Transportation Research Part D 38 (2015) 125–143

Contents lists available at ScienceDirect

Transportation Research Part D

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

Predictive usage mining for life cycle assessment

Jungmok Ma [a], Harrison M. Kim [b,*]

[a] Department of National Defense Science, Korea National Defense University, Seoul, Korea
[b] Department of Industrial and Enterprise Systems Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

ARTICLE INFO

Article history:
Available online 10 June 2015

Keywords:
Life cycle assessment
Usage modeling
Time series segmentation
Time series analysis

ABSTRACT

The usage modeling in life cycle assessment (LCA) is rarely discussed despite the magnitude of environmental impact from the usage stage. In this paper, the usage modeling technique, predictive usage mining for life cycle assessment (PUMLCA) algorithm, is proposed as an alternative of the conventional constant rate method. By modeling usage patterns as trend, seasonality, and level from a time series of usage information, predictive LCA can be conducted in a real time horizon, which can provide more accurate estimation of environmental impact. Large-scale sensor data of product operation is suggested as a source of data for the proposed method to mine usage patterns and build a usage model for LCA. The PUMLCA algorithm can provide a similar level of prediction accuracy to the constant rate method when data is constant, and the higher prediction accuracy when data has complex patterns. In order to mine important usage patterns more effectively, a new automatic segmentation algorithm is developed based on change point analysis. The PUMLCA algorithm can also handle missing and abnormal values from large-scale sensor data, identify seasonality, and formulate predictive LCA equations for current and new machines. Finally, the LCA of agricultural machinery demonstrates the proposed approach and highlights its benefits and limitations.

Introduction and background

Life cycle assessment (LCA) is an analytical assessment tool to quantify environmental impact of a product or system (Rebitzer et al., 2004; Finnveden et al., 2009). The potential environmental impact can be generated from all the stages of a product, i.e., manufacturing, usage, maintenance, and end-of-life. The LCA approach provides a holistic and systematic way to manage data associated with the product of interest. With the popularity of sustainable design and environmentally conscious design, LCA studies have reported various materials, electronics, automobiles, and complex systems (Kwak, 2012).

The LCA framework (Guinée, 2002; Reap et al., 2008a) consists of goal and scope definition, inventory analysis (LCI, life cycle inventory), impact assessment (LCIA, life cycle impact assessment) and interpretation. The goal and scope definition is the phase that defines the purpose, the systems or products, and the level of sophistication. The LCI is the phase that defines the system boundaries and the flow diagrams with unit processes (e.g., extraction of oil, refining, production of electricity, etc.). The main result from the LCI is the inventory table which quantifies inputs (e.g., raw material, land, energy, etc.) and outputs (e.g., pollutants such as CO_2, SO_2, NO_x, etc.) to the environment. The LCIA is the phase that translates the inventory table into relevant impact categories (e.g., carcinogens, climate change, acidification, etc.) and quantifies the environmental

* Corresponding author at: 104 S. Mathews Ave., Urbana, IL 61801, USA. Tel.: +1 (217) 265 9437; fax: +1 (217) 244 5705.
E-mail address: hmkim@illinois.edu (H.M. Kim).

http://dx.doi.org/10.1016/j.trd.2015.04.022
1361-9209/

impact using weighting and normalization. The interpretation is the phase that evaluates the results from the LCIA and makes recommendations of the LCA study.

Although the LCA approach is mature and has become a widely used method in various industries, it is usually *static* in that time is not considered in the assessment with the implicit assumption of steady-state processes. The necessity of considering time in LCA was discussed in literature. Reap et al. (2008b) provided insightful reviews on the temporal aspects of LCA. Temporal factors such as different rates of emissions over time and seasonal variation of their impacts can influence the accuracy of LCA. Levasseur et al. (2010) showed that the inconsistency in time frames can affect LCA results significantly. Memary et al. (2012) demonstrated that changes of environmental impact over time are useful information for assessing future technology and options. Collet et al. (2014) presented a method to find the most critical flows of information based on dynamic inventory data (i.e., LCI level) and sensitivity analysis. In addition to the aspect of time, spatial variation is another contributor that can significantly affect the accuracy of LCA (Reap et al., 2008b). Local, regional and continental differences can cause different results of LCA.

