

Sensitivity coefficients for matrix-based LCA

Reinout Heijungs

Received: 26 May 2009 / Accepted: 12 January 2010 / Published online: 20 February 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract

Background, aim, and scope Matrix-based life cycle assessment (LCA) is part of the standard ingredients of modern LCA tools. An important aspect of matrix-based LCA that is straightforward to carry out, but that requires a careful mathematical handling, is the inclusion of sensitivity coefficients based on differentiating the matrix-based formulas.

Materials and methods We briefly review the basic equations for LCA and the basic theory of sensitivity coefficients.

Results We present the complete set of sensitivity coefficients from inventory to weighting through characterization and normalization. We show the specific formulas for perturbation analysis, uncertainty analysis, and key issue analysis. We also provide an example using the ecoinvent data.

Discussion The limitations of the present approach include the restriction to small changes and uncertainties and the ignorance of correlation between input uncertainties. In contrast to common thinking, there is no restriction to normally distributed uncertainties: Every uncertainty distribution for which a variance can be defined can be submitted to the analytical uncertainty analysis.

Conclusions This paper provides a useful set of tables for a number of purposes related to uncertainty and sensitivity analysis.

Recommendations and perspectives Although the formulas derived are not simple, they are straightforward to implement in software for LCA. Once this is done, the use of these formulas can become routine practice, enabling a key issue analysis and speeding up perturbation and uncertainty analysis.

Keywords Derivatives · Life cycle interpretation · Matrix-based LCA · Sensitivity · Taylor series expansion · Uncertainty

1 Introduction

2 Background, aim, and scope

By now, the matrix approach to life cycle assessment (LCA) has received wide recognition. Proposed in the early 1990s (Möller 1992; Heijungs et al. 1992), it took a decade or so before this idea was embraced as a general accepted way of doing LCA (see Heijungs and Suh 2006 for a short history of the matrix approach to LCA).

The matrix method has been developed to solve the inventory problem in LCA. With the inventory problem, we refer to the task of scaling all unit processes in the system in such a way that they exactly produce the reference flow (or functional unit) and to use this scaling such that the inventory table can be computed. Formulas such as

$$\mathbf{g} = \mathbf{BA}^{-1}\mathbf{f} \quad (1)$$

(see Tables 1 and 2 for an explanation of the symbols involved) are now common in the specialized literature on life cycle inventory (LCI; Heijungs and Suh 2002; Peters 2007; Tan 2008).

Responsible editor: Ralph Rosenbaum

R. Heijungs (✉)
Institute of Environmental Sciences (CML), Faculty of Science,
Leiden University,
PO Box 9518, 2300 RA Leiden, The Netherlands
e-mail: heijungs@cml.leidenuniv.nl

Table 1 Overview of symbols representing input data in the LCA

Symbol	Name	Dimension (rows × columns)	Defined in
f	Final demand vector	Economic flows × 1	Goal and scope definition
A	Technology matrix	Economic flows × processes	Inventory analysis
B	Intervention matrix	Environmental flows × processes	Inventory analysis
Q	Characterisation matrix	Categories × environmental flows	Impact assessment
g	Intervention totals ^a	Environmental flows × 1	Impact assessment
h	Category totals ^a	Categories × 1	Impact assessment
w	Weighting factors	1 × categories	Impact assessment

^a See the note below on normalization

The matrix approach has also been extended in various directions. For instance, research has been devoted to add details on:

- the treatment of allocation and cutoff (Heijungs and Frischknecht 1998);
- how to connect a process-based LCI to an input–output table (Suh and Huppes 2005);
- how to efficiently compute an answer to the inventory problem (Peters 2007);
- how to analyze the feedback structure of the system (Suh and Heijungs 2007);
- how to calculate sensitivity coefficients (Heijungs 1994).

Most of these details indeed refer to exclusively inventory-oriented questions. The impact assessment follows the inventory results, so all issues related to cutoff, allocation, IO-based LCA, efficient algorithms, and structural path analysis are only interesting from an LCI point of view. For the last one mentioned, the sensitivity coefficients, this is different, however.

Sensitivity coefficients are important for both uncertainty analysis and sensitivity analysis (Heijungs 1994). In the context of uncertainty analysis, they serve to establish essential information for a Taylor series expansion, and for sensitivity analysis, they provide the multipliers that enable one to distinguish sensitive from non-sensitive parameters, the so-called key issues for refined data collection (Heijungs 1996). But uncertainty and sensitivity analyses are not only important in LCI but in impact assessment as well.

Moreover, the impact assessment adds additional uncertainty to the already uncertain results of the LCI. Likewise, not only the sensitivity of inventory results is of interest but also (or perhaps even more so) the impact assessment results.

The extension of the sensitivity coefficients from LCI to life cycle impact assessment was “left as an exercise” to the LCA practitioner. For instance, Heijungs and Suh (2002, p. 144) write that “In this way, all equations of LCA may be processed,” but they do not pursue this. It is a task that is not so often carried out, we guess, at least we have never seen the explicit results of such an exercise. This paper therefore aims to carry out this exercise and to make the results available. The formulas obtained can easily be implemented in matrix-based software for LCA. We have done so in CMLCA, a program for doing LCA, and some screenshots are shown at the end of the paper.

