Int J Life Cycle Assess (2014) 19:1110–1118
DOI 10.1007/s11367-013-0685-y

Application of robust optimization in matrix-based LCI for decision making under uncertainty

Ren Wang · Daniel Work

Received: 15 December 2012 / Accepted: 8 December 2013 / Published online: 4 April 2014

Abstract
Purpose When product systems are optimized to minimize environmental impacts, uncertainty in the process data may impact optimal decisions. The purpose of this article is to propose a mathematical method for *life cycle assessment* (LCA) optimization that protects decisions against uncertainty at the *life cycle inventory* (LCI) stage.
Methods A robust optimization approach is proposed for decision making under uncertainty in the LCI stage. The proposed approach incorporates data uncertainty into an optimization problem in which the matrix-based LCI model appears as a constraint. The level of protection against data uncertainty in the technology and intervention matrices can be controlled to reflect varying degrees of conservatism.
Results and discussion A simple numerical example on an electricity generation product system is used to illustrate the main features of this methodology. A comparison is made between a robust optimization approach, and decision making using a Monte Carlo analysis. Challenges to implement the robust optimization approach on common uncertainty distributions found in LCA and on large product systems are discussed. Supporting source code is available for download at https://github.com/renwang/Robust_Optimization_LCI_Uncertainty.
Conclusions A robust optimization approach for matrix-based LCI is proposed. The approach incorporates data uncertainties into an optimization framework for LCI and provides a mechanism to control the level of protection against uncertainty. The tool computes optimal decisions that protects against worst-case realizations of data uncertainty. The robust optimal solution is conservative and is able to avoid the negative consequences of uncertainty in decision making.

Keywords Life cycle assessment · Mathematical programming · Matrix-based LCI · Robust optimization · Uncertainty

1 Introduction

One application of *life cycle assessment* (LCA) is to improve the environmental performance of a product system (Azapagic and Clift 1999a). One option to improve the performance of a product system is achieved by making choices among functionally equivalent products or processes, with a goal to minimize environmental impacts. Because of the inherent uncertainty in life cycle assessment data, a key challenge for this application of LCA is to make optimal choices that are insensitive to this uncertainty.

This work addresses the above challenge by incorporating life cycle data uncertainty into optimization problems in LCA, so that when functionally equivalent products are compared with a goal to maximize environmental performance, the optimal solution will be robust to changes or errors in the data. Mathematical formulations are provided to model a deterministic *life cycle inventory* (LCI) optimization problem, and a robust formulation is introduced to model and solve the problem when the process data is uncertain.

Responsible editor: Andreas Ciroth

R. Wang (✉)
Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, 3117 Newmark Civil Engineering Laboratory, 205 North Mathews Ave., Urbana, IL, USA
e-mail: renwang2@illinois.edu

D. Work
Department of Civil and Environmental Engineering and Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, 1203 Newmark Civil Engineering Laboratory, 205 North Mathews Ave., Urbana, IL, USA

1.1 Optimization and LCA

Mathematical programming tools have been applied to optimize product systems over their life cycle. Azapagic and Clift (1999b) first proposed an *optimum LCA performance* methodology to identify the best alternative in a product system with an objective to minimize multiple conflicting environmental and socio-economic objectives. Linear programming was used to solve the problem and construct a Pareto frontier of optimal alternatives. Later, Tan (2005) applied *symmetric fuzzy linear programming* to solve multi-objective optimization problems in LCA.

The results of LCA studies have also been incorporated into optimization problems for system design, process design, and supply chain network planning (Hugo and Pistikopoulos 2005; Guillén-Gosálbez et al. 2008; Carvalho et al. 2011). In these studies, environmental concerns are modeled as constraints, and the resulting problems are solved with *mixed-integer linear programming* or *mixed-integer nonlinear programming*, due to various nonlinear features of the systems under consideration.

The problem of making optimal choices with life cycle consideration was proposed by Tan et al. (2008), which embedded a matrix-based LCA model as a constraint in the optimization problem. The resulting methodology can be used to identify the best system configuration to minimize environmental impacts. When multiple objectives are considered in the optimization problem, fuzzy linear programming is used to solve the resulting problem. A key feature of the formulation considered by Tan is that it is consistent with the matrix-based model of LCA (Heijungs and Suh 2002), and the optimization is performed over the full product system based on the LCI and *life cycle impact assessment* data, instead of optimizing over the final output of the LCA. We will follow a similar matrix-based LCI approach in the robust formulations proposed in this article.

1.2 Uncertainty in LCA

The sources and types of uncertainty in LCA have been investigated by several authors (e.g., Fava et al. 1994; US-EPA 1989; Morgan and Henrion 1990; Huijbregts 2002; Heijungs and Huijbregts 2004; Lloyd and Ries 2007). Various approaches to classify sources of uncertainty in LCA exist (e.g., due to model errors, linearity assumptions, spatial and temporal variability), see for example the review papers of Heijungs and Huijbregts (2004) and Lloyd and Ries (2007) for a detailed discussion.

A common theme in each of the categorizations proposed above is that data uncertainty is inherent and cannot be fully eliminated. This is especially true when LCA is used to predict the future environmental performance of a product system, which includes future processes and technologies that have not yet been developed.

