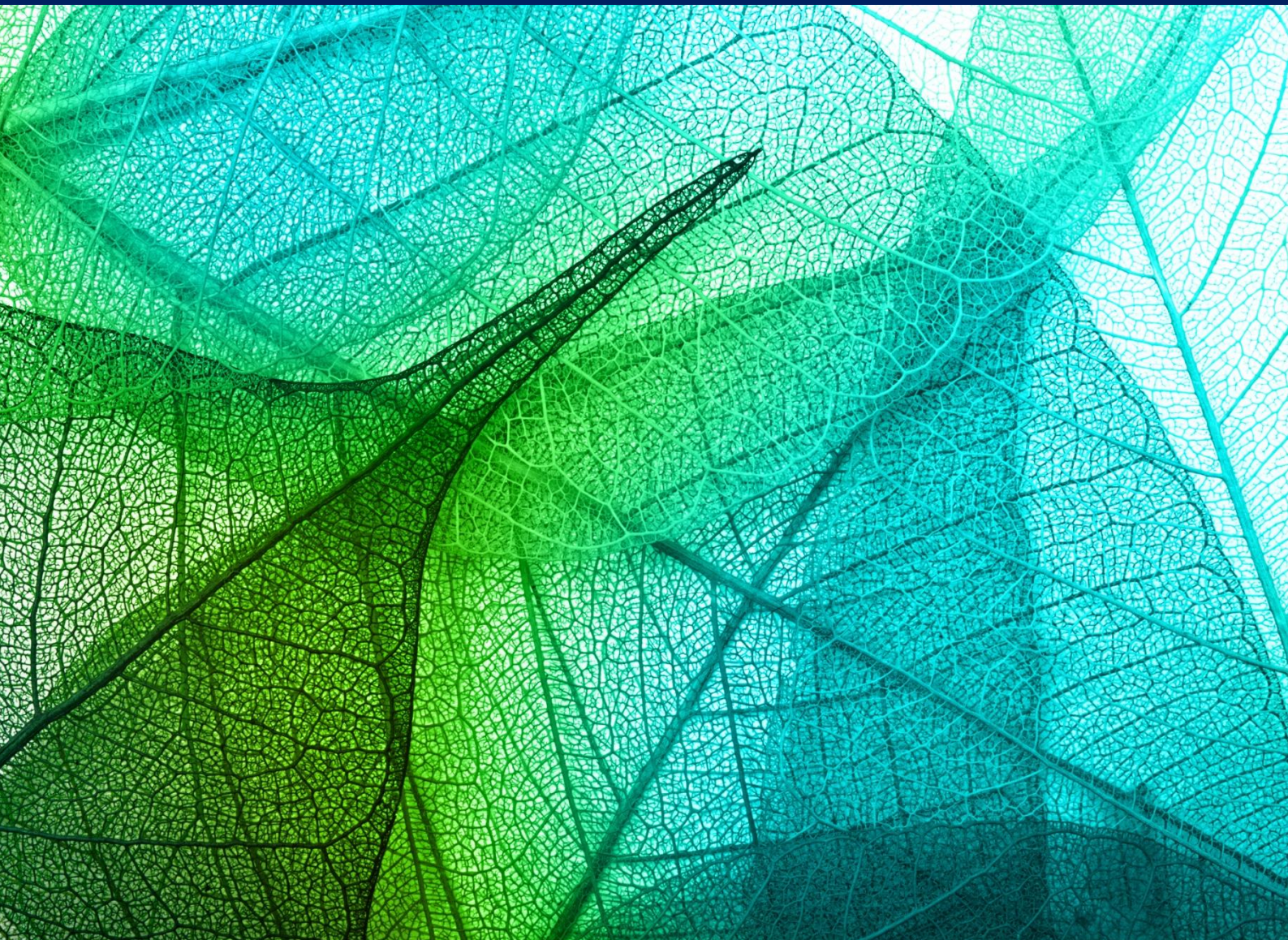# Data Quality Considerations for Estimating Financed Emissions

Discussion Paper, Sectoral Data Quality and Integrity Project

Authors: Kevin Tang, Gireesh Shrimali & Christophe Christiaen

6th January 2023

Oxford Sustainable Finance Group, Smith School of Enterprise and the Environment, University of Oxford

Aligning finance with sustainability is a necessary condition for tackling the environmental and social challenges facing humanity. It is also necessary for financial institutions and the broader financial system to manage the risks and capture the opportunities associated with the transition to global environmental sustainability.

The University of Oxford has world-leading researchers and research capabilities relevant to understanding these challenges and opportunities. The Oxford Sustainable Finance Group is the focal point for these activities and is situated in the University's Smith School of Enterprise and the Environment. The Group is multidisciplinary and works globally across asset classes, finance professions, and with different parts of the financial system. We are the largest such centre globally and are working to be the world's best place for research and teaching on sustainable finance and investment.

The Oxford Sustainable Finance Group is based in one of the world's great universities - the oldest university in the English-speaking world. We work with leading practitioners from across the investment chain (including actuaries, asset owners, asset managers, accountants, banks, data providers, investment consultants, lawyers, ratings agencies, stock exchanges), with firms and their managements, and with experts from a wide range of related subject areas (including finance, economics, management, geography, data science, anthropology, climate science, law, area studies, psychology) within the University of Oxford and beyond.

Since our foundation we have made significant and sustained contributions to the field, including in some of the following areas:

- Developing the concept of "stranded assets", now a core element of the theory and practice of sustainable finance.
- Improving the theory and practice of measuring environmental risks and impacts via new forms of geospatial data and analysis, including introducing the idea and importance of "spatial finance" and "asset-level data".
- Shaping the theory and practice of supervision as it relates to sustainability by working with the Bank of England, the central banks' and supervisors' Network for Greening the Financial System, and the US Commodity Futures Trading Commission, among others.
- Working with policymakers to design and implement policies to support sustainable finance, including through the UK's Green Finance Taskforce, Green Finance Strategy, the high-level Transition Plan Taskforce, and the Presidency of COP26.
- Nurturing the expansion of a rigorous academic community internationally by conceiving, founding, and co-chairing the Global Research Alliance for Sustainable Finance and Investment, an alliance of 30 global research universities promoting rigorous and impactful academic research on sustainable finance.

The Global Sustainable Finance Advisory Council that guides our work contains many of the key individuals and organisations working on sustainable finance. The Oxford Sustainable Finance Group's founding Director is Dr Ben Caldecott.

About the Sectoral Data Quality and Integrity Project

The Sectoral Data Quality and Integrity Project (SDQI), hosted by the Oxford Sustainable Finance Group and the UK Centre for Greening Finance and Investment, has been established to analyse and provide sector-specific insight into the quality of greenhouse gas emission datasets. SDQI will conduct empirical research comparing the quality of different available sectoral company and asset-level datasets. The project aims to develop practical sector-specific guidance and frameworks suitable for financial institutions to evaluate, apply, and communicate data quality. SDQI will tailor their inquiry to specific data use cases (including but not limited to: net zero target setting, financed emissions calculations, and risk management) and will aim to support the evolution and alignment of reporting and analytical frameworks. This discussion paper shares some initial observations, and outlines future areas of research for the project. The SDQI project started in 2022 and has been funded by the Wells Fargo Foundation.

