丁 宁,潘京津,崔 晗,等.LCA 数据质量评价研究现状及趋势 [J]. 中国环境科学, 2025,45(12):7097-7104.
Ding N, Pan J J, Cui H, et al. Research status and trends of LCA data quality evaluation [J]. China Environmental Science, 2025,45(12):7097-7104.

LCA 数据质量评价研究现状及趋势

丁 宁[1],潘京津[2],崔 晗[1,3],王梦思[1],吕 彬[1],杨建新[1,3*] (1.中国科学院生态环境研究中心,区域与城市生态安全全国重点实验室,北京 100085；2.中国电子技术标准化研究院 北京 100007；3.中国科学院大学,北京 101408)

摘要：系统梳理了数据质量评价(DQA)发展历程.目前,国内外 DQA 基本都基于谱系矩阵,在数据质量指标(DQI)基础上开展定性和定量评价;但整体还存在 DQI 不明确、评价方法不统一、量化评价困难等问题.当前,我国需要加大 DQA 研究,构建适合我国碳足迹评价数据现状的 DQA 流程、指标和方法,同时注重数据审核、数据文档要求,建立健全 DQA 体系,为我国碳足迹数据库的构建及 LCA 的开展提供保障.

关键词：生命周期评价；数据库；数据质量评价；谱系矩阵；生命周期清单
中图分类号：X820.3 **文献标识码**：A **文章编号**：1000-6923(2025)12-7097-08

Research status and trends of LCA data quality evaluation.

DING Ning[1], PAN Jing-jin[2], CUI Han[1,3], WANG Meng-si[1], LV Bin[1], YANG Jian-xin[1,3] (1.Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, State Key Laboratory for Ecological of Regions and Cities, Beijing 100085, China；2.China Electronics Standardization Institute, Beijing 100007, China；3.University of Chinese Academy of Sciences, Beijing 101408, China). *China Environmental Science*, 2025,45(12)：7097~7104

Abstract：This study systematically reviewed the history and the art of state of data quality assessment (DQA). DQA was implemented basically based on pedigree matrices and conducted qualitative and quantitative evaluations on the basis of data quality indicators (DQIs). However, there were still such problems as unclear DQIs, inconsistent evaluation methods, and difficulties in quantitative evaluation etc. It was recognized that DQA research needed to be increased, and DQA processes, indicators, and methods suitable for the current data status of carbon footprint evaluation had to be constructed. Additionaly, much attention should be paid to data review and data documentation requirements, so as to provide guarantees for the development China carbon footprint database, as well as the implementation of LCA.

Key words：life cycle assessment；database；data quality assessment；pedigree matrices；life cycle inventory

2023 年 11 月,国家发改委等五部委发布《关于加快建立产品碳足迹管理体系的意见》,明确了碳足迹管理体系建设的总体框架.2024 年 6 月,生态环境部等十五部委发布《关于建立碳足迹管理体系的实施方案》,提出分阶段推进碳足迹数据库建设目标:2027 年,产品碳足迹因子数据库初步构建;2030 年,覆盖范围广、数据质量高、国际影响力强的产品碳足迹因子数据库基本建成.在此背景下,生态环境部于 2025 年 1 月发布了我国电力碳足迹因子.同时,碳足迹因子数据库构建引起了各界的广泛关注和参与,在各部门、行业、企业持续推进中.

产品碳足迹核算及因子数据库构建的基础方法是生命周期评价(LCA),产品碳足迹因子数据库是 LCA 基础数据库的一部分.LCA 作为一种系统化工具,被广泛应用于识别产品或服务对环境的影响[1].由于目前 LCA 的开展严重依赖背景数据库,因此评价生命周期清单(LCI)数据库的数据质量至关重要.评估数据质量的过程称为数据质量评价(DQA)[2].ISO 14044 将数据质量定义为:数据在满足所声明的要求方面的能力特征.数据质量在 LCA 研究中起着根本性作用,并影响研究的整体结果[3-4].

LCA 领域已经开发了不同的方法来解决数据质量问题,并取得了一定进展,但仍然受到缺乏可重复的数据质量结果和文档的困扰[5-6].目前,应用最为广泛的DQA方法为谱系矩阵,该方法已被多次改编.尽管有研究者试图使用谱系矩阵等工具来提高 DQA 的可重复性,但评价方法标准化仍然面临较大挑战[2].

随着 LCA 的广泛应用,LCI 变得更加复杂,对有效DQA的需求也在不断增长.另外,我国碳足迹数据

收稿日期：2025-04-29
基金项目：国家重点研发计划项目(2023YFC3904700);国家自然科学基金面上项目(42477516)
* 通信作者, 教授, yangjx@rcees.ac.cn

库建设尚处于起步阶段,构建科学的DQA方法,可以指导数据收集、汇总计算、分析处理、审核入库等环节的开展,从而保证数据的可信度和准确性,推动数据库的高质量发展,为碳足迹数据库推动产业低碳发展提供支撑.本文将系统梳理 DQA 发展历程,解析 DQA 的研究现状和发展趋势,提出我国构建DQA的建议,为LCA数据体系构建及产品碳足迹管理体系建设提供支撑.