Because of this inherent uncertainty, several quantitative approaches have been introduced to assess how uncertainty in the data and inputs impacts the results of an LCI or LCA. *Monte Carlo analysis* (e.g., McCleese and LaPuma 2002; Sonnemann et al. 2003) can be used to numerically approximate the distribution of uncertainties in the LCI or LCA due to uncertainties in the process data. Another approach, based on *intervals*, has been used to model uncertain data in LCA (Chevalier and LeTeno 1996). Solving a set of interval equations (representing matrix-based LCI with uncertain coefficients modeled by intervals) provides a set of all possible environmental outputs under a given uncertainty set, and indicates all possible environmental outputs of the product system. *Fuzzy set theory* (Benetto et al. 2006; Tan 2008) was used to compute upper and lower bounds on emissions of a system when the input data is described as a fuzzy set. A comprehensive review of approaches to quantitative uncertainty in LCA can be found in the survey paper of Lloyd and Ries (2007).

1.3 Contributions of the article

While the above quantitative approaches can be used to assess the uncertainty in LCA and to interpret the influence of uncertainty on environmental outputs, an approach that explicitly computes optimal LCA choices that are insensitive to data uncertainties is wanting. It is well known that optimal solutions can be extremely sensitive to errors in the model or data if not explicitly treated in the optimization framework (Bertsimas and Thiele 2006), and in extreme cases, potentially lead to optimal decisions that become infeasible under small errors in the problem data.

Thus, the main contribution of this article is to link LCI uncertainty and LCI optimization through an approach known as robust optimization. In Section 2, the matrix-based LCI model is reviewed and a robust optimization formulation is proposed, which embeds the matrix-based LCI model as a constraint. In Section 3, the proposed approach is implemented for a simple numerical example involving electricity generation. The robust approach is then compared to a Monte Carlo simulation performed over the choice set. In Section 4, challenges to implement this approach with common uncertainty distributions and on large-scale commercial problems are discussed, and conclusions are presented in Section 5.

2 Methodology

2.1 Matrix-based LCI

In this section, we briefly review the mathematical formulation of matrix-based LCI, following the framework of Heijungs and Suh (2002). Later, this model will be embedded as a constraint in the proposed robust LCI optimization problem.

A product system is modeled as a set of n processes $\{p_1, p_2, \cdots, p_n\}$, where each process is encoded as a vector p. Each element in the vector quantifies inputs and outputs of the unit process with positive (negative) entries representing outflows (inflows) to the process. The vectors are structured such that upper entries describe economic flows (e.g., materials, energy), while the lower entries describe environmental flows (e.g., emissions to the environment). The product system can also be arranged in a process matrix $P=[p_1,p_2,\cdots,p_n]$. Because of the structure of each process vector, the process matrix P can be partitioned as follows:

$$P = \begin{bmatrix} A \\ B \end{bmatrix},$$

where $A\in\mathbb{R}^{v\times n}$ represents flows within the economic system, and is referred to as technology matrix, while B represents the environmental interventions, and is referred to as intervention matrix. We let m denote the number of environmental interventions (e.g., $B\in\mathbb{R}^{m\times n}$).

The matrix-based LCI model is given as:

$$\begin{cases} As = f \\ Bs = g, \end{cases}$$

where $s\in\mathbb{R}^n$ is the scaling vector, and $f\in\mathbb{R}^v$ represents the desired final demand of the product system. Thus, the scaling vector s encodes the amount of each unit process needed to produce the desired final demand f. The inventory vector $g\in\mathbb{R}^m$ describes the total environmental interventions of the product system, and is often the primary quantity of interest in LCI. When A is invertible, the scaling vector can be directly computed as $s=\mathrm{A}^{-1}f$, and the inventory vector can be computed as $g=\mathrm{BA}^{-1}f$.

2.2 Optimization in matrix-based LCI

When the number of processes in the economic system is larger than the dimension of the material and energy flows, the technology matrix is no longer invertible. This can occur, for example, if there are functionally equivalent processes in the product system, such as multiple suppliers of electricity. The resulting economic system is underdetermined, which means there are multiple ways to satisfy the specified final demand. Heijungs and Suh (2002) proposed to determine a scaling vector for underdetermined problems through linear programming, for example by minimizing environmental impact or cost.

To determine the scaling vector s and the inventory vector g, the following linear problem can be solved:

$$\begin{aligned} \underset{s,g}{\text{minimize}}: \quad & \lambda^T g \\ \text{subject to}: \quad & As \geq f \\ & Bs \leq g. \end{aligned} \tag{1}$$

The optimization problem (1) selects the scaling vector such that the required final demand is satisfied, and a linear objective function representing weighted environmental output, $\lambda^T g$, is minimized. The vector $\lambda\in\mathbb{R}^m$ is a vector of weighting coefficients and determines the relative importance of the elements in the inventory vector. If desired, the weighting coefficients can be chosen such that the optimal decision variables vector s and g in problem (1) are unique (Franklin 1987).

Freire et al. (2001) also proposed a variant of problem (1), called the *life cycle activity analysis* approach, which integrates life cycle balance equations into a linear programming framework, resulting in a joint consideration of monetary cost and environmental impacts.