In this paper, we first review the basic equations of LCA. Then, we proceed to review the theory of sensitivity coefficients in general and their form in LCI. This finally leads to a derivation and coherent presentation of the sensitivity coefficients for the entire LCA process.

3 Materials and methods

3.1 Basic equations for LCA

The basic equations for LCA have been discussed in a consistent notation by Heijungs and Suh (2002). Below, we

Table 2 Overview of symbols representing output results in the LCA and the formulas with which these results are obtained

Symbol	Name	Dimension (rows × columns)	Equation
s	Scaling factors	processes × 1	$\mathbf{s} = \mathbf{A}^{-1} \mathbf{f}$
g	Inventory results	Environmental flows × 1	$\mathbf{g} = \mathbf{B} \mathbf{s}$
h	Characterization results	Categories × 1	$\mathbf{h} = \mathbf{Q} \mathbf{g}$
Λ	Intensity matrix	Environmental flows × economic flows	$\mathbf{\Lambda} = \mathbf{B} \mathbf{A}^{-1}$
$\tilde{\mathbf{h}}$	Normalization results	Categories × 1	$\forall k : \tilde{h}_k = \frac{h_k}{\dot{h}_k}$
$\dot{\mathbf{h}}$	Category totals ^a	Categories × 1	$\dot{\mathbf{h}} = \mathbf{Q} \dot{\mathbf{g}}$
W	Weighted index	1 × 1	$W = \mathbf{w} \tilde{\mathbf{h}}$

^a See the note below on normalization

present two tables of symbols and equations connecting the fundamental concepts of LCA.

With respect to normalization, there are two situations which require separate treatment.

- A vector of intervention totals, $\dot{\mathbf{g}}$, can be defined for the reference situation. For instance, one can collect data on the emissions of CO₂, SO₂, etc., which then represent \dot{g}_1, \dot{g}_2 , etc. These then can be processed by the same characterization model to yield the vector of category totals, $\dot{\mathbf{h}}$. This then forms the basis of the normalization. Changing $\dot{\mathbf{g}}$ will induce a change in $\dot{\mathbf{h}}$, but $\dot{\mathbf{h}}$ itself will not be changed directly by the LCA practitioner. We will refer to this as normalization case 1.
- Alternatively, the vector of category totals $\dot{\mathbf{h}}$ can be known without a detailed specification of the underlying interventions. In that case, $\dot{\mathbf{h}}$ is not an output result (in the sense of belonging to Table 2), but input data (in the sense of belonging to Table 1). Thus, $\dot{\mathbf{h}}$ can be changed directly, and it will affect the normalization results and the weighted index. We will refer to this as normalization case 2.

Both approaches in fact appear in practice and are therefore elaborated below as separate cases.

3.2 General theory of sensitivity coefficients

There are various situations in which the stability of the results in terms of sensitivity for perturbations of the input data is of interest. In general, we may explore this issue as follows. Suppose that an output variable z depends on two input variables x and y and that the dependence is expressed by a function f :

$$z = f(x, y). \tag{2}$$

The crucial elements in determining sensitivity is the change of the result (Δz) caused by a marginal change in x (Δx) and by a marginal change in y (Δy). This is expressed using the partial derivatives:

$$\Delta z = \frac{\partial z}{\partial x} \Delta x + \frac{\partial z}{\partial y} \Delta y. \tag{3}$$

Coefficients such as $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ are referred to as sensitivity coefficients in the present context. Their evaluation requires a specification of the function f . Thus, with f specified, the sensitivity coefficient of f with respect to input parameter x is defined as:

$$\frac{\partial z}{\partial x} = \frac{\partial f(x, y)}{\partial x}. \tag{4}$$

4 Results

4.1 Sensitivity coefficients for LCA

Table 2 provides a specification of the functions f in the context of LCA. So we are to insert the equations of Table 2 into the general equation for calculating a sensitivity coefficient (Eq. 4). Example calculations have been provided by Heijungs and Suh (2002, Eqs. (6.21), (6.26), and (6.29)). For the scaling factors, s , we have:

$$\frac{\partial s_k}{\partial a_{ij}} = -(\mathbf{A}^{-1})_{ki} s_j \tag{5}$$

and for the inventory results, \mathbf{g} , we have:

$$\frac{\partial g_k}{\partial a_{ij}} = -\lambda_{ki} s_j \tag{6}$$

for the dependence on the elements of \mathbf{A} , and

$$\frac{\partial g_k}{\partial b_{ij}} = s_j \delta_{ik} \tag{7}$$

for the dependence on the elements of \mathbf{B} , where $\delta_{ik} = \begin{cases} 1 & \text{if } i = k \\ 0 & \text{otherwise} \end{cases}$ represents the Kronecker delta.

It is clear that the higher one moves in the sequence inventory–characterization–normalization–weighting, the larger the number of sensitivity coefficients there will be. Scaling factors only depend on the technology matrix, but the inventory results depend on the technology matrix and on the intervention matrix.

Table 3 gives a tabular overview of the derivatives for all expressions of matrix-based LCA.