Data Quality Considerations for Estimating Financed Emissions

Kevin Tang, Gireesh Shrimali & Christophe Christiaen

Oxford Sustainable Finance Group, Smith School of Enterprise and the Environment, University of Oxford

Abstract

Reporting standards for carbon accounting have been developed to assist financial institutions in compiling their financed emissions — the client emissions that are attributable to the institution's financing. Some of these include guidance on determining data quality for data sources used. For instance, the Partnership for Carbon Accounting Financials (PCAF) data quality hierarchy proposes reported and verified emissions disclosures as the most accurate and reliable source of emissions data, followed by various methods of inferred and estimated emissions. In this paper, we examine the extent to which guidance on data quality are aligned with the practical considerations and availability of emissions information for financial institutions. For financial institutions, an understanding of differences in data quality is particularly important, as they need to compile emissions data across multiple individual companies in their portfolio. However, the way these disclosures have been calculated by corporations can vary in methodology and accounting boundaries. This can result in several problems since variability in reporting methods or boundaries can amplify divergence in reporting errors, emissions estimation, and emissions that are attributable to the financial institution — ultimately impacting the way emission targets are set and progress or alignment monitored. We further explore revised guidance on assessing data quality, and the accuracy and precision of other methods of emissions estimation.

In this discussion paper we find that:

- Reporting standards' theoretical assessment and scoring of data quality does not match with the reality of availability, reliability, and confidence of the "highest quality" corporate reported emissions data.

- Financial institutions lack confidence in the standardized carbon accounting guidance that reported and verified disclosures are the most accurate and reliable source of emissions information.

- This has resulted in financial institutions relying on alternative data sources and methods of emissions estimation, inference, and data mixing to compile their financed emissions.

- However, this leads to non-standardization, inaccuracy, and inconsistency in emissions estimation and reporting between financial institutions, leading to concerns around comparability.
- We believe that "lower quality" granular asset and physical-activity data can be reasonably accurate and meet desirable criteria such as transparency, consistency, relevancy, completeness, and comparability.

Within our Sectoral Data Quality and Integrity project we propose to develop:

- Quantitative methods to assist financial institutions in evaluating data quality for the purposes of emission target setting, monitoring, and alignment, as well as in mixing data methods and sources, rather than relying on ad hoc approaches.
- Guidance to evaluate the importance of "comparability" as a data quality characteristic for emission target setting.
- Guidance on developing and interpreting asset-level emissions information, including on emission factors and activity levels.

Table of Contents

1 Estimating financed emissions: Data sources and data quality frameworks

In recent years financial institutions have launched several initiatives to support climate action and reduce their financed emissions by establishing high-level commitments, conducting climate scenario analysis, and setting emissions reduction targets. In order for these goals to be achieved, they need to know the attribution and links of their financial assets to specific emissions activity, which is primarily through standardized and regular reported emission disclosures by corporations.

However only a small share of corporations, operators, and producers comply with directly measured emissions reporting, with the majority of corporations not reporting at all.[1] In addition, even for the relatively small share of companies that do make voluntary emissions disclosures, the disclosed emissions that have been shown do not always or consistently reflect the company's breakdown of their own emitting activities.[2] Consequently financial institutions must rely on various methods of estimation and inference to compile emission data of corporations in each sector. Estimated or inferred emissions are derived from several different methodologies, including measures based on asset-level, production-level, or economic activity data. This set of statistical information is the basis for the compilation of emissions data, observed at different levels of granularity such as at the product, organizational, sectoral, national, regional, or supply chain level, which is subsequently used by financial institutions. Based on what data is available, what is directly observed, and how it is compiled, financial institutions have a choice in how their financed emissions are calculated and reported, and subsequently how they set and monitor their financed emissions alignment and targets.

The disparity in how financial institutions choose to measure, compile, and track the various set of data sources and methods to derive emissions data results in wide-ranging, inconsistent, unreliable, and largely subjective approaches.[3] Several frameworks for

---

[1] In 2018, Trucost reported on 8,446 companies with 15.94 per cent of those companies, or 1,346, having made voluntary carbon emissions disclosures; Patrick Bolton, Stefan Reichelstein, Marcin T. Kacperczyk, Christian Leuz, Gaizka Ormazabal, Dirk Schoenmaker. Mandatory corporate carbon disclosures and the path to net zero. *Management and Business Review* 1, no. 3 (2021), 21 – 28.

[2] Sergio Garcia Vega, Andreas G.F. Hoepner, Joeri Rogelj, and Frank Schiemann. Carbon disclosure quality: oil and gas. *University College Dublin Graduate Business School.*

[3] Malgorzata Paulina Olesieqicz, Jaakko Kooroshy, Sonja Greven. Navigating the corporate disclosure gap: Modelling of Missing not at Random C Data. *Global Research Alliance for Sustainable Finance and Investment Conference,* Beijing, China (2021).

standardizing this process have been developed in the practitioner-oriented literature to address the discrepancies in how derived emissions statistics are reported and measured amongst the variety of direct, indirect or inferred data sources. One widely adopted framework is the Partnership for Carbon Accounting Financials (PCAF), which has "310 financial institutions (…) committed to measure and disclose the greenhouse gas emissions associated with their portfolio of loans and investments", representing USD 80.0 trillion total financial assets. Ultimately the purpose of PCAF is to increase the confidence in the use of financed emissions data, allowing financial institutions to make commitments, assess climate risks, and set emissions targets more effectively.[4] One issue that PCAF seeks to address in their accounting and reporting standard, is the vast disparity of sources and methods from which the ultimate emission data observations and estimates are collected and used by financial institutions, with the questions and concerns of the quality and reliability of the data itself in the face of an inability to verify it against any known or observed "true value".[5]

Financial institutions draw from different data sources, including directly reported company-level or asset-level emissions data, or derived emissions observations from statistical information on production activity, and information on economic activity. While there are trade-offs associated with each data source, PCAF has developed guidance in the form of a data quality hierarchy to help financial institutions select data sources based on "data quality" when calculating financed emissions for different asset classes. The hierarchy assigns scores to different data sources, with the directly observed and reported emissions from the financial asset itself perceived as "highest data quality",. and estimation from various types of proxy data perceived as "lower" quality. The overall objective of these types of data quality frameworks are, as stated by the FTSE, to "improve accuracy, and reduce the risk of underestimating emissions, which has attracted greater scrutiny as the transition accelerates."[6] More generally, data quality frameworks are intended to be an objective assessment to guide financial institutions on which data is best suited for emissions estimation, irrespective of the use case for which the data is being compiled.[7]

---

[4] Ibid., 12.

[5] Global GHG Accounting and Reporting Standard for the Financial Industry. Partnership for Carbon Accounting and Reporting Standard for the Financial Industry. First Edition (2020), 15.

[6] John Simmons, Jakko Kooroshy, Edmund Bourne, Mallika Jain, Lee Clements. Mind the gaps: Clarifying corporate carbon. FTSE (May, 2022), 4.

[7] The Greenhouse Gas Protocol: A Corporate Accounting and Reporting Standard. World Business Council for Sustainable Development: World Resources Institute (2020), 4.

Guidance on the assessment and evaluation of data quality is a key concern for financial institutions, as they are forced to combine emissions data across their portfolios where companies do not always disclose their emissions, and those that do, provide their own calculation of emissions that may not be consistent or comparable with other companies in the portfolio. The lack of consistency and comparability in data sources, methods, and metrics for compiling emissions data within a sector have been widely cited as the most significant issue in a survey of financial institutions conducted by RMI.[8] As a result of the lack of consistency and comparability in emissions data, the potential for miscalculation and misattribution can significantly affect an institution's target setting, monitoring, and alignment activities. However, the application of data quality frameworks that are generalised across sectors and use cases introduces several issues, as they prioritise certain data quality characteristics over others, the reasoning and methodology for which have not been explicitly addressed in the literature, or in the development of these methodologies.

