DOI: 10.1111/jiec.13572

JOURNAL OF INDUSTRIAL ECOLOGY WILEY

RESEARCH ARTICLE

Data quality assessment of aggregated LCI datasets

A case study on fossil-based and bio-based plastic food packaging

Anna Carlesso | Lisa Pizzol | Antonio Marcomini | Elena Semenzin

Department of Environmental Sciences, Informatics and Statistics, Ca' Foscari University of Venice, Venice, Italy

Correspondence
Elena Semenzin, Department of Environmental Sciences, Informatics and Statistics, Ca' Foscari University of Venice, Via Torino 155, 30172 Venice, Italy.
Email: semenzin@unive.it

Editor Managing Review: Annie Levasseur

Environmental impacts resulting from plastic food packaging, made from both fossil-based and bio-based polymers, are increasingly analyzed in life cycle assessment (LCA) studies. However, the literature reveals significant variations in results for the same polymer within the same scope. To enhance the reliability of these assessments, data quality assessment (DQA) plays a relevant role. However, despite most of the LCA studies employing aggregated life cycle inventory (LCI) datasets, in the literature, DQA methods for aggregated processes are not available. To fill this gap, in this paper, a DQA for aggregated LCI datasets is proposed and demonstrated through its application to 101 aggregated LCI datasets, extracted from Ecoinvent and GaBi databases. The DQA method has been developed by adapting and integrating the pedigree matrix and the data quality ranking proposed by the recently published EC Plastic LCA method. The three data quality indicators (DQIs) used are technological, geographical, and time-related representativeness. The application of this method exhibits an overall positive evaluation of the selected datasets with differences among the three DQIs. Moreover, it highlights the role of metadata structure in adequately supporting a robust DQA. Indeed, in the absence of a common framework that defines, assesses, and provides access to data quality information, transparency must be assured by the operator in the metadata interpretation and related assumptions along the DQA process. Finally, although the proposed DQA method was developed for the plastic sector, its application can be extended to LCI aggregated datasets relevant to other sectors, materials, and products.

KEYWORDS
aggregated LCI datasets, data quality assessment, industrial ecology, life cycle inventory, plastic food packaging, representativeness

1 | INTRODUCTION

In Europe, approximately 40% of food is packaged in plastics (ING Economics Department, 2019). This packaging is criticized as a symbol of make–use–dispose economy due to its rather short lifespan, its dependency on fossil fuels, and its contribution to marine litter (Geyer et al., 2017; Schmidt & Laner, 2023; Sundqvist-Andberg & Åkerman, 2021). The European Union (EU) is attempting to address this problem in its European Strategy for

Plastics in a Circular Economy (EC, 2018) by supporting innovative materials and feedstocks for plastic, which demonstrate being more sustainable than their fossil-based (FB) counterparts. Among the alternative feedstocks, bio-based (BB) polymers have been presented as environmentally friendly compared to traditional polymers, despite their environmental impacts caused by significant energy demand and produced waste as highlighted in life cycle assessment (LCA) comparative studies evaluating the strengths and weaknesses of BB and FB plastics (Walker & Rothman, 2020). Indeed, LCA is a powerful tool used in environmental management to assess the impacts of materials, products, and technologies and thus implement policy strategies for sustainability (EC, 2020; Lewandowska et al., 2021; Lopes Silva et al., 2019). However, the results obtained by LCA show variation of the order of 200%–400% among studies on the same polymer with the same scope (Walker & Rothman, 2020). The reliability of the results of these studies is influenced by many factors. One of these is the quality of data of life cycle inventory (LCI) datasets (Lewandowska et al., 2021), that is, models that represent commercial and industrial processes and systems (GLAD WG2, 2023). Data quality assessment (DQA) of the existing LCI datasets of polymers is paramount since supply chains, processes, and technologies at different development and maturity levels and/or with different scales of production largely influence data quality (Nessi et al., 2021). The LCI datasets are the backbone of LCA (Coste et al., 2021) since they transform a product system into quantifiable unit processes and relevant input/output flows for environmental impact assessment (Zargar et al., 2022). The unit LCI datasets are the smallest element considered in the LCI analysis for quantifying input and output data. Sometimes, these are aggregated to protect confidential and proprietary information (Ciroth et al., 2020; Kuczenski, 2016). According to the Shonan Guidance Principles (UNEP, 2011), two aggregation procedures exist for generating aggregated LCI datasets: horizontal averaging and vertical aggregation. Horizontal averaging is used when multiple operation unit processes share a common function, each providing the same reference flow to create a new process dataset. In contrast, vertical aggregation refers to the combination of unit processes that succeed each other in a product life cycle. However, aggregated datasets are rarely purely horizontal averaging or vertical aggregation; they often comprise a mix of both. Unfortunately, information on the types of aggregation performed is usually not reported in metadata descriptors, which are the common basis for searching, filtering, and sorting information to determine fitness for the purpose of datasets by users (Ciroth et al., 2017). The absence of information on the type of aggregation performed, along with the aggregation itself, increases the level of uncertainty and reduces the transparency and accessibility to detailed dataset information. Therefore, the quality of data underlying the aggregated datasets cannot be analyzed in as much detail as the unit datasets.