One remark on these coefficients. Each of the formulas in Table 3 contains \mathbf{A}^{-1} or a symbol that depends on \mathbf{A}^{-1} , such as s , $\mathbf{\Lambda}$ or \mathbf{h} . Thus, we need to go through the process of matrix inversion to evaluate the sensitivity coefficients. Moreover, although Citroth et al. (2004) argue that one can solve LCAs without calculating a full matrix inverse, Table 3 shows that we need the full matrix inverse when we wish to extend the analysis to include sensitivity and analytical uncertainty studies.

4.2 Perturbation analysis

Heijungs and Kleijn (2001) and Sakai and Yokoyama (2002) describe perturbation analysis as a way of investigating which input data are most decisive for the results in terms of their relative sensitivity. That is, given a system with the prototypical form

$$z = f(x, y) \tag{8}$$

Table 3 Overview of the sensitivity coefficients that express how LCA output results (columns) change if LCA input data (rows) data change

	$\partial s_k / \dots$	$\partial g_k / \dots$	$\partial h_k / \dots$	$\partial \tilde{h}_k / \dots$	$\partial W / \dots$
$\dots / \partial a_{ij}$	$-(\mathbf{A}^{-1})_{ki} s_j$	$-\lambda_{ki} s_j$	$-s_j \sum_l q_{kl} \lambda_{li}$	$-\frac{s_j}{h_k} \sum_l q_{kl} \lambda_{li}$	$-s_j \sum_k \frac{w_k}{h_k} \sum_l q_{kl} \lambda_{li}$
$\dots / \partial b_{ij}$	0	$s_j \delta_{ik}$	$q_{ki} s_j$	$\frac{q_{ki} s_j}{h_k}$	$s_j \sum_k \frac{w_k}{(h_k)^2} q_{ki}$
$\dots / \partial q_{ij}$	0	0	$g_j \delta_{ik}$	$\left(\frac{g_j}{h_k} - \frac{h_k g_j^2}{(h_k)^2}\right) \delta_{ik}$	$w_i \left(\frac{g_j}{h_i} - \frac{h_i g_j^2}{(h_i)^2}\right)$
$\dots / \partial \dot{g}_i$ (normalization case 1)	0	0	0	$-\frac{h_k q_{ki}}{(h_k)^2}$	$-\sum_k \frac{w_k h_k q_{ki}}{(h_k)^2}$
$\dots / \partial \dot{h}_i$ (normalization case 2)	0	0	0	$-\frac{h_k}{(h_k)^2} \delta_{ik}$	$-\frac{w_i h_i}{(h_i)^2}$
$\dots / \partial w_i$	0	0	0	0	\tilde{h}_i

Example: row 3 and column 4 tell us that $\frac{\partial h_k}{\partial b_{ij}} = q_{ki} s_j$

Note that depending on the situation in normalization, either row 5 ($\dots / \partial q_{ij}$) or row 6 ($\dots / \partial \dot{h}_i$) applies

one investigates dimensionless multipliers, such as

$$\frac{\partial z}{\partial x} \frac{x}{z} \tag{9}$$

The idea is that a small change of either of the input parameters (say in x , hence a change Δx) leads to a change in the result (z , hence Δz), and that the relative change $\frac{\Delta z/z}{\Delta x/x}$ can be approximated by $\frac{\partial z/z}{\partial x/x}$. The results of Tables 2 and 3 can be combined to yield a complete overview of these multipliers (see Table 4).

4.3 Uncertainty analysis

Although the Monte Carlo method is right now the most frequently applied method for studying the propagation of input uncertainties into output uncertainties (Lloyd and Ries 2007), it has been recognized that this method may be too computationally intensive for application to large systems (Ciroth et al. 2004; Heijungs et al. 2005; Hong et al. 2008). Indeed, the present release of the ecoinvent data (v2.0, comprising almost 4,000 processes) has not been processed with Monte Carlo analysis, while the previous release (v1.3, 2,500 processes) contained such results, precisely for this reason. One way to address this is by means of smarter algorithms. Peters (2007) has proposed power series expansion, also in connection to Monte Carlo analysis. Moreover, there are more sophisticated sampling strategies than the naive Monte Carlo method, such as Latin hypercube sampling and response surface methods (e.g., Morgan and Henrion 1990). In this paper, we address it with a solution on the basis of analytical error propagation. The theory of analytical error propagation, using Taylor series expansion, has been proposed in LCA (Heijungs 1994; Giroth et al. 2004; Hong et al. 2008). Taylor series expansions are based on the approximation formula for calculating the variance of a stochastic result using

Table 4 Overview of the relative sensitivity coefficients (multipliers) that express how small changes in input data propagate into changes in output results

