

AutoPCF: A Novel Automatic Product Carbon Footprint Estimation Framework Based on Large Language Models

Biao Luo^{1#}, Jinjie Liu^{2#}, Zhu Deng¹, Can Yuan¹, Qingrun Yang¹, Lei Xiao¹, Yucong Xie¹, Fanke Zhou¹, Wenwen Zhou^{1*}, Zhu Liu³

¹ Alibaba Cloud, Hangzhou, Zhejiang, China

² School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, China

³ Department of Earth System Science, Tsinghua University, Beijing, China

luobiao.luo@alibaba-inc.com, jinjieliu@link.cuhk.edu.cn, {dengzhu.dz, yuancan.yc}@alibaba-inc.com, qingrunyang@gmail.com, {xiaolei.xiao, xie.xyc, fanke.zfk, zhoubo.zww}@alibaba-inc.com, zhuliu@tsinghua.edu.cn

Abstract

Estimating the product carbon footprint (PCF) is crucial for sustainable consumption and supply chain decarbonization. The current life cycle assessment (LCA) methods frequently employed to evaluate PCFs often encounter challenges, such as difficulties in determining the emission inventory and emission factors (EFs), as well as significant labor and time costs. To address these limitations, this paper presents AutoPCF, a novel automatic PCF estimation framework to conduct cradle-to-gate LCA for products. It utilizes deep learning models and large language models (LLMs) to automate and enhance the estimation process. The framework comprises five stages: Emission Inventory Determination (EID), Activity Data Collection (ADC), Emission Factor Matching (EFM), Carbon Emission Estimation (CEE), and Estimation Verification and Evaluation (EVE). EID generates production processes and activity inventory, while ADC collects comprehensive activity data and EFM identifies accurate EFs. Emissions are then estimated using the collected activity data and corresponding EFs. Experimental evaluations on steel, textile, and battery products demonstrate the effectiveness of AutoPCF in improving the efficiency of PCF estimation. By automating data collection and analysis, AutoPCF reduces reliance on subjective decision-making and enhances the consistency and efficiency of carbon footprint assessments, advancing sustainable practices and supporting climate change mitigation efforts.

Introduction

Efficiently estimating the carbon footprint of products plays a pivotal role in understanding their environmental impact and guiding informed decision-making toward sustainable consumption. By quantifying the emissions associated with

a product's life cycle, product carbon footprint (PCF) estimation enables carbon labeling and the development of effective strategies to reduce environmental impact (Jensen, 2012).

The traditional LCA approach to estimate PCF through a five-step process: determining system boundaries, emission inventory analysis, collecting activity data, identifying Emission Factors (EFs), and evaluating environmental impacts. However, these methods often face significant challenges and limitations. Determining the emission sources and constructing complete emission inventories require extensive research, often relying on a combination of primary data collection, literature reviews, and expert judgment (Curran et al., 2006). This process is time-consuming, resource-intensive, and subject to uncertainties. Moreover, the selection of EFs which directly influences the results of carbon footprint calculations, is highly dependent on expert knowledge and subjective decision-making, introducing potential biases and inconsistencies (McKinnon, 2010). These limitations necessitate the exploration of alternative approaches that can enhance the efficiency of carbon footprint estimation.

In recent years, machine-learning methods have emerged as promising approaches for carbon emission analysis (Algren et al., 2021; Balaji et al., 2023; Han et al., 2021; G. Liu et al., 2022; Shinde et al., 2022; Zhao et al., 2022). However, to the best of our knowledge, the only two studies that try to apply machine learning methods to conduct product carbon footprinting have not investigated the carbon estimation process of the products (Balaji et al., 2023; Zhao et al., 2022). In addition, the emergence of large language

These authors contribute equally.

* Corresponding: Corresponding: zhoubo.zww@alibaba-inc.com
Copyright © 2023, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

models (LLMs), such as GPT series (Brown et al., 2020; Radford et al., 2019, 2018) and GLM (Du et al., 2022; Zeng et al., 2022), has presented the potential for further advancements in carbon management. LLMs are powerful deep-learning models that have been trained on vast amounts of text data, enabling them to generate coherent and contextually relevant responses. Some papers have discussed the potential of LLMs in carbon management and environmental research (Agathokleous et al., 2023; Zhu et al., 2023). LLMs possess powerful language understanding capabilities, allowing them to generate detailed descriptions and simulate various production processes. This indicates a possibility of applying LLMs and deep learning models to estimate PCF and improve efficiency.