When data uncertainty is considered, the entries in the process vectors become uncertain. For economic flows, uncertainty indicates the material or energy flows between unit processes in the system which maybe unknown, have variability, or are subject to change. For environmental flows, uncertainty encodes the fact that the environmental impact associated with each unit process may similarly be unknown or fluctuate. For the j^{th} process vector, $p_j=[a_{1j},\cdots,a_{vj},b_{1j},\cdots,b_{mj}]^{\mathrm{T}}$, elements a_{ij} and b_{ij} are distributions. Often, only limited information is available to model the distribution of the inflows and outflows of each process. If the expected value of the distribution is known, and the distribution is symmetric with a known deviation, then symmetric intervals can be used to model the range of values each entry may take. In this case, we model the uncertainty set of i^{th} economic flow of the j^{th} process as $a_{ij}\in\left[\overline{a}_{ij}-\widehat{a}_{ij},\overline{a}_{ij}+\widehat{a}_{ij}\right]$, where $\overline{a}_{ij}$ is the expected value of the interval and $\widehat{a}_{ij}$ is the maximum deviation. The environmental flows can be modeled similarly; $b_{ij}\in\left[\overline{b}_{ij}-\widehat{b}_{ij},\overline{b}_{ij}+\widehat{b}_{ij}\right]$.

Since the entries of the technology matrix and environmental interventions matrix are described by uncertainty sets, so too, are the matrices themselves. We denote the uncertainty set of the technology matrix as $\mathcal{A}$, and the intervention matrix as $\mathcal{B}$, which are given as:

$$\begin{cases} \mathcal{A} = \{(a_{ij}) | a_{ij} \in [\bar{a}_{ij} - \hat{a}_{ij}, \bar{a}_{ij} + \hat{a}_{ij}], \forall i,j\} \\ \mathcal{B} = \{(b_{ij}) | \, b_{ij} \in [\bar{b}_{ij} - \hat{b}_{ij}, \bar{b}_{ij} + \hat{b}_{ij}], \forall i,j\} \end{cases}$$

This will be the uncertainty model used throughout the remainder of this work. When $A\in\mathcal{A}$ and $B\in\mathcal{B}$, the optimization problem (1) is no longer a standard linear programming problem.

2.3 Robust formulation of matrix-based LCI

The *robust* formulation of matrix-based LCI for decision making under uncertainty is described as follows:

$$\begin{aligned} \underset{s,g}{\text{minimize}}: \quad & \lambda^T g \\ \text{subject to}: \quad & As \geq f, \quad \forall A\in\mathcal{A} \\ & Bs \leq g, \quad \forall B\in\mathcal{B}. \end{aligned} \tag{2}$$

In this model, the optimal scaling vector s is chosen such that the final demand is satisfied for all possible realizations of A, and which minimizes the worst-case emissions created for any realization of environmental interventions $B \in \mathcal{B}$.

One feature of this formulation is that it models alternatives in the same product system instead of modeling them separately in the optimization problem. Since functionally equivalent products may share similar upstream processes, such modeling will automatically capture the fact that any uncertainty of the shared processes in the supply chain will impact the performance of all the alternatives simultaneously.

For convenience, we put problem (2) in standard form:

$$\begin{aligned} \underset{x}{\text{minimize}} : \quad & c^T x \\ \text{subject to} : \quad & \min_{K \in \mathcal{K}} \{Kx\} \geq l \end{aligned} \tag{3}$$

It can be verified that problem (3) is equivalent to problem (2) when $K = \begin{bmatrix} A & 0 \\ -B & I \end{bmatrix}$, $x = \begin{bmatrix} s \\ g \end{bmatrix}$, $l = \begin{bmatrix} f \\ 0 \end{bmatrix}$, $c = \begin{bmatrix} 0 \\ \lambda \end{bmatrix}$, and $\mathcal{K} = \left\{ \begin{bmatrix} A & 0 \\ -B & I \end{bmatrix} \mid A \in \mathcal{A}, B \in \mathcal{B} \right\}$ Note again that neither problem (2) or (3) are standard linear programming problems, since A and B belong to uncertainty sets $\mathcal{A}$ and $\mathcal{B}$.

The formulation (3) models choices at the life cycle inventory stage only. Extensions to model impact assessment can be developed by applying characterization factors. Three assumptions are made in this formulation. First, data uncertainties exist in both the technology matrix A and the intervention matrix B, which means the amount of economic flow and the values of environmental emissions of each process in the product system are not deterministic. Second, the uncertainty parameters are required to be specified as symmetric closed intervals. Uncertainty forms such as the uniform distribution or triangular distribution could be applied with this approach. Third, the uncertainty associated with the technology matrix A is assumed independent from the uncertainty in the interventions matrix B.

2.4 Robust optimization

The proposed model problem (3) is a standard form of a *robust optimization* problem. Robust optimization is a decision making tool, first proposed by Soyster (1973), to address decision making under uncertainty. When parameters in a problem are subject to uncertainty, robust optimization solves the problem by looking at the worst-case realization caused by these uncertain parameters to ensure the feasibility of the optimal solution under all perturbations. The approach proposed by Soyster solves the problem in a way that takes the worst possible value of each parameter k_{ij} of K in their uncertainty sets. The solution obtained by this approach ensures the constraint $Kx \geq l$ is always satisfied for any possible value that K may take in $\mathcal{K}$.