Consequently, this can lead financial institutions that implement standardized data quality guidance to rely on lower quality or unreliable data sources in the compilation of their attributable emissions, since such guidance does not take into consideration the use case for which emissions data is being compiled. Additionally, this can lead to a lack of confidence in the data quality framework, forcing institutions to develop their own framework of data quality or mix different data sources and methodologies in the compilation of attributable financed emissions. In either case, institutions have had to rely more on subjective, qualitative, and improvised measures in the determination of data quality which lacks standardization or comparability between institutions and over time. Therefore, since financial institutions face a unique set of challenges in the compilation of emissions data across companies within a sector, we believe a comprehensive assessment of data quality used for estimating financed emission is needed.

In this discussion paper, we present issues from practice on the universal application of data quality hierarchies, the margin of error in inferred emissions and mixed method estimation, and we discuss quantitative methods that can be implemented to more systematically address financial institutions' selection of data sources and methods for deriving attributable financed emissions data. As financial institutions are faced with the unique challenge of aggregating various and divergent emissions data and reporting across

---

[8] Kaitlin Crouch-Hess, Elizabeth Harnett, Lila Holzman, Eero Kekki, Alex Murray. Identification, access, and use of transition-relevant data and metrics. *RMI Centre for Climate Aligned Finance Insight Brief* (September, 2022).

companies for the purpose of building credible portfolios, engaging with clients to support their transition, and meeting their own climate goals, we assess data quality measures required by financial institutions for the purposes of conducting sector-specific target setting, alignment, and monitoring.

2   Data quality principles: Conceptual issues in standardized frameworks

PCAF defines data quality, in accordance with the GHG protocol, based on five accounting principles: completeness, consistency, relevance, accuracy, and transparency. While these data quality principles are generally applicable and desirable in all cases of carbon accounting, when practically applied, they frequently become subjective based on the specific circumstances of the financial institution, their portfolio, and the availability of data. The subjectivity in how each financial institution applies the PCAF accounting principles, and subsequently how emissions data is compiled, can lead to contradictory guidance in contrast to the intended standardization of carbon accounting principles for calculating financed emissions.

For example, the application of three of the PCAF principles, in this case, 'consistency", "accuracy", and "transparency", used to assess the sources and methods for calculating emissions data, shows how competing priorities can result in either neglecting other key traits or conveying contradictory guidance, either of which can consequently lead to less confidence and reliability in the emissions data.

"Consistency" refers to the need to "use consistent methodologies to allow for meaningful performance tracking of emissions over time" [and] "transparently document any changes to the data, inventory boundary, methods, or any other relevant factors in the time series." [9] However this definition can lead to accounting problems as it is focused on consistency over time, but does not include consistency across or between entities and sectors, to ensure that data reporting is comparable, a data quality requirement for financial institutions that need to compile emissions statistics between reporting and non-reporting corporations in their portfolio. While the GHG protocol and PCAF accounting principles include consistency in reporting as one of the main priorities, what is neglected in this definition is ensuring the importance of "comparability" through standardization of reporting inventory boundaries or accounting methodologies. Comparability of entities can be considered the

---

[9] "Global GHG Accounting and Reporting Standard for the Financial Industry". PCAF (2020), 34.

principal reason for the development of accounting standards.[10] The lack of comparability between reporting entities can introduce several other issues in the trust and utilization of reported data on its own but can also lead to lack of confidence in reported data as it relates to the application of the other GHG protocols and PCAF accounting principles of accuracy and transparency.

PCAF has defined accuracy as, "the quantification of GHG emissions that is systematically neither over nor under actual emissions." While the definition of "actual emissions" in this case is never explicitly discussed, the data quality hierarchy is intended to prioritise the data sources that will come closest to representing "actual emissions". To provide guidance on how firms should follow the principle of accuracy, PCAF refers to the reporting requirement of "data quality", which is premised on the data scoring hierarchy. The "data quality" requirement is that institutions should, "use the highest quality data available at each asset class and improve the quality of the data over time" in accordance with the data quality hierarchy.[11]

Finally, transparency is defined by PCAF as addressing "all relevant issues in a factual and coherent manner, based on a clear audit trail. Disclose any relevant assumptions and make appropriate references to the accounting and calculation methodologies and data sources used."[12]

PCAF has developed several data quality hierarchies to help financial institutions align with the principles set out in the GHG Protocol across different asset classes based on "generally accepted principles to ensure that an organization's disclosure represents an accurate, veritable, and fair account of its GHG emissions."[13] In general, PCAF guidance on data quality for bank lending has prioritized reported and verified emissions from company disclosures as the highest quality with a score of 1.[14] The second and third highest ranked data is for estimated emissions inferred through physical activity. This is based on taking observed energy consumption, score 2, or volume of production from an asset, score 3, then using a verified emissions factor to subsequently infer the emissions. PCAF score 4 and 5 are estimates based on economic activity data, taken from the company's revenues, score 4, or

---

[10] Barlev, B. & Haddad, J. R. Harmonization, "Comparability, and Fair Value Accounting". *Journal of Accounting, Auditing & Finance* 22, (2007), 493–509 .

[11] Ibid., 34.

[12] Ibid., 34.

[13] Ibid., 33.

[14] Global GHG Accounting and Reporting Standard for the Financial Industry. PCAF (2020), 66.

the company's asset turnover ratios, score 5, then deriving emissions estimates from assumptions on emissions factors for the sector per unit of revenue.

Based on the definition of these accounting principles and guidance as determined by PCAF, if a financial institution were to strive for the principles of "consistency", "accuracy", and "transparency" through selecting the highest quality data, then they would rely entirely on reported emissions data from companies in their portfolio**.** Thus, not considering any changes in inventory boundaries or accounting methodology by the reporting entity, foregoing "comparability".

For instance, if the inventory boundary changes from one year to the next, PCAF guidance would still encourage the use of the highest-quality reported data according to the hierarchy, while tracking changes in the boundary. However, in some cases, when the inventory boundary changes, not all and not the same emissions information is included, forcing the financial institution to make a subjective decision on how to account for the change in the inventory boundary. The institution is subsequently faced with a set of choices, neither of which satisfies all three accounting principles. They are forced to prioritise one principle over another, hence with any choice, the institution is left with less confidence in the final emissions compilation.

Financial institutions can continue to use the reported data since it is the highest quality, and "document any changes to the data". However, this still leaves the institution with a lack of trust and confidence in the accuracy of the emissions estimation since there are clear and known changes to the reporting that are inconsistently measured over time.

Alternatively, from the year the inventory boundary changes, they can replace the reported data with inferred statistics from different sources, resulting in a mixed approach to compiling emissions data. This would be aligned with guidance from other frameworks to financed emissions compilation, such as the FTSE multi-model approach to estimating emissions, but is not consistent with PCAF.[15] There are still other frameworks that could be utilised by the financial institution in these cases, however in doing so, the data is no longer consistent or transparent since it does not "use consistent methodologies to allow for meaningful performance tracking of emissions over time" and the choices made in mixing methods and sources for a given year or sector are subjectively decided, and hence do not, "make appropriate references to the accounting and calculation methodologies and data sources used" in deciding on the data source mix.

---

[15] John Simmons et al, op. cit.

Consequently, since PCAF's data quality hierarchy and principles provide general tacit guidance for financial institutions, rather than a clear analytical or evaluation process to determining data quality, the application of general guidance for institutions that rely on PCAF cannot be consistently applied in all cases of emissions reporting or compilation while ensuring . may try to follow general PCAF guidance, when it comes to target setting, alignment, and monitoring, the need to ensure trust and confidence in emissions compilation for specific sectors and across corporations leaves institutions without a sufficient or applicable data quality framework. Therefore, without guidance on how a financial institution should choose or adapt the method or source of reported emissions data with inferred data, the compilation of statistics becomes highly subjective to the methodological choices of the institution, thus not becoming a standardized method that is transparent or consistent over time. Hence, the PCAF methodology can convey contradictory guidance where the use of consistent, accurate data and transparent methodologies cannot be satisfied, forcing subjective, non-standardized, and non-transparent choices on the part of the financial institution in conducting emissions estimation. In either case, the result is less confidence and trust in the reliability of the data and the PCAF methodology.