In this paper, a novel framework that leverages the strengths of deep-learning methods and LLMs for automatic PCF estimation is proposed, named AutoPCF. It consists of multiple interconnected stages, including Emission Inventory Determination (EID), Activity Data Collection (ADC), Emission Factor Matching (EFM), Carbon Emission Estimation (CEE), and Estimation Verification and Evaluation (EVE). By integrating these stages, AutoPCF offers an efficient and automated solution to overcome the limitations of current estimation methods.

Method

In this paper, the estimation of PCF focuses on assessing the carbon emissions associated with the most significant sources of a product's lifecycle. This includes key stages, raw material extraction, and manufacturing and production, also known as cradle-to-gate LCA (Niccolucci et al., 2010). The PCF assessment takes into account seven greenhouse gases, namely carbon dioxide (CO₂), methane (CH₄), nitrous oxide (N₂O), hydrofluorocarbons (HFCs), perfluorocarbons (PFCs), sulfur hexafluoride (SF₆), and nitrogen trifluoride (NF₃), which are covered by the UNFCCC/Kyoto Protocol. To provide a standardized measure, the PCF is expressed in terms of CO₂ equivalent (CO₂-eq) based on their respective Global Warming Potentials (GWPs).

The proposed framework, AutoPCF, integrates deep-learning models and LLMs to optimize the process of LCA methods and enhance the efficiency of PCF estimation. The AutoPCF comprises an interconnected network of stages aimed at achieving comprehensive and automated PCF estimation, including EID, ADC, EFM, CEE, and EVE stages. The flowchart of the framework is shown in Fig. 1.

The efficiency of PCF estimation mainly depends on three processes: EID, ADC, and EFM. Below the improved methods for these three processes are described in detail. By utilizing the activity data and the corresponding EFs, the EE stage estimates the PCF, as represented by Eq. (1).

$$E = \sum AD_f \times EF_f \quad (1)$$

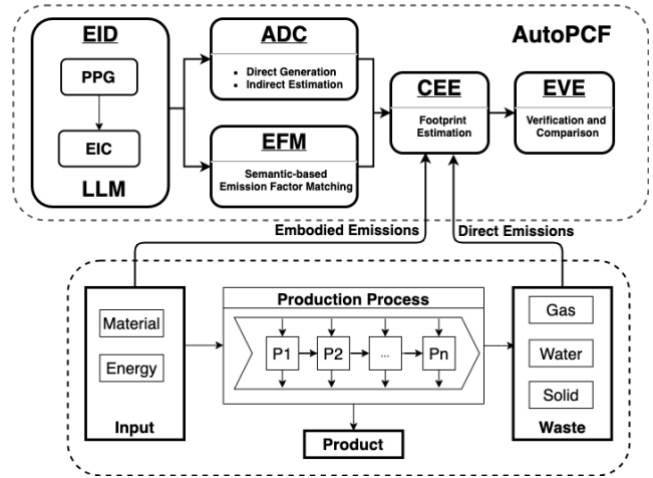


Figure 1: The framework of AutoPCF

where AD_f and EF_f are the activity data and EF of the activity f , e.g., the fuels consumed. Then EVE verifies the confidence and efficiency of the AutoPCF. By integrating these interconnected stages, AutoPCF offers a comprehensive and automated solution that addresses the limitations of existing estimation methods.

Emission Inventory Determination

Here LLMs are used to conduct emission inventory analysis by carefully designing prompts, including production process generation and activity inventory identification for each process, such as materials, energy, and waste. The production process of the product is acquired in a conversational way by providing prompts and requirements. The LLM is instructed to act as an expert in product production. The question asking about the production process of the product is first constructed. A requirement on the output is pre-defined to control the format of the outputted production process. Then a prompt is constructed with these settings and LLM outputs the production process. An example prompt for generating production processes is shown in Appendix Table A1. After the production process of the product is determined, the activity inventory of each process is to be generated in a similar way by appropriately providing prompts and in-context information.

Activity Data Collection

Activity data refers to the amount of raw materials or energy consumed and waste generated during the production process. Two methods are proposed to generate activity data required for PCF accounting.

- Direct generation approach: Similar to emission inventory generation, appropriate prompts are constructed to enable LLMs to automatically generate and output activity data corresponding to the inventory.

- Indirect estimation approach: A top-down approach was utilized, relying on industry-wide macro-statistics data, to create activity data for each process outlined in the emission inventory.