The features of robust optimization can be interpreted in the context of LCA. For the constraint $As \geq f$ in the robust formulation (3), robust optimization ensures that the constraint $As \geq f$ is satisfied when A is subject to any perturbation defined by the set $\mathcal{A}$. If the scaling vector is computed without considering uncertainty (for example by using only the expected value of A), the final demand may not be satisfied for some realizations of the system. The solution computed by robust optimization will prevent this scenario. It should be noted that the guaranteed feasibility of the constraints provided by robust optimization is achieved at the cost of possibly increasing the value of the objective function at optimality. Considering the constraint $-Bs+g \geq 0$, the robust optimal solutions of s and g will be determined in a way that guarantees the constraint is satisfied for all $B \in \mathcal{B}$. This constraint encodes that the solution is made based on the worst-case emissions each unit process could take for a given s. As a result, the solution by robust optimization minimizes the worst-case emissions of each alternative product or process in the system.

A criticism of the robust formulation (3) is that it is unlikely that all uncertain parameters will take their worst-case values simultaneously, and is therefore too conservative. To address this concern, Bertsimias and Sim (2004) and Bertsimas and Thiele (2006), proposed an alternative robust formulation that offers full control on the degree of conservatism for every constraint, which we will adopt for our robust LCI formulations. Like Soyster, this robust approach assumes coefficients k_{ij} of K are subject to uncertainty and belong to the symmetric uncertainty sets specified as $\left[\overline{k}_{ij} - \widehat{k}_{ij}, \overline{k}_{ij} + \widehat{k}_{ij}\right]$, where $\overline{k}_{ij}$ is the estimated nominal value, and the half-length $\widehat{k}_{ij}$ measures the precision of the estimate. A parameter z_{ij} is defined to describe the deviation of k_{ij} from its nominal value $\overline{k}_{ij}$ as:

$$z_{ij} = \frac{k_{ij} - \overline{k}_{ij}}{\widehat{k}_{ij}}.$$

Here, $\widehat{k}_{ij}$ is the maximum deviation the nominal value may take. Parameter z_{ij} is called the *scaled deviation*, and it always belongs to $[-1, 1]$. Note that the transformation of the uncertainty into a scaled deviation requires the deviation around the nominal value to be symmetric. A parameter Γ_i, called the *budget of uncertainty*, controls the degree of conservatism of constraint i and is related to z_{ij} by:

$$\sum_{j=1}^{n} \left| z_{ij} \right| \leq \Gamma_i, \forall i.$$

The budget of uncertainty Γ_i can be interpreted as the maximum number of parameters in constraint i that can deviate from their nominal values. When $\Gamma_i = 0$, the decision is made based on the nominal values of all parameters, while if $\Gamma_i = n$, the decision is made based on the worst-case realizations of all parameters. Intermediate values of $\Gamma_i \in (0,n)$ reflect the degree of conservatism on the i^{th} constraint. For example, if there are five uncertain parameters in a constraint and $\Gamma_i = 3$, perturbations to the three parameters which lead to the smallest $K_i x$ are considered. If $\Gamma_i = 2.5$, it means only 50 % of the uncertainty of the third most influential uncertain parameter is considered. Note the budget of uncertainty also can be interpreted in some cases as the probability of violating a constraint (Bertsimas and Thiele 2006), thus transforming the worst-case optimization into a problem where the designer has control over the probability of satisfying the final demand.

When adding a budget of uncertainty, the set $\mathcal{K}$ in problem (3) is redefined as:

$$\mathcal{K} = \left\{ (k_{ij}) \middle| k_{ij} = \overline{k}_{ij} + \widehat{k}_{ij} z_{ij}, \forall i, j, z \in \mathcal{Z} \right\},$$

with

$$\mathcal{Z} = \left\{ z \,\middle|\, |z_{ij}| \le 1, \forall i, j, \sum_{j=1}^{n} |z_{ij}| \le \Gamma_i, \forall i \right\}.$$

Thus problem (3) becomes:

$$\begin{aligned} \underset{x}{\text{minimize}}: \quad & c^T x \\ \text{subject to}: \quad & \overline{K}_i\, x + \min \left\{ \sum_{j=1}^{n} \widehat{k}_{ij} x_j z_{ij} \middle| z_i \in \mathcal{Z}_i \right\} \ge l_i, \quad \forall i. \end{aligned} \tag{4}$$

Problem (4) can be reformulated as a linear programming problem, called the *robust counterpart*, by applying strong duality, yielding:

$$\begin{aligned} \underset{x,v,q,y}{\text{minimize}}: \quad & c^T x \\ \text{subject to}: \quad & \overline{K}_i\, x - \Gamma_i v_i - \sum_{j=1}^{n} q_{ij} \ge l_i, \quad \forall i \\ & v_i + q_{ij} \ge \widehat{k}_{ij} y_j, \quad \forall i, j \\ & -y_j \le x_j \le y_j, \quad \forall j \\ & v_i, q_{ij} \ge 0, \quad \forall i, j \end{aligned} \tag{5}$$

where v, q, and y are dual variables. The solution x computed by solving the linear program (5) is the solution to robust optimization problem (4). A detailed introduction on this robust optimization method can be found in Bertsimas and Sim (2004).