1 数据质量评价发展历程及研究现状

1.1 发展历程

数据质量通常采用数据质量指标(DQI)来表征[5].DQA 研究始于 20 世纪 90 年代初,至今经历了3 个发展阶段.

第一个阶段:方法初步形成阶段(约 1994~2004年).1994 年,环境毒理学与化学学会(SETAC)发布报告,确定了数据质量的基本概念和原则;讨论了能源、材料、环境排放、生态健康和暴露、人类健康和暴露等特定领域的具体数据质量问题;探索了推进 DQA 的研究方向和技术[7].

1995 年,美国国家环境保护局(US EPA)发布报告,提出了包含 8 个 DQI 的评价方法,分析了 DQA 的关键问题[8],为LCI数据质量评估奠定了理论基础.从此,DQA 实践不断发展和演变,特别是 1996 年,Weidema 等[9]引入谱系矩阵方法,将数据质量以矩阵形式呈现,包含5个DQI.随着LCA研究的深入,谱系矩阵方法逐步被广泛应用,成为各类 LCA 和数据库指南导则中 DQA 的主要参考.

第二个阶段,各类 DQA 方法发展阶段(2005~2015 年).各类 LCA 和数据库指南纷纷制定自己的DQA 方法.2006 年,ISO 14044 标准对数据质量做出了定性要求,包括时间跨度、地域范围、技术覆盖面、数据来源等 10 类 DQI,但标准中并未提供量化的、针对 LCA 结果的 DQA.与此同时,数据库 Ecoinvent、Gabi 也发布了各自的 DQA 方法.

第三个阶段:方法丰富发展阶段(2016 年至今),体现在关注不同层次和不同行业的数据质量问题.例如:2016 年,US EPA 发布新的 DQA 方法,建立 LCI数据两个级别的 DQI—流和过程[5].另外,随着材料领域 LCA 的快速发展,研究者开始关注该领域的数据质量问题[10].这些研究不仅提升了 LCA 数据的可靠性,也推动了 DQA 方法在特定行业的深化应用.

表 1 不同方法中的 DQI

Table 1 DQI in different methods

序号	DQI	Weidema,1996	ISO,2006	ILCD,2010	Ecoinvnet,2013	GLAD	PEF	Gabi	EPA-flow	EPA-process	USDA	EPA,1995
1	时间跨度	✓	✓	✓	✓	✓	✓	✓	✓		✓	
2	地域范围	✓	✓	✓	✓	✓	✓	✓	✓		✓	
3	技术覆盖面	✓	✓	✓	✓	✓	✓	✓	✓		✓	
4	精度		✓	✓		✓	✓	✓			✓	✓
5	完整性	✓	✓	✓	✓	✓		✓			✓	✓
6	代表性		✓			✓						✓
7	一致性		✓			✓		✓				
8	可再现性		✓			✓						
9	数据源		✓			✓						
10	信息的不确定性		✓			✓					✓	
11	方法科学性和一致性			✓								
12	可靠性	✓			✓						✓	
13	数据收集方法								✓			✓
14	过程审核									✓		
15	过程完整性									✓		
16	可接受性											✓
17	误差											✓
18	可对比性											✓
19	参考											✓

1.2 主要数据质量评价方法及特征

1.2.1 国际参考生命周期数据系统(ILCD) 欧盟

制定的ILCD为LCA的一致性和高质量应用提供指导,并为制定具体产品的 LCA 标准、指南和工具提供技术基础.ILCD 支持双重数据质量评价方法:第一种方法通过 DQI 解决数据质量问题,包含了 6 个DQI(表1),用户根据数据与ILCD手册中列出标准的符合程度,按 1−5 的等级对每个 DQI 进行排序;第二种方法是评估数据质量的配套资料,如数据文档、数据审查和一致的命名法[11].

1.2.2　Ecoinvent 数据库数据质量指南　Ecoinvent 使用两种不确定性来评估数据质量:基本不确定性是由于测量不确定性、活动特定变化、时间变化等原因,产生的数值变化和随机误差;额外不确定性是由于使用估计值、缺乏验证、样本不完整或从时间、空间、技术上不同条件推断而产生的不确定性.Ecoinvent 使用包括 5 个指标的谱系矩阵(表 1).应用这些指标,Ecoinvent 提出了一种从谱系矩阵数字计算额外不确定性因子的方法.首先,按 1 到 5 的等级评估数据质量;然后,通过使用由Ecoinvent开发的一组不确定性因子计算额外不确定性;最后,由谱系矩阵计算出的额外不确定性加上基本不确定性确认总体不确定性[12].

