NOT PEER-REVIEWED

Original Article

A systematic transparency assessment framework for life cycle background database to address three-level black boxes

Lili Sun[1], Hang Yu[1], Yiping Zhang[1], Pengfei Wang[2], Hetian Zhu[1], Lingxi Xie[2], Hanchang Wang[2], Yang Wen[1], Yi Ding[1], Xiaoqian Liu[2]*

[1]CNOOC Research Institute Ltd., Beijing 100028, China.
[2]Carbon Neutrality Future Technology College, Sichuan University, Chengdu 610065, Sichuan, China.

Correspondence: Dr. Xiaoqian Liu, Carbon Neutrality Future Technology College, Sichuan University, Chengdu 610065, Sichuan, China. Email: xqliu1368@163.com

Received: 05 May 2026 | Approved: 15 May 2026 | Online: 15 May 2026

Abstract

Life cycle background database (LCBD) is central to carbon footprint and life cycle assessment (LCA) modeling, yet transparency remains inconsistent across mainstream (ecoinvent, GaBi, USLCI) and emerging (TianGong, HiQLCD) databases. A lack of unified disclosure standards results in a pervasive “three-layer black box” across product coverage, model traceability, and data quality documentation. This opacity undermines the reproducibility of the results and hinders database selection. To address this, we systematically surveyed publicly available database information and constructed a full-process transparency chain: “results → model → unit processes/datasets → input and output inventory → raw data and processing.” The findings are as follows. Specifically, first, product quantity claims are frequently inflated through repetitive combinations of processes, masking limited actual diversity. Second, most databases lack reverse traceability from results to raw data, resulting in

DOI: 10.20517/scierxiv202605.0317.v1

poor model reproducibility, reflecting a deficiency in basic standards. Data quality assessment methods are often misaligned with model logic, rendering the assessments themselves opaque and impractical for integration into user models. We propose the first end-to-end transparency assessment framework for LCBD, identifying key gaps and defining disclosure requirements. The framework includes a fact sheet for developers and a questionnaire tool for users, enabling evidence-based database selection. This work advances the methodological foundation for systematic transparency assessment, directly contributing to enhanced credibility, reproducibility, and standardization in the development and application of LCBD.

Keywords: Life cycle assessment, life cycle database, transparency, traceability, data quality

INTRODUCTION

Life cycle assessment (LCA) is defined as the compilation and evaluation of the inputs, outputs, and potential environmental impacts of a product system throughout its life cycle-spanning raw material extraction, production, transportation, use, and end-of-life disposal[1,2]. Presently, LCA has been widely deployed in product design[3], technological performance evaluation[4], and supply chain management[5] and has been established as a pivotal instrument for driving the global green and low-carbon transition[6]. Concurrently, the quantification of product carbon footprints via LCA methodologies has emerged as a focal point of international regulatory frameworks[7].

Mandatory carbon footprint requirements for products have been stipulated in key legislative instruments, including the EU Regulation 2023/1,542 concerning batteries and waste batteries, the draft revised Construction Products Regulation, and France’s public procurement policies for photovoltaic modules. The enactment of these policies signals a new phase in which LCA and carbon footprint assessment are evolving from voluntary evaluation tools to compulsory regulatory mandates[8]. This shift may further develop into technical barriers that shape the landscape of international trade rules and market access conditions[9].

Life cycle background database (LCBD) is developed by investigating thousands of unit processes associated with the production of primary energy carriers and raw materials. Raw data from these unit processes are collected, processed, and cross-validated to generate unit process datasets (UPDs)[10]. These UPDs are then integrated with the background database to construct complete life cycle models, and dedicated LCA software aggregates and calculates the corresponding Life Cycle Inventory (LCI) results. When stored in database, these results form aggregated process datasets (APDs)[11]. In turn, such datasets can function as background databases for upstream processes, supporting life cycle modeling, computation, and analysis for downstream products[12].

The reliability of LCA and carbon footprint outcomes is highly dependent on the underlying life cycle background database[13]. Nevertheless, carbon footprint results derived from different databases and research studies exhibit substantial discrepancies[14,15]. Taking China's average grid carbon footprint factor in 2023 as an illustrative case, the China Electricity Council (CEC) reports a value of 0.6205 kg CO_2 e/kWh[16]. It was found through the database that the China Life Cycle Database (CLCD) yields a result of 0.617 kg CO_2 e/kWh, whereas the ecoinvent and GaBi produce values of 0.962 kg CO_2 e/kWh and 0.7611 kg CO_2 e/kWh, respectively, representing deviations of 55.04% and 22.66% relative to the CEC figure. As a fundamental energy source, electricity is extensively employed in industrial product manufacturing[17]. However, marked inconsistencies persist between current grid carbon footprint factors and realistic values, and no scientifically rigorous justification or detailed elaboration has been provided by the relevant database in publicly available information. This directly casts doubt on the credibility of life cycle assessment results for all products, thereby posing a fundamental challenge to the scientific robustness of the LCA methodology[18]. More critically, persistent opacity in database information may further trigger trade disputes in international commerce arising from divergent carbon footprint accounting outcomes[19]. Accordingly, the establishment of a standardized information disclosure and traceability system for LCBD is urgently required to enhance database transparency[20].

However, the achievement of transparency in LCBD is confronted with substantial challenges. Primarily, the ISO 14040 and ISO 14067 standards, which serve as the cornerstones of the LCA methodological framework, explicitly acknowledge the importance of transparency. However, the standards fail to stipulate concrete methodologies for attaining database transparency. They neither delineate specific operational protocols nor establish systematic and comprehensive regulations for LCBD construction workflows, data quality control, and traceability requirements[1,2]. Second, a systematic review of the LCA literature reveals that contemporary research predominantly focuses on environmental impact quantification[21], uncertainty assessment[22], and the application of LCA to renewable energy systems and related domains[23]. Terms such as background database and transparency have not yet emerged as core thematic concerns within the field[24,25], and few studies have investigated what specific information should be disclosed to improve database transparency[26]. More significantly, the current LCBD landscape is characterized by a supply structure dominated by a small number of key databases[27]. Within this paradigm, strong market competition and strict intellectual property controls are lacking. As a result, major databases face neither sufficient external pressure nor adequate internal motivation. This environment prevents major databases from systematically disclosing their full construction rationale, data processing workflows, and detailed quality control procedures[28].

Recently, database transparency has attracted increasing attention, and relevant research has gradually emerged. For instance, Guo[29] adopted the findability and accessibility of raw data as core indicators to assess traceability and transparency and analyzed six major global LCI databases, including ecoinvent, GaBi, and CLCD. This study provided actionable quantitative criteria for evaluating LCBD traceability and partially supplemented transparency-related research. Nevertheless, it simplified database transparency to mere findability and accessibility of data sources. In essence, it reduced the complex transparency issue to a one-dimensional assessment of traceability, and a systematic interpretation of database transparency remains absent. Database transparency is not confined to traceability at the data source level. At the database level, model construction logic, data sources, and data quality assessment and control procedures should be clearly defined to guarantee transparency and accuracy[30]. At the individual model level, dataset documents should disclose the full calculation logic from raw data collection and unit process modeling to the final results[31]. Research indicates that only 35% of LCA studies disclose complete LCI inventories, while fewer than 2% make their models publicly available[32]. As an essential component of database transparency, completeness also demands considerable attention, yet the current completeness disclosure of the LCI database is far from satisfactory. Some datasets suffer from missing input-output data and incomplete unit processes, which significantly undermines database transparency and credibility[33]. Therefore, database transparency is not limited to data source traceability but is a comprehensive concept covering overall model and data quality disclosure, the full-cycle calculation logic of individual models, and inventory completeness.