3 Numerical example

A simple numerical example is used to illustrate the main ideas presented in this work. The example involves the comparison of two electricity generation strategies, and it is a robust extension of a problem proposed in Heijungs and Suh (2002). The objective is to identify the optimal electricity generation strategy (encoded in the scaling vector) to satisfy a final demand for electricity while minimizing the total carbon dioxide emissions, in the presence of data uncertainty.

3.1 Process descriptions

The process data and uncertainties for this example are presented in Table 1. Data uncertainties are assumed to be uniformly distributed, and the final demand is 100 kWh of electricity. The uncertainty range on the CO_2 emissions for *electricity by oil* and *oil production* are modeled as 40 % of the mean while CO_2 emissions for *electricity by coal* and *coal production* are modeled as 8.3 % and 20 % of the mean, respectively. Note the CO_2 emissions ranges for the processes associated with electricity generated by oil are modeled relatively larger compared to coal, to illustrate the main features of the robust optimization approach in a simple numerical example.

The technology matrix A, intervention matrix B, final demand vector f, and vector λ are constructed as follows:

$$A = \begin{bmatrix} -2 & 100 & 0 & 0 \\ 10 & 0 & 10 & 0 \\ 0 & 0 & -5 & 50 \end{bmatrix}, B = [\,10 \quad 5 \quad 12 \quad 1\,], f = \begin{bmatrix} 0 \\ 100 \\ 0 \end{bmatrix}, \lambda = [1].$$

With the range defined by:

$$\widehat{A} = \begin{bmatrix} 0.3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0.2 & 0 \end{bmatrix}, \widehat{B} = [\,4 \quad 2 \quad 1 \quad 0.2\,]$$

Table 1 Process data

Process	Economic outflows	Economic inflows		Environmental outflows	
		Mean value	Range	Mean value	Range
Electricity by oil	10 kWh of electricity	2 l of oil	[1.7, 2.3]	10 kg of CO_2	[6, 14]
Oil production	100 l of oil	–	–	5 kg of CO_2	[3, 7]
Electricity by coal	10 kWh of electricity	5 kg of coal	[4.8, 5.2]	12 kg of CO_2	[11, 13]
Coal production	50 kg of coal	–	–	1 kg of CO_2	[0.8, 1.2]

Accordingly, $\widehat{k}_{ij}$ in problem (4) can be read as the elements of $\widehat{K}$, given by:

$$\widehat{K} = \begin{bmatrix} 0.3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.2 & 0 & 0 \\ 4 & 2 & 1 & 0.2 & 0 \end{bmatrix}$$

The last parameter to be specified is the budget of uncertainty Γ_i, which reflects the decision maker's attitude towards uncertainty for each constraint in optimization problem (3). The effect of this parameter on the solution is detailed next. After this parameter is set, the solution to the LCA robust optimization problem (4) is computed by solving the dual problem (5) with a standard linear program solver.

3.2 Numerical results

When budget of uncertainty values are defined in this case study, full protection against uncertainty is given for all economic flow constraints. This is to ensure that the solution derived from this optimization problem will always satisfy material and energy constraints for any possible value of A in the given uncertainty set $\mathcal{A}$. As a result, the analysis of the level of protection against data uncertainties is focused on uncertainties in intervention matrix B. Since there are four uncertain elements in matrix B that correspond to CO_2 emissions, the budget of uncertainty $\Gamma \in [0,4]$. When Γ is 0, the solution is computed using the mean value of CO_2 emissions, and when it equals to 4, the decision is made for the case when all the four processes take the largest CO_2 emission values specified in the range. When Γ is set between 0 and 4, it means the solution is derived based on the intermediate case between the previous two. In this study, the value of Γ is gradually increased from 0 to 4, and the resulting optimal generation strategy (i.e., the amount of electricity generated from oil and coal) is shown in Fig. 1. In the figure, the value of Γ is scaled to a level of protection against uncertainty using the formula: $\Gamma/4 \times 100\%$.

Figure 1 shows that if the decision maker does not consider any uncertainty in environmental output data, electricity should be generated by oil. However, the strategy changes dramatically when protection against more than 12 % of the uncertainty is considered. When protecting against more than 68 % of the data uncertainty in B, there is a complete switch to electricity generation by coal. Since there are four uncertainty parameters in the constraint, an uncertainty budget of 25 % corresponds to full protection against the most influential parameter. A switch in the solution occurs when protection against 12 % of the uncertainty is considered (in other words, when protection against 50 % uncertainty of the most influential parameter is considered).