1.2.3　US EPA 数据质量评价指南　US EPA 制定 DQA 方法始于 1995 年,并在不断更新.2016 年,US EPA 制定了最新指南,以进一步支持可重复的 LCI 数据质量结果并告知用户如何正确应用 US EPA 的数据质量系统(DQS).最新指南定义了流和过程的数据质量指标.流指标进行了重新构建并明确定义,较为完善;过程指标较为新颖,仅定义了两个指标(表 1).同时,US EPA 指南强调了数据文档的重要性,但没有提供汇总不同指标的默认程序,而是由 LCA 从业者自行决定[5].

1.2.4　欧盟 PEF 数据质量要求　产品环境足迹(PEF)是欧盟委员会的一项倡议,它始于2010年左右,其理念是“确保消费者获得有关产品环境性能的可靠信息”.欧盟 PEF 的 DQA 包括两个基础要求:完整性和方法论的适当性、一致性;四个 DQI;三方面质量保障:数据文档、命名和审查.DQA 计算基于 4 个DQI的平均值,每个DQI分为5个等级进行打分.PEF要求数据符合ILCD格式,并满足PEF兼容数据集指南中提供的元数据信息的要求[13].

1.2.5　Gabi 数据库数据质量评价方法　Gabi 数据库采用了 PEF 和 ILCD 中的 DQA 方法.Gabi 主要将DQI(表 1)作为标准,涉及由可用信息级别决定的评分过程或标签系统.Gabi 中 DQI 评价的非数值数据(非常好/差)需要转换为质量等级(1 到 5),因为根据ILCD 手册中开发的类似方法,质量等级使用标签系统(好/差),然后按照 PEF 或 ILCD 指南中的 DQA 对其进行加权[14].

1.2.6　生命周期评价数据库的全球指导原则(GLAD)　GLAD 由联合国环境规划署和环境毒理学和化学学会(UNEP/SETAC)于 2014 年发起,为增强全球范围内 LCA 数据集和数据库的可交换性和互连性提供基础.该指南提供了 10 个 DQI 的指导原则,与ISO14044保持一致,但修改了第9个指标:数据来源是指定义数据来源的文档资料.与其他 DQA 相比,由于建模方面的问题比时间或区域更加多样化,GLAD 更侧重于解决 LCI 建模中的差异性,以提高数据可用性[15].

1.2.7　美国农业部 (USDA) 数据质量评价方法　USDA 制定了可应用于流指标的 DQA.针对 7 项指标中的每一个指标都制定了标准声明.它采用两个等级,如果指标符合标准,则获得“A”,如果不及格,则获得“B”;该方法简化了数据质量评分并提高可重复性[16].

1.3　数据质量指标含义

1.3.1　指标类型及含义　基于以上 DQA 发展历程及主要方法,本节汇总主要 DQI(表 2).这些指标基本分为 3 类:第一类是数据表征和一致性方面(时间、地理、技术代表性等);第二类是建模方面(方法科学性和一致性、选定的命名法、可再现性等);第三类是测量方面(完整性、来源的可靠性、数据的不确定性等);第四类是程序方面(过程审核等).

1.3.2　重点 DQI 释疑　不同的 LCA 或数据库指南使用不同的 DQI,有些指标是常用的,含义已经基本达成共识,但有些指标含义会略有不同或是某个指南特有的,以下将讨论这些指标.

(1)数据收集方法

数据收集方法是US EPA定义的指标,是指对所用采样方法(即样本量)的鲁棒性和数据收集期的稳健性的评估.LCI 数据收集中的样本量通常受数据可用性的限制,因此该指标衡量样本量和与所需样本群体相关的限制因素[5].

表 2 DQI 类型及含义

Table 2 Types and meanings of DQI

序号	DQI	涵义
1	时间跨度	数据年份及收集数据的跨度[15]
2	地域范围	为实现研究目的所收集的单元过程数据的地域[15]
3	技术覆盖面	具体的技术及技术组合[15]
4	精度	对每一个数据值的变动的度量[15]
5	完整性	测量或测算的流所占比例[15]
6	代表性	更广义上指数据集如何很好地代表目标数据集的时间、位置和技术[15], 狭义上,代表性是指数据集的代表性,通过考虑代表给定和所述市场的数据份额
7	一致性	方法学是否能统一应用到不同的分析内容中而进行的定性评价[15]
8	可再现性	对于其他独立从业人员采用同一方法学和数据值信息获取相同研究结果的定性评价[15]
9	数据源	数据来源的记录[15]
10	信息的不确定性	数据、模型和假设[15]
11	方法科学性和一致性	应用的 LCI 方法和方法论选择(例如分配,替换等)符合数据集的目标和范围,尤其是其预期的应用程序和决策支持环境[5]
12	可靠性	来源的可靠性[9]
13	数据收集方法	对所使用的采样方法的鲁棒性的评估(即样本量)和数据收集期[5]
14	过程审核	评估数据集在单位过程级别上进行的审查级别[5]
15	过程完整性	所包含的流能够代表实际情况的程度[5]
16	可接受性	数据来源通过审核;误差通过专家判断[8]
17	误差	与真实值的差距[8]
18	可对比性	不同的方法、数据集等能够带来相似的结果[8]
19	参考	数据源的完整参考记录[8]

