International Scientific Journal of Engineering and Management
ISSN: 2583-6129
Volume: 04 Issue: 03 | Mar – 2025
DOI: 10.55041/ISJEM02353
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Opportunity for scaling Product Carbon Footprinting using Large Language Models

Zaid Thanawala
Sr. Sustainability Scientist

Introduction

Product carbon footprinting (PCF) is a method used to quantify the greenhouse gas (GHG) emissions associated with the life cycle of a product. This includes emissions from raw material extraction, production, distribution, use, and disposal. PCFs increasingly important for industries aiming to reduce their environmental impact and comply with regulatory requirements. It serves as a critical business indicator, influenced by life cycle assessment (LCA) methodologies, and is essential for making informed decisions about sustainability practices. PCFs are crucial for meeting regulatory requirements, such as the EU's Battery Passport Initiative, which mandates transparency in carbon footprint documentation for electric vehicle batteries starting in 2026 (Gutwald et al., 2024). Companies use PCFs to identify emission hotspots and implement reduction strategies, which can lead to significant reductions in emissions and provide a competitive edge in markets where sustainability is a purchasing criterion (Rüdele & Wolf, 2023)(Rüdele & Wolf, 2023). Organizations like BASF have developed ISO-conformant methodologies to calculate PCFs, aiming to provide maximum transparency to consumers and stakeholders (Paliwal, 2022).

This goal of this paper is to present key issues with scaling of PCFs in the industry and identify opportunities where Large Language Models (LLMs) can help in order to scale PCFs to millions of products.

Methodology for PCFs

PCFs play a pivotal role in decarbonizing supply chain emissions by providing a systematic measure of greenhouse gas emissions associated with the production and lifecycle of products.. By breaking down emissions across the supply chain, companies can identify key areas for emission reduction, such as upstream suppliers or manufacturing processes (Meinrenken et al., 2020)(Chen et al., 2017). Companies can also optimize their product portfolios by eliminating high-emission products and introducing low-carbon alternatives, thus achieving full decarbonization across the supply chain (Bai et al., 2024). Product carbon footprint assessments guide manufacturers in implementing green supply chain measures, such as improving energy efficiency and selecting low-carbon suppliers (Chen et al., 2017).

LCA is a comprehensive method used to evaluate the PCF, though it requires significant effort and interdisciplinary knowledge (Lang et al., 2024). It helps companies achieve carbon reductions by improving processes throughout a product's value chain (Meinrenken et al., 2022). Accurate PCF calculation requires complex data collection and accounting for GHG emissions, which is often lacking, especially for indirect emissions (Scope 3) that can constitute a significant portion of a product's carbon footprint (Boukherroub et al., 2024)

ISSN: 2583-6129
DOI: 10.55041/ISJEM02353

(Jaeger et al., 2022). There is no single mandatory standard for PCF accounting, leading to challenges in comparing results across different products and industries (Shams et al., 2023). The calculation of PCF involves numerous assumptions and uncertainties about the supply chain, which can introduce uncertainties and affect the reliability of the PCF as a standalone value (Quernheim et al., 2023).

Traditional process for creating a PCF

Estimating the PCF generally involves three main steps: defining system boundaries, performing a life cycle inventory analysis, and assessing environmental impacts. LCA practitioners create a life cycle inventory (LCI) based on their expertise and the established system boundaries related to the carbon footprint of a product. Practitioners typically break down the processes associated with each stage of the product's life cycle to identify the necessary inputs and outputs for constructing a comprehensive life cycle model (Muthu, 2014).

Subsequently, they assess the environmental impacts, particularly greenhouse gas emissions, associated with all identified inputs and outputs throughout the product's life cycle. The input components usually encompass raw materials and energy, while the GHG linked to the production of these raw materials, consumables, and secondary energy sources, such as electricity, are incorporated into the product's carbon footprint, specifically within Scope 2 and Scope 3 upstream categories. The output components generally consist of waste gases, wastewater, and solid waste materials, including greenhouse gas emissions generated from the combustion of fossil fuels, classified under Scope 1 (Sfez et al., 2019). Ultimately, these greenhouse gas emissions are converted into carbon dioxide equivalents (CO2-eq) and aggregated to calculate the overall carbon footprint of the product.

This assessment process can be lengthy, as it necessitates a detailed examination of each life cycle stage based on industry production knowledge or relevant scholarly literature to ascertain the inputs and outputs for each process. Values for the input-output inventory are determined using actual production conditions, leading to the final calculations. While there are various life cycle modeling tools and databases available that may enhance the efficiency of these calculations, completing a carbon footprint assessment for a product can still require several days or even months.