Multiplier	Definition	Formula
$\sigma_k(a_{ij})$	$\frac{\partial s_k / s_k}{\partial a_{ij} / a_{ij}}$	$-\frac{a_{ij}}{s_k} (\mathbf{A}^{-1})_{ki} s_j$
$\gamma_k(a_{ij})$	$\frac{\partial g_k / g_k}{\partial a_{ij} / a_{ij}}$	$-\frac{a_{ij}}{g_k} \lambda_{ki} s_j$
$\gamma_k(b_{ij})$	$\frac{\partial g_k / g_k}{\partial b_{ij} / b_{ij}}$	$\frac{b_{ij}}{g_k} s_j \delta_{ik}$
$\eta_k(a_{ij})$	$\frac{\partial h_k / h_k}{\partial a_{ij} / a_{ij}}$	$-\frac{a_{ij}}{h_k} s_j \sum_l q_{kl} \lambda_{li}$
$\eta_k(b_{ij})$	$\frac{\partial h_k / h_k}{\partial b_{ij} / b_{ij}}$	$\frac{b_{ij}}{h_k} q_{ki} s_j$
$\eta_k(q_{ij})$	$\frac{\partial h_k / h_k}{\partial q_{ij} / q_{ij}}$	$\frac{q_{ij}}{h_k} g_j \delta_{ik}$
$\tilde{\eta}_k(a_{ij})$	$\frac{\partial h_k / h_k}{\partial a_{ij} / a_{ij}}$	$-\frac{a_{ij}}{h_k} \frac{s_j}{h_k} \sum_l q_{kl} \lambda_{li}$
$\tilde{\eta}_k(b_{ij})$	$\frac{\partial h_k / h_k}{\partial b_{ij} / b_{ij}}$	$\frac{b_{ij}}{h_k} \frac{q_{ki} s_j}{h_k}$
$\tilde{\eta}_k(q_{ij})$	$\frac{\partial h_k / h_k}{\partial q_{ij} / q_{ij}}$	$\frac{q_{ij}}{h_k} \left(\frac{g_j}{h_k} - \frac{h_k g_j^2}{(h_k)^2}\right) \delta_{ik}$
$\tilde{\eta}_k(\dot{g}_i)$ (normalization case 1)	$\frac{\partial h_k / h_k}{\partial \dot{g}_i / \dot{g}_i}$	$-\frac{\dot{g}_i}{h_k} \frac{h_k q_{ki}}{(h_k)^2}$
$\tilde{\eta}_k(\dot{h}_i)$ (normalization case 2)	$\frac{\partial h_k / h_k}{\partial \dot{h}_i / \dot{h}_i}$	$-\frac{\dot{h}_i}{h_k} \frac{h_k}{(h_k)^2} \delta_{ik}$
$\omega(a_{ij})$	$\frac{\partial W / W}{\partial a_{ij} / a_{ij}}$	$-\frac{a_{ij}}{W} s_j \sum_k \frac{w_k}{h_k} \sum_l q_{kl} \lambda_{li}$
$\omega(b_{ij})$	$\frac{\partial W / W}{\partial b_{ij} / b_{ij}}$	$\frac{b_{ij}}{W} s_j \sum_k \frac{w_k}{(h_k)^2} q_{ki}$
$\omega(q_{ij})$	$\frac{\partial W / W}{\partial q_{ij} / q_{ij}}$	$\frac{q_{ij}}{W} w_i \left(\frac{g_j}{h_i} - \frac{h_i g_j^2}{(h_i)^2}\right)$
$\omega(\dot{g}_i)$ (normalization case 1)	$\frac{\partial W / W}{\partial \dot{g}_i / \dot{g}_i}$	$-\frac{\dot{g}_i}{W} \sum_k \frac{w_k h_k q_{ki}}{(h_k)^2}$
$\omega(\dot{h}_i)$ (normalization case 2)	$\frac{\partial W / W}{\partial \dot{h}_i / \dot{h}_i}$	$-\frac{\dot{h}_i}{W} \frac{w_i h_i}{(h_i)^2}$
$\omega(w_i)$	$\frac{\partial W / W}{\partial w_i / w_i}$	$\frac{w_i}{W} \tilde{h}_i$

Example: $\sigma_k(a_{ij})$ tells how many percent s_k will change if a_{ij} changes 1%

stochastic data (Bevington and Robinson 1994; Morgan and Henrion 1990). For a system of the form

$$z = f(x, y) \tag{10}$$

it assumes the form

$$\begin{aligned} \text{var}(z) = & \left(\frac{\partial f}{\partial x}\right)^2 \text{var}(x) + \left(\frac{\partial f}{\partial y}\right)^2 \text{var}(y) \\ & + 2 \frac{\partial f}{\partial x} \times \frac{\partial f}{\partial y} \text{cov}(x, y) \end{aligned} \tag{11}$$

where $\text{var}(x)$ is the variance of the variable x and $\text{cov}(x, y)$ represents the covariance between the stochastic variables x and y .

In most cases, no covariance data are available, or the covariance can be assumed to be negligible as the uncertainties are in many cases independent. In those cases, we set covariance to zero and obtain

$$\text{var}(z) = \left(\frac{\partial f}{\partial x}\right)^2 \text{var}(x) + \left(\frac{\partial f}{\partial y}\right)^2 \text{var}(y) \tag{12}$$

Such equations are for a general function $f(x, y)$, and they require the evaluation of the derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$. In the present case, we have specified the function (Table 2), and we have found equations for the derivatives (Table 3). So all ingredients are available to complete the structure of the uncertainty analysis.

Heijungs and Suh (2002, Eqs. (6.73) and (8.87)) elaborate this only for the inventory vector, and then even only partially, and with a typo. The complete and correct expression for this is

$$\text{var}(g_k) = \sum_{ij} (s_j \lambda_{ki})^2 \text{var}(a_{ij}) + \sum_j (s_j)^2 \text{var}(b_{kj}). \tag{13}$$

The remaining expressions for the uncertainty of the impact assessment results are elaborated in Table 5.

Again, mind that there are two expressions for $\text{var}(\tilde{h}_k)$ and two expressions for $\text{var}(W)$ for both normalization variants.