(2)方法科学性和一致性

方法的选择,尤其是生命周期建模方法通常会对结果产生很大的影响(例如,归因建模与归果建模).因此,选择最合适的建模原则和方法,并在整个分析过程中保持一致性,对于结果的合理性和可重复性非常重要[5].

(3)不确定性

不确定性是 DQI 中含义最广泛的指标.方法中区分了模型不确定性、输入不确定性和数据可变性.模型不确定性是指不同建模者对于建模因素(系统边界、分配规则等)选择不相同,因此具有不确定性.输入不确定性是指 LCA 模型中的输入数据不完全确定,这种不确定性也称为参数不确定性.可变性是指在现实中发生变化但未反映在 LCA 模型中的事件[17].

在 US EPA 系统中,不确定性不作为指标;PEF 数据质量系统和 Gabi 数据库中使用精度代替不确定性,并根据来源的可靠性进行评估(测量、文献、计算、估计等方式)[18].在 Ecoinvent 中,不确定性的定量概率分布函数结果基于数据质量指标序数得分,并与专家评价相结合.Ciroth 等[19]指出 Ecoinvent 数据库并未记录将指标得分转化为定量不确定性的具体方法,这引发了对提供的定量不确定性的讨论.

1.4 数据质量评价其他要素

为了更好地开展 DQA 以及 LCA 研究的审查,数据质量通过两种互补的方法来解决:首先是严格意义上的数据质量,即确定清单数据和相关 LCIA 结果质量方面.其次是与数据质量文档、数据审查、命名和术语一致性相关的方面[11].所以,DQA 除了开展指标评价和定量不确定性分析外,还应包括以下要素.

(1)数据格式

有的研究在DQA中提出数据格式,指导LCA各种基本元素的命名和分类,例如流、单位等.欧盟提出使用 ILCD 作为统一数据集格式[11].

(2)数据审查

多个指南中提到 DQA 需要开展数据审查.审查范围应包含:数据集的目标和范围定义、原始数据、LCI 方法、适用的 LCIA 方法、单元流程清单、汇总清单、数据集文档等.数据集开发人员有责任在数据集审查过程中准备自我评估报告,以评估数据集的质量[12].

(3)数据文档

文档的范围和细节是支持数据透明度并确保结果可复制的关键.同时,文档对于 LCA 从业者了解数据集的代表性及适用性很重要.文档报告信息表格(报告,数据集)和格式(ILCD 参考格式,ILCD 报告模板等),可以通过电子方式交换[12,15].

(4)命名法

命名法是与基本流和其他元素的命名和结构有关的问题.统一的命名法则可确保不同从业人员能够始终如一地使用数据(例如,基本流二氧化碳是通过名称和识别号确定),并且 LCI 数据可以与该命名正确链接[15].

1.5　数据质量量化评价方法

在 DQI 基础上,需要对指标进行量化,常用方法如下:(1)谱系矩阵.此方法采用矩阵形式展示数据质量评分结果,即根据定量化表格,对 DQI 进行打分,最终表示为 $n\times1$ 阶矩阵.目前 Ecoinvent 数据库即采用此方法进行评价,并在数据集的具体信息中记录各输入、输出项的谱系矩阵[12].(2)分数累加法.此方法对数据质量设置一个满分,对不同来源、不同类型、不同时间的数据赋分,如最新的现场测量数据分数最高.该方法以计算单元过程所有数据的平均得分来判断数据质量.(3)加权法.通过权重赋值规则,对各项 DQI 赋权重,将 DQI 得分和权重相乘汇总 DQI 评价结果为一个指数.

1.6　数据质量评价流程

对于数据质量控制而言,除了 DQI 和定量不确定性评价外,还需要科学统一的评估流程.

SETAC 定义了 DQA 的流程,提出将 DQA 应用于 LCA 实施过程.首先,在目标和范围界定过程中建立数据质量目标(DQG),必要时还需针对特定生命周期阶段和数据类别建立相应 DQG.其次,开展数据采集.随后,对数据质量进行判断,如果数据质量满足要求,继续进行;否则应采取措施.可能的措施包括获取更好的数据、明确定义所收集数据的局限性、进行敏感性以不确定性分析、修改研究目的/范围或停止研究[5].