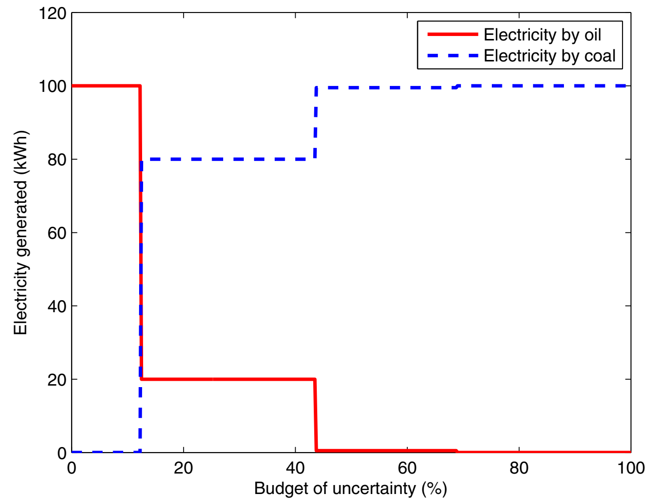

Fig. 1 Electricity generation strategy with different uncertainty budget

Next, the objective function values (total CO_2 emissions) are compared between the robust optimization approach and a deterministic approach under different worst-case realizations of B, parameterized by the budget of uncertainty. For the robust optimization approach, optimal electricity generation strategy is displayed in Fig. 1. For the deterministic approach, the electricity generation strategy is derived using the expected values of all parameters, which means that oil is always used to generate electricity. The total CO_2 emissions for the deterministic strategy are computed using the perturbed realization of B. This deterministic approach mimics the worst-case situation in reality when one optimizes based on nominal values—the true emissions are computed based on a realization of B, which is not a priori known to the decision maker. The CO_2 emissions associated with these two approaches are shown in Fig. 2. It shows that when B takes a worst-case deviation more than 12 % of the total data uncertainty in B, the robust approach will achieve a lower CO_2 emissions compared to the deterministic approach.

3.3 Comparison with Monte Carlo method

The Monte Carlo method is a widely accepted method to analyze uncertainty in LCA (e.g., McCleese and LaPuma 2002; Sonnemann et al. 2003), and it is implemented in commercial LCA software such as Simapro. In this study, a Monte Carlo simulation is performed separately on the two electricity generation product systems, using 5,000 samples. A uniform distribution is assumed for data in the uncertainty set summarized in Table 1. As can be seen from the results in Fig. 3, electricity derived by oil has much higher emission uncertainty compared to coal. Thus, although the expected emissions is lower for electricity generation by oil, coal is a better choice to protect against possibly high CO_2 emissions due to variability in emissions by oil, which is consistent with the results obtained by using robust optimization.

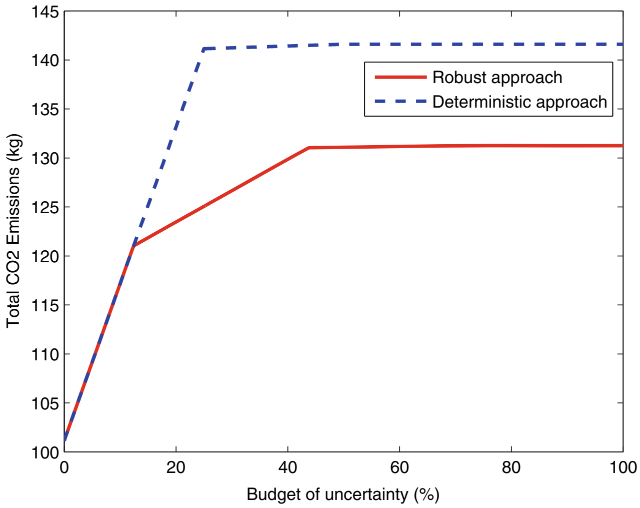

Fig. 2 Comparison between robust optimization and the nominal deterministic solution