4.4 Key issue analysis

Key issue analysis has been defined by Heijungs (1996) as the decomposition of the uncertainty of a result in terms of the contribution of the uncertainties of the input data. Morgan and Henrion (1990) refer to it as uncertainty importance, while Saltelli et al. (2000) use the (perhaps confusing) term sensitivity analysis. All expressions for the variance in Table 4 are the result of the weighted aggregation of a large number of variances of input data. For instance, the expression for $\text{var}(g_k)$ comprises two summations, one over i, j and one over j . Adding the

interpretation of these indices, we see summations over all economic flows and twice over all processes. Take the case of ecoinvent v2.0 where the number of processes and economic flows is almost 4,000; this variance is the result of almost 12,000 terms. Each of these terms is positive, so we may really consider $(s_j \lambda_{ki})^2 \text{var}(a_{ij})$ and $(s_j \lambda_{ki})^2 \text{var}(b_{ij})$ as the contribution that one individual $\text{var}(a_{ij})$ or $\text{var}(b_{ij})$ makes to the total $\text{var}(g_k)$. We therefore define a number of dimensionless coefficients ζ , namely:

$$\zeta(g_k, a_{ij}) = \frac{(s_j \lambda_{ki})^2 \text{var}(a_{ij})}{\text{var}(g_k)} \tag{14}$$

and

$$\zeta(g_k, b_{ij}) = \frac{(s_j \delta_{ik})^2 \text{var}(b_{ij})}{\text{var}(g_k)} \tag{15}$$

as the relative contributions by each $\text{var}(a_{ij})$ and $\text{var}(b_{ij})$ to the total $\text{var}(g_k)$. Naturally,

$$\sum_{ij} \zeta(g_k, a_{ij}) + \sum_j \zeta(g_k, b_{kj}) = 1 \tag{16}$$

so these indeed represent relative contributions to the variance of the total.

Table 6 provides an overview of the expressions for the different types of key issues.

4.4.1 Example

We have programmed these equations into CMLCA (<http://www.cmlca.eu/>). As an example, we loaded the ecoinvent v1.3 (2,630 unit processes) and calculated the system for a reference flow of 1-kWh Swiss electricity, low voltage, at grid. In the example below, we have restricted the analysis to the inventory analysis and to the emission of Carbon dioxide, fossil, emitted to air, low population density.

We performed a perturbation analysis, concentrating on the perturbations of the technology matrix. Figure 1 shows a screenshot of the results, hence of the coefficients $\gamma_k(a_{ij})$ for k is CO₂ and for all i (economic flows) and all j (processes). A not-too modern computer was able to complete the calculations within a few minutes.

We also performed an uncertainty analysis. Figure 2 shows a screenshot of these results, so tabulating $\text{var}(g_k)$ (along with some other statistics) for k is CO₂. Again, results are obtained within a few minutes.

We finally performed a key issue analysis. Figure 3 shows a screenshot of these results, so tabulating $\zeta(g_k, a_{ij})$

Table 5 Overview of the variance of output results as a function of the variance of the input data

Uncertainty	Equation
$\text{var}(s_k)$	$\sum_{ij} (s_j)^2 \left((\mathbf{A}^{-1})_{ki} \right)^2 \text{var}(a_{ij})$
$\text{var}(g_k)$	$\sum_{ij} (s_j \lambda_{ki})^2 \text{var}(a_{ij}) + \sum_j (s_j)^2 \text{var}(b_{kj})$
$\text{var}(h_k)$	$\sum_{ij} \left(\sum_l q_{kl} \lambda_{li} \right)^2 \text{var}(a_{ij}) + \sum_{ij} (s_j q_{ki})^2 \text{var}(b_{ij}) + \sum_j (g_j)^2 \text{var}(q_{kj})$
$\text{var}(\tilde{h}_k)$	$\left\{ \begin{aligned} & \left[\sum_{ij} \left(\sum_l q_{kl} \lambda_{li} \right)^2 \text{var}(a_{ij}) + \sum_{ij} (s_j q_{ki})^2 \text{var}(b_{ij}) + \sum_j \left(g_j - \frac{h_k g_j}{h_k} \right)^2 \text{var}(q_{kj}) + \sum_l \left(\frac{h_k g_l}{h_k} \right)^2 \text{var}(g_l) \right] \quad (\text{normalization case 1}) \\ & \frac{1}{(h_k)^2} \left[\sum_{ij} \left(\sum_l q_{kl} \lambda_{li} \right)^2 \text{var}(a_{ij}) + \sum_{ij} (s_j q_{ki})^2 \text{var}(b_{ij}) + \left(\frac{h_k}{h_k} \right)^2 \text{var}(h_k) \right] \quad (\text{normalization case 2}) \end{aligned} \right.$
$\text{var}(W)$	$\left\{ \begin{aligned} & \sum_{ij} (s_j)^2 \left(\sum_l \frac{w_l}{h_l} \sum_k q_{kl} \lambda_{ki} \right)^2 \text{var}(a_{ij}) + \sum_{ij} (s_j)^2 \left(\sum_l \frac{w_l}{h_l} q_{li} \right)^2 \text{var}(b_{ij}) + \sum_{ij} \left(\frac{w_i}{h_i} \right)^2 \left(g_j - \frac{h_k g_j}{h_k} \right)^2 \text{var}(q_{ij}) + \sum_l \left(\frac{w_l g_l}{h_l} \right)^2 \text{var}(h_l) + \sum_l \tilde{h}_l^2 \text{var}(w_l) \quad (\text{normalization case 1}) \\ & \sum_{ij} (s_j)^2 \left(\sum_l \frac{w_l}{h_l} \sum_k q_{kl} \lambda_{ki} \right)^2 \text{var}(a_{ij}) + \sum_{ij} (s_j)^2 \left(\sum_l \frac{w_l}{h_l} q_{li} \right)^2 \text{var}(b_{ij}) + \sum_{ij} \left(\frac{w_i}{h_i} \right)^2 \left(g_j - \frac{h_k g_j}{h_k} \right)^2 \text{var}(q_{ij}) + \sum_l \left(\frac{w_l g_l}{h_l} \right)^2 \text{var}(h_l) + \sum_l \tilde{h}_l^2 \text{var}(w_l) \quad (\text{normalization case 2}) \end{aligned} \right.$

