Development and application of basis database for materials life cycle assessment in china

Xiaoqing Li\textsuperscript{1a}, Xianzheng Gong\textsuperscript{1b} and Yu Liu\textsuperscript{1e}

\textsuperscript{1}Beijing University of Technology, Beijing 100124, China

E-mail: \textsuperscript{a}lixiaoqing@bjut.edu.cn, \textsuperscript{b}gongxianzheng@bjut.edu.cn, \textsuperscript{c}liuyu@bjut.edu.cn

\textbf{Abstract.} As the data intensive method, high quality environmental burden data is an important premise of carrying out materials life cycle assessment (MLCA), and the reliability of data directly influences the reliability of the assessment results and its application performance. Therefore, building Chinese MLCA database is the basic data needs and technical supports for carrying out and improving LCA practice. Firstly, some new progress on database which related to materials life cycle assessment research and development are introduced. Secondly, according to requirement of ISO 14040 series standards, the database framework and main datasets of the materials life cycle assessment are studied. Thirdly, MLCA data platform based on big data is developed. Finally, the future research works were proposed and discussed.

\textbf{Keywords:} LCA database, material life cycle inventory, data acquisition, big data

1. Introduction

With the rapid growth of economy, the contradiction among economic development, resources and environment has become increasingly acute. Life cycle assessment (LCA) is considered to be one of the most important tools to analyze and improve environmental performance of materials and products, which provides technical and decision making support in many areas (e.g., achieve energy-saving and emission reducing targets) [1]. High quality environmental burden data is an important premise of carrying out materials life cycle assessment (MLCA), and the reliability of data directly influences the credibility of the life cycle assessment results and its application performance.

From the point of LCA methodology, the researches of data quality are mainly embodied in the uncertainty analysis of the life cycle inventory (LCI) results and the sensitivity analysis of the LCI data. Then the critical data with high uncertainty and high sensitivity are identified in the LCA model. Thus the key points of control and improve the data quality are pointed out. While in the field of information technology, the consistency and correctness of the data can be guaranteed when the data is described by the integrity constraints of the database. Therefore, establishing a reliable LCA database and develop the corresponding LCA software are important methods to ensure the reliability and quality of inventory data. Due to the unbalanced distribution of resources in different countries and regions, the different levels of science and technology and the asymmetry of energy consumption, the impact of materials (products) on the environment in the whole life cycle is very different. Many countries, research institutes and enterprises are struggling to establish generic or domestic professional LCA database.

Currently, in China, materials environmental burden data mainly come from survey data and published statistical data [2]. At present, there are a lot of problems in collecting LCI data: most time
and effort consuming, long period, data missing, updating slowly and so on. These problems will inevitably affect the reliability of the LCA results. Therefore, it is necessary to accumulate certain field monitoring data and carry out statistical analysis, and study materials life cycle inventory data acquisition based on big data processing technology.

2. Related Work

It is well-known that LCA is very data intensive. In order to facilitate this work and to avoid duplication in data compilation, many databases have been developed in the last decade around the world. Over the past 20 years, different generic LCA databases have been developed, such as SPINE@CPM database in Sweden, LCI database in Australia, Gabi database in Germany, NREL-USLCI in America, ECOINVENT in Swiss and ELED database in Europe [3-6].

Research works on LCA data in China are started relatively late. With the support of the National 863 Program, environmental burden data of main materials and products were collected and processed (Steel, cement, aluminum, engineering plastics, architectural coatings, ceramics, etc.). Based on these data, the MLCA foundation database (SinoCenter) and the network platform have been developed at BJUT. With the continued support of National 863 Program, National 973 Program, National Sci-Tech Support Plan and Beijing natural science fund, presently, more than 120,000 records are incorporated in the Sino Center database [7-10]. In addition, researches on LCA database are ongoing at China standardization institute and Sichuan university [11].

At present, the development and application of domestic professional LCA software tools are not mature. SimaPro and Gabi are used mostly in related researches and practical applications. The performance of these two kinds of software is compared through the case in [12] and [13]. For LCA researchers and practitioners, they often can only get a part of the data, other data cannot be obtained because of a variety of reasons in a practical application. Through the analysis of the data structure of production process, a kind of completion of nonnegative matrices model is proposed to solve the data missing problem in [14], and the model and corresponding algorithms have been applied to ECOINVENT database.