3 Discrepancies in the implementation of data quality frameworks: Evidence from financed emissions reporting

The contradictory guidance that can arise in attempting to maintain the PCAF principles of consistency, accuracy, and transparency can be understood when there are changes in the boundary of the reported emissions data. This is a not uncommon occurrence, as demonstrated in the following example from the Southern Company. This is a primary US energy utility company serving nine million customers across several US states, with 43,000 megawatts (MW) of electricity generating capacity, and a commitment to net zero GHG emissions by 2050, with an interim 2030 target of 50% emissions reduction from 2007 levels.[16] An overview of the company's self-reported emissions disclosure by scope is illustrated in table 1.

From the table, this company shows effective progress towards GHG emissions reductions between 2017 to 2020. For any institution that is calculating their financed emissions from the Southern Company, the directly reported emissions data shows progress

---

[16] The Southern Company CDP Climate Change Questionnaire. (August, 2022), 1.

towards their goal of achieving net zero by 2050. As the Southern Company has self-reported and verified emissions information, this would qualify under the PCAF data scoring hierarchy as the most accurate, with a PCAF score of 1, and should be the preferred data source according to the GHG Protocol and PCAF guidelines. Additionally, since this type of self-reported data has been consistently reported each year, it also satisfies the carbon accounting principle of consistency. Therefore, the emissions statistics summarized in table 1 should be the highest quality and should be the data used by financial institutions in emissions calculation.

TYPE	2017	2018	2019	2020	UNIT	DATA QUALITY
REDUCTION IN GHG EMISSIONS SINCE 2007	35	34	44	52	Percent	
SCOPE 1	101.2	102.6	88.2	75.1	Million metric tons CO2e	Third-party verification or assurance process in place
SCOPE 2	-	-	0.2	0.2	Million metric tons CO2e	Third-party verification or assurance process in place
SCOPE 3	37.0	39.4	38.8	36.6	Million metric tons CO2e	No third-party verification

Table 1: The Southern Company, enterprise-wide energy use and emissions[17]

However, taking a closer look at the Southern Company's reporting methods indicates changes in the reporting boundary, thus bringing into question a financial institution's reliance and trust in the self-reported, verified, and consistently observed data as the highest-quality data source. Table 1 shows reduction in scope 1 emissions over time, with no measurable changes in scope 2 and scope 3 emissions, and the data under scope 2 emissions omitted for

---

[17] Data has been collected across reports from the Southern Company and is based on its own summary of annual emissions. Scope 2 emissions for 2017 and 2018 are not included because they were not reported by the company and did not undergo "third-party verification or assurance process." The Southern Company, ESG Data Table (September 2021), 2.

2017 and 2018. Closer inspection of the company's self-reported data — particularly in the omission of company reporting under scope 2 emissions in 2017 and 2018 — suggests that the increase in scope 3 emissions in 2018 was potentially due to the result of a change in the reporting boundary, as the company states:

> Scope 2 emissions were previously calculated and reported as emissions from generation of purchased power to serve our customer load (2,142,130 metric tons CO_2e in 2018). However, scope 2 emissions are now reported as emissions from the electricity purchased for our own company use (35,568 metric tons CO_2e in 2019); and purchased power to serve our customer load is now reported as scope 3 emissions.[18]

Based on this report, the reallocation of over 2 million metric tons CO_2e from scope 2 to scope 3 emissions due to changes in the reporting boundary accounts for the vast majority of the Southern Company's entire scope 2 emissions, as these in 2019 amounted to only 0.2 million metric tons from the time the boundary change was implemented.

In circumstances such as the one illustrated by the Southern Company, PCAF and GHG Protocol do not offer clear guidance on how financial institutions should treat changes in the reporting boundary. According to the principle of accuracy, since the Southern Company's emissions data is both self-reported and verified, this should be of the highest quality. However, for the purposes of estimating their financed emissions and tracking progress on climate action in their portfolio, the change in reporting boundary for the Southern Company does not effectively capture the company's emissions, hence it introduces a lack of confidence and trust in the PCAF data quality hierarchy, the principles of GHG protocol, and the assessment of their own financed emissions.

Similar issues concerning the inability of financial institutions to effectively implement the PCAF data quality hierarchy and satisfy the GHG Protocol principles have been suggested from other emissions reporting. As a member of PCAF, Barclays' carbon footprint reporting is generally aligned to PCAF guidelines. However, for the purposes of calculating and setting climate targets, and measuring progress and alignment, Barclays has developed its own methodology called "BlueTrack", which is a mixed-method and mixed-source approach to calculating financed emissions. Part of the reason it has done so, despite being a member of PCAF, is surmised to be because of lack of trust and confidence in the data quality hierarchy,

---

[18] The Southern Company CDP Climate Disclosure (2020), 37.

and consequently an inability to apply the guidelines to their financed emission calculation. It says that "company-level disclosure has improved significantly in recent years … nonetheless, we continue to find that the data is still not sufficiently robust to be used as the primary source for BlueTrack. Because of this, the methodology estimates company emissions by combining external fossil-fuel production databases with assumptions about emission factors."[19]

While PCAF's data quality hierarchy suggests that company-level disclosures are the most accurate, and should be preferred over other sources, Barclays has suggested it does not have sufficient confidence or trust in these disclosures, either as the most accurate, or the most fit-for-purpose data source. Instead, they cite an alternate methodology that draws from data that is not as highly scored on the PCAF data quality hierarchy as company-level reports, and they use a mixed-method approach. As a result, the Barclays method is not fully transparent, since the combination of databases with assumptions on emission factors is not known, nor is it standardized, since the methodology could not be replicated by other financial institutions, and is highly subjective, as the choice of external fossil-fuel production databases are not methodologically assessed, and the assumptions on emission factors are subjectively selected. While the methodology may have internal consistency for Barclays, it is not consistent nor standardized with other institutions.

Similar issues have been raised by Wells Fargo. Like Barclays, Wells Fargo has found inconsistencies in company reported emissions data, and an inability to adhere to the PCAF and GHG Protocol guidelines premised on data quality and accuracy. It says that "while the quality of voluntary emissions reporting has dramatically improved in recent years, there remains great variability in reporting both across and within sectors. Some inventories may be audited and generally include all emissions scopes, (Scopes 1, 2, and 3) while others may be unverified, incomplete, or both."[20] Although Wells Fargo is not a member of PCAF, the bank's lack of confidence in emissions reporting as a result of the variability in company disclosures would make it in any case unable to follow PCAF guidance. Instead, Wells Fargo has developed its own mixed method approach that does not prioritize company-level reported data. Its methodology "does not take a 'one size fits all approach'" preferring company-reported data to third-party production data or vice versa. Rather, we evaluate the data sources for each

---

[19] About BlueTrack: An update on our methodology for reducing our financed emissions. Barclays PLC, 2.B.1 (March, 2022), 4.

[20] CO2eMission: Net-zero alignment and target-setting methodology. Wells Fargo Bank (May, 2022), 14 – 15.

sector separately."[21] In its mixed-method approach, Wells Fargo has highlighted key issues in data quality frameworks that are overlooked by the universal framework of PCAF, which are that universal standards cannot be applied to all use cases of carbon accounting, and that such a standard cannot be applied across sectors.