Example: $\text{var}(g_k)$ represents the variance of the inventory result for the k th environmental flow

Table 6 Overview of the contributions to the variance of output results by the individual variance of the input data

Contribution	Formula
$\zeta(s_k, a_{ij})$	$\frac{((A^{-1})_{ki} s_j)^2 \text{var}(a_{ij})}{\text{var}(s_k)}$
$\zeta(g_k, a_{ij})$	$\frac{(s_j \lambda_{ki})^2 \text{var}(a_{ij})}{\text{var}(g_k)}$
$\zeta(g_k, b_{ij})$	$\frac{(s_j \delta_{ik})^2 \text{var}(b_{ij})}{\text{var}(g_k)}$
$\zeta(h_k, a_{ij})$	$\frac{\left(s_j \sum_l q_{kl} \lambda_{li} \right)^2 \text{var}(a_{ij})}{\text{var}(h_k)}$
$\zeta(h_k, b_{ij})$	$\frac{(s_j q_{ki})^2 \text{var}(b_{ij})}{\text{var}(h_k)}$
$\zeta(h_k, q_{ij})$	$\frac{(g_j)^2 \text{var}(q_{ij})}{\text{var}(h_k)}$
$\zeta(\tilde{h}_k, a_{ij})$	$\frac{\left(s_j \sum_l q_{kl} \lambda_{li} \right)^2 \text{var}(a_{ij})}{(\tilde{h}_k)^2 \text{var}(\tilde{h}_k)}$
$\zeta(\tilde{h}_k, b_{ij})$	$\frac{(s_j q_{ki})^2 \text{var}(b_{ij})}{(\tilde{h}_k)^2 \text{var}(\tilde{h}_k)}$
$\zeta(\tilde{h}_k, q_{ij})$	$\frac{\left(g_j - \frac{h_k g_j}{h_k} \right)^2 \text{var}(q_{ij})}{(\tilde{h}_k)^2 \text{var}(\tilde{h}_k)}$
$\zeta(\tilde{h}_k, \hat{g}_i)$ (normalization case 1)	$\frac{\left(\frac{h_k q_{ki}}{h_k} \right)^2 \text{var}(\hat{g}_i)}{(\tilde{h}_k)^2 \text{var}(\tilde{h}_k)}$
$\zeta(\tilde{h}_k, \hat{h}_i)$ (normalization case 2)	$\frac{\left(\frac{h_k}{h_k} \right)^2 \text{var}(\hat{h}_i)}{(\tilde{h}_k)^2 \text{var}(\tilde{h}_k)}$
$\zeta(W, a_{ij})$	$\frac{(s_j)^2 \left(\sum_l \frac{w_l}{h_l} \sum_k q_{lk} \lambda_{ki} \right)^2 \text{var}(a_{ij})}{\text{var}(W)}$
$\zeta(W, b_{ij})$	$\frac{(s_j)^2 \left(\sum_l \frac{w_l}{h_l} q_{li} \right)^2 \text{var}(b_{ij})}{\text{var}(W)}$
$\zeta(W, q_{ij})$	$\frac{\left(\frac{w_i}{h_i} \right)^2 \left(g_j - \frac{h_i g_j}{h_i} \right)^2 \text{var}(q_{ij})}{\text{var}(W)}$
$\zeta(W, \hat{g}_i)$ (normalization case 1)	$\frac{\left(\frac{w_i h_i}{h_i h_i} \right)^2 \text{var}(\hat{g}_i)}{\text{var}(W)}$
$\zeta(W, \hat{h}_i)$ (normalization case 2)	$\frac{\left(\frac{w_i h_i}{h_i h_i} \right)^2 \text{var}(\hat{h}_i)}{\text{var}(W)}$
$\zeta(W, w_i)$	$\frac{(\tilde{h}_i)^2 \text{var}(w_i)}{\text{var}(W)}$

Example: $\zeta(g_k, a_{ij})$ represents the contribution to the variance of g_k by the variance of a_{ij}

and $\zeta(g_k, b_{ij})$ for k is CO₂ and for all i (economic flows) and all j (processes). Again, a few minutes suffice.