Emission Factor Matching

The semantic similarity between the emission inventory and the EF database (Ecoinvent database (Wernet et al., 2016)) is calculated based on sentence transformers, and the EF with the highest similarity is assigned to each emission inventory. First, the EF names, process names, and EF classification information in the EF database are encoded into vectors of length 768 by a pre-trained language model(msmarco-bert-base-dot-v5) and stored in the vector library. Second, the emission inventory is input into the same encoder to obtain the corresponding sentence vector when performing EF matching for a specific emission inventory. Third, the cosine similarity between the emission inventory sentence vector and the vector in the vector library is calculated, and the EF with the highest similarity is determined.

Experiments and Results

This study evaluated the effectiveness of the proposed framework by the reliability of PCF results and the time cost of PCF estimation, which is verified on products from three industries, hot rolled round steel (HRRS), printed and dyed fabric (PDF), and lithium iron phosphate battery (LIPB). Five LLMs are employed to generate the required information and compared, including GPT-3.5, GPT-4, Tongyi Qianwen, GLM-130B, and ChatGLM-6B. Not only the effectiveness of the AutoPCF is analyzed, but the uncertainty of deep learning models used for various estimation stages is discussed.

Emission Inventory Generation Performance

The production processes of the three products and the emission inventories in each process have been obtained through various LLMs. Different LLMs may yield varying outputs based on different prompt words. However, for the purpose of comparing the effects between different models, the prompts used were generally kept consistent and were not specifically optimized for individual models.

An illustrative example of the results generated by GPT-3.5 is shown in Appendix Table A2. In this instance, the production of printed and dyed fabric included 9 processes, which covered the entire process as per expert experience, including steps such as Spinning, weaving, printing, dyeing, and finishing (see Appendix Table A3). In addition, the model also yielded 17 different emission inventories involved in each process according to the prompt requirements.

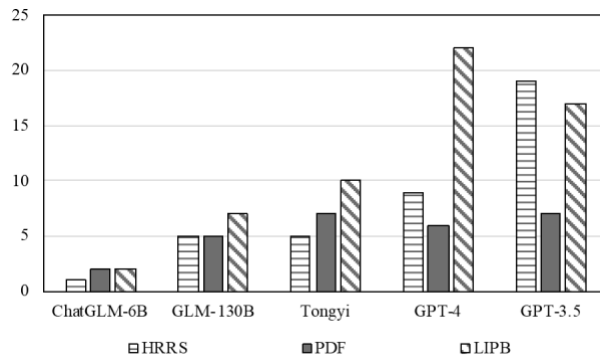


Figure 2: Performance of LLMs in inventory analysis (y-axis: number of raw materials generated)

To assess the comprehensiveness of the emission inventories produced by the model, Figure 2 compares the quantities of raw materials generated by the different LLMs and manually confirmed to be correct. Noticeable variations exist in the generative capacity across different models. ChatGLM-6B yielded a comparatively smaller number of inventories, which could be attributed to the smaller scale of the training corpus and model parameters employed. Conversely, GPT-3.5 and GPT-4 demonstrate superior performance, showcasing greater alignment and comprehensiveness akin to expert-derived inventories. This observation underscores the pivotal role of the model's knowledge base in achieving effective generation.

PCF Estimation Effectiveness

We proceed to examine the efficiency of PCF estimation achieved by AutoPCF. This assessment encompasses both result reliability and cost-effectiveness. The PCF results estimated by AutoPCF based on direct activity data generation with various LLMs are compared in Table 1. In order to reflect the effectiveness of the results, we use the data provided by the actual production enterprises as the basis for carbon footprint accounting using the Energy Expert platform (<https://energy.alibabacloud.com/>), and compare and analyse the results of the accounting as an expert value with the results output by AutoPCF.

While the model may occasionally generate emission inventories with omissions or irrelevant items in comparison to the comprehensive inventories compiled by experts, the ensuing PCF outcomes remain within an acceptable range. Specifically, AutoPCF using GPT 3.5 exhibited consistent estimation results for HRRS and PDF, aligning closely with expert estimations and showcasing its substantial comprehension of LCA principles. In contrast, the application of GLM-130B within AutoPCF yielded notably extreme estimations for these same three products. This suggests that GLM-130B might lack pertinent training data within its pretraining stages.