Challenges to scale PCFs

The task of scaling PCFs involves various challenges that encompass technical, methodological, and organizational aspects. Accurately measuring and reporting the carbon emissions linked to products throughout their entire lifecycle is a complex endeavor that necessitates a thorough and systematic approach. This complexity is further intensified by the requirement for standardized methods, dependable data, and effective communication within supply chains.

There are significant gaps in PCF data within companies and life cycle assessment (LCA) databases, which hinder the accurate calculation of carbon footprints (Jaeger et al., 2022). Ensuring high-quality and consistent data is a major challenge, as data must be reliable and comparable across different products and supply chains (Jaeger et al., 2022) (Smith, 2012). Sharing detailed carbon footprint data can raise confidentiality issues, making companies hesitant to disclose sensitive information (Shams et al., 2023).

Conducting a life cycle assessment requires detailed knowledge and expertise, which can be a bottleneck in scaling PCF efforts (Balaji et al., 2024). The lack of standardized methods for calculating and reporting PCFs leads to inconsistencies and difficulties in comparing results across different products and industries (Pedersen & Remmen, 2022)

ISSN: 2583-6129
DOI: 10.55041/ISJEM02353

(Klockenhoff, 2009). Inadequate functional units and evolving impact categories, such as those for biodiversity and land use change, complicate the fair comparison of products (Pedersen & Remmen, 2022).

The process of measuring and reporting PCFs can be resource-intensive, although costs may decrease with increased familiarity and scale (Smith, 2012). The dynamic and complex nature of supply chains requires continuous updates and adjustments to PCF calculations, which can be difficult to manage (Shams et al., 2023). Engaging suppliers and ensuring their cooperation in providing necessary data is crucial but often challenging (Smith, 2012). Effectively communicating PCF results to stakeholders and ensuring transparency in the process is essential for credibility and trust (Shams et al., 2023).

Although these challenges are considerable, they also offer avenues for innovation and enhancement. Developments in AI and LLMs can assist in automating and refining the life cycle assessment process, thereby increasing its accessibility and efficiency.

Leveraging LLMs

One promising avenue for leveraging LLMs in the realm of PCF is their ability to enhance data integration and analysis. By employing natural language processing capabilities, these models can sift through vast amounts of unstructured data—from supplier reports to regulatory documents—enabling organizations to extract relevant emissions information more efficiently. This capability not only streamlines the life cycle inventory process but also mitigates some of the challenges associated with data gaps and inconsistencies that currently hinder accurate PCF calculations (Waddell et al., 2011).

Introduction to Large Language Models

LLMs represent a significant leap in artificial intelligence (AI), particularly in the field of natural language processing (NLP). These models are designed to understand and generate human-like text by capturing complex linguistic patterns and contextual subtleties. LLMs have revolutionized various domains, from conversational systems and machine translation to complex system modeling and software engineering. LLMs are built on deep neural networks with billions of parameters, enabling them to capture intricate language patterns. Models like GPT-3 and GPT-4 exemplify this capability, using massive datasets to train and understand language comprehensively (Sreerakuvandana et al., 2024)

(Haque, 2024). The architecture of LLMs often involves attention mechanisms, which allow the models to focus on relevant parts of the input text, enhancing their ability to generate coherent and contextually appropriate responses (Atkinson, 2024).

In complex systems research, LLMs are integrated into Generative Agent-Based Models (GABMs) to simulate human behavior and model complex interactions, such as social dynamics and epidemic modeling (Lu et al., 2024). In education, LLMs are used to enhance reading, writing, and speaking skills, as well as to develop intelligent tutoring systems, offering new opportunities for personalized learning experiences (Alhafni et al., 2024). In software engineering, LLMs assist in code generation and interpretation, streamlining workflows and fostering innovation in software development (Haque, 2024). Similar to the use of LLMs in other domains, this study aims to explore the use of LLMs in assisting LCAs and PCFs.

ISSN: 2583-6129
DOI: 10.55041/ISJEM02353

Life cycle inventory creation

Life cycle practitioners build a life cycle inventory (LCI) by systematically collecting and compiling data on the elementary flows of materials and energy within a product system. This process is integral to LCA and involves several methodical steps to ensure accuracy and comprehensiveness. Practitioners rely on a combination of data sources, including public databases, proprietary software, and hybrid methods, to construct a robust LCI. There are portions of the process that can be sped up by the use of LLMs.