Here, we see that only four coefficients make up 80% of the uncertainty of the CO₂ emission. In other words, we can

obtain a more reliable, less uncertain result by trying to find more accurate data for these four coefficients.

5 Discussion

Three important restrictions must be born in mind.

First, the formulas are for the “normal” LCI. This means that the equations become more complicated once we incorporate other features and developments, such as allocation and IO-based LCA. Heijungs et al. (2006) show how the formula for $\frac{\partial g_k}{\partial a_{ij}}$ changes when the inventory is done using a hybrid method, combining process-based and IO-based LCA. This modification may be propagated to the impact assessment level as well.

Second, some of the formulas (namely those in Tables 4, 5, and 6) are based on a first-order Taylor series approximation. That means that they are correct for small changes, uncertainties, and perturbations, but not necessarily for larger ones. To fix this, one may either include second-order terms (or even further). For instance, for the perturbations, we may improve by going from

$$\Delta z = \frac{\partial z}{\partial x} \Delta x \tag{17}$$

into

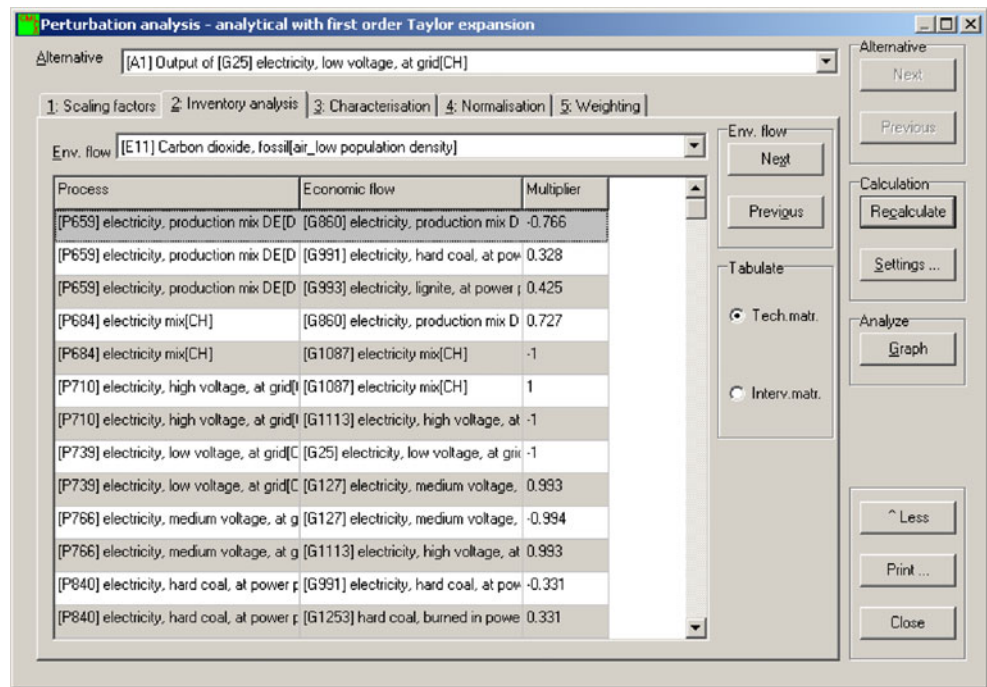
$$\Delta z = \frac{\partial z}{\partial x} \Delta x + \frac{1}{2} \frac{\partial^2 z}{\partial x^2} (\Delta x)^2 + \frac{1}{6} \frac{\partial^3 z}{\partial x^3} (\Delta x)^3 + \dots \tag{18}$$

This requires the evaluation of not only $\frac{\partial z}{\partial x}$ but also of the second derivative $\frac{\partial^2 z}{\partial x^2}$ or even beyond. Table 3 might be extended to include expressions for $\frac{\partial^2 s_k}{\partial (a_{ij})^2}$ and similar. Alternatively, the approach by Sherman and Morrison (1950) can be used to provide an exact answer to this question.

Third and finally, the formulas for uncertainty (Tables 5 and 6) are based on the ignorance of the covariance between input variables. As we noted above, the complete first-order expression includes an additional term for the covariance between input data:

$$\begin{aligned} \text{var}(z) &= \left(\frac{\partial f}{\partial x} \right)^2 \text{var}(x) + \left(\frac{\partial f}{\partial y} \right)^2 \text{var}(y) \\ &+ 2 \frac{\partial f}{\partial x} \times \frac{\partial f}{\partial y} \text{cov}(x, y) \end{aligned} \tag{19}$$

Fig. 1 Screenshot of CMLCA showing the results of a perturbation analysis, the sensitivity of an output result for small changes of the input data. Only multipliers smaller than -0.3 and larger than 0.3 have been shown



It is possible to carry out the program of this paper including covariance, but the expressions become much more complicated. For instance, for $\text{var}(g_k)$, we obtain

$$\begin{aligned}
 \text{var}(g_k) &= \sum_{i,j} (s_j \lambda_{ki})^2 \text{var}(a_{ij}) + \sum_j (s_j)^2 \text{var}(b_{kj}) \\
 &+ 2 \sum_{i,j,l,m} s_j \lambda_{ki} s_m \lambda_{kl} \text{cov}(a_{ij}, a_{lm}) \\
 &+ 2 \times \sum_{j,l} s_j s_l \text{cov}(b_{kj}, b_{kl}) \\
 &+ 2 \times \sum_{i,j,l} s_j \lambda_{ki} s_l \text{cov}(a_{ij}, b_{kl}). \quad (20)
 \end{aligned}$$