Accordingly, this study aims to develop a systematic framework for evaluating database transparency. From the perspective of publicly accessible information, six representative global and emerging cross-sector databases are selected, namely, ecoinvent, GaBi, CLCD, TianGong, USLCI, and HiQLCD. Based on life cycle model outputs and computational logic, this study assesses public information disclosure across each database and proposes relevant recommendations. The findings reveal several critical issues. First, no clear and transparent statistical definitions govern the overall scale and product coverage of LCBD. Although global LCBD is widely perceived to cover tens of thousands of products, only database from six countries include more than 1,000 basic commodities, and the maximum product coverage for any single region is approximately 4,000 items, with the large reported number derived mainly from repeated combinations of identical unit process datasets. Second, in terms of traceability and completeness, most LCBDs fail to support stepwise tracing from final results back to the raw data and processing procedures, leading to poor result reproducibility, whereas some databases suffer from conceptual errors and substantial inventory gaps, indicating a lack of unified basic standards worldwide. Third, current data quality methods and documentation formats are incompatible with life cycle model structures, as quality assessment results remain in a ‘black box’, lacking traceability and cannot be integrated with user models, resulting in ineffective quality governance across the field.
This study contributes three key innovations. First, in terms of data coverage, it significantly advances the breadth and depth of transparency information collection for the life cycle database, overcoming the limitations of prior research that relied on narrow and superficial information. Most existing studies assess only the transparency

of raw data[31]. This work further gathers comprehensive information covering dataset volume, model structure, documentation, and quality evaluation, supporting a better understanding of real database transparency. Second, this study innovatively assesses LCA database transparency from the user perspective as well as the LCA results and calculation logic of life cycle models. It thereby addresses the limitations of prior research that relies primarily on a developer perspective. Existing research often focuses on data format optimization, AI applications, and platform tools[28]. In this study, the results are traced back to the original data to identify transparency shortcomings, providing a scientific basis for database selection and encouraging more robust and trustworthy database development. Third, in terms of research content, this study establishes the first operational definition and characterization of traceability-based transparency for LCBD, filling the gap in unclear and non-standardized assessment procedures. Earlier studies adopted semi-quantitative scoring without detailed practical criteria[34]. This work proposes a traceability chain from calculation results and life cycle models to unit process datasets, inventory data, and raw data with processing procedures, with specific indicators designed for each level to form an operational analytical framework for future research. Based on these findings, a traceability-focused scoring system is developed, which offers clear improvement guidance for developers and rational selection support for users, ultimately promoting more standardized and transparent life cycle data management.

COMPARISON AND ANALYSIS OF GLOBAL DATABASE

Database overview

LCDBs are core infrastructure for quantifying the environmental performance of products. A key characteristic is their coverage of diverse industries and product categories[27]. Database limited to one sector cannot independently complete full life cycle assessment and relies on other databases for background data[35]. This study selects the most widely used mainstream database, including CLCD × WebLCA, ecoinvent, GaBi, IDEA, and USLCI, and incorporates two emerging databases, HiQLCD and TianGong, with their basic features systematically summarized in Table 1. By compiling key information such as dataset scale, documentation completeness, and data quality assessment methods, this study comprehensively profiles each comprehensive database and lays the groundwork for subsequent transparency evaluation.

Table 1. Systemic characteristics of the database.

Database	CLCD × WebLCA	ecoinvent	GaBi	IDEA	USLCI	HiQLCD	TianGong
Institution	Sichuan University, Yike Environmental Technology Co.,Ltd	the Swiss research institutions, EPFL, Empa, Agroscope, the Paul Scherrer Institute	IKP, University of Stuttgart	JEMAI, AIST	National Renewable Energy Laboratory	EC Digit	Tsinghua University, *et al.*
Country/Region	China	Global	Global	Japan	USA	China	China
Number of APDs	4,000	0	20,577	3,800	0	85,775	0
Number of UPDs	0	25,000	0	0	995	0	1,839
Fee-based	√	√	√	√	×	√	×
Documentation Format	WebLCA,ILCD-compliant	ecoSPOLD	ILCD-compliant	ILCD-compliant	JSON-LD	ILCD-compliant	TIDAS, ILCD-compliant
Data Quality Assessment Method	WebLCA-DQR,Pedigree matrix + uncertainty analysis	Pedigree matrix + uncertainty simulation	PEF-DQR-like	Unknown	USEPA-DQR	Unknown	Unknown

Methodology Guide Publicly Available	√	√	×	×	×	×	√
Development Tool/Platform	eFootprint (SaaS)	ecoEditor (Desktop)	GaBi (Desktop)	Unknown	openLCA (Desktop)	HiQEditor (SaaS)	TianGong (SaaS)
Open Participation	√	×	×	×	×	×	√

The compilation of publicly available information reveals clear differences in database type and dataset volume. Ecoinvent, USLCI, and TianGong are unit process databases containing only UPDs, while the remaining databases are aggregated process databases dominated by APDs. Dataset scales vary widely across databases, and some do not disclose unit process data at all, indicating inconsistent transparency at the dataset level. This inconsistency and ambiguity in disclosing datasets complicates horizontal comparisons of the actual data modeling depth and core capacity of the database[36].

Number of databases per country

Given the requirement for comprehensive regional data coverage in LCA database, this study analyzed the geographic distribution of datasets. For the multi-country database including ecoinvent and GaBi, the 20 countries with the highest dataset counts were examined. For national databases such as USLCI, IDEA, CLCD, TianGong, and HiQLCD, the total number of domestic datasets was compiled. All results were benchmarked against open-access data from the Global LCA Data Access network (GLAD).

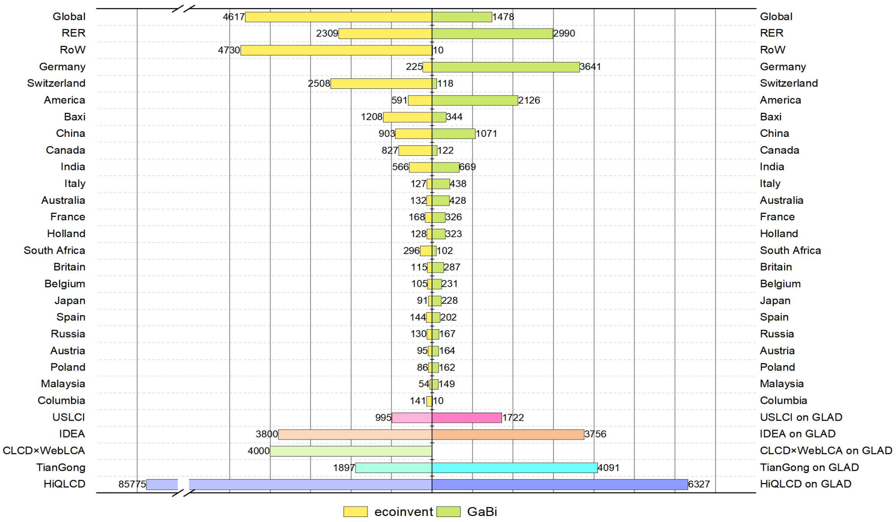

Figure 1. Number of countries in the datasets for each database.

Figure 1 reveals a severe imbalance in global life cycle data resources. Only a few countries, such as Germany, Switzerland, the United States, China, and Japan, possess foundational inventories with more than 1,000 datasets in mainstream databases, preliminarily enabling them to support LCA analysis for their primary industries. The vast majority of countries, especially developing nations, have extremely scarce data records, which constrain the completeness and accuracy of product carbon footprint accounting. Consequently, this study recommends that the database regularly publish reports on dataset counts and sectoral distribution by country or region, transparently communicating their data accumulation progress and gaps, thereby providing guidance for international data collaboration and development priorities.