In literature, DQA methods for aggregated processes are not available, since data quality systems currently implemented in major databases provide indicators only at either flow (e.g., Ecoinvent) or unit process level (e.g., ILCD, Environmental Footprint) (Edelen & Ingwersen, 2018), and LCA study level (e.g., data quality ranking in Plastics LCA method). DQA in the context of LCA has been defined as a comparative analysis of the characteristics of data to be assessed based on data quality indicators (DQIs) against data quality goals or qualitative statements defining specifications for the adequacy of data (Bakst et al., 1995). According to Edelen and Ingwersen (2018), the methods or data quality systems currently in use can be classified as qualitative or semi-quantitative. Focusing on semi-quantitative methods, which are the topic of this study, the two dominant semi-quantitative methods are represented by the pedigree matrix approach used by Ecoinvent (Weidema et al., 2013) and the quality ranking system used by ILCD (EC et al., 2010). The pedigree matrix approach consists of five DQIs and a five-point scale for scoring. Reliability, completeness, temporal, geographical, and technological correlation are the indicators that are performed at the flow level or, in other words, to individual exchange in a unit process. The quality ranking system applied at the unit process level consists of six indicators: (i) technological, (ii) geographical, (iii) time-related representativeness, (iv) completeness, (v) precision/uncertainty, and (vi) methodological appropriateness and consistency. All indicators are evaluated according to a five-point scale scoring system. The scores of each DQI are then aggregated to formulate an overall data quality ranking score for the process. The DQA method developed by the *ILCD Handbook* has been partially adopted into the Product Environmental Footprint (PEF) and the Organizational Environmental Footprint (OEF). However, as highlighted by Masoni et al. (2014), ILCD and PEF data quality requirements have many common aspects (e.g., application at flow and unit process levels, ranking score, six DQIs, five-point scale for scoring), but they differ in the degree of strictness of the methodological and data quality requirements. Within this frame, in 2021, the European Commission issued the Plastics LCA method (Nessi et al., 2021), a specific LCA-based method to consistently evaluate the potential environmental impacts of plastic products based on alternative and conventional feedstocks. Compared to the PEF method, it provides specific rules to better and/or properly cover relevant methodological or modeling issues for plastic products and the related variety of feedstocks. Herein, the DQA can follow two distinct ways depending upon whether the datasets are company-specific or are secondary datasets (i.e., including aggregated processes). In the latter case, the Plastics LCA method provides a DQA, from now on in this paper called data quality ranking, that can be applied at both flow and (unit or aggregated) process levels to assess LCA studies rather than (unit or aggregated) LCI datasets per se. In the case of aggregated process evaluation, the data quality ranking follows the PEF system but considers only three DQIs: technological, geographical, and time-related representativeness. The reduction of DQIs is determined by the nature of aggregated information, which does not allow the analysis of "completeness" and "precision" of the process. This brief state-of-the-art description confirms that, to the best of the authors' knowledge, a method for evaluating aggregated LCI datasets independently from a specific LCA study is still lacking. This gap exists despite the common necessity to work with aggregated LCI datasets, often driven by patent-related considerations. The first objective of this study was to develop a DQA method to assess aggregated LCI datasets, independently from a specific LCA study and only on the basis of metadata descriptors, which adapts and integrates the data quality ranking proposed by the Plastics LCA method and the pedigree matrix. The newly proposed DQA is demonstrated through its application to a set of aggregated processes available for FB and BB polymers used in lightweight food packaging, to highlight its strengths and potential limitations.

TABLE 1 Matrix scoring table providing the combination of data quality indicators (DQIs) adapted from the pedigree matrix and the data quality ranking.

		Data quality indicators		
Quality rating	**Quality level**	**TeR**	**GeR**	**TiR**
1	Excellent	The technology used in the process is the same as the one in the scope of the dataset.	Data from an area under study	Less or equal to 3 years' difference to the period of the initial data collection
2	Good	The technologies used in the process are included in the mix of technologies in the scope of the dataset.	Average data from a larger area in which the area under study is included	Less or equal to 6 years' difference to the period of the initial data collection
3	Fair	The technologies used in the process are only partly included in the scope of the dataset.	Data from a smaller area with similar production conditions in a larger area	Less or equal to 10 years' difference to the period of the initial data collection
4	Poor	The technologies used in the process are similar to those included in the scope of the dataset.	Data from an area with slightly similar production conditions	Less or equal to 15 years' difference to the period of the initial data collection
5	Very poor	The technologies used in the process are different from those included in the scope of the dataset.	Data from an unknown area or an area with very different production conditions	Age of data unknown or more than 15 years' difference
		Not evaluated; data cannot be verified.	–	–

Abbreviations: TeR, technological representativeness; GeR, geographical representativeness; TiR, time-related representativeness.

TABLE 2 Overall data quality level of a life cycle inventory (LCI) dataset according to the achieved data quality rating.

Data quality level	Data quality rating (DQR)
Excellent	1 < DQR ≤ 1.33
Good	1.33 < DQR ≤ 2.33
Fair	2.33 < DQR ≤ 3.33
Poor	3.33 < DQR ≤ 4.33
Very poor	4.33 < DQR ≤ 5

2 | METHODS

2.1 | DQA method development

The proposed DQA method is an adaptation of the two semi-quantitative DQA methods: the pedigree matrix and the data quality ranking. These two methods were merged into a single one to allow the evaluation of data quality of aggregated LCI datasets per se, that is, independently from a specific LCA study. The three pivotal elements composing this method are DQIs, data quality rating (DQR), and a matrix scoring table.

The three adopted DQIs are technological (TeR), geographical (GeR), and time-related (TiR) representativeness. To each DQI, a score is assigned according to the description provided in the matrix scoring table illustrated in Table 1. The scoring is the DQR, that is, a five-point scale, where 1 is the best (corresponding to an excellent data quality level), and 5 is the worst (corresponding to a very poor data quality level) achievable score. This determines that each evaluated dataset receives three scores, one for each DQI.

To obtain a single DQR for a dataset, the three scores are then averaged and the result is classified according to the data quality levels in Table 2, where the DQR ranges were defined following the principle outlined in the *ILCD Handbook* (European Commission et al., 2010): If one indicator receives a lower-level rating compared to the other two (e.g., one "good" and two "excellent"), the overall evaluation remains unchanged ("excellent" in the same example); if two indicators receive a lower-level rating compared to the highest, then the overall evaluation is downgraded accordingly.

The three DQIs in Table 1 are adopted from the data quality ranking, while the rating level definitions in the matrix scoring table are a combination of the data quality ranking and the pedigree matrix methods, adapted to the aim of this research work.

TABLE 3 Interpretative scheme for rating level definition for technological representativeness (TeR).

Rating	Level	Class descriptions	Interpretation	Associated metadata
1	Excellent	The technology used in the process is the same as the one in the scope of the dataset.	The data used for modeling the described technology derive from a company.	The indicator receives this level when the process data source is: • EPD report; • LCI report of producer; • company name.
2	Good	The technologies used in the process are included in the mix of technologies in the scope of the dataset.	The data used for modeling the described technology derive from more companies with similar technologies and the process model is validated by expert judgment.	The indicator receives this level when the process data source is a questionnaire of the dataset and expert judgment or when "The data set covers all relevant process steps/technologies over the supply chain of the represented cradle-to-gate inventory with a good overall data quality."
3	Fair	The technologies used in the process are only partly included in the scope of the dataset.	The data used for modeling the described technology derive from similar average technologies to produce polymers.	The indicator receives this level when the process data source is another dataset and the metadata provide this information: "It is the average production mix, composed of several different processes/technologies" or "It might not reflect the actual production mix and should hence only be used as a proxy."
4	Poor	The technologies used in the process are similar to those included in the scope of the dataset.	The data used for modeling the described technology derive from the different technologies used to produce polymers.	The indicator receives this level when: (i) the process is a child dataset of a parent dataset, that is, a global dataset that was derived from a European dataset built on collected data; or (ii) no process data source is indicated, and metadata provide explicit indications on different technologies used.
5	Very poor	The technologies used in the process are different from those included in the scope of the dataset.	The data used for modeling the described technology derive from the different technologies used different technologies to produce polymers.	The indicator receives this level when no process data source is indicated, and metadata do not provide indications on different technologies used.
		Not evaluated; data cannot be verified.	No description of the technology is provided	The indicator receives this level when no process data source is indicated, and metadata do not provide any information on technologies used.