US EPA 提出所有 DQG 均应在目标和范围阶段确定,并应指导数据收集过程.在清单分析期间,应根据数据与已建立的 DQG 相比的方式评估数据质量.LCA 结果的解释应包括对 DQA 的解释.DQG 用于描述数据代表性和过程完整性,并应描述所建模的日期、位置、技术、系统边界、数据收集的技术领域,以避免数据波动[8].

US EPA 提出 DQA 涉及五个步骤:(1)确定 DQG 和数据收集设计;(2)进行初步数据描述:审查质量检查报告,计算基本统计数据并生成数据图;(3)进行统计测试:根据 DQG,对采样设计和初步数据描述进行审查,选择最合适的程序来汇总和分析数据;(4)验证统计测试的假设:鉴于有关研究的实际数据和其他信息,评估基本假设是否存在,或者是否可以接受;(5)得出结论:开展统计测试所需的计算,并记录由于这些计算而得出的推论.

1.7　不确定性和数据质量指标的关系

现有 DQA 可分为定性或定量,这是因为该过程依赖于用户判断来分配标签,例如 Gabi 中的好和差,或如 Ecoinvent 中的谱系矩阵,数字分数从 1 到 5.在这些方法中,只有 Ecoinvent 将 DQA 分数与定量不确定性评估联系起来[20].谱系矩阵单独使用时不能提供有关结果质量的充分信息,为了克服这一限制,必须结合数据不确定性的信息来看待谱系矩阵.

不确定性和数据质量评价具有不同的属性[21].在 DQI 基础上,研究者提出进一步将单元过程清单数据的 DQI 评分转化为以统计参数形式表示的不确定性[12,22],并且提出了诸如误差传递算法[23]、区间算法[24]、随机模拟[25]、泰勒级数展开[26-27]、修改后的 Beta 分布等[28-29],以上方法计算不确定性在 LCA 模型中的传递,并最终得出 LCA 结果的不确定性.Ecoinvent 数据库给出了数据库中各条汇总清单数据结果的不确定性[12].

Ecoinvnet 提出,如果样本数据可用,则可以直接计算样本的概率分布和标准差.如果样本较小,则可以根据范围(最大观测值与最小观测值之间的差值)计算出近似的标准差.对于正态分布,当样本量分别为 10、30 和 100 时,范围大约是标准差的 3、4 和 5 倍.生命周期清单数据通常来自少量观测,因此当观测数量未知时,使用因子 3 更为合理[12].

1.8　数据质量评价的前沿方法和发展趋势

近年来,DQA 在方法与技术路径上不断突破,整体呈现出融合化、智能化、行业化的发展趋势,已成为当前 LCA 领域研究的关键热点.在此背景下,国际上如瑞士 Ecoinvent 中心、US EPA 和欧盟联合研究

中心(JRC)等研究机构,通过构建标准化谱系矩阵方法、提出多层次 DQI 体系和统一数据格式规范,持续推动DQA方法的发展.在国内,中国科学院生态环境研究中心、中国电子技术标准化研究院、北京工业大学等单位,积极开展 DQA 标准化和指标体系构建等方面的研究,旨在建立契合我国国情的高质量 LCA 数据管理体系[30].

具体而言,当前 DQA 研究主要聚焦于以下几个方向:一是谱系矩阵方法的深化与不确定性建模的耦合.Ecoinvent 数据库团队提出了一种创新方法,将 DQI 评分转化为概率分布参数,结合蒙特卡模拟、区间算法和泰勒级数展开等方法开展不确定性传导计算,使谱系矩阵由定性工具逐步向定量工具转型[12].基于此,Kim 等[31]利用 Ecoinvent 数据库开展了敏感性分析,系统量化了生命周期背景数据不确定性对 LCA 结果的影响.此外,针对产品碳足迹评价,建立了基于敏感度和不确定度分析的数据质量模型.该模型结合谱系矩阵和蒙特卡罗模拟,量化 DQI 并识别关键影响因子,有效提升了碳足迹评价结果的可靠性[30].

二是人工智能辅助 DQA 的研究正稳步推进.MacMaster等提出利用大语言模型模拟专家评分流程,基于时间、地域、技术等多维度谱系矩阵,自动生成数据质量等级,从而提高了评估过程的一致性与效率.该方法已在多流程、多数据源背景下开展了初步的应用研究,展现出人工智能在推动 DQA 流程标准化方面的应用潜力[2].

三是 DQA 方法体系正逐步向行业化方向发展.例如,Notarnicola 等[3]针对意大利农业食品行业的典型产品开展研究,指出现有数据库在特定行业应用中的局限性,并提出构建行业定制化的数据清单及产品级的 DQA 体系,以应对通用数据库适用性不足的问题.Moutik 等[32]面向纤维增强聚合物复合材料,开发了能够区分碳纤维、玻璃纤维等不同材料类型的行业专属 DQA,并根据制造工艺环节定制相应的数据质量等级和不确定性阈值.