Granularity and duplication of products

While country coverage reflects overall data availability, the granularity of product datasets plays an equally important role in balancing representativeness and practical usability[37]. To explore in depth the granularity strategies of each database and their impact on the dataset scale, this study selected the electricity power sector for micro-level analysis, with the results shown in Table 2.

Table 2. Statistical overview of power products and their datasets across various databases.

Database		**CLCD**	**ecoinvent**	**GaBi**	**IDEA**	**USLCI**	**HIQLCD-1.4.0**	**TianGong**
Total Entries		4,000+	2,508	3,641	3,800+	995	85,775	1,897
Electricity	Total Entries	70	60	67	-	11	3,301	18
	Technology Typology (*n*)	30	7	35	-	7	24	16
	Specification Typology (*n*)	4	3	2	-	4	16	6
	Region Typology (*n*)	41	1	1	-	4	38	5

Note: “-” indicates missing data.

The statistics indicate that all databases have subdivided power products to varying degrees across dimensions such as technology, specification, and region. However, the subdivision strategies of some databases may lead to a combinatorial explosion in dataset counts. For instance, HiQLCD generates as many as 85,775 datasets for electricity products alone through such approaches. While this practice of generating numerous derivative datasets through attribute cross-combination theoretically enhances model adaptability to specific contexts, if the underlying unit process data are not correspondingly increased or substantially adjusted for each derivative scenario, it results in a large number of highly similar datasets or those with redundant core data. This phenomenon can easily mislead users who focus solely on total dataset counts, leading to misunderstandings about the actual capacity and unique information content of a database. The discrepancy in dataset counts reported by HiQLCD on its own platform versus the GLAD platform, differing by an order of magnitude, further highlights the issue of scale transparency.

To overcome the limitations of current database users who rely solely on entry counts and to drive database development from pursuing scale expansion toward enhancing substantive data quality and transparency, this paper proposes constructing an "information density index" as a core evaluation tool. The index aims to quantify the effective information content, which is represented by three sub-indicators.

$$I_1 = \frac{N_{pt}}{N_{UPDs}} \tag{1}$$

The indicator I_1 represents the ratio of the number of independent product types to the number of UPDs, which reflects the average diversity of products supported by each core data unit in the database. A lower I_1 generally indicates that the database builds product models based on more fundamental, highly reusable unit processes, with original data concentrated at the foundational level.

$$I_2 = \frac{N_{APDs}}{N_{UPDs}} \tag{2}$$

The indicator I_2 represents the ratio of the number of APDs to the number of UPDs, which reveals the extent to which a database generates derivative datasets by incorporating different upstream background data. A high I_2 suggests that a large volume of aggregated process datasets may be built upon a limited core library of unit process datasets, indicating potential information redundancy through combinatorial derivation.

$$I_3 = \frac{N_{DS}}{N_{UPDs}} \tag{3}$$

The indicator I_3 represents the ratio of the number of cited data sources to the number of UPDs. It assesses the richness of supporting documentation for UPDs, serving as a key measure of data reliability, robustness, and the thoroughness of data development. A higher I_3 generally indicates that the data have undergone more extensive cross-validation across multiple independent sources.

Model traceability and integrity

Model tracing methods

To address the black box of life cycle models that leads to the non-reproducibility of the LCA results, this section proposes a standardized traceability chain derived from the process-based life cycle model and its computational logic, which is consistent with the transparent traceability chain from the LCA results to the raw data and processing procedures. The full-process transparency chain developed in this study follows results → model → unit processes/datasets → input and output inventory → raw data and processing. As illustrated in Figure 2, the dashed lines represent the forward flow of data processing, whereby raw data are collected, processed, and incorporated into model development, ultimately generating the final LCA results. In contrast, the solid lines indicate the reverse traceability sequence for LCA database transparency, which allows for the backtracking of LCA results to their original raw data sources.

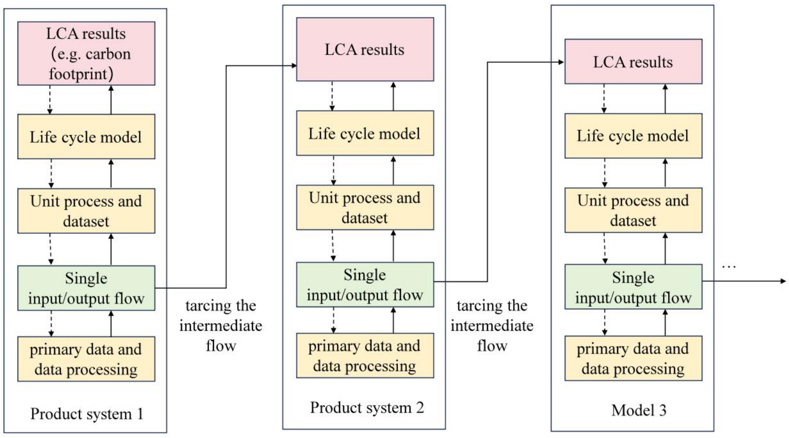

Figure 2. Model tracing methods and sequences; product systems 1-3 are examples of lifecycle models for different product systems.

Transparency in the life cycle database means that database providers should disclose sufficient information to ensure the traceability of the entire process of data collection, processing and calculation, which constitutes the fundamental and paramount transparency requirement for such database. In practical life cycle database development, however, full information disclosure is not always feasible, as the complete traceability of data is subject to two key constraints. First, some unit process data involve corporate trade secrets or third-party intellectual property rights and thus cannot be disclosed to the public. Second, data processing workflows are highly complex; documenting all processing algorithms and procedures in full requires considerable time and relies on robust functional support from specialized software tools. Given these constraints, complete end-to-end traceability in life cycle database is difficult to achieve. Therefore, this study proposes differentiating the degree of transparency (DOT) for each link in the traceability chain based on specific disclosure recipients and scenarios. According to the disclosure scope and accessibility of information at each link in the traceability chain, we define a five-level transparency scale for LCBD, with transparency decreasing progressively from a to e.

Applying the aforementioned traceability methodology and transparency level definitions, this study takes the accessible perspective of database users as the core and collects publicly disclosed information for each link in the traceability chain across six target LCBD. Data collection is conducted through systematic access to the official websites of each database and manual retrieval and collation of open information, with the transparency status of the database derived and presented in Table 3.

Table 3. Database traceability and transparency status reports

Database		**CLCD ×WebLCA**	**ecoinvent**	**GaBi**	**USLCI**	**TianGong**	**HiQLCD**
Results and life cycle models	1) Results transparency	3	3	3	n.a.	n.a.	3
	2) Document transparency	a (Model documentation)	a (Unit process documentatio	a (Model documentatio	a (Unit process docu	a (Unit process documentation)	a (Model documentation)

			n)	n)	mentation)		
	3) Subroutines included in the document disclosure model	a	n.a.	a*	n.a.	n.a.	unknown
	4) Document navigation	a	b	a *	a *	n.a.	unknown
	5) Visibility of input and output lists	a	a	unknown	a	a	unknown
	6) Visibility of list values	b	a	unknown	a	a	unknown
Unit Process and list data	7) Searchability of the unit process data list	b	Mentioned in documentation/no list on unit process level	Listed on model level, but not unit process level	n.a.	n.a.	Mentioned in documentation/no list on model level
	8) Data processing for individual line items	b	d?	d?	n.a.	n.a.	d?