	HRRS	PDF	LIPB
Expert	2.2 (2.2-3.0)	20.9 (16.2-27.2)	10.8 (4.0-22.0)
GPT-3.5	2.4 (2.1-3.3)	19.0 (12.1-30.9)	25.0 (21.0-27.1)
GPT-4	5.7 (0.1-8.7)	13.1 (0.9-31.1)	25.0 (5.7-73.6)
Tongyi	20.9	37.4	31.4
Qianwen	(-)	(24.3-38.7)	(29.7-41.5)
GLM-130B	47 (0-403.0)	61.0 (2.6-198.6)	114.1 (2.3-1023)
ChatGLM-6B	2.2 (1.9-3.8)	4.2 (1.8-5.7)	3.1 (0-5.7)

Table 1: PCF estimation results of the three selected products. Unit: kgCO₂-eq/kg. Values in parentheses are uncertainty ranges.

To rigorously assess the stability and confidence of the proposed AutoPCF, we examined the uncertainty ranges of both experts' estimations and AutoPCF outcomes, as depicted in Table 1. Among experts, uncertainties primarily emanate from the determination of activity data and the selection of EFs. Variations can arise due to missing data or subjective judgments introduced during the estimation process. For AutoPCF, the outcomes produced by LLMs inherently exhibit stochastic behavior.

To further demonstrate the generalisability of AutoPCF, we performed a GPT3.5 analysis on a further 10 products. The results, as displayed in Figure 3, showed that the ratio of test result values to the reference value ranged from 25% to 211%. While some of the outcomes exhibited considerable variance, some products showed values that were in close proximity to the reference value. This confirms that AutoPCF can be utilised to determine the carbon footprint of other products.

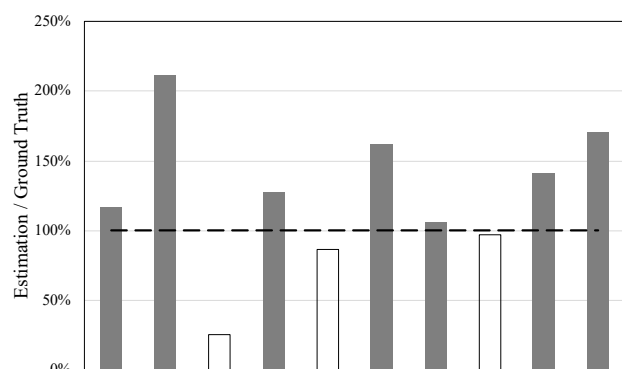


Figure 3: Comparison between tested and reference values for 10 types of products

Furthermore, AutoPCF has effectively demonstrated significant cost reduction by automating the estimation process through programming. The carbon footprint estimation for a product can now be accomplished within a matter of minutes, a stark contrast to the expert method that consumes hours or days. It is essential to recognize that even experts cannot achieve absolute consistency in PCF estimation, given the intricate and variable nature of production processes and raw materials. The primary objective here is to attain the most comprehensive and inclusive estimation of a product's carbon footprint. In light of the estimated PCF, standardized default values for carbon footprints can be derived and utilized for environmental disclosure. This is particularly valuable for producers who may encounter challenges in furnishing credible and transparent PCF calculations.

Conclusion and Future Work

In this paper, AutoPCF is proposed for PCF estimation, aiming to enhance efficiency and reduce costs of LCA-based PCF estimation in the following way:

1) Leveraging the exceptional capabilities of LLMs in language comprehension and knowledge extraction to assist LCA practitioners in quickly identifying emission inventories during the PCF accounting process.

2) A deep learning-based matching model is developed to replace the experience of experts, which can be well applied to the generation of activity data and the matching of EFs.

3) By employing the automatic PCF modeling approach, it becomes possible to rapidly estimate the carbon footprint of a specific product category with minimal data input.

Future improvements on the models and implications for emission reduction and carbon neutrality path are worth to be conducted. Future work can focus on enhancing the performance and accuracy of the underlying deep-learning models and LLMs used in AutoPCF. This may involve exploring advanced architectures, incorporating more diverse training data, and fine-tuning the models to specific industry sectors. Especially, the LLM models can be improved by constructing a local knowledge base and training an LLM model specially designed for PCF estimation. In addition, a PCF calculator is expected to be launched in the future, containing a wide range of products for various industries.

References

- Agathokleous, E., Saitanis, C.J., Fang, C., Yu, Z. 2023. Use of ChatGPT: What does it mean for biology and environmental science? *Science of The Total Environment* 888, 164154. doi.org/10.1016/j.scitotenv.2023.164154.
- Algren, M., Fisher, W., Amy E. 2021. Machine learning in life cycle assessment. doi.org/10.1016/C2018-0-02415-9.