总体来看,当前 DQA 研究正沿着方法深化、技术集成与行业扩展持续推进,不仅提升了数据评价的准确性和一致性,也为多源、多尺度、多行业数据环境下的 LCA 研究提供了方法基础.

1.9 基于不同数据来源的数据质量评价重点

根据数据来源的差异可将 DQA 分为三个方面:商业数据库、文献资料数据和前端过程数据的质量分析.商业数据库可以提供标准化、结构化的背景数据,其数据覆盖度广但可能存在时间滞后或区域代表性不足等问题,排放依赖行业标准,本地化程度低,其 DQA 的研究重点在于数据权威性、时间/地理覆盖度等问题.文献资料作为二手数据,其数据碎片化程度高,基于实验或模型的排放受限于原始研究条件,其 DQA 需关注文献权威性、数据采集方法及适用性.前端数据作为一手资料,本地化程度高,但采集成本较高且易受测量误差影响,DQA 核心在于采样代表性、测量精度及不确定性量化.目前可以通过多源数据共同分析提高 DQA 的准确性,最终确保研究结论的科学性和可靠性[33-35].

2 数据质量评价存在问题及建议

2.1 存在问题

2.1.1 缺乏统一 DQA 方法及流程 目前,DQA 尚未形成统一的评价方法及流程.各类指南中的 DQA 方法各具特征,部分指标虽然达成了共识,但是评判方法却存在主观性.另外,LCA 领域对于如何处理清单和数据集中的不确定性还远未达成一致.LCA 中的不确定性因缺乏基于经验的数据而受到影响,也很少在原始数据收集中被解决,且通常不会呈现在最终报告中[18].

2.1.2 缺乏针对不同层面的DQA方法 DQA不仅与建模过程的每个数据集相关,还与建模过程连接以及整个生命周期清单相关,因此也与整体 LCA 研究相关[36].LCA 研究需要综合系统的数据质量管理,而不仅仅是作为 LCA 研究的一部分进行被动和定期的管理[37].所以,目前还需要进一步明确在流、过程数据及 LCA 整体结果的 DQA 方法[20].

2.1.3 未体现不同行业数据质量要求差异 虽然 DQA 在不断发展,但是不同行业对数据需求及数据质量关注点差距较大.早在 1994 年,SETAC 就提出针对不同行业需要关注的数据质量问题,并针对不同行业和不同类型数据提出了数据质量的具体要求[7].但目前针对不同行业的数据质量评价研究较为滞后,尚不完善.

2.2　构建 DQA 方法的建议

数据质量评价是未来 LCA/碳足迹作为决策支持工具,以及在行业内更广泛应用的重要支撑.在政策引导下,我国将会开展碳足迹评价及认证,构建 DQA、保障数据质量是需解决的首要问题.

2.2.1　建立数据质量评价标准　尽管 LCA 是一个成熟的领域,我国已经发布相应标准和其他实践的指导文件,包括评估和管理数据质量,但是,这些标准和指导文件并没有为 DQA 提供详细内容.我国目前正处于碳足迹数据库构建的起步阶段,后期会有大量 LCA/碳足迹工作开展,目前在碳足迹实践及标准制定过程中迫切需要 DQA 方法支撑.所以,制定标准化的具有指导性的 DQA 方法能够为碳足迹评价推广提供基础.

2.2.2　建立完善科学的数据审查程序　数据审查是保证数据质量的重要措施.科学家和行业专家应积极参与,明确数据审查方法和程序[10].数据审查必须相对于一系列质量标准进行,提供有关审查过程的详细信息.审查需要规定审查的方式、标准、流程、文档等.在数据集审查过程中,有两个关键问题:首先,在 LCI 数据集入库之前,需进行审核验证;其次,在 LCI 数据集包括在公开可用的 LCI 数据库中之前,需经过外部独立评估[15].

2.2.3　明确数据格式及数据文档要求　LCI 数据集格式和数据文档要求是数据审查的基础,也是数据互信互通的基础.统一物质目录、规定数据格式与信息,能够帮助规范数据质量.建立我国碳足迹数据库统一的数据格式及数据文档要求,兼顾国际互通,也要兼顾我国的 LCA/碳足迹数据信息情况、产业现状及保密性要求.

2.2.4　采用新技术手段提升数据质量评价效率　随着对 LCA 需求的增长,DQA 工作规模也非常巨大.由于将 DQA 应用于 LCA(例如,包括数百个流程)中的每条数据是一项耗时的工作.人工智能具有规范从业者的工作流程、提高工作效率的能力.有研究证明大语言模型可以成功执行 DQA[2].所以,积极寻求新技术手段提高 DQA 的效率和质量至关重要.