Notes: "a" means the corresponding feature is publicly accessible. "b" reveals that access is limited to subscription users. "c" indicates that the feature is available only to the reviewers. "d" means that access is restricted exclusively to the author. "e" shows that the feature is not accessible even to the author. The notation "n.a." represents that the feature is not currently available. "Unknown" reveals that feature accessibility cannot be determined from publicly available information, and the asterisk "*" indicates that the feature has been incompletely implemented. All assessments are based solely on public information obtained from the official platforms of each LCBD, and the actual transparency status may require further disclosure by database developers for definitive verification.

Traceability aggregate results and models

Building on the traceability methodology and transparency grading criteria proposed in Section 3.1, this section presents a systematic comparative analysis of the transparency performance of six target LCBD across all the links in the traceability chain. The analysis identifies the disparities among the database in terms of model traceability, document disclosure and navigation functionality, thereby providing empirical data to underpin the subsequent formulation of transparency disclosure specifications for LCBD.

With respect to the transparency of life cycle results, as presented in row 1 of Table 3, CLCD, ecoinvent, GaBi and HiQLCD have developed life cycle models and generated corresponding calculation outcomes, which are accessible only to paying users. In contrast, TianGong and the USLCI have not established complete life cycle models; thus, no relevant LCA calculation results have been produced for these two databases. Turning to documentation, as shown in row 2 of Table 3, all the aforementioned databases make their documentation available to users: CLCD, GaBi and HiQLCD offer model documentation, while ecoinvent, USLCI and TianGong provide unit process documentation. Model documentation integrates relevant information across multiple unit processes included in a life cycle model, whereas unit process documentation focuses solely on the information of individual unit processes.

Beyond such categorical differences, the content structure of the documentation, listed in rows 3 and 4 of Table 3, varies across the database. For model documentation, CLCD automatically extracts all model unit processes to generate a unit process table while providing representative descriptions and input-output data for each unit process with accessible data sources. Hyperlinks are also embedded for each unit process input, enabling navigation to the corresponding upstream background life cycle models for in-depth traceability. By comparison, GaBi displays model structures via static diagrams or flowcharts, failing to fully present all integrated unit processes or support viewing their input-output data. Unlike CLCD and GaBi, HiQLCD does not provide unit process tables or hyperlinks to upstream background processes in its public documentation, making it impossible to verify the effective linkage between background data and life cycle models. For unit process documentation, ecoinvent calculates the life cycle results for each unit process and its corresponding products and provides only individual unit process documents that allow navigation between one another. Neither the USLCI nor TianGong has linked respective unit processes to construct complete life cycle models or produce LCA results, with only unit process documentation accessible. Accordingly, these three databases provide information only on individual unit processes, with no access to unit process tables; thus, the subroutines included in the document disclosure model are marked as “n.a.”. With respect to document navigation functionality, partial connections have been established between unit processes in the USLCI, enabling navigation between some of them. In contrast, in TianGong, no interconnections exist between unit processes, precluding any such navigation.

This section identifies varying transparency levels across traceability links in life cycle database, with inconsistent disclosure content across different platforms. Most unit process databases do not integrate into complete life cycle models or generate LCA results; thus, they are unable to disclose model-included unit processes or enable cross-model traceability navigation. Unlike unit process database, the aggregated process database is built on life cycle models, yet most fail to fully disclose their model-included unit processes or provide traceability access to all upstream background processes, falling short of the basic transparency requirements for LCBD. Based on the above findings, relevant stakeholders are recommended to refer to the transparency assessments in Table 3 and formulate differentiated tiered traceability disclosure specifications aligned with LCBD types (unit/aggregated process database) and user group needs (paying users, general users, reviewers). For the aggregated process database, mandate full disclosure of model-included unit process information and traceability links to all upstream background processes. For the unit process database, standardize the content structure of unit process documentation, promote interconnection between unit processes, and promote the gradual development and disclosure of complete life cycle models. Database developers are also required to disclose relevant information accurately and comprehensively in accordance with these specifications to enhance LCBD transparency, thereby supporting users in selecting appropriate databases and further alleviating the prevalent black box issue in LCBD.

Traceability unit process datasets

The traceability of UPDs is a critical challenge impacting the transparency of the current LCBD. Specifically, methodological hurdles exist regarding the visibility of input-output inventories, detailed in row 5 of Table 3. Furthermore, the numerical transparency of inventory values, as shown in row 6 of Table 3, presents another challenge. Under the public information disclosure framework, marked discrepancies exist among the leading LCBDs. GaBi and HiQLCD typically categorize their unit process inventory data and specific values as ‘unknown’, rendering their detailed underlying inventories inaccessible to external users. Ecoinvent, in contrast, permits paying subscribers to access inventory tables and their associated values. Free UPDs, such as the USLCI and TianGong, publicly display their unit process inventory tables and specific values. CLCD×WebLCA offers input-output inventory tables for each unit process, complemented by descriptive information on their technical, geographical, and temporal representativeness within its model documentation; however, it does not fully disclose detailed inventory values, which are generally restricted to authorized users who have purchased the unit process or successfully completed a review process. These differing disclosure practices directly contribute to the varying degrees of “black box” characteristics among LCBD, affecting model traceability and completeness.

Significant disparities are also evident in the accessibility of UPD sources listed in row 7 of Table 3 across various LCBD. GaBi typically provides only a generalized description of referenced data sources at the model level and lacks detailed source lists for specific unit processes. The construction of UPDs in USLCI and TianGong still has room for improvement in terms of multi-source data cross-validation and completeness of source inventories. Ecoinvent and HiQLCD offer relatively limited access to data lists; they generally mention the types of data sources employed in their documentation but fail to provide clear, accessible lists within specific unit process documents. This disclosure method “Mentioned in documentation/no list on unit process level” impedes users’ ability to trace the origin of specific data. Conversely, CLCD×WebLCA enables users to access the multiple data sources underpinning unit process construction, establishing a robust foundation for in-depth data verification and source evaluation.

Differences in transparency at the level of individual inventory data and data processing in row 8 of Table 3 are most pronounced, directly influencing data quality and the reproducibility of the LCA results. Some unit process databases have room for improvement in data source support and processing workflow disclosure. The derivation basis for their unit process data is relatively homogeneous, and the processing flow from raw data to final inventory values is not fully public, which increases the difficulty of ensuring data integrity and technical representativeness to a certain extent. Although USLCI and ecoinvent compile individual inventory entries and data from multiple sources, they do not furnish detailed explanations of data source composition or the complete processing path, thus hindering users' ability to trace and reproduce the data processing logic. GaBi offers only a general description of dataset-level data sources and processing, without deconstructing these for individual unit process and their corresponding inventory data. HiQLCD does not elucidate the processing path detailing how specific inventory data are generated from raw inputs through screening, conversion, and calculation. In contrast, the CLCD×WebLCA system meticulously records and discloses the complete processing workflow from raw data to standardized inventory data, ensuring that the generation of each inventory value is traceable and reproducible. In summary, the construction of UPDs within LCBD should be grounded in the cross-validation of multiple sources, accompanied by the disclosure of specific data integration methodologies and processing pathways to guarantee data reliability and reproducibility. The CLCD×WebLCA system exemplifies an open and collaborative model, facilitating the participation of multiple stakeholders in database development. It mandates that each inventory data entry be systematically integrated and processed from multiple source materials. Its data processing procedures are accessible to authorized users, thereby achieving a significant degree of transparency in the database construction process.