- Balaji, B., Vunna, V.S.G., Guest, G., Kramer, J. 2023. CaML: Carbon Footprinting of Household Products with Zero-Shot Semantic Text Similarity. doi.org/10.1145/3543507.3583882.
- Brown, T.B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D.M., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., Amodei, D. 2020. Language Models are Few-Shot Learners. arXiv:2005.14165.
- Curran, M.A., Laboratory (US), N.R.M.R. 2006. life cycle assessment: principles and practice. U.S. Environmental Protection Agency, Washington, DC, EPA/600/R-06/060, 2006. <https://nepis.epa.gov/Exec/zyPDF.cgi/P1000L86.PDF?Dockey=P1000L86.PDF>. Accessed: 2023-08-23.
- Du, Z., Qian, Y., Liu, X., Ding, M., Qiu, J., Yang, Z., Tang, J. 2022. GLM: General Language Model Pretraining with Autoregressive Blank Infilling. 320-335. doi.org/10.18653/v1/2022.acl-long.26.
- Han, Y., Gopal, A., Ouyang, L., Key, A. 2021. Estimation of Corporate Greenhouse Gas Emissions via Machine Learning. arXiv.2109.04318.
- Kronborg Jensen, J. 2012. Product carbon footprint developments and gaps. *International Journal of Physical Distribution & Logistics Management*, 42(4): 338-354. doi.org/10.1108/09600031211231326.
- Li, Z., Ning, H. 2023. Autonomous GIS: the next-generation AI-powered GIS. arXiv.2305.06453
- Liu, G., Liu, J., Zhao, J., Qiu, J., Wu, Z., Mao, Y., Wen, F. 2022. Real-Time Corporate Carbon Footprint Estimation Methodology Based on Appliance Identification. *IEEE Transactions on Industrial Informatics*. 1-1. doi.org/10.1109/TII.2022.3154467.
- Liu, Z., Deng, Z., Davis, S.J., Giron, C., Ciais, P. 2022. Monitoring global carbon emissions in 2021. *Nature Reviews Earth and Environment* 3, 217-219. doi.org/10.1038/s43017-022-00285-w.
- McKinnon, A.C. 2010. Product-level carbon auditing of supply chains: Environmental imperative or wasteful distraction? *International Journal of Physical Distribution and Logistics Management* 40, 42-60. doi.org/10.1108/09600031011018037.
- Niccolucci, V., Rugani, B., Botto, S., Gaggi, C. 2010. An integrated footprint based approach for environmental labelling of products: the case of drinking bottled water. *International Journal of Design & Nature and Ecodynamics* 5(1) : 68-75. doi.org/10.2495/DNE-V5-N1-68-75
- Radford, A., Narasimhan, K., Salimans, T., Sutskever, I. 2018. Improving Language Understanding by Generative Pre-Training. https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf. Accessed: 2023-08-23.
- Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., Sutskever, I. 2019. Language Models are Unsupervised Multitask Learners. https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf. Accessed: 2023-08-23.
- Reimers, N., Gurevych, I. 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. arXiv: 1908.10084.
- Shinde, R., Froemelt, A., Kim, A., Hellweg, S. 2022. A novel machine-learning approach for evaluating rebounds-associated environmental footprint of households and application to cooperative housing. *Journal of Environmental Management* 304, 114205. doi.org/10.1016/j.jenvman.2021.114205.
- Wernet, G., Bauer, C., Steubing, B., Reinhard, J., Moreno-Ruiz, E., Weidema, B. 2016. The ecoinvent database version 3 (part I): overview and methodology. *Int J Life Cycle Assess* 21: 1218-1230. doi.org/10.1007/s11367-016-1087-8.
- Zeng, A., Liu, X., Du, Z., Wang, Z., Lai, H., Ding, M., Yang, Z., Xu, Y., Zheng, W., Xia, X., Tam, W.L., Ma, Z., Xue, Y., Zhai, J., Chen, W., Zhang, P., Dong, Y., Tang, J. 2022. GLM-130B: An Open Bilingual Pre-trained Model. arXiv: 2210.02414.
- Zhao, B., Shuai, C., Qu, S., Xu, M. 2022. Using Deep Learning to Fill Data Gaps in Environmental Footprint Accounting. *Environ. Sci. Technol.* 56, 11897-11906. doi.org/10.1021/acs.est.2c01640.
- Zhu, J.-J., Jiang, J., Yang, M., Ren, Z.J. 2023. ChatGPT and Environmental Research. *Environ. Sci. Technol.* doi.org/10.1021/acs.est.3c01818.