Practitioners begin by identifying the processes that constitute the product system under study. This involves defining the system boundaries and functional units, which guide the scope of the inventory (Bjørn et al., 2018) (Arvidsson & Ciroth, 2021). Planning for data collection is crucial, as it involves determining the types of data needed and the sources from which they will be obtained. This step may include consulting existing LCI databases, such as those maintained by US federal agencies, which provide a wealth of data but may present interoperability challenges (Ingwersen, 2015). Data collection involves gathering information on the inputs and outputs of each process within the system. This can include raw materials, energy consumption, emissions, and waste products (Bjørn et al., 2018). Practitioners often use a combination of process-based and input-output methods to compile data. Hybrid LCI methods can address limitations of traditional approaches by integrating different data types and sources (Crawford et al., 2018).

Once data is collected, practitioners construct unit processes, which are the building blocks of the LCI model. Each unit process represents a specific activity or operation within the product system (Bjørn et al., 2018). Quality checking is essential to verify the accuracy and consistency of the data. This may involve cross-referencing with other databases or conducting sensitivity analyses to assess the impact of data variability on the LCI results (Bjørn et al., 2018)

(Wu & Su, 2020). The LCI model is constructed by linking unit processes according to the flow of materials and energy. This model is then used to calculate the overall inventory of elementary flows for the product system (Bjørn et al., 2018). Practitioners may use software like SimaPro to perform these calculations, although limitations in data exchange and documentation features may necessitate additional tools or methods (Coste et al., 2021). The final step involves reporting the LCI results, which includes documenting the data sources, assumptions, and methodologies used. This transparency is crucial for the credibility and reproducibility of the LCA study (Bjørn et al., 2018). Managing uncertainty is also important, as it allows practitioners to understand the potential variability in the results and make informed decisions based on the LCI data (Bjørn et al., 2018).

Opportunity to utilize LLMs

While the process of building a life cycle inventory is methodical and data-driven, it is not without challenges. Practitioners must navigate issues such as data availability, quality, and interoperability, particularly when using diverse data sources. LLMs can significantly enhance the data collection and curation process for creating an LCI by addressing key challenges such as missing data and inconsistency in data matching. LLMs can automate the extraction and organization of data from various sources, reducing the manual effort required in traditional LCI data collection methods. This automation is particularly useful in handling large datasets and ensuring that all relevant data is captured efficiently (Tu, 2024). By leveraging retrieval augmented generation (RAG) techniques, LLMs can improve the accuracy of data retrieval, ensuring that the most relevant and up-to-date information is used in the LCI process (Li et al., 2024

ISSN: 2583-6129
DOI: 10.55041/ISJEM02353

. LLMs can help standardize data formats and terminologies across different datasets, addressing the issue of inconsistency in background data matching. This standardization is crucial for ensuring that the data used in LCI is comparable and reliable. The integration of LLMs with knowledge graphs can further enhance data consistency by providing a structured framework for organizing and linking data from various sources. LLMs can process and analyze data from multiple modalities, such as text, numerical data, and images, providing a comprehensive view of the data landscape. This capability is essential for capturing the full scope of environmental impacts in LCI (Tu, 2024). The use of LLMs in semantic parsing and text-to-SQL methodologies can facilitate the querying and analysis of complex datasets, making it easier to extract meaningful insights from the data (Zhu et al., 2023).

LLMs can be used to track data provenance, ensuring transparency in how data is collected, processed, and used in LCI. This tracking is vital for maintaining trust in the data and the conclusions drawn from it (Lauro et al., 2024). By automatically rewriting user-defined pipelines and storing detailed descriptions of data processing activities, LLMs can provide a clear audit trail of data transformations, supporting the explanation and validation of LCI results (Lauro et al., 2024).

Conclusion

In conclusion, the integration of Large Language Models (LLMs) into the process of Product Carbon Footprinting (PCF) presents a transformative opportunity to address the significant challenges associated with scaling PCFs across various industries. The current landscape of PCF is characterized by complexities in data collection, inconsistencies in methodologies, and the need for accurate and standardized reporting. By leveraging the advanced capabilities of LLMs, organizations can enhance data integration and analysis, streamline the life cycle inventory process, and improve the overall efficiency of carbon footprint assessments. The potential for LLMs to automate data extraction, standardize terminologies, and enhance data provenance tracking will not only facilitate more accurate PCF calculations but also foster greater transparency and trust among stakeholders. As industries increasingly prioritize sustainability and regulatory compliance, the adoption of LLMs in PCF processes could significantly contribute to the decarbonization of supply chains and the broader goal of reducing greenhouse gas emissions on a global scale. Therefore, this paper underscores the importance of exploring and implementing LLM-driven solutions as a means to enhance the scalability and effectiveness of product carbon footprinting initiatives.