More specifically, for the TeR, the rating level definition proposed by the data quality ranking was slightly modified by substituting "Environmental Footprint study" with the term "process," to support the assessment of data quality independently from a specific LCA study. As noted by Edelen and Ingwersen (2018), the terminology around the class definitions in both data quality ranking and pedigree matrix is vague and unclear; "similar," "slightly similar," "partially," "some," and "many" are the terms provided by the two abovementioned methods to classify technological representativeness based on the information provided by the metadata descriptors, which also usually lack completeness and clarity. As a result, individual evaluators must apply subjective choices. Table 3 was developed to provide an interpretative and applicative scheme on how to evaluate TeR according to the available metadata information, with the final aim of minimizing the discretionality of individual evaluators.

As far as GeR and TiR are concerned, their rating level definitions were adopted from the pedigree matrix. For GeR, no text modifications were proposed but, similarly to TeR, an interpretative scheme was developed to guide its application (Table 4).

Finally, for TiR, the rating level definition was modified by adding the term "equal" (obtaining "less or equal" instead of "less" only) and "the period of the initial data collection" to the original version provided by the pedigree matrix. Such modifications implement the suggestions proposed by Ciroth et al. (2020) to avoid ambiguity in the interpretation of the time intervals.

2.2 | Case study for DQA method's demonstration

The proposed DQA was demonstrated through its application to selected LCI datasets for plastic food packaging. The dataset selection followed three steps: (i) identification of the most suitable LCI databases used for plastic food packaging; (ii) identification of the most relevant FB and BB

TABLE 4 Interpretative scheme for rating level definition for geographical representativeness (GeR).

Rating	Level	Class descriptions	Interpretation	Associated metadata
1	Excellent	Data from an area under study	Data were collected in the geographic area for which the process was developed.	E.g., data gathered in Italy contribute to modeling the Italian datasets.
2	Good	Average data from a larger area in which the area under study is included	Data were collected in the larger geographic area for which the process was developed. Data were not collected in the field since datasets are literature-based.	Datasets based on global sources of information are employed as database for elaborating and modeling a dataset for a smaller area (e.g., Quebec).
3	Fair	Data from a smaller area with similar production conditions in a larger area	Data were collected in the smaller geographic area for which the process was developed.	E.g., data collected in a single European nation, but deemed suitable for the whole European Union.
4	Poor	Data from an area with slightly similar production conditions	Data has a global geographic boundary.	E.g., data collected in Europe but the dataset was suited for the global boundary representativeness.
5	Very poor	Data from an unknown area or an area with very different production conditions	Neither info on geographical area nor expert judgment.	

polymers for a specific food packaging category, that is, flexible food packaging; and (iii) identification of the available LCI processes for the identified polymers in the selected databases.

For the databases' identification, two recent review papers were considered: Walker and Rothman (2020) and Bishop et al. (2021). Both reviews discuss issues affecting LCA studies comparing FB and BB polymers. Overall, they included in the analysis 69 LCA studies where the Ecoinvent database was used in 64% of the studies, GaBi in 7%; in 6%, a mixed use of Ecoinvent and GaBi was found, while the remaining 23% applied other databases (e.g., DEAM, MTEC Thailand, and IDEMAT). Hence, for the dataset collection, it was decided to use Ecoinvent version 3.9.1 and GaBi by consulting the online format of LCA datasets in 2022.

To identify the most relevant FB and BB polymers for flexible food packaging, a rapidly growing category of plastic packaging (World Economic Forum, 2016), two different procedures were followed. FB polymers were selected based on a combination of data from scientific literature (Jariyasakoolroj et al., 2020; World Economic Forum, 2016) and industry data from Plastics Europe (2022), while BB polymers were chosen considering both market demand (European Bioplastics, 2022) and the list of BB polymers used in flexible packaging pointed out by (Jariyasakoolroj et al., 2020).

Then, the list of relevant FB and BB polymers for flexible food packaging was cross-checked with those included in the Ecoinvent and GaBi databases. More specifically, it was decided to focus on LCI aggregated processes for the production of granulates entirely composed of either FB or BB materials. Polymers made from recycled materials or a mixture of FB and BB materials, as well as processes related to waste treatments, were excluded from consideration. As a result, the FB polymers assessed in this study include high-density polyethylene (HDPE), low-density polyethylene (LDPE), linear low-density polyethylene (LLDPE), polyethylene terephthalate (PET), and polypropylene (PP), while as BB polymers thermoplastics starch (TPS) and polylactic acid (PLA) were selected.

3 | RESULTS AND DISCUSSION

The three-step procedure described in Section 2.2 allowed us to select 101 LCI aggregated datasets to work with for demonstrating the application of the proposed DQA method. Out of them, 84 were related to FB polymers and 17 to BB polymers for flexible food packaging. This difference was expected as only two BB polymers were investigated out of 7. GaBi contributed the most with 80 datasets (68 for FB polymers and 12 for BB), while Ecoinvent provided 21 datasets (16 for FB and 5 for BB). The complete list of selected LCI aggregated processes is reported in Section 1 of the SI (Table 1 and Table 2).

Overall, 90% of these datasets received a positive evaluation, falling into excellent (2.97%), good (78.22%), and fair (8.91%) quality levels; 9.90%, instead, were rated as poor and none achieved the very poor quality level. Figure 1 reports the results for FB and BB datasets separately, highlighting that 90.48% of FB and 88.24% of BB processes achieved a positive evaluation.

More specifically, a small percentage of aggregated LCI datasets achieved an excellent quality level: 2.38% for FB and 5.88% for BB; 83.33% of FB datasets and 52.94% of BB datasets were evaluated as good followed by 4.76% of FB and 29.41% of BB datasets rated as fair. The lowest quality level reached in the DQA application was poor, with 9.52% of FB and 11.76% of BB datasets classified.

FIGURE 1 Results of the data quality assessment (DQA) of the life cycle inventory (LCI) datasets divided by fossil-based and bio-based polymers. The data used to calculate the DQA results can be found in Supporting Information.