The concept of "traceability" in life cycle methods encompasses two fundamental directions. First, life-cycle process chain traceability involves hierarchical tracing backward from a target process along the product life-cycle path to its upstream associated processes, ensuring the inclusion of all relevant processes within the system boundary. Second, data drill-down traceability pertains to the micro-level examination of data construction for a specific unit process. This entails sequentially tracing the visibility of its input-output inventories, the provenance of their values, and the accessibility of source lists, ultimately aiming for complete transparency of the original source of a single inventory data entry and its associated processing and conversion logic. Comparative analysis of major LCBD shows that most databases offer public visibility in input-output inventory inclusivity and numerical transparency. These aspects meet basic data accessibility requirements. However, deeper traceability is needed for data lists, individual inventory data items, and their processing procedures. As the depth of traceability increases, the granularity of the involved information becomes finer, and the volume of data requiring recording and disclosure expands substantially. This characteristic typically leads to a corresponding reduction in the level of public disclosure. However, it is precisely the transparency of this in-depth information that forms the cornerstone for ensuring the reliability, reproducibility, and scientific rigor of LCA results, directly determining the practical value of LCBD and the credibility of LCA outcomes. Consequently, future LCBD development should transcend mere compliance with basic data format and superficial structural requirements, focusing instead on establishing a fully transparent, end-to-end system. Such a system necessitates not only data visibility but also the independent verifiability and reproducibility of its generation logic.

Based on the established information disclosure standards, database providers are strongly encouraged to clearly define the specific content to be disclosed at different levels of traceability. Given the substantial volume and heterogeneous nature of information required for deep traceability, coupled with the significant associated workload and technical complexity, reliance on manual recording alone is unsustainable. Therefore, the development and integration of relevant software tools are imperative for providing technical support for standardized data recording, systematic processing, and end-to-end traceability.

Life cycle completeness

Building upon advancements in the transparency and traceability of LCBD, the assessment of life cycle completeness (LCC) has emerged as a pivotal indicator for evaluating the systemic quality of the database. The achievement of LCC is contingent upon two fundamental levels. First, at the unit process level, LCC necessitates the comprehensive inclusion of input-output inventories for each unit process. Achieving this necessity is done primarily through the systematic collection and cross-verification of multi-source data, which serves to identify and rectify any inventory gaps inherent in a single data source.

To illustrate this, we conducted a comparative analysis of the LCC for petroleum and gas production across various LCBD, as presented in Table 4. GaBi and HiQLCD present their aggregation processes as either nondisaggregated “black boxes” or provide highly generalized technical classifications, precluding the examination of relevant inventory entries. Consequently, the integrity of their processes cannot be substantiated through inventory analysis. In contrast, ecoinvent’s petroleum and gas unit process inventory system demonstrates a high degree of completeness, encompassing the entire material flow from energy and raw materials to capital equipment and a wide spectrum of emissions, thereby facilitating process integrity verification. However, the USLCI oil and gas unit process, however, lists 119 output entries, detailing an excessive number of specific pollutants. While this granularity might appear comprehensive, it significantly increases data collection difficulty and cost, diminishes data universality and accessibility, and consequently limits its practical utility in subsequent LCA aggregation and computation processes. For oil and gas production, USLCI and TianGong have certain room for optimization in terms of inventory scope and content presentation. This marked disparity in inventory scope and content reveals serious completeness issues within UPDs, underscoring the imperative for a cross-validation mechanism employing multiple data sources within the context of LCC. Specifically, through the systematic integration of multiple data streams, complementary verification of critical material flows and environmental emissions can be performed, thereby mitigating systematic biases arising from single-source limitations and ensuring robust completeness.

Table 4. Completeness of oil and gas production unit processes in the database

Database	CLCD×WebLCA	ecoinvent	USLCI	TianGong
Unit Process	petroleum and gas	petroleum and gas production, onshore	Crude oil, on-shore domestic, at	Production stage; Oil and gas; Drilling

				extraction	
Reference product		petroleum and gas	petroleum	Crude oil	Oil and gas
Input Inventory Count	Energy consumption	2	6	6	0
	Raw materials	2	2	0	1
	Infrastructure	4	2	0	0
	Natural resources	1	2	1	0
Output Inventory Count	Disposal waste	1	4	2	1
	Emissions to air	3	2	99	0
	Emissions to water	2	9	18	0
	Emissions to soil	2	1	0	0

Comparative analysis indicates significant discrepancies in the inventory composition for the same unit process across different databases, posing challenges to guaranteeing unit process completeness:

1. Conceptual ambiguity of unit process definition: The database exhibits differing interpretations of the fundamental concept of unit process. This conceptual ambiguity can lead to highly simplified unit process inventory entries that fail to present complete input-output lists.

2. Completeness of the unit process inventory type: A complete unit process should encompass all essential input inventories, including raw materials, energy consumption, and natural resources, as well as output inventories such as waste awaiting disposal and environmental emissions to the atmosphere, water bodies, and soil.

3. Comprehensiveness of unit process input–output content: The content of unit process

input–output must be comprehensive. Taking ecoinvent as an example, it typically employs multi-source data fusion and cross-validation to construct a relatively complete material balance.

Second, at the model system level, establishing a complete life-cycle modeling chain is paramount. This involves effectively linking the intermediate flows of each unit process to corresponding background data to trace their upstream environmental impacts. The associated unit process input-output inventories must also be clear and complete. Only when the environmental impact of a particular intermediate flow meets predefined cutoff criteria can its linkage be interrupted without significantly compromising the credibility of the final LCA results. CLCD×WebLCA facilitates this by enabling users to publicly view input and output inventories and their names. Paying users can further access inventory values and drill down to the underlying unit process level, thereby verifying process completeness.

Therefore, it is recommended that database providers systematically collect and compare multiple datasets for the same product or production technology to ensure the completeness of process input and output inventories. Furthermore, the collaborative development and promotion of standardized unit process templates are essential to ensure the comprehensive characterization of critical material flows and environmental emission data.

DATA QUALITY ASSESSMENT AND DATABASE DOCUMENTATION

Data quality assessment (DQA) represents a critical component in ensuring the traceability and completeness of LCBD. Concurrently, database documentation serves as the primary medium for conveying data quality information. These two elements collectively form the foundation for transparency concerning the ‘data quality’ dimension of LCBD. This chapter systematically analyzes the current status and limitations of DQA methodologies for LCBD. By examining the practical implementation reflected in database documentation, targeted recommendations for methodological optimization are proposed, thereby contributing to the refinement of the LCBD transparency framework.

Data quality assessment

Following a systematic review of the qualitative requirements for model traceability and completeness, following a systematic review of the qualitative requirements for model traceability and completeness, the challenge of scientifically quantifying and evaluating these attributes to assess overall data quality emerges as a core element in ensuring the reliability and credibility of LCA results[38]. Within the LCA field, various interrelated DQA methods have been developed. The ISO standards establish principled requirements for data quality assessment, providing a foundational theoretical framework. However, they lack specific quantitative indicators or scoring mechanisms, limiting their practical applicability. Building upon the ISO framework, the international life cycle data (ILCD) system pioneered a systematized, semi-quantitative assessment framework across six dimensions[39]. It refined scoring rules and quantification methods for data quality indicators (DQIs) and introduced the 'magnify the weakest link' assessment principle. The product environmental footprint (PEF) method draws on the core principles of the ILCD framework, further streamlining the assessment process with a focus on specific product categories[40]. At the database application level, GaBi's assessment methods include ILCD, PEF, and its own six-dimensional indicator averaging approach[41]. Ecoinvent transforms quality scores into uncertainty coefficients via a pedigree matrix and quantifies overall uncertainty through Monte Carlo simulation[42]. This approach links data quality to statistical uncertainty, in contrast to methods centered on expert scoring, such as the EU's data quality rating (DQR). CLCD×WebLCA employs a combination of pedigree matrix weighting and error propagation models to identify key sources of uncertainty, enabling a dynamic quantitative analysis of data impact within specific models [43]. Despite differences in dimensionality, scoring mechanisms, and results presentation, these methods share the common objective of ensuring the reliability and credibility of LCA outcomes, as depicted in Figure 3.

