Journal of Cleaner Production 43 (2013) 166–173

Contents lists available at SciVerse ScienceDirect

Journal of Cleaner Production

journal homepage: www.elsevier.com/locate/jclepro

A hybrid Data Quality Indicator and statistical method for improving uncertainty analysis in LCA of complex system – application to the whole-building embodied energy analysis

Endong Wang, Zhigang Shen*

The Durham School of Architectural Engineering and Construction, University of Nebraska-Lincoln, 113 NH, Lincoln, NE 68588-0500, USA

ARTICLE INFO

Article history:
Received 25 February 2012
Received in revised form
3 December 2012
Accepted 4 December 2012
Available online 27 December 2012

Keywords:
Data uncertainty
Whole-building embodied energy
LCA
Data quality indicator
Statistical
Stochastic
Monte Carlo simulation

ABSTRACT

Uncertainty analysis has been recommended when using LCA for choosing sustainable products. The existing uncertainty analysis methods are helpful but have more or less inherent deficiency. The goal of this paper is to present a hybrid stochastic method to improve the uncertainty estimate in LCA with data limitations. This method can be a valuable tool especially to evaluate deterministic results of LCA of complex product system (e.g. building) when uncertain information is needed for decision-making. Compared to deterministic results, probabilistic results were often considered more reliable when large data uncertainties existed, such as data uncertainties in embodied energy coefficients of building materials. Both the statistical and Data Quality Indicator methods have been used to estimate data uncertainties in LCA. However, neither of those alone is adequate to address the challenges in LCA of complex product system, due to the large quantity of material types and data scarcity. This paper presents a hybrid method, which combines Data Quality Indicator and the statistical method by using a prescreening process based on Monte Carlo rank-order correlation sensitivity analysis. By optimizing the utilization effect of the available statistical data, this hybrid method can increase the reliability of the uncertainty estimate compared to the pure data indicator method. In the presented case study which performed the stochastic estimating of whole-building embodied energy, improved results from the hybrid method were observed compared to the pure Data Quality Indicator method. In conclusion, the presented hybrid method can be used as a feasible alternate for evaluating deterministic LCA results like whole-building embodied energy, when more reliable results are desired with limited data availability. Although this approach is presented in the context of building embodied energy uncertainty analysis, it can be used for LCA uncertainty analysis for conveniently making more reliable decision in the case of choosing complex "greener" products in other fields.

1. Introduction

Estimating whole-building embodied energy (WBEE) is a significant part in whole-building life cycle assessments (WB-LCA) (Ortiz et al., 2009). WBEE is defined as the summation of the embodied energy of each individual building material (Eq. (1)). Embodied energy coefficient (EEC) was defined as the energy consumed per quantity unit of a building material during its production process (Alcorn and Baird, 1996). Traditionally, WBEE was estimated with deterministic approach which used a fixed point value to represent EEC and generated a single fixed point result. Due to differences in the production processes and the lack of detailed production data, there are significant variations in EEC values among many different life cycle inventory (LCI) databases (Costanza, 1980; Sugiyama et al., 2005). The variations can affect the results of WB-LCA significantly. This type of variations is generally termed "data uncertainty", which is a typical LCA uncertainty category (Huijbregts, 1998). Incorporating analysis of data uncertainty of EEC is considered to be an important improvement to the deterministic approach because it can provide more information for decision-making (Kennedy et al., 1996; Steen, 1997; Tan et al., 2002b; Acquaye et al., 2011; Sonnemann et al., 2003). However, the uncertainty analysis has rarely been conducted in building LCA (Acquaye et al., 2011).

$$\mathrm{WBEE} = \sum_i Q_i * \mathrm{EEC}_i \tag{1}$$

* Corresponding author. Tel.: +1 402 472 9470; fax: +1 402 472 4087.
E-mail addresses: zshen3@unl.edu, shen@unl.edu (Z. Shen).

0959-6526/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jclepro.2012.12.010

Where Q_i is the quantity of material i (tons); EEC_i is the EEC of material i (GJ/ton)

Data quality indicator (DQI) and statistical methods were often used to estimate data uncertainty in LCA with differing advantages and shortcomings (Sugiyama et al., 2005; EPA, April 1995; Junnila and Horvath, 2003; Hanssen and Asbjørnsen, 1996). DQI estimates the reliability and uncertainty of data based on the descriptive metadata and expert knowledge, such as data's age, source, etc. It can be used both qualitatively (Junnila and Horvath, 2003) and quantitatively (Coulon et al., 2011) in LCA studies. The statistical method, on the other end, fits data samples with the goodness of fit test to characterize data range with probabilistic distributions if enough data samples are available. Although less-accurate than the statistical method (Tan et al., 2002b; Venkatesh et al., 2011), DQI costs less when compared to the statistical method. Although the statistical method is desirable when high accuracy is required (Sugiyama et al., 2005), due to its high implementation cost, DQI is widely applied when the high accuracy of uncertainty estimate is not critical, or the data sample size is not large enough for meaningful statistical analysis. Data scarcity is also the driving force of some other DQI quasi methods, e.g. fuzzy logic (Tan et al., 2007), interval theory (Chevalier and Le Téno, 1996), possibility theory (Tan et al., 2002a), so they can also be utilized for uncertainty analysis in the situation of data scarcity in LCA with the aid of expert judgment.

