Risk Factor (EMI-'Efficient Minus Intensive'). *ET Index Spec. Rep.* **2015**. [CrossRef]
11. Jacob, A. *Carbon Efficiency: A Strategic Look*; S&P Dow Jones Indices: New York, NY, USA, 2015.
12. Matsumura, E.M.; Prakash, R.; Vera-Munoz, S. Firm-Value Effects of Carbon Emissions and Carbon Disclosures. *Account. Rev.* **2014**, *89*, 695–724. [CrossRef]
13. Yongqing, L.; Eddie, I.; Jinghui, L. Impact of Carbon Emissions on Asset Values and Operating Cash Flows: Evidence from Australian Listed Companies. *J. Mod. Account. Auditing* **2013**, *9*, 94–111.
14. Gaspar, J.; Gurvich, A. *Low Carbon Solutions for Long-Term Portfolios*; Commonfund Whitepaper: Wilton, CT, USA, 2015.
15. Gurvich, A. Three Essays on Carbon Energy Finance and Fund Performance Screening. Ph.D. Dissertation, Stevens Institute of Technology, Hoboken, NJ, USA, 24 November 2020.
16. Derwall, J.; Guenster, N.; Bauer, R.; Koedijk, K. The Eco-Efficiency Premium Puzzle. *Financ. Anal. J.* **2005**, *61*, 2. [CrossRef]
17. Konar, S.; Cohen, M.A. Does the Market Value Environmental Performance? *Rev. Econ. Stat.* **2001**, *83*, 281–289. [CrossRef]
18. Heaps, T.; Balkissoon, K. Performance and Impact—Can Low Carbon Equity Portfolios Generate Healthier Financial Returns? 2014. Available online: <https://ssrn.com/abstract=2519803> (accessed on 20 October 2021).
19. Gurvich, A.; Creamer, G.G. Carbon Risk Factor Framework. 2021. Available online: <https://doi.org/10.2139/ssrn.3901956> (accessed on 20 October 2021).
20. MSCI Global Investable Market Indexes Methodology. *MSCI Ind. Methodol.* **2020**. Available online: <https://www.msci.com/index-methodology> (accessed on 20 October 2021).