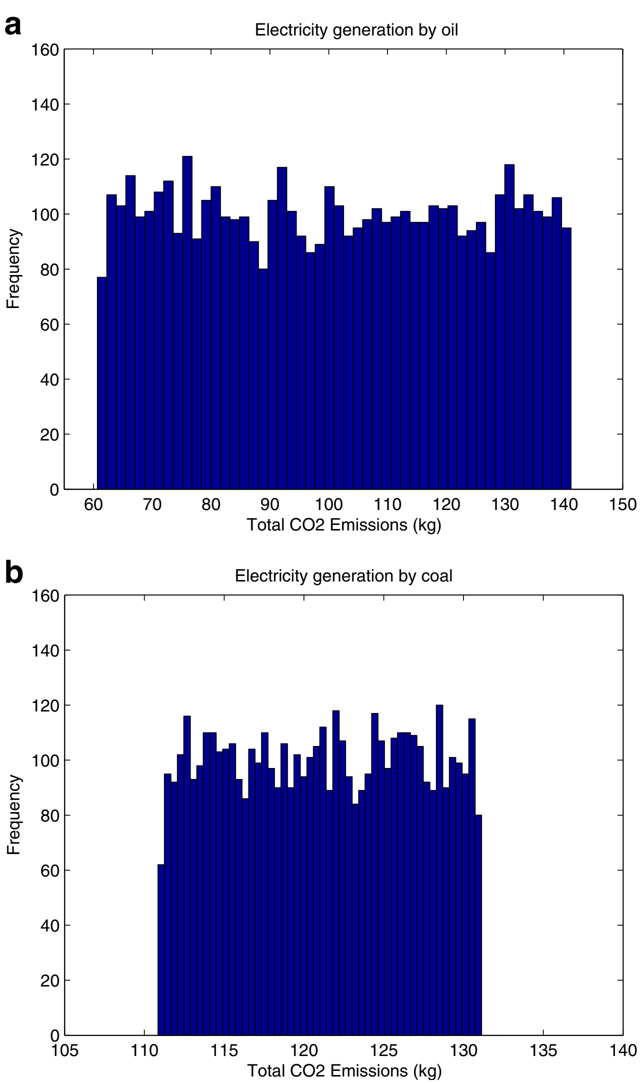

Fig. 3 Monte Carlo simulation

Therefore, both the robust optimization approach and Monte Carlo method could conclude the fact that electricity generation by coal is a better worst-case choice under uncertainty. However, an important benefit of the robust optimization approach is that it is computationally cheap compared to Monte Carlo analysis, especially as the number of alternative choices grows large. Specifically, when a decision across many alternatives under uncertainty is needed, one would need to enumerate each alternative, perform a Monte Carlo simulation on each design, and then compare the uncertainty distributions of each choice. Alternatively, one could perform a Monte Carlo analysis directly on (1), which would require solving a linear program for each sample. In contrast, the benefit of the Monte Carlo method is that it provides a more complete picture of the performance of the product systems, while the robust approach only concerns worse case realizations.

4 Discussion

4.1 Application with common uncertainty distributions in LCA

One assumption of the proposed robust optimization approach is that the uncertainty set is defined as a symmetric, closed interval. The interval must be closed in order to achieve strong guarantees of a feasible solution under any realization of the distribution, while the interval is assumed to be symmetric in order to introduce the concept of an uncertainty budget and to derive the robust counterpart (5). For practical implementation in commercial software, it is desirable that the robust optimization approach can accommodate a wide range of uncertainty distributions commonly used in LCA.

Characterization of the uncertainty distribution is itself an active area of research. For example, a survey article by Lloyd and Ries (2007) identifies the normal, triangle, uniform, and log normal distributions as common uncertainty models in the literature. Moreover, these distributions are also used in the ecoinvent commercial database. Note, only the triangle and uniform distribution immediately fit the robust optimization framework without modification.

To implement the approach on normally distributed data, truncating the distribution at two or three standard deviations can close the uncertainty set. The truncation would lose the guarantee of feasibility under all realizations of the original distribution, but this may not be critical, especially if an uncertainty budget is later introduced. Implementation on log normally distributed data is the most challenging, because the distribution is neither closed nor symmetric. The distribution can be truncated to close the uncertainty set. However, because the distribution is not symmetric, the nominal value (used in the robust optimization when the parameter is not allowed to vary) might be quite different from the mean value of the distribution, possibly resulting in an overly optimistic or pessimistic analysis. Further work is needed to extend the approach to non-symmetric data.

A second assumption for our methodology is the independence of the uncertainty in *A* and *B* matrices, which may not always hold in real product systems. Following the example presented in Section 3, 4.8 to 5.2 kg of coal is required to generate 10 kWh of electricity. Part of this uncertainty is due to fluctuations in the carbon content of the coal, since coals with higher carbon content provide more energy per kilogram. Simultaneously, the corresponding CO_2 emissions per kilogram of coal also depend on the carbon content. The optimal solution based on the worst-case technology matrix scenario (5.2 kg of coal required) is too pessimistic, because the corresponding CO_2 emissions/kilogram coal can no longer vary within the full range, once the required amount of coal is fixed. While the worst-case guarantees of robust optimization still holds, the resulting analysis may be overly conservative when correlations are ignored.

4.2 Challenges for large-scale implementation

A complete product system may involve several thousand processes in commercial software. For example, the ecoinvent database indicates about 2,000 processes for each product system, resulting in a sparse technology matrix. To implement the robust optimization approach on a product system of this size, two issues must be addressed.

The first issue concerns the numerical stability of the optimization problem. As part of our initial experiments, we extracted data from the ecoinvent database and constructed the matrix *A* for several invertible product systems. Each resulting system was ill-conditioned and numerical errors are introduced when solving the equation $s=A^{-1}f$. For example, we constructed the technology matrix for the process of bus manufacturing based on the ecoinvent database, the largest discrepancy between the final demand calculated from the computed scaling vector, and the given final demand, was on the order of 10^{-4}. To solve this problem, LU factorization with partial pivoting was applied following the suggestion of Frischknecht et al. (2007), which reduced the error to the order of 10^{-9}. Depending on the specific linear programming code used, the numerical pre-conditioning of the product system might be necessary to get accurate results from the robust optimization algorithm.

The second issue concerns the size of the robust counterpart (5), which is larger than the deterministic problem (1). If the number of constraints in (1) is m, and the number of decision variables is n, then the robust counterpart will contain $(m+nm+2n)$ constraints and $(m+nm+2n)$ variables. This will generate a large-scale linear programming problem, and commercial optimization software such as CPLEX must be used.