While Wells Fargo's methodology addresses key issues in carbon accounting that have been similarly observed by other financial institutions by adapting data quality frameworks to the particular use case and sector, the primary problem with these approaches is that, while they may be internally consistent, the methodology cannot be compared to other banks, is not replicable, and is not transparent. Similar to the Barclays methodology, the mixed method approach of Wells Fargo strays from PCAF and GHG Protocol guidance as it is not consistent, since data sources are essentially evaluated on a sector-by-sector basis. While the sector-based approach to emissions estimation taken by Wells Fargo highlights some of the key issues in PCAF's universal data quality hierarchy irrespective of the particular data issues for each sector, the use of alternate methods in place of PCAF's standardized method lacks comparability and consistency between financial institutions. Internal methods may increase trust and confidence in estimated emissions for the particular institution, but the resulting financed emissions are not comparable with other institutions as they rely on a set of assumptions that are not disclosed, and employ a methodology for selecting data sources that is highly subjective.

4 Analysis of data availability for PCAF disclosures: Low quality data scores and incomplete coverage

Difficulties with following PCAF guidelines based on the data quality hierarchy are also suggested from the disclosure of PCAF data quality scores used to calculate financed emissions by PCAF partner institutions. In contrast to banks such as Barclays and Wells Fargo that have developed their own methodology for emissions estimation, several institutions have followed the PCAF guidance. In figure 1, we have compiled the average frequency of each PCAF data quality score used for a sample of 28 institutions that have made public disclosures of their financed emissions assessments in partnership with PCAF. Similar to Barclays and Wells Fargo, no institution in figure 1 has been able to compile emissions estimates entirely from directly reported company-level emissions data. Instead, half of all firms used the lowest

---

[21] Ibid., 15.

quality data score 5, which is for emissions inferred from economic activity. These findings are similar to a study by the 2 Degrees Investing Initiative (2DII) that has also looked at public disclosures of financial institutions that are partnered and aligned with PCAF. This found a mix in reporting results among 70 financial institutions that have publicly disclosed financed emissions, with 30 firms reporting unknown data quality, and less than 30 per cent reaching a data quality score of 3 or higher.[22] The weighted averages of the data quality that institutions were able to use to compile emissions estimates show that no firms in the sample were able to achieve an average data quality at the highest score, and that averages tended toward the lowest quality. In addition to the lack of trust and confidence cited by Wells Fargo and Barclays in using the highest quality data with a PCAF score 1, evidence from figure 1 suggests a wide discrepancy between the PCAF guidelines and the realities of data availability. PCAF has instructed financial institutions to develop a data waterfall to optimize as low a score as possible based on data availability: "high quality data is often not available to the financial institution for all asset classes. In these instances, the institution should use the best available data in accordance with the data hierarchy."[23] Hence, since the lowest quality is the most widely used, this suggests a discrepancy between the PCAF guidelines, and what is available in practice to financial institutions when compiling financed emissions.

---
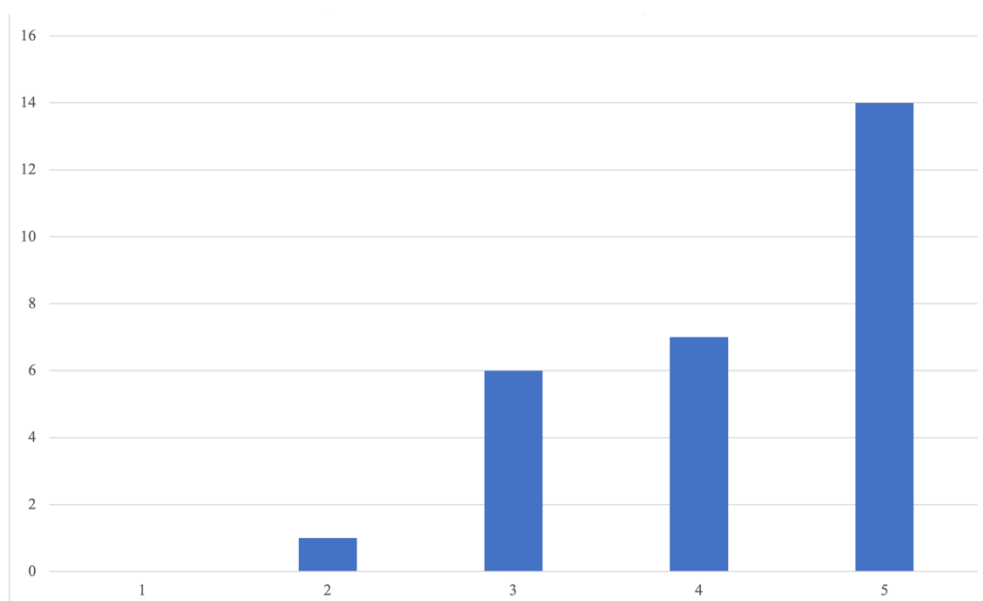
[22] Jakob Thoma. 0% of PCAF signatories comply with PCAF reporting requirements. 1in1000 program of the 2 degree Investing Initiative.

[23] Global GHG Accounting and Reporting Standard for the Financial Industry. PCAF (2020), 39.

Figure 1: PCAF data quality score disclosures from select financial institutions[24]

This is further illustrated in figure 2, which shows the percentage of emissions by asset class on the balance sheet that banks were able to account for using the PCAF methodology against the data quality score used, with the size of the nodes corresponding to the amount of emissions covered. Data in figure 2 is taken from a subset of the banks from figure 1 where public disclosures also include information on the PCAF score for the specific asset class. From the figure, there was only one case in which a firm was able to achieve accurate, high-quality data and near completeness in accounting for emissions in their portfolio. This was the case only for a portfolio of motor vehicle loans, which has a unique type of data framework and guidelines outlined by PCAF.[25] Otherwise, firms in the top right quadrant were able to have completeness in their carbon accounting with their portfolio coverage at more than 75 per cent, but using low quality data. There were still several firms that did not perform well on either of the two carbon accounting principles of accurate data and completeness, as can be observed from the figure from the firms in the bottom right quadrant.

---

[24] The selected 28 firms include a global cross-section of commercial banks, asset management companies, and other financial institutions that have made public disclosures on their financed emissions using the PCAF methodology. Banks that have been previously discussed, including Wells Fargo and Barclays, did not use the PCAF methodology for their emissions estimation, so are not included in figure 1.

[25] PCAF guidance for motor vehicles has a unique framework compared to other types of commercial lending, since emissions compilation is based on the specific make and model of the motor vehicle, information that is more widely accessible and reported by car manufacturers. Motor vehicle loans emission factors: Database methodology. *Partnership for Carbon Accounting and Financials* (April 2020).

Figure 2: PCAF data quality scores against percentage of balance sheet accounted[26]

The overall finding showing that lower data quality scores are used to compile incomplete portfolios of financed emissions has been similarly acknowledged by banks. For instance, HSBC follows PCAF's data quality hierarchy in its financed emissions estimation, but has cited the need to revert to lower data quality to compile portfolio emissions "not all companies in our oil and gas, and power and utilities portfolio report emissions publicly… of these exposures only approximately 1/3 of clients by exposure report scope 1 and 2 emissions data. In addition, scope 3 emissions data is only available for approximately 10% of our oil and gas portfolio. This means that we need to use proxies to estimate emissions when reported emissions data is unavailable."[27] Hence, the majority of HSBC's financed emissions are derived from various forms of estimation, with the final portfolio-level emissions being a mix of methodologies**.**

---

[26] HSBC is not included in the figure since—although it has stated that only 1/3 of clients directly report scope 1 and 2 emissions data, and only 10% of scope 3 emissions are reported in its oil and gas portfolio—rather than state the data quality score that was used for estimating the rest of its portfolio emissions it says only that it is a combination of physical and economic activity data, in line with the PCAF data quality hierarchy.