2.3　数据质量评价基础框架

基于目前的 DQA 现状,本文提出 DQA 的基础方法框架,包括 DQA 流程和评价方法两方面(图 1).

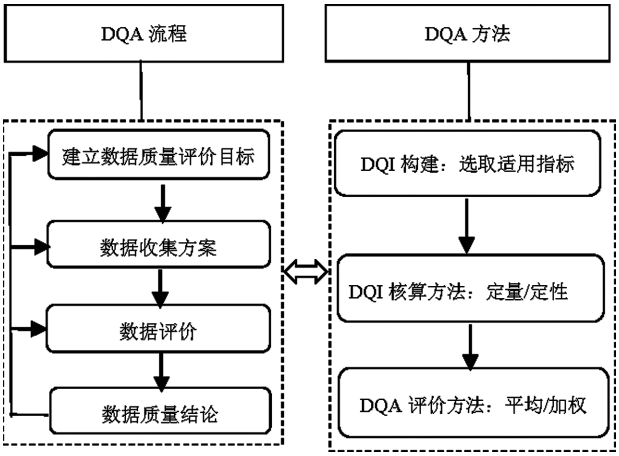


图 1　DQA 流程及评价方法框架

Fig.1　Framework of DQA process and assessment method

3　总结

在我国构建产品碳足迹因子数据库、开展碳足迹认证等政策引导下,各行业和企业将会开展大量 LCA/碳足迹研究工作.其中,DQA 是制定相关标准、收集数据、选择背景数据的重要支撑.目前,国内外 DQA 基本都基于谱系矩阵,在 DQI 的基础上开展定性和定量评价;但整体还存在 DQI 不明确、评价方法不统一、量化评价困难等问题.当前,我国需要加大 DQA 研究,构建适合我国碳足迹评价数据现状的 DQA 流程、指标和方法,同时注重数据审核、数据文档等要求,为我国产品碳足迹管理体系建设提供重要支撑.

参考文献:

[1] Hellweg S, Milà I, Canals L. Emerging approaches, challenges and opportunities in life cycle assessment [J]. Science, 2014,344(6188): 1109−1113.

[2] Macmaster S, Sinistore J. Testing the use of a large language model (LLM) for performing data quality assessment [J]. The International Journal of Life Cycle Assessment, 2024,https://doi.org/10.1007/s11367-024-02405-8.

[3] Notarnicola B, Tassielli G, Renzulli P A, et al. Life cycle inventory data for the Italian agri−food sector: background, sources and methodological aspects [J]. The International Journal of Life Cycle Assessment, 2024,29(8):1430−1445.

[4] Guinée J. Handbook on life cycle assessment — operational guide to the ISO standards [J]. The International Journal of Life Cycle Assessment, 2001,6(5):255.

[5] Edelen A, Ingwersen W. Guidance on data quality assessment for life cycle inventory data [M]. US Environmental Protection Agency, Washington, DC. 2016.

[6] Edelen A, Ingwersen W W. The creation, management, and use of data quality information for life cycle assessment [J]. The International Journal of Life Cycle Assessment, 2018,23(4):759−772.

[7] Fava J, Jensen A A, Lindfors L, et al. Life−cycle assessment data quality: A conceptual framework [M]. Society of Environmental Toxicology and Chemistry and SETAC Foundation for Environmental Education, 1992.

[8] Bakst J S, Lacke C J, Weitz K A, et al. Guidelines for assessing the quality of life−cycle inventory analysis [C]. EPA (NTIS), 1995.

[9] Weidema B P, Wesnæs M S. Data quality management for life cycle inventories—an example of using data quality indicators [J]. Journal of Cleaner Production, 1996,4(3):167−174.

[10] Zhang C. 23 − Life cycle assessment (LCA) of fibre reinforced polymer (FRP) composites in civil applications [M]. Eco−efficient Construction and Building Materials, Pacheco−Torgal F, Cabeza L F, Labrincha J, et al, Woodhead Publishing, 2014:565−591.

[11] Pennington D W, Chomkhamsri K, Pant R, et al. ILCD handbook public consultation workshop [J]. The International Journal of Life Cycle Assessment, 2010,15(3):231−237.

[12] Bauer C, Hischier R, Mutel C, et al. Overview and methodology−Data quality guideline for the ecoinvent database version 3 [R]. 2013.

[13] Lehmann A, Bach V, Finkbeiner M. Product environmental footprint (PEF). Fortschritt oder Rückschritt für die Ökobilanzforschung? [J]. uwf Umwelt Wirtschafts Forum, 2016,24:83–87.

[14] Nada salbi, Databases & Modelling Principles 2022 [M]. Sphera, 2022,2:1−3.