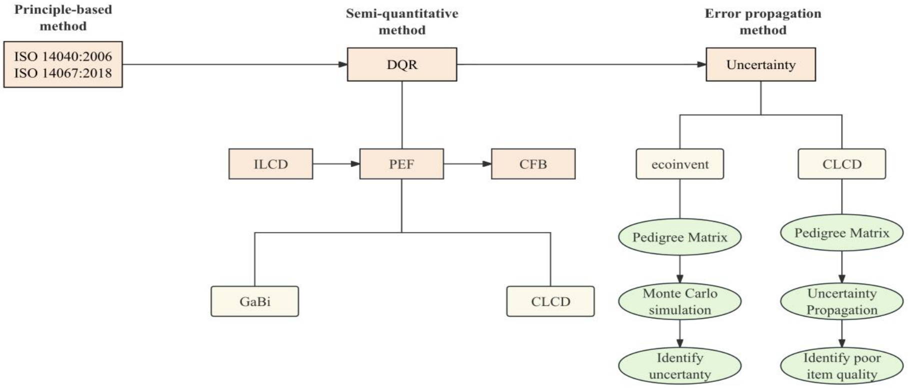

Figure 4. Hierarchy Diagram of Data Quality Assessment Methodologies.

However, current DQA methods face significant challenges regarding theoretical clarity and practical transparency. In methods represented by DQR, substantial heterogeneity exists in how different implementing bodies interpret and apply the same methodology. For instance, GaBi has not explicitly disclosed its weighting calculation logic within the DQR process, whereas the EU New Battery Regulation mandates a weighted average for unit processes contributing ≥ 80% to environmental impacts. In practical applications of DQR across different LCBD, regulatory contexts, and time periods, significant and evolving disparities exist in specific scoring algorithms, weight allocation, and compliance interpretations. Furthermore, DQR has yet to establish a clear, transferable pathway linking front-end qualitative requirements (i.e., traceability and completeness) to back-end quantitative scoring, resulting in a disconnect between qualitative stipulations and quantitative assessment. Conversely, uncertainty quantification methods based on Monte Carlo simulation, while capable of generating probability distributions and confidence intervals for overall results, often fail to identify and pinpoint the specific sources of major uncertainties within the system. Consequently, they provide limited actionable guidance for targeted data improvement, making it difficult to implement precise optimization measures focused on the low-quality life cycle stages with the greatest impact on the final results. A pertinent example is the significant deviation (up to 55%) observed between the electricity carbon footprint factors calculated by ecoinvent and GaBi on the one hand and those from the China electricity council (CEC) and CLCD×WebLCA on the other. Existing DQA systems struggle to effectively address critical questions arising from such discrepancies: Can the underlying mechanism of this deviation be explained through data quality scores? Can the specific low-quality data stages causing deviation be precisely identified? Can the quality of electricity carbon footprint factors across different datasets be scientifically distinguished?

In summary, current DQA methods exhibit shortcomings in theoretical consistency, methodological maturity, and operational transparency. It is therefore recommended that database providers establish detailed, standardized, and actionable DQA scoring rules. Such rules should enable users not only to assess overall data quality and identify critical weak points but also to receive direct guidance on priorities for data quality improvement. Only through the development and adoption of such transparent and diagnostic DQA methodologies can the “black box” between front-end data development and back-end database applications be effectively dismantled, thereby enhancing the scientific robustness and practical utility of LCA.

Database documentation

As the principal conduit for the storage, presentation, and communication of data quality information, the design and specifications of LCBD documentation must adhere to multi-dimensional transparency requirements. Specifically, documentation should facilitate the systematic disclosure of data sources, calculation methods, and critical assumptions. Furthermore, it must demonstrate traceability, completeness, and robust data quality assessment while providing standardized metadata that complies with international benchmarks[44].

Examining current documentation practices in LCBD highlights an issue. Databases such as GaBi, HiQLCD, and TianGong adopt the ILCD format directly, a point illustrated in Table 5. Despite this, empirical verification indicates that their inventory documentation is plagued by pervasive content gaps. This results in a significant deficit in database documentation transparency. The ILCD format, primarily designed to standardize and unify data structures[45,46], prioritizes structural conformity over the intuitive presentation of internal logic. Consequently, its direct application without bespoke adaptation can lead to a relative reduction in information readability, comprehensibility, and overall transparency, thereby impeding users' practical needs for tracing the underlying logic of data generation. Ecoinvent and the USLCI have achieved a rudimentary level of transparency through their documented fields. However, the detailed disclosure of the processing logic for individual inventory data entries remains deficient. The scope of their document fields is more oriented toward meeting basic data visibility requirements, falling short of the deeper transparency necessary for the traceability and reproducibility of data generation logic.

In contrast, CLCD×WebLCA has developed a dynamic model documentation system leveraging efootprint software. This system automatically generates standardized documentation directly from LCA models. This documentation not only furnishes essential background information, including temporal, geographical, and technical specifications, but also incorporates input and output inventory tables. Crucially, it integrates built-in inventory contribution analysis and uncertainty analysis results. This enables users to assess model quality, identify key inventories significantly influencing carbon footprint outcomes, and scientifically evaluate the robustness of the entire LCA result on the basis of quantitative uncertainty metrics. From the perspectives of both document structure and field coverage, CLCD×WebLCA's database documentation demonstrates a preliminary yet comprehensive fulfillment of multi-dimensional transparency requirements.

Based on these observations, it is recommended that database providers establish a standardized, structured database documentation format. This format clearly defines a series of core metadata fields to provide critical information that enhances data transparency.

Table 5. Transparency of information in database documentation fields

Transparency information	**Database documentation**					
	WebLCA（CLCD×WebLCA）		**Ecospold（ecoinvent）**		**ILCD（GaBi）**	
Document transparency	√	Model documentation	√	Documentation	√	Process dataset

Subroutines included in the document disclosure model	√	Name of the main background dataset	√	Exchanges	√	Flow diagram(s) or picture(s) (source dataset)
Document navigation	√	Upstream process	√	Exchanges	√	Included datasets (process dataset)
Input/Output inventory visibility	√	Inventory and parameters	√	Exchanges	×	/
Inventory data visibility	√	Inventory and parameters	√	Exchanges	×	/
Availability of unit process data list	√	Reference	√	Data source	×	/
Data and information processing for inventory	√	Algorithms and assumptions	×	Exchange details	×	/
Data quality assessment	√	Data quality uncertainty	√	Data quality uncertainty	√	Data quality indicator

Note: “/” indicates that this information field is not defined in this database document format.

CONCLUSIONS

Taking typical global LCBD as its subject, this study establishes a full-chain transparent traceability system using public information and develops the first systematic framework for evaluating LCBD transparency across the entire process. From the dimensions of product coverage, model traceability and data quality, three prevalent black boxes in current LCBD are identified, and standardized disclosure requirements are specified, providing methodological support for addressing transparency challenges and improving the robustness and reliability of LCA and product carbon footprint accounting. The results show that first, the absence of a standardized product counting system is the main cause of the scale black box in LCBD. Only six national databases worldwide cover more than 1,000 basic products, and the maximum coverage in a single region is approximately 4,000 items. The tens of thousands of widely claimed products are mostly generated by combined variations in region, technology, and specification rather than new unit processes, meaning that they do not genuinely expand basic product coverage. This also reveals a global shortage of large-scale functional LCBD, which cannot meet the LCA demands of diverse regions and industries. Second, the lack of model traceability leads to a second black box, as the core computational logic cannot be verified or reproduced. Most LCBDs do not support backward tracing from the final results to the raw data and processing procedures, and no standardized traceability mechanism exists. Some databases contain definitional errors and missing unit process inventories; aggregated databases often fail to disclose unit process details and upstream pathways, whereas unit process databases lack proper linkage to form complete life cycle models, meaning that disclosure and functionality generally fall below basic transparency requirements. Third, a data quality black box arises from the misalignment between current quality assessment methods and model logic. Existing DQR approaches do not disclose the weighting logic or the linkage between traceability, completeness, and quality scores. Moreover, uncertainty methods such as Monte Carlo simulation cannot identify specific sources of variation or provide targeted improvements, resulting in weak quality control across the LCA field.