However, neither DQI nor the statistical method alone is practical for stochastic WB-LCA. The major challenges include the large number of different types of building materials involved and the scarcity of data. Using DQI method alone may propagate significant errors considering the large number of different building materials and the large variations of EEC in building materials; on the other end, the statistical method will be too expensive when applied to every building material.

Heijungs (1996) introduced the concept of using *uncertainty* and *contribution* as two parameters to categorize life cycle inventory (LCI) data and recommended using different methods to estimate uncertainty based on the categorization. Both Maurice et al. (2000) and Canter et al. (2002) adopted Heijungs's data categorization concept in their works of LCI uncertainty analysis, in which deterministic approach was used in calculating parameter contributions. Huijbregts et al. (2001) pointed out the difficulty and importance of estimating the uncertainty distributions of LCI inputs for Monte Carlo simulations (MCS). Huijbregts et al. also suggested a deterministic sensitivity analysis to determine the importance of parameters' contributions.

Based on the above-mentioned studies, especially Heijungs (1996) framework, considering the trade-offs between accuracy and cost of implementation, the authors present an alternative stochastic solution using a hybrid DQI-Statistical (HDS) approach to improve the quality of pure DQI method while reducing the cost of the pure statistical method in WB-LCA. The major departure from previous works is the stochastic prescreening process using quantitative DQI and MCS to determine the influence of the parameters' contributions.

After the categorization, the statistical method is adopted for critical parameters, and the DQI based distributions are used for non-critical parameters. An application case is presented in the paper to validate the presented solution.

The goal of this paper is to present a hybrid stochastic method to improve uncertainty estimate of whole-building embodied energy with data limitations. This method can be a valuable tool to evaluate deterministic results of whole-building embodied energy when uncertain information is needed for decision-making.

2. Methods

2.1. The DQI method

DQI characterizes data quality using descriptive indicators that are often formatted as a data quality pedigree matrix (DQPM) (Table 1). Columns in the matrix represent data quality indicators, such as reliability of data source, age of data, and etc. Rows represent an ordinal quality scale, such as from 1 to 5 or from 1 to 10 (Junnila and Horvath, 2003). An overall quality of a data can be characterized by an aggregated number which takes into account all the individual indicators. As a simple example, Fig. 1 shows a three-indicator DQI to evaluate the data quality of the parameter of steel EEC 35 MJ/kg (Zabalza Bribián et al., 2009) in the specific application context. For example, if (2, 3, 4) are assigned to the three indicators respectively, and all indicators are treated equal in weight, the aggregated DQI score for the parameter is $T = 2 \times 1/3 + 3 \times 1/3 + 5 \times 1/3 = 3.3$. According to previous studies (Junnila and Horvath, 2003; Maurice et al., 2000), the "acceptable data quality" and "fair data quality" can be respectively set to the DQI score of 2 and 3 in the case of 1–5 scale, so it indicates that this parameter is "good" in this application (Junnila and Horvath, 2003). This equally weighting approach has been commonly used in the previous DQI aggregation (Maurice et al., 2000; May and Brennan, 2003; Kennedy et al., 1997). Although this traditional approach has arguable deficiency, at this stage, it is difficult to decide which indicator is more important for a parameter due to the information scarcity. The agreed method for weighting different DQI indicators lacks but may need further research. It is also directly adopted here with two considerations: 1) DQI is only used to categorize the parameters in the proposed hybrid approach. 2) The focus of this paper is to combine the advantages of this traditional DQI and the pure statistical method to develop a more practical approach.

2.2. Quantitative DQI

Quantitative DQI transforms the aggregated DQI scores to probability density functions to enable uncertainty quantification (Kennedy et al., 1996; Tan et al., 2002b; Maurice et al., 2000; May and Brennan, 2003; Weidema and Wesnæs, 1996). The basic idea is to characterize data of different quality by distinct probability density functions based on the "rule of thumb" (Finnveden and Lindfors, 1998). The DQI transformation matrix (Table 2) was often used (Kennedy et al., 1996; Tan et al., 2002b; Canter et al., 2002; May and Brennan, 2003; Kennedy et al., 1997) to convert the aggregated DQI scores into Beta functions (Eq. (2)) :

$$\begin{aligned} f(x;\alpha,\beta,a,b) &= [1/(b-a)]*\{\Gamma(\alpha+\beta)/[\Gamma(\alpha)*\Gamma(\beta)]\} \\ &\quad *[(x-a)/(b-a)]^{\alpha-1}*[(b-x)/(b-a)]^{\beta-1} \\ &\quad (a \le x \le b) \end{aligned} \tag{2}$$

Where α, β are distribution shape parameters; a, b are selected range endpoints.

The Beta probability function is used herein primarily due to the fact that "the shape parameters and range end points allow virtually any shape probability distributions to be represented. The shape parameters establish the shape of the distribution and thus the location of the probability mass, whereas the endpoints limit the range of possible values" (Canter et al., 2002). The rationale of adopting Beta function when the actual probabilistic data distributions were difficult to obtain was also discussed in (Kennedy et al., 1996; Tan et al., 2002b). Given that DQI = 1 and 5 were assigned to the parameter of steel EEC 35 MJ/kg (Zabalza Bribián