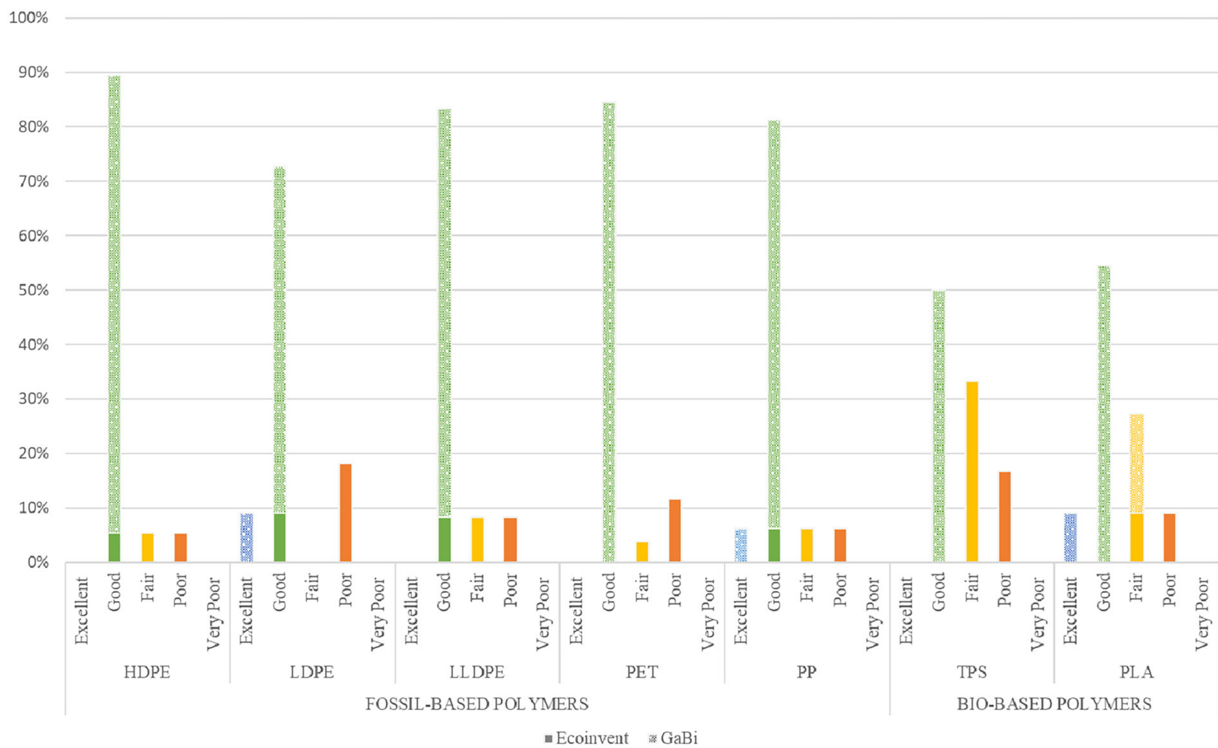

FIGURE 2 Data quality assessment (DQA) of each analyzed polymer. The color code indicates the five data quality levels: blue (excellent), green (good), yellow (fair), orange (poor), and red (very poor). Ecoinvent datasets have a solid-color background filling while GaBi datasets have a dotted background filling. The data used to calculate the DQA results can be found in Supporting Information. HDPE, high-density polyethylene; LDPE, low-density polyethylene; LLDPE, linear low-density polyethylene; PET, polyethylene terephthalate; PP, polypropylene; TPS, thermoplastics starch; PLA, polylactic acid.

Figure 2 allows further exploration of the results at both polymer and database levels. Among FB LCI datasets, the five selected polymers showed similar trends; only small differences can be highlighted. Regarding the excellent quality level, only 9.09% of LDPE and 6.25% of PP LCI datasets achieved it. The main assigned quality level was good, with the following percentages: 89.47% for HDPE, 72.73% for LDPE, 83.33% for LLDPE, 84.62% for PET, and 81.25% for PP. Fair quality level achieved percentages always below 10% (5.26% for HDPE, 0% for LDPE, 8.33% for LLDPE, 3.85% for PET, and 6.25% for PP). LDPE instead obtained the higher percentage of LCI datasets classified as poor (18.18%) compared to the other FB polymers (5.26% for HDPE, 8.33% for LLDPE, 11.54% for PET, and 6.25% for PP).

Focusing on the BB polymers, TPS and PLA LCI datasets showed a similar trend, but different from the FB polymers. Indeed, their classification was characterized by a lower percentage of LCI datasets reaching a good quality level and a higher percentage of fair quality. More specifically, only PLA showed LCI datasets of excellent quality (9.09%); good quality was achieved by 50.00% of TPS and 54.55% of PLA LCI datasets. The remaining

TPS LCI datasets were classified as fair (33.33%) and poor (16.67%), and the remaining PLA datasets were classified as fair (27.27 %) and poor (9.09%).

Figure 3 reports the classification obtained by the BB and FB polymer LCI datasets according to the three DQI (TeR, GeR, and TiR) ratings. Starting from the TeR indicator, the evaluations were largely positive. Although FB LCI datasets never achieved the excellent quality level (Figure 3a), 86.90% reached the good quality level. The remaining datasets were rated fair (5.95%), poor (1.19%), and very poor (5.95%). In contrast, BB LCI datasets (Figure 3b) received 88.24% excellent and good ratings, while the remaining 11.76% were rated as very poor. Following the scheme proposed in Table 3, among BB LCI datasets, excellent rating was assigned to those characterized by data sources such as LCI reports (e.g., for PLA) or Environmental Product Declaration (EPD) documents (e.g., for TPS). Regarding FB datasets, excellent level was never achieved because those LCI datasets resulted from data collection conducted by industry associations (e.g., PlasticsEurope), which consist of industry data that are completed, where necessary, by secondary data. This condition allowed them to be rated as of good quality, which is the same rating assigned to BB LCI datasets under the same conditions. Fair (3) and poor (4) ratings were assigned to FB datasets resulting from data aggregation at multiple levels. Fair (3) was assigned to all the datasets, such as "Polypropylene, granulate {RoW}| production," for which one of the two sentences reported in the last column of Table 3 was present in the metadata, whereas poor (4) was assigned to a specific dataset for Québec (Canada) "polyethylene terephthalate production, granulate, amorphous, CA-QC" that, according to the metadata descriptors, resulted to be a child dataset of a parent dataset, that is, a global dataset which, in turn, was derived from a European dataset built on data collected by Plastics Europe for Eco-profiles. According to the scheme in Table 3, very poor level was assigned when neither the technology description nor the data source provided information about the technologies used (e.g., in the case of the dataset "Polyethylene production, high density, granulate {RoW}").