5 Conclusions and recommendations

A robust optimization approach to the matrix-based LCI model has been proposed. This approach incorporates data uncertainties into the optimization framework of LCA and provides control over the level of protection against uncertainty. The proposed approach computes a conservative solution that protects against worst-case realizations of data uncertainty and is able to avoid the negative consequences of uncertainty in decision making. Consequently, under worst-case realizations of the product system, lower environmental impacts can be achieved using robust decisions compared to decisions obtained from deterministic optimization approaches that do not consider uncertainty.

In the future, this work might be combined with a sensitivity analysis (Sakai and Yokoyama 2002; Heijungs 2010) to protect against potential risks in decision making due to sensitive parameters, in the absence of known distributions for process data.

Acknowledgments The authors thank the anonymous reviewers for suggesting meaningful improvements to the manuscript, including the example of correlated uncertainty in *A* and *B* presented in Section 4.1.

References

Azapagic A, Clift R (1999a) Life cycle assessment and multiobjective optimisation. J Clean Prod 7:135–143

Azapagic A, Clift R (1999b) The application of life cycle assessment to process optimisation. Comput Chem Eng 23:1509–1526

Benetto E, Dujet C, Rousseaux P (2006) Fuzzy-sets approach to noise impact assessment. Int J Life Cycle Assess 11(4):222–228

Bertsimas D, Sim M (2004) The price of robustness. Oper Res: 35–53

Bertsimas D, Thiele A (2006) Robust and data-driven optimization: modern decision making under uncertainty. In: Johnson M (ed) Tutorials in operations research: models, methods, and applications for innovative decision making. INFORMS, Catonsville, pp 95–122

Carvalho M, Serra LM, Lozano MA (2011) Optimal synthesis of trigeneration systems subject to environmental constraints. Energy 36:3779–3790

Chevalier JL, Téno JFÇL (1996) Life cycle analysis with ill-defined data and its application to building products. Int J Life Cycle Assess 1: 90–96

Fava JA, SETAC (Society) and SETAC Foundation for Environmental Education (1994) Life-cycle assessment data quality: a conceptual framework: workshop report. Society of Environmental Toxicology and Chemistry and SETAC Foundation for Environmental Education

Franklin JN (1987) Methods of mathematical economics: linear and nonlinear programming, fixed-point theorems, vol 37, Classics in applied mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia

Freire F, Thore S, Ferrao P (2001) Life cycle activity analysis: logistics and environmental policies for bottled water in Portugal. OR Spectr 23(1):159–182

Frischknecht R, Jungbluth N, Althaus HJ, Doka G, Heck T, Hellweg S, Hischier R, Nemecek T, Rebitzer G, Spielmann M and others (2007) Overview and methodology. Ecoinvent report: 51–53

Guillén-Gosálbez G, Caballero JA, Jiménez L (2008) Application of life cycle assessment to the structural optimization of process flowsheets. Ind Eng Chem Res 47:777–789

Heijungs R (2010) Sensitivity coefficients for matrix-based LCA. Int J Life Cycle Assess 15:511–520

Heijungs R, Huijbregts MA (2004) A review of approaches to treat uncertainty in LCA. In: iEMSs 2004 International Congress: "Complexity and Integrated Resources Management". International Environmental Modelling and Software Society, Osnabrueck, Germany

Heijungs R, Suh S (2002) The computational structure of life cycle assessment, vol 11, Eco-efficiency in industry and science. Springer, New York

Hugo A, Pistikopoulos EN (2005) Environmentally conscious long-range planning and design of supply chain networks. J Clean Prod 13: 1471–1491

Huijbregts M (2002) Uncertainty and variability in environmental life cycle assessment. Int J Life Cycle Assess 7(3):173–173

Lloyd SM, Ries R (2007) Characterizing, propagating and analyzing uncertainty in life cycle assessment: a survey of quantitative approaches. J Ind Ecol 11(1):161–179

McCleese DL, LaPuma PT (2002) Using Monte Carlo simulation in life cycle assessment for electric and internal combustion vehicles. Int J Life Cycle Assess 7:230–236

Morgan MG, Henrion M (1990) Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge University Press, Cambridge

Sakai S, Yokoyama K (2002) Formulation of sensitivity analysis in life cycle assessment using a perturbation method. Clean Techn Environ Policy 4:72–78

Sonnemann GW, Schuhmacher M, Castells F (2003) Uncertainty assessment by a Monte Carlo simulation in a life cycle inventory of electricity produced by a waste incinerator. J Clean Prod 11:279–292

Soyster AL (1973) Convex programming with set-inclusive constraints and applications to inexact linear programming. Oper Res: 1154–1157

Tan RR (2005) Application of symmetric fuzzy linear programming in life cycle assessment. Environ Model Softw 20:1343–1346

Tan RR (2008) Using fuzzy numbers to propagate uncertainty in matrix-based LCI. Int J Life Cycle Assess 13:585–592

Tan RR, Culaba AB, Aviso KB (2008) A fuzzy linear programming extension of the general matrix-based life cycle model. J Clean Prod 16:1358–1367

US-EPA (1989) Exposure factors handbook. EPA, Washington