3. Development and Application of Chinese Database for MLCA

3.1. SinoCenter database

Life cycle assessment database system is a complex and huge application system. In order to maintain the integrity and consistency, some of the basic objects must be standardized, which are usually the most basic data in the database, such as basic units, flows, locations and classifications of information. Therefore, the SinoCenter database is divided into basic data set and application data set, which facilitates the management and maintenance of the database, and makes the database having a good compatibility and expansibility. The structure of the database is shown in figure 1.

The basic dataset includes the entities of basic flows, units, sites and classifications, which store all material information data, measurement units and the corresponding conversion relationships, geographic information data and classification information data in the whole database system.

The application dataset includes three parts, the management information dataset, the product system dataset and the LCA methodology dataset. The management information dataset main storage data related to the management and modeling validation, and description of the data sources and their characteristics in LCA practice to increase the transparency of data. The core data of LCA practice are stored in the product system dataset, includes plans, processes and exchange flows and so on. The evaluation method and corresponding evaluation factor data are stored in the LCA methodology dataset.
3.2. Main dataset

At present, there are more than 120,000 records in the SinoCenter database, includes 68 kinds of materials and process inventories, such as electric power products [15-17], fossil energy products [18-20], transportation [21,22], ferrous material [23,24], building materials [25-28], non-ferrous metal [29-32], polymer materials [33-35], bonding materials and so on [36].

LCA characteristics are analyzed and localized in the LCA methodology dataset, which includes Eco-indicator 99, CML, IPCC, etc. Based on the resources characteristics of China, the localization resource damage factor and normalization factor about 42 kinds of metal minerals and 58 kinds of non-metallic minerals resources are established. At the same time, the land resources damage factor and the normalization factor of different land use and conversion are established based on the situation of land resources in China. Details of the environmental burden datasets and data sources are shown in table 1.

Table 1. Datasets of material life cycle environmental burden data.

| Materials          | Data sources                                      | Datasets number |
|--------------------|---------------------------------------------------|-----------------|
| 1. Building materials | Cement: 32 NSP and 5 Shaft kiln cement production enterprises | More than 90    |
|                    | Glass: 12 float glass production enterprises       | More than 40    |
|                    | Ceramics: 4 sanitary wares and 24 construction tiles production enterprises | More than 60    |
|                    | Wall: 38 wall materials production enterprises     | More than 40    |
| 2. Ferrous metals  | Steel: More than 70 steel production enterprises   | More than 200   |
|                    | Alloy: Enterprises investigation and statistics    | 2               |
| 3. Nonferrous metals | Magnesium: Enterprises investigation            | 1               |
|                    | Aluminum: Enterprises investigation                | 1               |
|                    | Plumbum: Enterprises investigation                 | 1               |
|                    | Zinc: Enterprises investigation and statistics     | 1               |
|                    | Cuprum: Enterprises investigation and statistics   | 1               |
| 4. Energy          | Electricity: Enterprises investigation and statistics | 8              |
|                    | Fossil energy: Enterprises investigation and statistics | 10             |
| 5. Transportation  | Road: Enterprises investigation and statistics     | 10              |
6. LCA methods

| Method     | Localization | Note |
|------------|--------------|------|
| CML        | 1            |      |
| EPS        | 1            |      |
| Ecoindicator | 1        |      |

In addition, some famous foreign LCA environmental burden database are introduced as important parts of the SinoCenter database, such as Simapro, Umberto, Gabi, DEAM, IVAM and Ecoinvent, in which mainly related to daily necessities, paper, iron and steel, petrochemical, organic (inorganic) materials, energy, transport and electronic products and so on. In particular, the localization of the comparative researches and development for core content in Ecoinvent database is carried out. These database provide basic data for the comparative study of materials (products) LCA practice in China.

3.3. Application software development

Based on SinoCenter database, integrating the large number of researches in the field of models and methods, such as characteristic model of mineral resource depletion, environmental impact characteristic model of land use, etc., in view of the actual situation of energy conservation and emission reduction and ecological design of typical materials industry, the corresponding software tool kit is developed (2016SR001217, 2013SR038488, 2010SRBJ1357, 2007SRBJ2464, etc.).

These software can help companies identify the energy consumption in the product life cycle, and then find effective ways to improve production efficiency, provide data and methods for enterprises to carry out energy-saving emission reduction and eco-design products.

3.4. Internet Oriented Application Development

Material (product) life cycle environmental burden data has become one of the national materials science data sharing network node: material environmental burden data resource node (http://cnmlca bjut.edu.cn). In the data platform, the classification query based on query conditions, such as the type of material, the scale of production and the time range, can be realized.