For the GeR indicator, the assessment of FB LCI datasets was largely positive (i.e., results are similar to those obtained by the TeR indicator), while the assessment of BB LCI datasets, although mainly positive, presented a wider percentage of negative results. As shown in Figure 3c,d, the excellent rating was achieved by both FB and BB LCI datasets at different percentages. In FB datasets, excellent rating accounts for an overall 11.90%, good level alone covered 76.19% of the LCI datasets, fair quality was never reached, while poor and very poor accounted for 7.14% and 4.76%, respectively. In BB LCI datasets, excellent and good quality levels were reached by 5.88% and 58.82% of the datasets, respectively. Fair quality level constituted 5.88% of the total, while poor and very poor quality levels were assigned to 17.65% and 11.76% of the datasets, respectively.

A parent dataset attained an excellent quality level when metadata descriptors reported that it was built on data collected in the geographic area for which the process was developed (according to the scheme in Table 4). However, in case this dataset underwent further aggregation to provide child datasets with higher geographical coverage (e.g., RoW and GLO), the achievable quality level was reduced, that is, RoW and GLO are rated 4. For example, an LCI report produced by NatureWorks from its production plant located in Nebraska was used in the dataset "Ingeo Polylactide (PLA) biopolymer production," with a specific recommendation to use solely for representing NatureWorks Ingeo production. Such a dataset was rated excellent (1) according to Table 4. In another case, the same LCI report was further elaborated to provide the dataset "Polylactide, granulate {GLO}| production," which was consequently rated poor (4).

In general, datasets were mainly classified as of good quality. Regardless of whether a dataset was modeled with primary or literature-based data, if it was described with the following statement: "The data set represents the country-specific situation in [a specific state, e.g., Germany], focusing on the main technologies, the region-specific characteristics, and/or import statistics," it received a score equal to 2, corresponding to the good quality level.

Concerning the TiR indicator, similarly to the GeR indicator, the assessment of FB LCI datasets was largely positive, while the assessment of BB LCI datasets, although mainly positive, presented a wider percentage of negative results. However, among positive results, while GeR datasets were mainly rated as good, for TiR they were mainly rated as excellent. More specifically, for FB polymers (Figure 3e), the excellent quality level was achieved by 77.38% of the LCI datasets, good quality level by 3.57%, fair by 13.10%, poor by 2.38%, and very poor by 9.57%. For BB polymers (Figure 3f), no datasets received a good quality rating: The excellent quality level was achieved by 58.82% of the LCI datasets, fair quality level by 11.76%, poor by 11.76%, and very poor by 17.65%. These results demonstrate that FB and BB polymers are mainly represented by up-to-date datasets. The achieved excellent quality shows that the age of most of the datasets is about less or equal to 3 years from the period of the initial data collection. Focusing on the type of polymers, the average age of FB datasets is 5, while BB datasets are 8 years old on average. This difference can probably be ascribed to two reasons: (i) BB datasets are lower in number and therefore a smaller amount of "aged" datasets is sufficient to increase the average age compared to FB datasets; and (ii) for FB polymers, LCI datasets can frequently be updated according to the information that PlasticsEurope questionnaires periodically provides, while the update of BB LCI datasets mainly depends on information and reports made available directly by manufacturing companies.

3.1 | Role of metadata in DQA

DQA highly depends on metadata available for aggregated LCI datasets, which can be classified as qualitative or quantitative. According to the DQA application performed in this study, qualitative data are usually available for assessing TeR) and GeR. They provide descriptions of the polymer's