Implication

To enhance LCBD transparency and promote industry standardization, database developers should tailor tiered traceability disclosure guidelines based on database type and user needs. They should shift focus from maximizing dataset quantity to improving actual data quality and effective information content while developing specialized software to support standardized data recording, systematic processing, and end-to-end traceability. Collaboration on standardized unit process templates is essential to ensure the completeness of input-output lists and transparent disclosure of data sources, processing workflows, and quality evaluation results. With respect to LCA standard-setting bodies, existing standards such as ISO 14040 and 14044 should be updated to address LCBD transparency gaps. Comprehensive, operational standards for LCBD construction, data quality control, and traceability disclosure should be established to unify the requirements for product coverage counting, model tracing, and quality assessment, thereby fostering global consistency in fundamental regulations. For LCBD users, the fully transparent traceability chain and indicators proposed in this study enable targeted database selection based on research needs. They can scientifically identify a high-transparency, reliable-quality database and validate the transparency and quality of the utilized database, ultimately improving the robustness and reliability of LCA research outcomes.

Limitations

This study retains scope for further improvement. For instance, its transparency assessment relies solely on official public database information, and future work could involve collaboration with developers to obtain more accurate data for refined evaluation. In addition, the scoring system may be enhanced through large-scale user surveys and the Delphi method to refine indicator weights, whereas sensitivity analysis using the analytic hierarchy process and entropy weight method can be applied to examine how different weighting schemes influence database rankings, thereby improving the quantification, adaptability, and rationality of the framework. Finally, this research focuses on a static evaluation of current transparency conditions; future studies could integrate big data and artificial intelligence to establish a dynamic quantitative assessment system for LCBD transparency, enabling real-time updates alongside database evolution and LCA standard development, thus providing sustained methodological support for global transparency governance of LCBD.

DECLARATIONS

Authors' contributions

Made substantial contributions to the conception and design of the study and performed data analysis and interpretation, Performed data acquisition and technical and material support and Performed data acquisition and administrative approval: XQ.L, LL.S, H.Y, YP.Z, PF.W, LX.X, HC.W;

Performed data acquisition and material support: XQ.L, HT.Z, H.T, W.Y, D.Y.

Availability of data and materials

Data and materials are available from the corresponding authors upon reasonable request.

AI and AI-assisted tools statement

Not applicable.

Financial support and sponsorship

This work is supported by the “Oil and Gas Product Carbon Footprint Data Measurement and Calculation” project (No. CCL2024RCPS0198RSN), the Natural Science Foundation of Sichuan Province (2025NSFSC1987), and the Postdoctoral Fellowship Program of CPSF (GZB20240484).

Conflicts of interest

All the authors declare that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

REFERENCES

1.International Organization for Standardization, 2006a. ISO 14040, Environmental Management-Life Cycle Assessment Principles and Framework.https://www.iso.org/standard/37456.html

2.International Organization for Standardization, 2006b. ISO 14044, Environmental Management-Life Cycle Assessment Requirements and Guidelines.https://www.iso.org/standard/38498.html

3.Turner, C., Oyekan, J., Garn, W., Duggan, C., Abdou, K. Industry 5.0 and the circular economy: utilizing LCA with intelligent products. *J. Sustainability.* 2022, 14, 14847.[DOI: 10.3390/su142214847]

4.Hu, Z., Li, P.C., Zhang, Z.Z., Chen, G.Y., Song, C.F. Microalgae fixed flue gas CO_2 into biomass: comparative of life cycle assessment and technical-economic analysis of different technologies. *J. Environ. Chem. Eng.* 2025,13, 119467.[DOI: 10.1016/j.jece.2025.119467]

5.Cordero, P. Carbon footprint estimation for a sustainable improvement of supply chains: state of the art. *J. Ind. Eng. Manage.* 2013,6, 805-13.[DOI: 10.3926/jiem.570]

6.Fernández-González, J., Rumayor, M., Domínguez-Ramos, A., Irabien, A., Ortiz, I. The relevance of life cycle assessment tools in the development of emerging decarbonization technologies. *JACS Au.* 2023, 3, 2631-9.[DOI: 10.1021/jacsau.3c00276]

7.Finnveden, G., Hauschild, M.Z., Ekvall, T., et al., Recent developments in Life Cycle Assessment. *J. Environ. Manage.* 2009, 91, 1-21.[DOI: 10.1016/j.jenvman.2009.06.018]

8.Sala, S., Amadei, A.M., Beylot, A., Ardente, F. The evolution of life cycle assessment in European policies over three decades. *Int. J. Life Cycle Assess.* 2021, 26, 2295-314.[DOI: 10.1007/s11367-021-01893-2]

9.Guo, Y.J. Green Trade Barriers under the Developing Country Perspective. *Future Human Image.*2024, 21, 4-15.[DOI: 10.29202/fhi/21/1]

10.Liu, X. L., Wang, H. T., Chen, J., et al., Method and basic model for development of Chinese reference life cycle database. *J. Acta Sci. Circumst.* 2010, 30, 2136-44.[DOI: 10.13671/j.hjkxxb.2010.10.028]

11.UNEP, Global Guidance Principles for Life Cycle Assessment Databases: A Basis for Greener Processes and Products. 2011, France.

12.Kuczenski, B. Partial ordering of life cycle inventory databases. *J. Life Cycle Assess.* 2015, 20, 1673-83.[DOI: 10.1007/s11367-015-0972-x]

13.Liu, T., Liu, Y., Song Z. Research and practice on method of LCA background database development in Chinese steel industry. *steel and iron.* 2025, 60, 262-70.[DOI: 10.13228/j.boyuan.issn0449-749x.20250242]

14.Pauer, E., Wohner, B., Tacker, M. The influence of database selection on environmental impact results. Life cycle assessment of packaging using GaBi, ecoinvent 3.6, and the environmental footprint database. *Sustainability.* 2020, 12, 9984.[DOI: 10.3390/su12239948]

15.Seckar, M., Schwarz, M., Pochyba, A., Polgar, A. A comparative analysis of the environmental impacts of wood-aluminum window production in two life cycle assessment software. *Sustainability* 2024,16, 9581.[DOI: 10.3390/su16219581]

16.Ministry of Ecology and Environment of the People's Republic of China. Announcement No. 19 of 2025: Release of 2024 power carbon footprint factors. 2025. Available from: http://www.mee.gov.cn/xxgk2018/xxgk/xxgk01/202510/t20251024_1130734.html. [Last accessed on 10 March 2026]

17.Yi, J., Sun, H.R., Lin, W.F., et al., Comparative analysis of carbon emission accounting standards for power systems. *Power Syst. Technol.* 2025, 49, 920-33.[DOI: 10.13335/j.1000-3673.pst.2024.0386]

18.Zhu, G., Tian, Y., Xiong, J., et al., High-resolution data unveils overestimation of China's electricity carbon footprint in international LCA databases. *The Innovation Energy.* 2026, 3, 100132.[DOI: 10.59717/j.xinn-energy.2026.100132]