[27] Financed emissions methodology. *HSBC Holdings PLC* (2022), 11.

There are potential estimation errors that can arise simply from the lack of consistency in mixing data sources within a portfolio, however these errors can be further propagated as the process of mixing data sources and quality is conducted irrespective of sector-specific considerations. While banks such as Barclays and Wells Fargo have developed their own approach to sector-specific emissions estimation that is not aligned with PCAF's guidelines, HSBC claims that its methodology generally aligns with PCAF by calculating emissions "based on reported emissions, physical activity-based emissions, or economic activity-based emissions, and data is prioritised according to source and robustness."[28] However, it still needed to rely on more subjective, qualitative, inconsistent, and non-transparent assessments of data sources and methods, which it has not disclosed, and for which PCAF has not provided guidance.

Evidence illustrated here suggests that there is a wide discrepancy between how PCAF guidelines are followed, and the practical issues facing financial institutions in estimating financed emissions. The potential for such contradictory guidance and subjective choices made by the financial institutions has been illustrated from the frequent use of low-quality data, incomplete coverage in portfolio emissions, and firms such as Barclays and Wells Fargo developing their own carbon accounting methodology. While several financial institutions have developed their own methodologies for setting and measuring alignment targets, the need to do so has generally been explained as a consequence of the lack of confidence and trust in reported and verified data. Although institutions that develop their own methodologies may attempt to align with PCAF's data quality hierarchy even if they are not able to entirely adopt it, others may still diverge and develop their own assessment of data quality based on characteristics that are different from PCAF and based on the particular profile or needs of the bank. However, in either case, while this may address the bank's specific needs, and may be internally consistent, these methodologies are not entirely transparent and cannot be reproduced. This can potentially lead to inconsistencies and lack of comparability in financed emissions over time, thus having the opposite effect of standardizing reporting methodologies.

5   Implications of low quality and low granularity data: Increased variability in financed emission estimations

---

[28] Ibid.

The inability of financial institutions to effectively apply PCAF guidance and principles in their carbon accounting is not just an issue of consistency in reporting disclosures, or transparency and standardization in methodology, it is also an issue of accuracy.

First, as reported from figures 1 and 2, the reliance on the PCAF data hierarchy has resulted in financial institutions relying on low scored data, which can have a wide margin of error. A study by FTSE has shown that data categorised with approximately a score of 3 or lower according to PCAF methodology, which encompasses a wide range of inferred methods including estimation based on total revenue or number of assets by sector and broad geographic averages, has an error margin of more than +/- 200% relative to reported data, thus significantly lacking any accuracy, leading to a lack of confidence and trust in any firm's financed emissions estimates.[29] Additional studies support this wide margin of error, where 70% of PCAF disclosures are based on a data quality score of 3 or lower: emissions compilation at this level "effectively provides no meaningful insight into lending and investment strategies of the disclosing entities beyond their regional focus, and their sectoral split."[30]

Second, in addition to a wide margin of error for low PCAF scored data, a study by Bain & Company finds that different levels of granularity for methods of inferred emissions can lead to large differences in the calculation of financed emissions in the portfolio. Compiling a representative dataset of emissions in different sectors and levels of granularity, Bain & Company finds that the level of granularity in emissions estimation can lead to large over- and under- estimation of emissions depending on the sector.[31] In the case of an electric power portfolio, estimated emissions from data that is aggregated to a high-level, such as large geographic or sectoral averages, can overestimate emissions by 90 per cent compared to granular data that is more specific to a country, region, or sector. For a metallurgy portfolio, high-level emissions estimation underestimate emissions by 120 per cent relative to granular

---

[29] The study finds that nearly half of all methods of estimated emissions diverged from the reported data by 100%, and over a quarter of the estimated data, primarily methods of estimation at the lower end of the PCAF data quality spectrum, were off by at least 200%. Analysis of accuracy in methods of emission compilation compares emissions data from companies directly reported to FTSE All World Index against inferred emissions compiled via the FTSE Russell's estimation model. The FTSE Russell model incorporates several different methods based on the type of data that is available, including extrapolation from historical reported emissions data, inferences from energy production, sectoral-regional median, input-output strategy, inverse distance weighted interpolation, and carbon intensity regression. John Simmons et al, op. cit.

[30] Jakob Thoma, op. cit..

[31] Camille Goossens, Christian Graf, Avishek Nandy, Michael Kochan, Rocco D'Acunto & Amira Boualla. Bank's great carbon challenge. *Bain & Company Brief* (06 June 2022).

data.[32] Such differences are observed even when the portfolio emissions are compiled from intermediate data quality, rather than most or least granular, corresponding to approximately a PCAF score of 2 or 3. The study highlights not only the wide-ranging estimates that can result from applying various methods of inferred emissions at different levels of data quality or granularity, but also the wide range of sectoral variation at each level of granularity.

As a result of the wide margin of error when using lower quality data, the wide discrepancy in portfolio emissions depending on the level of granularity in the data, and the large sectoral variation at each level of granularity, a financial institution's calculation of financed emissions for target setting, monitoring and alignment can vary significantly. Whether the institution adheres to PCAF guidelines, which forces it to rely on lower quality and high-level data with a wide margin of error as a result of limitations on data availability, or whether it develops its own methodology for mixing data quality, resulting in non-standardized, subjective, and non-transparent methods of determining data quality and data sources, the result is a compilation of financed emissions with a wide margin of error that lacks comparability and transparency between financial institutions.

6 Alternative methods to estimating financed emissions: "Lower quality" but more consistently defined and comparable datasets

Previous sections have demonstrated issues for financial institutions in applying the PCAF data quality framework for the purposes of target setting, monitoring, and alignment, and have shown how the use of lower and mixed data quality can lead to lower accuracy and wider margins of error in emissions estimation. While PCAF has developed a universal data quality hierarchy that should be applied across sectors irrespective of sectoral variation, and premised on reported and verified emissions data as the highest quality and most reliable, we instead propose that "lower quality", but more consistently defined, comparable, and highly granular sector-specific datasets can be more useful than relying on reported data for financed emissions estimation.

While financial institutions in previous examples have used a mixed-method approach, other banks that are not members of PCAF have used methods that are entirely based on sector-specific, physical-activity data. For example, the Development Bank of Singapore

---

[32] For automotive loans, high-level data sources can overestimate emissions by 40 per cent relative to granular data. For home mortgages, the high-level data sources can underestimate emissions by 70 per cent relative to highly granular data.

(DBS) has not followed PCAF data quality guidance, but instead has compiled estimated emissions based on asset-level data combined with physical and economic activity. The bank does not draw from reported data or use a mixed method approach: "to derive a portfolio baseline, we have adopted a bottom-up approach mapping steel plants to their respective companies. We then assigned individual mills and emissions intensity based on their proportion of crude steel production from three key steelmaking methods, considering the GHG emissions from electricity use in their respective country of operation. Therefore, this approach accounts for both scope 1 and scope 2 emissions from steel production."[33] The bank has done this for four sectors: steel, automotive, aviation, and power. The reasons it gives for doing this rather than following the PCAF methodology are that "such analysis also enabled us to more accurately reflect emissions profiles when financing specific legal entities or providing ring-fenced financing."[34] This example by DBS suggests that using asset-level data in combination with other physical or economic activity-based data that is lower scored by PCAF, but more consistently observed, could be more useful for financial institutions in their measurement of financed emissions and target setting.