References

Gutwald, B., Baumann, N., Funk, F., Reichenstein, T., Albayrak, B., & Franke, J. (2024).*Sustainable manufacturing practices: A systematic analysis and guideline for assessing the industrial Product Carbon Footprint*. https://doi.org/10.1109/epts61482.2024.10586733

Rüdele, K., & Wolf, M. (2023a). Identification and Reduction of Product Carbon Footprints: Case Studies from the Austrian Automotive Supplier Industry.*Sustainability*. https://doi.org/10.20944/preprints202309.1047.v1

Rüdele, K., & Wolf, M. (2023b). Identification and Reduction of Product Carbon Footprints: Case Studies from the Austrian Automotive Supplier Industry.*Sustainability*. https://doi.org/10.3390/su152014911

Paliwal, D. (2022). Product Carbon Footprints - zur Vergleichbarkeit der Produkte die wir kaufen?*Controlling*. https://doi.org/10.15358/0935-0381-2022-4-44

ISSN: 2583-6129
DOI: 10.55041/ISJEM02353

Meinrenken, C. J., Chen, D., Esparza, R. A., Iyer, V., Paridis, S. P., Prasad, A., & Whillas, E. (2020). Carbon emissions embodied in product value chains and the role of Life Cycle Assessment in curbing them.*Scientific Reports*. https://doi.org/10.1038/S41598-020-62030-X

Chen, J., Sun, L., & Guo, H. (2017).*Product carbon footprint assessment supporting the green supply chain construction in household appliance manufacturers*. https://doi.org/10.1088/1755-1315/94/1/012142

Bai, C., Ma, X., & Zhu, Q. (2024). Product portfolio decarbonization: Deleting hot products for a cooler supply chain.*Business Strategy and The Environment*. https://doi.org/10.1002/bse.3958

Lang, S., Engelmann, B., Schiffler, A., & Schmitt, J. (2024). A simplified machine learning product carbon footprint evaluation tool.*Cleaner Environmental Systems*. https://doi.org/10.1016/j.cesys.2024.100187

Meinrenken, C. J., Chen, D. S. J., Esparza, R. A., Iyer, V., Paridis, S. P., Prasad, A., & Whillas, E. (2022). The Carbon Catalogue, carbon footprints of 866 commercial products from 8 industry sectors and 5 continents.*Scientific Data*. https://doi.org/10.1038/s41597-022-01178-9

Boukherroub, T., Bouchery, Y., Tan, T., Fransoo, J. C., & Corbett, C. J. (2024).*Carbon Footprinting in Supply Chains: Measurement, Reporting, and Disclosure*. https://doi.org/10.1007/978-3-031-45565-0_3

Jaeger, F. A., Saling, P., Otte, N., Steidle, R., Bollen, J., Golembiewski, B., Dencic, I., Letinois, U., Rehl, T., & Wunderlich, J. (2022). Challenges and requirements of exchanging Product Carbon Footprint information in the supply chain.*E3S Web of Conferences*. https://doi.org/10.1051/e3sconf/202234907005

Shams, S. B., Kind, A., Jaeger, F. A., Beitinger, G., Leonte, I. A., Weinhold, M., & Pham, V. (2023).*Trustworthy Supply Chain Exchange for Product Carbon Footprint*. https://doi.org/10.1109/aibthings58340.2023.10292452

Quernheim, N., Anderl, R., & Schleich, B. (2023). Framework for comparison of product carbon footprints of different manufacturing scenarios.*Proceedings of the Design Society*. https://doi.org/10.1017/pds.2023.194

Muthu, S. S. (2014).*Life cycle assessment (LCA) and product carbon footprint (PCF) modelling of textile products*. https://doi.org/10.1533/9781782421122.132

Sfez, S., De Meester, S., Vlaeminck, S. E., & Dewulf, J. (2019). Improving the resource footprint evaluation of products recovered from wastewater: A discussion on appropriate allocation in the context of circular economy.*Resources Conservation and Recycling*. https://doi.org/10.1016/J.RESCONREC.2019.03.029

Smith, S. (2012).*Large scale product carbon footprinting of consumer goods*. https://doi.org/10.1007/978-94-007-3010-6_59