19.Luo, B., Gu, A., Chen, X., et al., EU carbon border adjustment mechanism and international industrial landscape: Impact assessment based on a global computable general equilibrium model. *J. Tsinghua Univ. (Sci. Technol.).* 2024,64, 1492-501.[DOI: 10.16511/j.cnki.qhdxxb.2023.26.050]

20.Gao, H. B., Yan, K., Zhao, L. H., et al., Research on localized product carbon footprint accounting in China: Taking lithium battery industry as an example. *CAE.* 2025, 27, 1-12.[DOI: 10.15302/J-SSCAE-2025.01.014]

21.Bluhm, H., Wohlschlager, D., Pohl, J., et al., Understanding digitalization's environmental impact: why LCA is essential for informed decision-making. *NPJ Clim Action.* 2025, 4, 41.[DOI: 10.1038/s44168-025-00246-1]

22.Tan, E., Tu, Q., Martins, A.A., et al., Uncertainty in inventories for life cycle assessment: state-of-the-art, challenges, and new technologies. *Environ. Prog. Sustain. Energy.* 2025, 44, 14644.[DOI: 10.1002/ep.14644]

23.Li, J., Wang, J., Hao, Y., et al., Global evolution of research on life cycle assessment: a data-driven visualization of collaboration, frontier identification, and future trend. *Environ. Impact Assess. Rev.* 2026, 116, 108093.[DOI: 10.1016/j.eiar.2025.108093]

24.Isah, M.E., Zhang, Z., Matsubae, K., et al., Bibliometric analysis and visualization of research on life cycle assessment in Africa. *Int. J. Life Cycle Assess.* 2024, 29, 1339-51.[DOI: 10.1007/s11367-024-02313-x]

25.Moutik, B., Summerscales, J., Graham-Jones, J., et al., Life cycle assessment research trends and implications: a bibliometric analysis. *Sustainability.* 2023, 15, 13408.[DOI: 10.3390/su151813408]

26.Nair, R. R., Chougule Mallesh, K., Gomez, J. C., et al., A generalized schema to publish and share life cycle inventories (LCI): Exemplary case of an aviation fuel supply chain. *J. Clean. Prod.* 2024, 520: 146120.[DOI: 10.1016/j.jclepro.2025.146120]

27.Kalverkamp, M., Helmers, E., Pehlken, A. Impacts of life cycle inventory databases on life cycle assessments: A review by means of a drivetrain case study. *J Clean Prod.* 2020, 269, 121329.[DOI: 10.1016/j.jclepro.2020.121329]

28.Amon, F., Dahlbom, S., Blomqvist, P. Challenges to transparency involving intellectual property and privacy concerns in life cycle assessment/costing: A case study of new flame retarded polymers. *Clean. Environ. Syst.* 2021, 3, 100045.[DOI: 10.1016/j.cesys.2021.100045]

29.Guo, J., Li, R.Q., Zhang, R.R., et al., Shedding light on the shadows: Transparency challenge in background life cycle inventory data. *J. Ind. Ecol.* 2025, 29, 766-76.[DOI: 10.1111/jiec.70010]

30.Ecos. Proposals for minimum requirements on LCI database's quality & transparency. 2024, Brussels, Belgium.

31.Zampori, L., Wolf, M. A., Cenian, K., et al., Modeling requirements on LCI models under the Environmental footprint for interoperable data exchange via the eILCD format. 2020, European Platform on Life Cycle Assessment.

32.Wu, S. R., Wang, L. Higher transparency: A desideratum in environmental life cycle assessment research. *J. Clean. Prod.* 2022, 374, 134074.[DOI: 10.1016/j.jclepro.2022.134074]

33.Rolinck, M., Khakmardan, S., Cerdas, F., et al., Completeness evaluation of LCI datasets for the environmental assessment of lithium compound production scenarios. *Procedia CIRP.* 2023, 116, 726-31.[DOI: 10.1016/j.procir.2023.02.122]

34.Martínez-Rocamora, A., Solís-Guzmán, J., Marrero, M. LCA databases focused on construction materials: a review. Renew. *Sust. Energ.* Rev.2016, 58, 565-73.[DOI: 10.1016/j.rser.2015.12.243]

35.Valente, A., Vadenbo, C., Fazio, S., et al., Elementary flow mapping across life cycle inventory data systems: a case study for data interoperability under the global life cycle assessment data access (glad) initiative. *Int. J. Life Cycle Assess.* 2024, 29, 789-802.[DOI: 10.1007/s11367-024-02286-x]

36.Teng, Y., Li, C.Z., Shen, G.Q.P., Yang, Q., Peng, Z. The impact of life cycle assessment database selection on embodied carbon estimation of buildings. *Build. Environ.* 2023, 243, 110648.[DOI: 10.1016/j.buildenv.2023.110648]

37.Clayton, R., Kirk, J., Banford, A., Stamford, L. A review of radioactive waste processing and disposal from a life cycle environmental perspective. *Clean Technol. Environ. Policy.* 2025, 27, 665-82.[DOI: 10.1007/s10098-024-02998-6]

38.Baitz, M., Piotrowski, M. Appropriateness and reliability of life cycle assessment results in relation to data quality: avoiding result discrepancy while improving decision certainty via use of adequate inventory data. Environ. *Res.: Infrastruct. Sustain.* 2025, 5, 3.[DOI: 10.1088/2634-4505/adec1a]

39.European Commission JRC. International Reference Life Cycle Data System (ILCD) Handbook: General guide for life cycle assessment-Detailed guidance. Institute for Environment and Sustainability.2010, Ispra, Italy.[DOI: 10.2788/38479]

40.European Commission. Recommendation (EU) 2021/2279 of 15 December 2021 on the use of the Environmental Footprint methods to measure and communicate the life cycle environmental performance of products and organisations. 2021. Available from: https://eur-lex.europa.eu/eli/rec/2021/2279. [Last accessed on 15 March 2026]

41.Kupfer, T., Baitz, M., Makishi Colodel, C., et al., GaBi Databases & Modeling Principles. *Sphera,* 2021.

42.Bamber, N., Turner, I., Arulnathan, V. et al., Comparing sources and analysis of uncertainty in consequential and attributional life cycle assessment: review of current practice and recommendations. *Int. J. Life Cycle Assess.* 2020, 25, 168-80.[DOI: 10.1007/s11367-019-01663-1]

43.Xu, G., Luo, Y., Zhang, Y., et al., Comparison on environmental impacts of cereal and forage production in the Loess Plateau of China: Using life cycle assessment with uncertainty and variability analysis. *J. Cleaner Pro.*2020, 380, 135094.[DOI: 10.1016/j.jclepro.2022.135094]

44.Wolf, M., Kusche, O., Düpmeier, C. The International Reference Life Cycle Data System (ILCD) Format-Basic Concepts and Implementation of Life Cycle Impact Assessment (LCIA) Method Data Sets. *Int. Confer. Informatics Environ. Pro.* 2011.https://dl.gi.de/handle/20.500.12116/26142

45.Kusche, O., Düpmeier, C., Recchioni, M., et al., Creating LCA Data Exchange Networks. *Int. Confer. Informatics Environ. Pro.* 2012.https://api.semanticscholar.org/CorpusID:18451738

46.Cardoso, V.E., Sanhudo, L., Silvestre, J.D., et al., Challenges in the harmonization and digitalization of Environmental Product Declarations for construction products in the European context. *Int. J. Life Cycle Assess.* 2024, 29, 759-88.[DOI: 10.1007/s11367-024-02279-w]