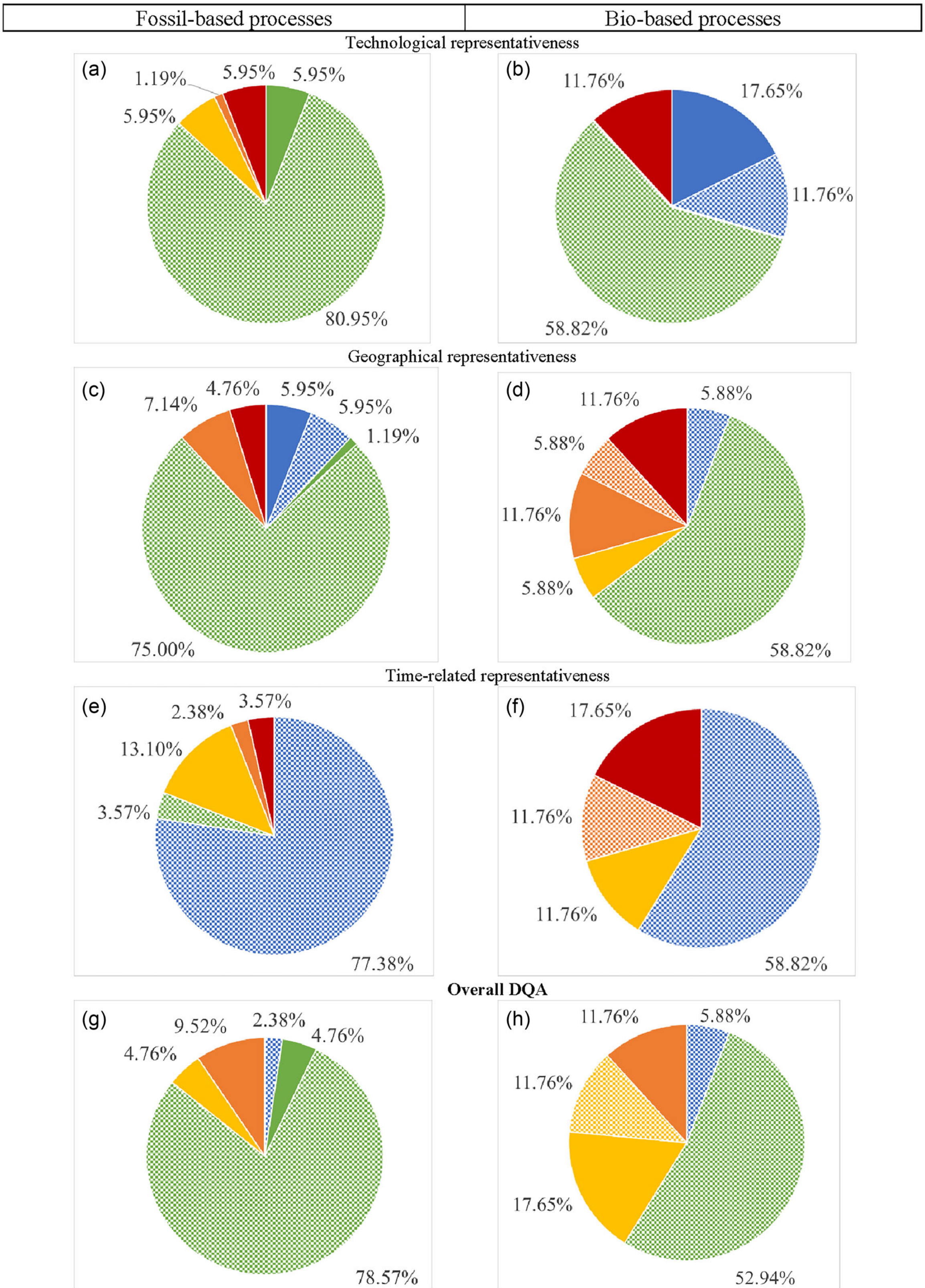

FIGURE 3 Data quality assessment (DQA) for single data quality indicators divided by fossil-based (FB) and bio-based (BB) polymer datasets. The data used to calculate the DQA results can be found in Supporting Information. (a) DQA of technological representativeness (TeR) for FB polymers; (b) DQA of TeR for BB polymers; (c) DQA of geographical representativeness (GeR) for FB polymers; (d) DQA of GeR for BB polymers; (e) DQA of time-related representativeness (TiR) for FB polymers; (f) DQA of TiR for BB polymers; (g) overall DQA for FB polymers; (f) overall DQA for BB polymers. The color code indicates the five data quality levels: blue (excellent), green (good), yellow (fair), orange (poor), and red (very poor). The databases are indicated as follows: Ecoinvent has a solid-color background filling and GaBi has a dotted background filling.

production technology and the geographical area of reference. Quantitative data, instead, are available for assessing TiR (e.g., the date indicating the end of the dataset's temporal validity).

A no-standardized structure of metadata within LCI datasets belonging to different databases emerged during data collection, confirming what Edelen and Ingwersen (2018) observed in their study. In GaBi, metadata descriptors are organized into four main fields, which are further divided into sections and subsections (e.g., "Location" is contained in subsection "Key data set information" within the "Process information" section; for the complete list, see Supporting Information, Section 2). In Ecoinvent, instead, all metadata are within the "Documentation" section, structured into 18 subsections, with "Comment" being the primary repository for relevant information. "Comment" encompasses both free-text and structured fields as outlined in Section 3 in SI.

Accordingly, data collection followed two different (i.e., database-specific) procedures and allowed to highlight that more structured and detailed metadata would minimize the need for subjective interpretation by the DQA operator. As an example, in the Ecoinvent "Comment" subsection, the free text typically includes data sources used to model the LCI dataset (e.g., EPD report, LCI report of producer, company name), descriptions of polymer production, and additional notes on updates of the LCI dataset compared to previous versions. However, such information is not consistently provided for each LCI dataset. For instance, the qualitative metadata regarding polymer production can vary from detailed to concise, introducing some uncertainty in the related DQI assessment performed by the operator according to the criteria presented in Section 2.1.

In this case study, the DQI mostly affected by the absence of a consistent metadata structure was TeR. Indeed, to assess it, the DQA operator needs information on whether the data used for modeling the described technology derive: (i) from a company, (ii) from more companies with similar technologies, (iii) from similar average technologies, (iv) from similar technologies, or (v) from different technologies (Table 3). At present, in Ecoinvent, such data must be searched in the "Comment" free text, which includes the data source and a description of polymer production. The procedure followed to obtain the requested information was to cross-reference the details provided in the description with the data source. Conversely, in GaBi, the available metadata comprise the data source, the technology description provided in the "Technology description including background system" (contained in the "Process information" section), and an internal assessment of technological representativeness (in five classes) done by the developers, which is provided in the "Modeling and validation" section. In this case, the procedure followed to obtain the desired information was to cross-reference the internal evaluation of technological representativeness with the provided description and the data source. The internal evaluation of technological representativeness was surely a robust basis, which facilitated the TeR assessment. Indeed, only in one case (i.e., for the "Polylactic Acid (PLA) TH 2018 System" dataset), the internal evaluation, equal to good, was upgraded to excellent according to our method. This happened because the data source was an LCA study of a plant of Total Corbion in Thailand, and therefore, it represented the technology used in a specific company's operations. According to the proposed DQA method (Table 3), this matched the highest possible rating (DQR = 1).

In other cases, the data source played a crucial role: when both GaBi and Ecoinvent databases provided datasets based on the same data source. Once GaBi datasets were assessed for TeR (following the procedure described above), the same rating could be assigned to the Ecoinvent datasets based on the same data source. This allowed circumventing the problem of poor or unclear metadata in Ecoinvent to inform TeR evaluation. It happened for all the FB datasets based on data collected by PlasticsEurope questionnaires for HDPE, LDPE, LLDPE, PET, and PP, and for the PLA production processes, where the data source for both GaBi and Ecoinvent datasets was the same report from NatureWorks, a key manufacturer of PLA.

TeR assessment would surely benefit from a more consistent metadata structure. For example, the four criteria (i.e., process design, operating conditions, material quality, and process scale) proposed by Edelen and Ingwersen (2016) for unit datasets could be considered. Edelen and Ingwersen (2016) suggest that when using aggregated data, the user should not attempt to apply quality scores to the data unless supplementary documentation detailing the needed information about the data generation is included. This means that if sufficient supplementary documentation (on process design, operating conditions, material quality, and process scale) is present, then the data quality assessor could use it to apply Edelen and Ingwersen (2016) TeR evaluation; if not (as in all the datasets investigated in this study), Edelen and Ingwersen (2016) TeR evaluation should only be completed by the originator/aggregator. Therefore, the adoption of the four proposed criteria as metadata descriptors of aggregated datasets would allow to move from the interpretative scheme proposed in Table 3 to the Edelen and Ingwersen (2016) approach for TeR evaluation, reducing subjectivity and difficulties in DQA application.