4. MLCA Data platform based on Big Data

Big data has played an important role in the retail industry. So it is with the materials industry, whether the “Made in China 2025” proposed by China or “Industry 4.0” proposed by Germany, to implementation of the major materials country to the world materials power, it is needed to analyze and consider how to realize the innovation and reshaping of the traditional materials business model based on big data.

In the new century, government and enterprise generally use the information-based and digital technology to collect, processing, transmission and analysis the various data in the process of production. Simultaneously, data resources platform based on wide-area internet has entered a rapid development period, which provides good opportunities for the automatic collection of materials environment burden data. Nevertheless, there is not a unified data format in various materials industries. And the data collection process have the features of massive high concurrency, reading and writing. Therefore, materials environmental burden data show the characteristics of big data. For traditional data collection methods and relational data management system, it is tough to deal with these massive, heterogeneous and high concurrent read and write data, especially for data acquisition, cleaning, storage and management.

The objective of this section is to develop a data automatic acquisition and analysis solution for materials environmental burden data based on big data processing technology.

4.1. Architecture

---

**Table:**

| Method     | Localization | Note |
|------------|--------------|------|
| CML        | 1            |      |
| EPS        | 1            |      |
| Ecoindicator | 1        |      |

---
MLCA data platform adopts three layers of structure: client layer, application layer and data layer. Based on component based software development, MLCA platform with customizability, expandability and maintainability is developed. The architecture with core technologies is shown in figure 2.

4.2. Data layer

The data from web, App and sensor contains structured data, semi-structured data and unstructured data. In order to better compatible with relational databases, the unified storage platform design concept was used in this solution. New inventory data would be obtained from the original data by data mining, and used to update corresponding inventory data in SinoCenter for MLCA practice.

As the types and size of materials environmental burden data from different clients are ever-increasing, classification rules and metadata schemas need to keep learning and extending to adapt this situation. The framework of materials environmental burden data acquisition based on big data and classification rules learning process are shown in figure 3 and figure 4 respectively.

4.3. Experiments
Experiments were performed on servers with an 8 core, 2.67GHz Intel processor, 4GB of main memory and 4TB hard disk storage running on windows server 2008 operating system.

Three kinds of data types (1 million for each type of data) were continuous uploaded by three clients to simulate materials environmental burden data acquisition. The results is that the approximate number of three kinds of data respectively are 1, 0.9 and 1.1 million. Figure 5 shows the processing time and results.

Figure 6 shows the actual situation of classification. Accuracy rates respectively are 98%, 96.9% and 90%. Comparing with actual data, we find that semi-structured data are easy to be mistaken for unstructured data, which does not affect the later storage.

We compared our solution (notated as MLCADB) with MongoDB. The results in figure 7 show that MLCADB takes more processing time than MongoDB for big data storage and read operation. Nevertheless, MLCADB is developed for adapting relational database system better, and there are opportunities for optimization in a later work.

5. Conclusion
The establishment and application of Chinese life cycle inventory database provides scientific and basic data support for energy conservation and emission reduction, carbon footprint analysis, clean production audit and so on. At the same time, the development and application of big data in material industry is also facing some difficulties, in general, industry data acquisition is relatively difficult.

The development direction of Chinese LCA database mainly includes the following three aspects: The first aspect is the further study of data acquisition, data management and data quality assurance and other related technology, which provide technical support to database update and expansion. The second is on the method to expand the database structure and data need to be further studied, which can be used to support dynamic life cycle assessment. And finally, under the guidance of the concept of materials genome, the material environmental burden genomic needs to be established, the potential relationship among material structure, performance and environmental friendliness should be studied.

Acknowledgement
This work was both supported in part by the Beijing Science and Technology plan Project (Grant No. D161100002416001) and the Beijing Natural Science Foundation (Grant No. 2141001).