Similar evidence has been observed by the National Australia Bank (NAB). While it has broadly attempted to follow PCAF's methodology for compiling and attributing emissions, it does not follow the data quality hierarchy. Instead, the bank has "reported our financed emissions at a more granular industry sector level as we believe this provides a more detailed and meaningful representation of our lending portfolio." In addition to using more granular industry sector data, NAB uses a mixed method approach, "based on the availability of data (i) a bottom-up approach – based on individual company GHG data, and (ii) a top-down approach – based on industry level data where bottom-up information was unavailable."[35] While the NAB approach uses a mixed method top-down and bottom-up approach, in contrast to the DBS asset-level approach, in both cases the banks have not followed the PCAF data quality hierarchy. Instead they have relied on asset-level data and lower PCAF scored activity data to compile emissions at the sectoral level, presumably since these methods provide more reliable and accurate estimates than reported data or other mixed methods.

Asset-level data combined with lower PCAF scored physical-activity data have also been used by several other banks for certain sectors. Wells Fargo has compiled emissions

---

[33] Our path to net zero: Supporting Asia's transition to a low-carbon economy. Development Bank of Singapore. (2022), 61.

[34] Ibid., 26.

[35] NAB Group's 2021 attributable financed emissions methodology. National Australia Bank (2022), 2.

estimates using physical-activity data with a PCAF score 2 for the power sector. ING has developed its own Terra approach towards target setting, alignment, and monitoring of its loan book, which draws from different methodologies based on each sector. [36] For most sectors, including power, cement, aviation, automotive, oil and gas, ING uses the PACTA methodology, developed along with 2DII, which is premised on asset-level physical-activity data for emissions estimation.[37] The reasons for prioritising asset-level data in their emissions calculation methodology are cited by ING as providing "the most granular and accurate measurement on our clients' impact, which mainly comes from the buildings, aircraft, ships, power-generation plants that they own or operate. When we provide general purpose loans rather than asset financing, we make use of company level data. This is also the case when asset level data is not available."[38] Based on evidence cited from Wells Fargo, ING, and other banks' experiences, the confidence and trustworthiness of emissions estimation can be improved considerably by using more granular, more consistently observed physical-activity and asset-level data which is not rated as highly in PCAF **scores**

In addition to the use of asset-level or inferred emissions data, other methods of compiling emissions statistics to improve accuracy and consistency have been developed to fillin missing emissions data from lower quality, but consistently collected, data sources. Machine learning approaches have demonstrated greater accuracy in predicting emissions to fill in gaps in the data. Nguyen et al. (2021) have developed a two-step, mixed method framework to combine reported and inferred emissions predictions across models, which they say results in an increase in emissions prediction accuracy of 30% compared to other models. Additionally, incorporating other lower quality but more consistently observed data, such as energy production, consumption, sectoral and regional information can further improve emissions prediction accuracy.[39] Similarly, Serafeim and Caicedo (2022) have used machine learning models to make scope 3 emissions predictions using a mixed method approach that draws on a broad set of information on corporates, financial metrics, and scope 1 and 2 data. They find that machine learning models can improve on emissions predictions to fill in gaps in scope 3 emissions by 78% compared to other linear prediction methods, which improve

---

[36] ING Climate Report: Our integrated report to climate action. ING Group (2021), 121.
[37] The Disclosure Puzzle: The role of PACTA. *Paris Agreement Capital Transition Assessment* (2022), 6.
[38] ING Climate Report: Our integrated report to climate action, ING Group (15 September 2022), 53.
[39] Quyen Nguyen, Ivan Diaz-Rainey, Duminda Kuruppuarachchi. Predicting corporate carbon footprints for climate finance risk analyses: A machine learning approach. *Energy Economics* 95 (2021).

emissions predictions by 46%.[40] These machine learning models have been used to provide greater information on emissions when not otherwise available, while also addressing the issue of confidence and trustworthiness in the data for financial institutions by firms such as Emmi Solutions.[41]

These methods illustrate how there are other quantitative methods that can be used to increase accuracy in emissions estimation that draw from a wider variety of data sources, often from perceived lower quality data sources, rather than adherence to a universal data quality as has been suggested by PCAF. In line with the evidence presented by some financial institutions in their financed emissions disclosures, and as illustrated from machine learning algorithms, we believe that the use of physical-activity data, in combination with asset-level data, could provide a reasonable level of accuracy, while also being more transparent, consistent, reliable, and trustworthy compared to the overall reliance on reported and verified emissions data as suggested by PCAF. While this has been suggested in several cases presented in this paper, we will further explore the extent to which granular asset-level and physical-activity data can be a better estimator of emissions for the compilation of financed emissions.

7 Understanding data quality trade-offs: Quantitative methods for comparing financed emissions data sources

Despite the potential for other types of data sources and methods that are not typically treated as the highest quality to be more practical and trustworthy, there remain issues in the use of these data sources for financed emissions. Whether financial institutions follow PCAF, develop their own method, rely on lower quality data, or take a mixed method approach, there is not a systematic process for these alternative methods to be implemented in the same way that PCAF has laid out a standard for compiling emissions data. As shown throughout this paper, financial institutions have taken a variety of different approaches that are best suited to their own needs. While PCAF has attempted to create a universal standard method, we found evidence that many financial institutions are not able to meet this guidance, or have not been able to implement it at all. While PCAF and GHG Protocol have set a qualitative standard for financial institutions to broadly follow, the current state of emissions data can lead to outcomes

---

[40] George Serafeim and Gladys Velez Caicedo. Machine learning models for prediction of scope 3 emissions". *Harvard Business School Working Paper* 22-080 (2022), 1 – 36.

[41] Ben McNeil, Can we fix the carbon data problem for investors? (10 March 2022).

that are inconsistent with PCAF and GHG Protocol principles which are not fit for purpose. It is on this basis that we believe these definitions and guidance can be further supplemented with quantitative measures.

PCAF guidelines for mixing data sources are aligned with the data quality hierarchy, based on the replacement of missing data points at the asset or production level, and aggregated up to the portfolio-level.[42] However, the mixing of data sources and methods can introduce greater variability in the compilation of portfolio-level emissions, as previously illustrated in studies observing a wide margin of error and a wide discrepancy based on levels of data granularity. The increasing variance in the process of mixing data sources and methods has been cited by Morgan Stanley. From its own assessment of data quality for emissions calculation, it has found that "estimates of the starting point created by combining different sources of emissions data with different sector-specific output data generate a range of results. This contributed to further uncertainty about the actual portfolio footprint and limited our confidence in the ability to effectively manage the target over time. In addition, specific sub-industries would need to be included or excluded based on the availability of output data."[43] Hence, while banks such as HSBC have followed PCAF's data quality hierarchy while mixing data sources to fill in gaps when the highest quality data is not available, and other banks such as Barclays and Wells Fargo have adapted the PCAF approach based on data limitations and lack of confidence in reported data to develop a mixed method approach to their emissions data compilation, evidence still suggests that this results in a wider range of error, greater uncertainty, and limited confidence for financial institutions. While Morgan Stanley has cited this issue only with reference to baseline emissions, these estimation errors would be further exacerbated as portfolio emissions are aggregated across data sources, methods, geographic averages, and over time. Yet, as cited in several instances, several banks do use a mixed method approach to emissions calculations which is based on a combination of reported and inferred emissions data.

Since PCAF guidelines on the replacement of missing emissions data in the portfolio are guided only by the data hierarchy, this leaves a great deal of discretion for the financial institution to subjectively choose the data point that needs to be replaced, and from which method of estimation and source it is inserted at the portfolio-level. For example, as previously

---

[42] "Due to data limitations, financial institutions might use Options 1 or 2 for certain companies and Option 3 for others. The data quality mix shall be reflected in the average data quality score. PCAF, 53.