Focusing on GeR DQI, its assessment requires the following information: (i) modeled geographical area; (ii) geographical area of reference and information regarding whether the collected data were used to model the geographical area of reference or processed to cover broader or narrower areas, and (iii) whether the data came from literature-based datasets. The first one is the easiest to find in the metadata. In Ecoinvent, it is clearly reported in the name of the dataset, according to five geographical levels: regional (e.g., Québec), national (e.g., Italy), supra-national/regional (e.g., European Union), continental (e.g., Europe—RER), and global (RoW, GLO). The global level in turn distinguishes between Global (GLO) and Rest-of-the-World (RoW), where GLO datasets are created individually to accurately reflect the global average conditions based on international data, while RoW datasets represent the world (GLO) minus all local geographies for which a process exists in the database (ecoinvent.org). In GaBi, such information is provided in the geographical representativeness description (i.e., a subsection of "Process information"), according to three geographical levels: national (e.g., Italy), supra-national/regional (e.g., EU-28), and global (e.g., GLO). However, knowing the modeled geographical area is not sufficient to assess GeR. Indeed, as reported in the interpretative scheme presented in Table 4, it is crucial to collect the information previously described at points (ii) and (iii), to guide the operator in the assessment by reducing the degree of subjectivity. Such additional information can be found in the metadata of both databases: in specific structured fields of the "Comment" subsection for Ecoinvent and the "General comment on data set" subsection of "Process information" for GaBi.

Finally, the TiR DQI was not affected by the different metadata structures. To inform it, the following information is needed: (i) the period of the initial data collection and (ii) the end-of-time validity of LCI datasets. Despite the differences in the nomenclature used in Ecoinvent and GaBi (further details are provided in Section 4, Table 3 in SI), the information was clearly supplied by both databases.

In addition to the specific comments on metadata informing the three DQIs, it is important to highlight a general lack of information on the applied aggregation procedures. As explained in Section 1, each aggregated dataset has undergone an aggregation procedure that can be classified as either horizontal averaging or vertical aggregation and, most often, comprises a mix of both. In case vertical aggregation (i.e., the combination of unit processes that succeed each other in a product life cycle) is applied, it is possible that the reported metadata are related to the final unit process only, which could be characterized by a different data quality compared to upstream processes. As the information on which data were aggregated and how this aggregation was carried out is currently not available in metadata (as confirmed by the investigated datasets), the DQA operator can only assume that the reported metadata are representative of the entire dataset (i.e., of all the unit processes aggregated in it). The inclusion of such metadata information in the future would allow to further improve the proposed DQA method and reduce uncertainty in its application.

4 | CONCLUSION

This paper introduces a DQA method developed to evaluate the quality of aggregated LCI datasets for the plastic sector. This method is an adaptation and integration of the data quality ranking proposed by the Plastics LCA method and the pedigree matrix, which includes three DQIs: TeR, GeR, and TiR. The DQA has been demonstrated through its application to a total of 101 datasets, selected from the Ecoinvent and GaBi databases, on FB (i.e., HDPE, LDPE, LLDPE, PET, and PP) and BB (i.e., TPS and PLA) polymers. Overall, the findings indicate that 90.48% of FB and 88.24% of BB datasets are at excellent, good, or fair quality levels. The best performing DQI is TiR, showing the highest percentage of excellent evaluations, followed by TeR, which reaches slightly better evaluations for BB datasets, and GeR, with slightly better evaluations for FB datasets. However, TeR is the DQI mostly affected by different metadata structures. Indeed, in some cases, the metadata reported in the databases are not complete and consistent, making it challenging for the operator to adequately inform the DQA.

As already pointed out by Edelen and Ingwersen (2018), the assessment of data quality would highly benefit from a framework that defines, assesses, and provides access to data quality information. Two suggestions were proposed in our paper. The first in regard to TeR, as its assessment would surely benefit from the adoption of the four Edelen and Ingwersen (2016) criteria (i.e., process design, operating conditions, material quality, and process scale) as metadata descriptors of aggregated datasets. The second in regard to the need to include in the metadata of aggregated datasets detailed information on which data were aggregated and how this aggregation was carried out, as DQA results are currently potentially affected by the lack of such information in case vertical aggregation is applied.

Nevertheless, despite the absence of a common framework and the need to enhance metadata descriptors, it is still possible to conduct a DQA of aggregated datasets, but each interpretation and assumption must be explicitly declared to ensure process transparency.

Finally, while the DQA method presented in this study was developed for the plastic sector, its application can be extended to LCI aggregated datasets relevant to other sectors, materials, and products.

ACKNOWLEDGMENTS

The scientific activity was performed within the PhD project of Anna Carlesso, financially supported by a fellowship co-funded by COREPLA, the National Consortium for the Collection and Recycling of Plastic Packages, in Italy.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Data are subject to third-party restrictions. The data that support the findings of this study are available from the Ecoinvent and GaBi databases. Data are available directly from the third party.

ORCID

Anna Carlesso https://orcid.org/0000-0001-7571-5252

REFERENCES

Bakst, J., Lacke, C., Weitz, K., & Warren, J. (1995). Guidelines for assessing the quality of life-cycle inventory data. https://nepis.epa.gov/Exe/ZyNET.exe/10000VPN.txt?ZyActionD=ZyDocument&Client=EPA&Index=1995%20Thru%201999&Docs=&Query=&Time=&EndTime=&SearchMethod=1&TocRestrict=n&Toc=&TocEntry=&QField=&QFieldYear=&QFieldMonth=&QFieldDay=&UseQField=&IntQFieldOp=0&ExtQFieldOp=0&XmlQuery=&File=D%3A%5CZYFILES%5CINDEX%20DATA%5C95THRU99%5CTXT%5C00000000%5C10000VPN.txt&User=ANONYMOUS&Password=anonymous&SortMethod=h%7C-&MaximumDocuments=1&FuzzyDegree=0&ImageQuality=r75g8/r75g8/x150y150g16/i425&Display=hpfr&DefSeekPage=

Bishop, G., Styles, D., & Lens, P. N. L. (2021). Environmental performance comparison of bioplastics and petrochemical plastics: A review of life cycle assessment (LCA) methodological decisions. *Resources, Conservation and Recycling*, *168*, 105451. https://doi.org/10.1016/j.resconrec.2021.105451

Ciroth, A., Arbuckle, P., Cherubini, E., Ugaya, C., & Edelen, A. (2017). Task 3: Core meta-data descriptors and guidance on populating descriptors. *WG3 of the Global Life Cycle Data Access Network (GLAD)*, *0*(June), 80. https://wedocs.unep.org/bitstream/handle/20.500.11822/22539/ReportWG3_final_Nov.17%281%29.pdf?sequence=1&isAllowed=y