References
[1] Z. R. Nie 2013 Int. J. Life Cycle Assess 18 1435–1439
[2] X. Z. Gong, Z. R. Nie, Z. H. Wang et al. 2006 Rare Metals 25 101–104
[3] Swiss Centre for Life Cycle Inventories. The Ecoinvent Database [EB/OL]. (2015-09). http://www.ecoinvent.org/
[4] Industrial Environmental Informatics(IMI).The SPINE@ CPM Database[EB/OL].(2015-09).http://cpmdatabase.cpm.chalmer.se
[5] National Renewable Energy Laboratory (NREL). The USA Life Cycle Inventory Database [EB/OL].(2015-09).http://www.nrel.gov/lci
[6] European Commission-DG Joint Research Centre. The European LCI Database [EB/OL].(2015-09).http://eplca.jrc.ec.europa.eu/
[7] X. Z. Gong, Z. R. Nie, Z. H. Wang et al. 2006 Rare Metals 25 101–104
[8] X. Z. Gong, Q. Zhang, Y. Liu et al. 2009 Journal of Beijing University of Technology 35(7) 997–1001
[9] X. Z. Gong, Z. R. Nie and Z. H. Wang 2004 Journal of WuHan University of Technology 26(3) 12–14
[10] X. Z. Gong, Z. R. Nie, Z. H. Wang et al. 2011 Materials China 30(8) 1–7
[11] X. L. Liu, H. T. Wang, J. Chen et al. 2010 ActaScientiae Circumstantiae 30(10) 2136–2144
[12] I. T. Herrmann and A. Moltesen 2015 Journal of Cleaner Production 86 163–169
[13] E. Martinez, J. Blanco, E. Jimenez et al. 2015 Renewable Energy 74 237–246
[14] F. F. Xu 2014 Studies on Models Algorithms and Application for Matrix Completion (Shanghai: Shanghai Jiao Tong University)
[15] X. H. Di 2006 Several Fundamental Researches in the Life Cycle Assessment for Mineral Resources and Materials (Beijing: Beijing University of Technology)
[16] X. H. Di, Z. R. Nie and T. Y. Zuo 2005 China Environmental Science 25(5) 632–635
[17] X. H. Di, Z. R. Nie, B. R. Yuan et al. 2007 The International Journal of Life Cycle Assessment 5 331–336
[18] B. R. Yuan 2006 Measurement Method for Sustainable Development of Chemical Industry and Its Application (Beijing: Beijing University of Technology)
[19] B. R. Yuan, Z. R. Nie, X. H. Di et al. 2006 Modern Chemical Industry 4 59–61
[20] B. R. Yuan, Z. R. Nie, X. H. Di et al. 2006 Chemical Industry and EngineeringProgress 25(3) 334–336
[21] L. P. Ma 2007 The localization research on road transportation in China for material life (Beijing: Beijing University of Technology)
[22] Y. F. Zhang, X. Z. Gong 2011 Science Technology and Industry 11(2) 64–67
[23] H. M. Zhou 2001 Life Cycle Assessment on Iron and Steel Processes (Beijing: Beijing University of Technology)
[24] X. Z. Gong, Z. R. Nie and Z. H. Wang 2006 Research on Iron and Steel 4 49–52
[25] W. J. Chen 2007 The Study on Life Cycle Assessment of Flat Glass Production (Beijing: Beijing University of Technology)
[26] X. P. Li 2009 The Life Cycle Assessment of Sanitary ware in China (Beijing: Beijing University of Technology)
[27] N. Luo 2009 Research on Environmental Impact of Sintered Brick Production in China (Beijing: Beijing University of Technology)
[28] Q. Zhang 2009 The Research and Development of Life Cycle Analysis Network Database of Materials (Beijing: Beijing University of Technology)
[29] F. Gao 2008 Research on Life Cycle Assessment and the Application in China Magnesium
[30] F. Gao, Z. R. Nie, Z. H. Wang et al. 2009 *The International Journal of Life Cycle Assessment* **14** 480–489

[31] F. Gao, Z. R. Nie, Z. H. Wang et al. 2009 *Science in China, Series E: Technological Sciences* **52**(8) 2161–2166

[32] F. Gao, Z. R. Nie, Z. H. Wang et al. 2006 *Transactions of Nonferrous Metals Society of China* **18**(3) 749–754

[33] H. Chen, W. C. Hao, F. Shi et al. 2004 *Acta Scientiae Cicumdantiae* **24**(3) 545–549

[34] X. C. Meng 2010 *Life Cycle Assessment of Pycarbonante and Polylactic Acid* (Beijing: Beijing University of Technology)

[35] Y. J. Li 2007 *Life Cycle Assessment of CO2-Based Biodegradable Polymers* (Beijing: Beijing University of Technology)

[36] X. X. Wang 2008 *Life Cycle Assessment of Lead-Free Solders* (Beijing: Beijing University of Technology)
Abstract. Start your abstract here…

6. The first section in your paper
The first paragraph after a heading is not indented (Bodytext style).
Other paragraphs are indented (BodytextIndented style).

7. Another section of your paper
The first paragraph after a heading is not indented (Bodytext style).
Other paragraphs are indented (BodytextIndented style).

7.1. A subsection
Some text.

7.1.1. A subsubsection. The paragraph text follows on from the subsubsection heading but should not be in italic.

References
[37] A reference
This reference has two entries but the second one is not numbered (it uses the ‘Reference (no number)’ style.

[38] Another reference

[39] More references