Balaji, B., Guest, G., Vunnava, V. S. G., Kramer, J., Srinivasan, A., & Taptich, M. N. (2024).*Scaling Carbon Footprinting: Challenges and Opportunities*. https://doi.org/10.1609/aaaiss.v2i1.27646

Pedersen, E., & Remmen, A. (2022). Challenges with product environmental footprint: a systematic review.*International Journal of Life Cycle Assessment*. https://doi.org/10.1007/s11367-022-02022-3

ISSN: 2583-6129
DOI: 10.55041/ISJEM02353

Klockenhoff, J. (2009). Product Carbon Footprinting und der Kohlendioxid-bewusste Konsument.*Journal Fur Verbraucherschutz Und Lebensmittelsicherheit-Journal of Consumer*. https://doi.org/10.1007/S00003-009-0489-Y

Waddell, C. J., Hurteau, M. D., & Huntzinger, D. (2011). Product carbon footprinting: A proposed framework to increase confidence, reduce costs and incorporate profit incentive.*Carbon Management*. https://doi.org/10.4155/CMT.11.68

Sreerakuvandana, S., Pappachan, P., & Arya, V. (2024). Understanding Large Language Models.*Advances in Computational Intelligence and Robotics Book Series*. https://doi.org/10.4018/979-8-3693-3860-5.ch001

Haque, Md. A. (2024).*LLMs: A Game-Changer for Software Engineers?*https://doi.org/10.48550/arxiv.2411.00932

Atkinson, J. (2024).*Large Language Models*. https://doi.org/10.1201/9781003517245

Lu, Y., Aleta, A., Du, C., Shi, L., & Moreno, Y. (2024). LLMs and Generative Agent-Based Models for Complex Systems Research.*Physics of Life Reviews*. https://doi.org/10.1016/j.plrev.2024.10.013

Alhafni, B., Vajjala, S., Banno, S., Maurya, K. K., & Kochmar, E. (2024). LLMs in Education: Novel Perspectives, Challenges, and Opportunities.*arXiv.Org*. https://doi.org/10.48550/arxiv.2409.11917

Bjørn, A., Bjørn, A., Moltesen, A., Laurent, A., Owsianiak, M., Corona, A., Birkved, M., & Hauschild, M. Z. (2018).*Life Cycle Inventory Analysis*. https://doi.org/10.1007/978-3-319-56475-3_9

Arvidsson, R., & Ciroth, A. (2021).*Introduction to "Life Cycle Inventory Analysis."*https://doi.org/10.1007/978-3-030-62270-1_1

Ingwersen, W. W. (2015). Test of US federal life cycle inventory data interoperability.*Journal of Cleaner Production*. https://doi.org/10.1016/J.JCLEPRO.2015.03.090

Crawford, R. H., Bontinck, P.-A., Stephan, A., Wiedmann, T., & Yu, M. (2018). Hybrid life cycle inventory methods – A review.*Journal of Cleaner Production*. https://doi.org/10.1016/J.JCLEPRO.2017.10.176

Wu, Y., & Su, D. (2020).*Life cycle inventory management*. https://doi.org/10.1007/978-3-030-39149-2_8

Coste, G., Biard, Y., Roux, P., & Helias, A. (2021). ELDAM: A Python software for Life Cycle Inventory data management.*The Journal of Open Source Software*. https://doi.org/10.21105/JOSS.02765

Tu, Q. (2024). Mitigating Grand Challenges in Life Cycle Inventory Modeling through the Applications of Large Language Models.*Environmental Science & Technology*. https://doi.org/10.1021/acs.est.4c07634

Li, G., Zhou, X., & Zhao, X. (2024). LLM for Data Management.*Proceedings of The Vldb Endowment*. https://doi.org/10.14778/3685800.3685838

ISSN: 2583-6129
DOI: 10.55041/ISJEM02353

Zhu, J., Bazaz, S. A., Dutta, S., Anuraag, B., Haider, I., & Bandopadhyay, S. (2023).*Talk to your data: Enhancing Business Intelligence and Inventory Management with LLM-Driven Semantic Parsing and Text-to-SQL for Database Querying*. https://doi.org/10.1109/icdabi60145.2023.10629374

Lauro, L., Lazzaro, P. L., Lazzaro, M. D., Missier, P., & Torlone, R. (2024).*An LLM-guided Platform for Multi-Granular Collection and Management of Data Provenance*. https://doi.org/10.21203/rs.3.rs-5369549/v1