Ciroth, A., Foster, C., Hildenbrand, J., & Zamagni, A. (2020). Life cycle inventory dataset review criteria—A new proposal. *International Journal of Life Cycle Assessment*, *25*(3), 483–494. https://doi.org/10.1007/s11367-019-01712-9

Coste, G., Biard, Y., Roux, P., & Hélias, A. (2021). ELDAM: A Python software for life cycle inventory data management. *Journal of Open Source Software*, *6*(57), 2765. https://doi.org/10.21105/joss.02765

EC, E. C. (2018). Strategy for plastics in a circular economy, *3*. https://doi.org/10.4325/seikeikakou.30.577

EC, E. C. (2020). Chemicals strategy for sustainability towards a toxic-free environment. Brussels, 14.10.2020 COM(2020) 667 Final. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52020DC0667

EC, E. C., JRC, R. C., & IES, I. E. S. (2010). *International Reference Life Cycle Data System (ILCD) handbook—General guide on LCA—Detailed guidance*. Publications Office of the European Union. https://doi.org/10.2788/38479

Edelen, A., & Ingwersen, W. W. (2018). The creation, management, and use of data quality information for life cycle assessment. *International Journal of Life Cycle Assessment*, *60*(218), 1–24. https://doi.org/10.1007/s11367-017-1348-1.The

Edelen, A., & Ingwersen, W. (2016). Guidance on data quality assessment for life cycle inventory data (EPA/600/R-16/096). U.S. Environmental Protection Agency. https://cfpub.epa.gov/si/si_public_record_Report.cfm?Lab=NRMRL&dirEntryId=321834

European Bioplastics. (2022). Continued growth: global production capacities of bioplastics 2022-2027 Bioplastics. Bioplastic Market Development Update 2022, 2022, 2019–2020. https://docs.european-bioplastics.org/publications/market_data/2022/Report_Bioplastics_Market_Data_2022_short_version.pdf

Geyer, R., Jambeck, J. R., & Law, K. L. (2017). Production, use, and fate of all plastics ever made. *Science Advances*, *3*(7), 25–29. https://doi.org/10.1126/sciadv.1700782

GLAD WG2. (2023). *Guidance on Glad'S metadata descriptors* (Version 1.0). https://www.globallcadataaccess.org/sites/default/files/guidance_metadata_v10.pdf

ING Economics Department. (2019). *Plastic packaging in the food sector—Six ways to tackle the plastic puzzle*. ING Economics Department. https://think.ing.com/uploads/reports/ING_-_The_plastic_puzzle_-_December_2019_%28003%29.pdf

Jariyasakoolroj, P., Leelaphiwat, P., & Harnkarnsujarit, N. (2020). Advances in research and development of bioplastic for food packaging. *Journal of the Science of Food and Agriculture*, *100*(14), 5032–5045. https://doi.org/10.1002/jsfa.9497

Kuczenski, B. (2016). *Privacy-preserving aggregation in life cycle assessment. Environment systems and decisions*. https://doi.org/10.1007/s10669-016-9620-7

Lewandowska, A., Joachimiak-Lechman, K., & Kurczewski, P. (2021). A dataset quality assessment—An insight and discussion on selected elements of environmental footprints methodology. *Energies*, *14*(16), 1–19. https://doi.org/10.3390/en14165004

Lopes Silva, D. A., Nunes, A. O., Piekarski, C. M., Da Silva Moris, V. A., De Souza, L. S. M., & Rodrigues, T. O. (2019). Why using different Life Cycle Assessment software tools can generate different results for the same product system? A cause–effect analysis of the problem. *Sustainable Production and Consumption*, *20*, 304–315. https://doi.org/10.1016/j.spc.2019.07.005

Masoni, P., Porta, P., Tarantini, M., Zamagni, A., Recchioni, M., Mathieux, F., & Pant, R. (2014). Evaluation of ILCD data network entry-level and PEF data quality requirements in LCA databases. Science across bridges, borders and boundaries; SETAC EUROPE; 2014. p. 137. https://publications.jrc.ec.europa.eu/repository/handle/JRC86490

Nessi, S., Sinkko, T., Bulgheroni, C., Garcia-Gutierrez, P., Giuntoli, J., Konti, A., Sanye-Mengual, E., Tonini, D., Pant, R., Marelli, L., & Ardente, F. (2021). *Life cycle assessment (LCA) of alternative feedstocks for plastics production. Part 1 : The Plastics LCA method*. Publications Office of the European Union. https://doi.org/10.2760/271095

Plastics Europe. (2022). Plastics Europe Association of Plastics Manufacturers Plastics—The Facts 2022. The Facts 2022, October, 81. https://plasticseurope.org/knowledge-hub/plastics-the-facts-2022/

Schmidt, S., & Laner, D. (2023). The environmental performance of plastic packaging waste management in Germany: Current and future key factors. *Journal of Industrial Ecology*, 1–14. https://doi.org/10.1111/jiec.13411

Sundqvist-Andberg, H., & Åkerman, M. (2021). Sustainability governance and contested plastic food packaging—An integrative review. *Journal of Cleaner Production*, *306*, 127111. https://doi.org/10.1016/j.jclepro.2021.127111

UNEP. (2011). Global guidance principles for life cycle assessment databases—"Shonan Guidance Principles". *International Journal of Energy, Environment, and Economics*, *2*(February), 69–75.

Walker, S., & Rothman, R. (2020). Life cycle assessment of bio-based and fossil-based plastic : A review. *Journal of Cleaner Production*, *261*, 121158. https://doi.org/10.1016/j.jclepro.2020.121158

Weidema, B. P., Bauer, C., Hischier, R., Mutel, C., Nemecek, T., Reinhard, J., Vadenbo, C. O., & Wernet, G. (2013). Overview and methodology: Data quality guideline for the ecoinvent database version 3. *Swiss Centre for Life Cycle Inventories*, *3*(1), 1–23.

World Economic Forum. (2016). Ellen MacArthur Foundation and McKinsey & Company, The New Plastics Economy — Rethinking the future of plastics. Rethinking the future of plastics. http://www.ellenmacarthurfoundation.org/publications

Zargar, S., Yao, Y., & Tu, Q. (2022). A review of inventory modeling methods for missing data in life cycle assessment. *Journal of Industrial Ecology*, *26*(5), 1676–1689. https://doi.org/10.1111/jiec.13305

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Carlesso, A., Pizzol, L., Marcomini, A., & Semenzin, E. (2024). Data quality assessment of aggregated LCI datasets - A case study on fossil-based and bio-based plastic food packaging. *Journal of Industrial Ecology*, 1–12. https://doi.org/10.1111/jiec.13572