[43] Methodology for Morgan Stanley's 2030 interim financed emissions targets on the path to Net-Zero. Morgan Stanley (November, 2021), 12.

cited, HSBC compiles its portfolio emissions using a mixed-method approach, but primarily relying on the PCAF data quality hierarchy. When filling in data gaps from reported emissions, HSBC has developed its own framework to supplement the PCAF one for using proxy data: "proxies are selected based on: availability, transparency, accuracy, simplicity, and relevance."[44] However again, as with the broad qualitative principles and guidance outlined by PCAF and GHG Protocol, each of these measures are subjectively defined by the financial institutions, which leaves extensive discretion in emissions data compilation for an individual practitioner or institution. For HSBC the question is, how relevant is the calculation for different types of clients While this and other qualitatively defined measures for mixing methodologies may prove effective to the specific needs of the clients, it can lead to increased variability, inconsistency, and lack of transparency when compared over time or between sectors.

Therefore, we propose a more structured, quantitative, and consistent way of compiling portfolio emissions that is aligned with the PCAF and GHG Protocol guidelines for carbon accounting. For example, the root means squared error (RMSE) test can be used to compare data quality in inferred emissions relative to reported and verified emissions data.[45] These quantitative measures, among other methods, can assist financial institutions in the process of mixing and selecting the highest quality reported data with various forms of proxy data, for cases in which various forms of inferred and estimated emissions statistics are used. The use of quantitative measures will hopefully lead to greater transparency, accuracy, simplicity, and objectivity for financial institutions in their emissions compilation, rather than the various set of frameworks that have been created internally by financial institutions to subjectively assess proxy data.

---

[44] Financed emissions methodology, *HSBC Holdings PLC* (2022), 11.

[45] The root mean squared error is a measure of accuracy that can be used to compare two datasets in terms of the difference between observations from predicted data in relation to a reference dataset, or to compare methods of estimated emissions in relation to reported emissions data. It is a simple exercise that subtracts the difference in the estimated data relative to the reference reported data for each unit of observation, squares each difference, takes the average of all squared differences, and then the square root of the average. This is summarized in the equation below, where *e* is the estimated data, *r* is the reported data, and *N* is the size of the dataset.

$$RMSE_{e,r} = \left[\sum_{i=1}^{N} \frac{\left(z_{e,i} - z_{r,i} \right)^2}{N} \right]^{\frac{1}{2}}$$

8 Enhancing data quality frameworks and financed emissions reporting: Next steps for the SDQI project

As demonstrated in this paper, financial institutions have attempted to either follow the PCAF data quality hierarchy by using a mixed method approach of combining lower quality data when the highest quality is not available, or develop a methodology not premised on PCAF's hierarchy which does not treat reported and verified emissions as the best data source. In both cases, estimated emissions from physical activity have been used to supplement reported data, or has been the primary data source used to compile portfolio emissions. To this extent, it suggests that physical-activity data can supplement or provide more information for financial institutions in the compilation of emissions compared to reported data. While this has so far been suggested independently by disclosures from various financial institutions, we will study the extent to which this approach can systematically improve accuracy in emissions compilation. We will review the extent to which physical-activity data at different levels of granularity, within and between different geographies, and for key sectors including power, steel, and cement, can provide more accurate and consistent emissions estimates for financial institutions compared to reported data, in contrast to prevailing PCAF guidance.

Second, while the PCAF framework has set out five principles to assist financial institutions in carbon accounting and data quality, we will provide further guidance to support financial institutions by evaluating the importance of comparability in data quality. The comparability of emissions data between reporting entities has not been widely addressed, as the emphasis in the literature has primarily been on consistency over time, rather than between entities. However, if reported data cannot be established or verified due to lack of consistency between entities, then this can have significant implications for ratings, indices, and investment vehicles, which can result in the misallocation of funds. We will examine the extent to which reported data lacks comparability, whether the standardization of reporting boundaries or narrowing the scope of reporting addresses the issue and whether the use of other methods of inferred emissions provides better comparability. We will then develop a set of principles to assist financial institutions in addressing comparability as an additional principle to carbon accounting for estimating financed emissions.

Third, even if physical-activity data may be more accurate, it will still depend on the level of granularity in the observation. While activity-based data has led to greater trust and confidence for some financial institutions, studies have suggested that if the data is not sufficiently granular, then it can introduce a greater magnitude of variance in the resulting

estimates. Morgan Stanley contends that at a sectoral level, activity-based estimation is not reliable: "the underlying sector-specific output data, especially in the energy sector, varies widely across data vendors due to different assumptions about upstream production rights, conversion factors, and so on. Carbon data is often estimated, and Scope 3 emissions estimates are derived from different vendor models. Combining these metrics yields uncertainty in the denominator of the intensity calculation, which is further compounded if combined with other variables."[46] This suggests that physical-activity data alone may not be a better data source than reported data, but may depend on the level of granularity, and the ability to control and extract specific variables. To this extent, asset-level data can be a highly detailed data source that, along with highly granular activity data, can contribute to increased accuracy and confidence in emissions estimation. Having such highly granular data can help financial institutions in their target setting and monitoring activities by allowing them to set accounting boundaries and emission factors consistently across counterparties, whereas these boundaries and factors can vary when relying on emissions data from other third-party vendors and sources for sector-specific activity and emissions data.

However, the coverage and availability of this type of data remains low. As posited by Morgan Stanley, sectoral output data is often observed at an aggregated level, and is not sufficiently granular to be used as a data source in emissions compilation. This issue persists across financial institutions. A survey of 1,290 companies by BCG has shown that 49% of respondents found that sufficiently granular operational activity data was "hard" or "very hard" to find. In addition to activity data, 55% of respondents found sufficiently granular emissions factors as "hard" or "very hard" to find.[47] While banks such as DBS, NAB, Wells Fargo, and ING have conducted asset-level emissions compilation from operational activity and emissions factors with sufficient levels of granularity, the improved granularity and availability of this type of activity data and emissions factors, along with other machine learning algorithms for filling-in missing emissions data, can greatly increase the level of accuracy and confidence in emissions compilation. However, for highly granular activity and emissions data to be used by financial institutions, equally granular data on the location and ownership of an emitting asset is also needed.

While this data has so far been limited in availability, as evidenced by respondents to BCG's survey, third-party data providers are increasingly developing highly granular asset-

---

[46] Morgan Stanley, op. cit, 12.

[47] Use AI to measure emissions exhaustively, accurately, and frequently. *Carbon measurement survey report 2021.* BCG Gamma (13 October 2021).

level data, operational activity and emissions factor data, and leveraging new technologies such as remote sensing and artificial intelligence.[48] Asset-level data can be used in combination with highly granular activity and emissions data by institutions as an alternative method to compiling financed emissions. Taking inferred emissions from asset and activity data in specific sectors (e.g., power, cement), we will evaluate the accuracy, precision, and consistency of this type of data compared to a reference dataset of reported and verified emissions in order to determine how well these methods can approximate reported data for carbon accounting. We will develop guidance for developing and interpreting asset-level emissions information, including on the use of emission factors, and physical and economic activity data. This has implications in assessing the extent to which these methods of estimation can be more practically applied by financial institutions instead of relying on reported data, both for filling in gaps in missing reported data, and to increase consistency, transparency, relevance, accuracy, and comparability in emissions compilation for a reporting entity.

---

[48] Both in the profit and non-profit sector, where organisations such as Global Energy Monitor, Climate Trace and our own Spatial Finance Initiative are generating highly granular asset-level datasets and physical-activity insights.