In this paper, a new perspective of dynamic LCA is proposed to consider time in LCA, especially the modeling of the usage stage. Among the life cycle stages of a product, the manufacturing stage, which is the chosen stage in the majority of LCA studies, can be considered as a one-time event, i.e., time-independent event. Although the dynamic inventory approach (Collet et al., 2014) attempted to relax this (e.g., the impact from material x or process y can be changed over time), the inventory data is considered constant in this study. On the other hand, the usage stage (with maintenance and end-of-life stages) is a time-dependent event, which means the lifespan of a product has a large impact on LCA. Many studies showed that the majority of environmental impact can come from the usage stage over life cycle (e.g., more than 60% for cars (Sullivan and Cobas-Flores, 2001), more than 80% for off-load machinery (product of interest in this paper) (Kwak et al., 2012), and 80–90% for some small electronics (Telenko and Seepersad, 2014)). Therefore, *how to model the usage stage in LCA* is critical and one of the main questions of this work.

Even though the importance of the usage modeling has been recognized among LCA researchers and practitioners, it is rarely discussed in literature. LCA studies in literature usually utilized a constant rate (Lee et al., 2000; Choi et al., 2006; Kwak et al., 2012; Kwak and Kim, 2013; Li et al., 2013) of usage information (hereinafter constant rate method) with the implicit assumption of steady-state processes (e.g., average fuel consumption rate in kg/h, fixed operating hours per month, etc.). This method is simple and easy to apply, but if data has complex patterns (e.g., trend, seasonality and segments), the prediction accuracy of the constant rate method can be significantly reduced. The constant rate method only allows us to calculate life cycle impact in a nominal time horizon, e.g., 10 years as a whole instead from October 2014 to December 2024. This can be an important issue to policy makers and manufacturers when they want to estimate the environmental impact of the future. Fig. 1 shows the expected result from both the proposed model and the constant rate method. Based on the available historical data, a usage (e.g., diesel fuel consumption) model should be built and used for predicting the future usage profile. It can be seen in Section 'Numerical prediction tests for PUMLCA' that the constant rate method can misinterpret the upcoming usage profile while the proposed model is expected to provide higher prediction accuracy with lower variance predictions.

One exception is Telenko and Seepersad (2014) who proposed a usage context modeling technique in LCA using Bayesian network models. The usage context includes human, situational, and product variables. Based on a pre-defined probabilistic network of relevant usage patterns (e.g., weather → usage of electric kettle with probability of x), a usage profile and its variability can be modeled as a form of distribution. However, in order to apply this approach, causal relationships among different usage contexts should be known, which is expressed as a probabilistic network. For example, the usage of agricultural machinery (e.g., crop sprayer, harvester, nutrient applicator, etc.) can be affected by a various usage context (e.g., weather, soil, experience of farmers, price of fuel and crops, machine deterioration). It will be difficult to correlate these variables with specific usage information (e.g., diesel fuel consumption and operating hours). Furthermore, Telenko and Seepersad (2014) did not consider time in LCA.

Alternatively, this study proposes a time series usage modeling technique, predictive usage mining for life cycle assessment (PUMLCA), as shown in Fig. 2. Companies such as Caterpillar (PRODUCT Link™) and John Deere (JDLink™) have

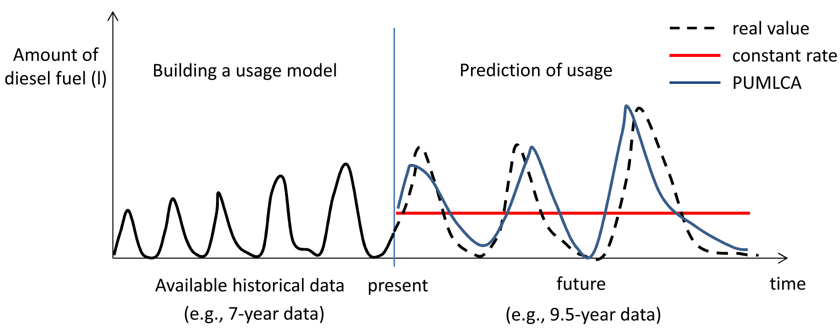

Fig. 1. A prediction scenario of PUMLCA and constant rate method.