[15] Sonnemann G, Vigon B, Broadbent C. Global guidance principles for life cycle assessment databases: A basis for greener processes and products [M]. Paris, France: United Nations Environment Programme, 2010.

[16] Cooper J S, Kahn E. Commentary on issues in data quality analysis in life cycle assessment [J]. The International Journal of Life Cycle Assessment, 2012,17(4):499−503.

[17] Huijbregts M A J. Application of uncertainty and variability in LCA [J]. The International Journal of Life Cycle Assessment, 1998,3(5): 273−280.

[18] Hunkeler D. Life cycle inventory analysis−methods and data [J]. The International Journal of Life Cycle Assessment, 2020,25(6):1168−1170.

[19] Ciroth A, Muller S, Weidema B, et al. Empirically based uncertainty factors for the pedigree matrix in ecoinvent [J]. The International Journal of Life Cycle Assessment, 2016,21(9):1338−1348.

[20] Balcioglu G, Fitzgerald A M, Rodes F A M, et al. Data quality and uncertainty assessment of life cycle inventory data for composites [J]. Composites Part B: Engineering, 2025,292:112021.

[21] Coulon R, Camobreco V, Teulon H, et al. Data quality and uncertainty in LCI [J]. The International Journal of Life Cycle Assessment, 1997, 2(3):178−182.

[22] Frischknecht R, Jungbluth N, Althaus H J, et al. Overview and methodology, Data v2.0 (2007). Ecoinvent report No. 1 [J]. Environ Report No.1, 2007,(1):1−77.

[23] Heijungs R. Identification of key issues for further investigation in improving the reliability of life−cycle assessments [J]. Journal of Cleaner Production, 1996,4(3):159−166.

[24] Chevalier J, Téno J L. Life cycle analysis with ill−defined data and its application to building products [J]. The International Journal of Life Cycle Assessment, 1996,1(2):90−96.

[25] Maurice B, Frischknecht R, Coelho−Schwirtz V, et al. Uncertainty analysis in life cycle inventory. Application to the production of electricity with French coal power plants [J]. Journal of Cleaner Production, 2000,8(2):95−108.

[26] Heijungs R. Sensitivity coefficients for matrix−based LCA [J]. The International Journal of Life Cycle Assessment, 2010,15(5):511−520.

[27] Ciroth A, Fleischer G, Steinbach J. Uncertainty calculation in life cycle assessments [J]. The International Journal of Life Cycle Assessment, 2004,9(4):216−226.

[28] Yoo W, Ozer H, Ham Y. System−level approach for identifying main uncertainty sources in pavement construction life−cycle assessment for quantifying environmental impacts [J]. Journal of Construction Engineering and Management, 2019,145(2):4018137.

[29] Kennedy D J, Montgomery D C, Quay B H. Data quality − Stochastic Environmental Life Cycle Assessment Modeling [J]. The International Journal of Life Cycle Assessment, 1996,1(4):199−207.

[30] 陈 莎,李燚佩,曹 磊,等.产品碳足迹评价中不确定度与敏感度相结合的数据质量分析 [J]. 中国环境科学, 2014,34(4):1067−1072.
Chen C, Li Y P, Cao L, et al. Data quality analysis with combination uncertainty and sensitivity for carbon footprint assessment of products [J]. China Environmental Science, 2014,34(4):1067−1072.

[31] Kim A, Mutel C L, Froemelt A, et al. Global sensitivity analysis of background life cycle inventories [J]. Environmental Science & Technology, 2022,56(9):5874−5885.

[32] Moutik B, Summerscales J, Graham−Jones J, et al. Quality assessment of life cycle inventory data for fibre−reinforced polymer composite materials [J]. Sustainable Production and Consumption, 2024,49:474−491.

[33] Batini C, Cappiello C, Francalanci C, et al. Methodologies for data quality assessment and improvement [J]. Association for Computing Machinery Computing Surveys, 2009,41(3):16.

[34] Wang R Y, Storey V C, Firth C P. A framework for analysis of data quality research [J]. IEEE transactions on knowledge and data engineering, 1995,7(4):623−640.

[35] Bergdahl M, Ehling M, Elvers E, et al. Handbook on data quality assessment methods and tools [M]. Ehling, Manfred Körner, Thomas, 2007.

[36] Bretz R. SETAC LCA Workgroup: data availability and data quality [J]. The International Journal of Life Cycle Assessment, 1998,3(3): 121−123.

[37] Bicalho T, Sauer I, Rambaud A, et al. LCA data quality: A management science perspective [J]. Journal of Cleaner Production, 2017,156:888−898.

作者简介：丁 宁(1983−),女,河北保定,副研究员,博士,主要从事产业生态学研究.发表论文 50 余篇.ningding@rcees.ac.cn.