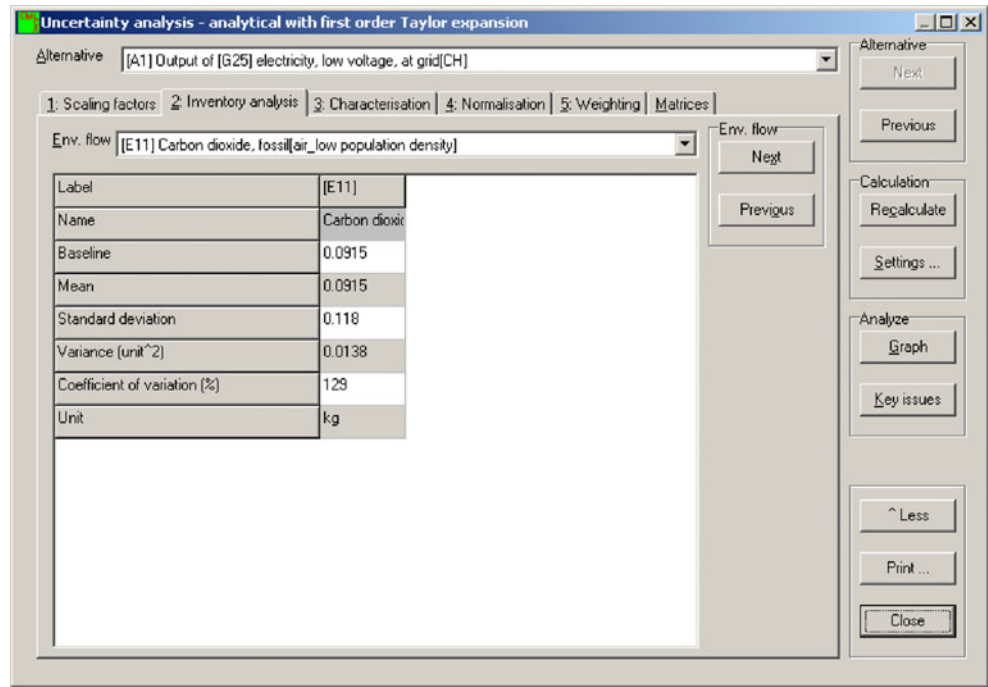
The last three terms requires us to specify $\text{cov}(a_{ij}, a_{lm})$, $\text{cov}(a_{kj}, b_{kl})$, and $\text{cov}(a_{ij}, b_{kl})$ for all combinations of processes, economic flows, and environmental flows. Although some uncertainties will be definitely correlated (for instance, the fuel input of a combustion process and the CO₂ emission of the same process may have a correlation close to 1), most uncertainties will be uncorrelated, or any information on such correlations is lacking. The infrastructure needed (lots of memory for storing a number of covariance matrices, much more complicated formulas, much more data collection and estimation) will probably not offset the relatively limited gain of having a slightly more accurate computation. In the end, there is always something perverse about knowing the uncertainty with certainty.

A question that always arises in connection to analytical error propagation is whether it only works for normally distributed uncertainties. This is not the case. The theory of analytical error propagation through Taylor series expansions (Morgan and Henrion 1990, p. 183 ff.) nowhere contains the assumption of normally distributed distributions. The only assumption is that for a first-order approximation, the function should be sufficiently close to a linear function within the range of uncertainty. This point has been addressed above. A practical issue is of course that the formulas require a specification of the variance, whereas most distributions are specified without an explicit variance. For instance, a uniform distribution is often specified in terms of its width and a lognormal distribution in terms of the geometric standard deviation (or its square, as inecoinvent). Heijungs and Frischknecht (2005) provide formulas to easily calculate a variance from the standard parameters of a normal, lognormal, uniform, and triangular distribution. As we can see in Table 5, the propagated variances are the sum of a large number of terms. Following the central limit theorem (e.g., Morgan and Henrion 1990), such a propagated uncertainty will tend to become normally distributed provided that the input uncertainties are independent.

6 Conclusions

This paper has carried out the “exercise” that was left over by Heijungs and Suh (2002) and related work on deriving the complete set of sensitivity coefficients for matrix-based LCA.

Fig. 2 Screenshot of CMLCA showing the results of an analytical uncertainty analysis consisting of a number of characteristic parameters for the uncertainty of an output result



7 Recommendations and perspectives

For some, the formulas in Tables 4, 5 and 6 are intimidating and may offer little insight. But they are straightforward to implement in computer code. Once implemented, doing an uncertainty analysis is as easy as doing a Monte Carlo analysis: click a button and wait for the results. Likewise, doing a key issue analysis is as easy as doing a classical contribution analysis.

We hope that the availability of these equations will stimulate developers of software, commercial or not, to implement analytical, Taylor series-based approaches toward uncertainty and sensitivity analysis. In particular for the key issue analysis, no good Monte Carlo approach is available, and the analytical solution using Table 6 provides an extremely powerful way of reducing the uncertainties in LCA. For perturbation and uncertainty analysis, numerical approaches can be used as well, but

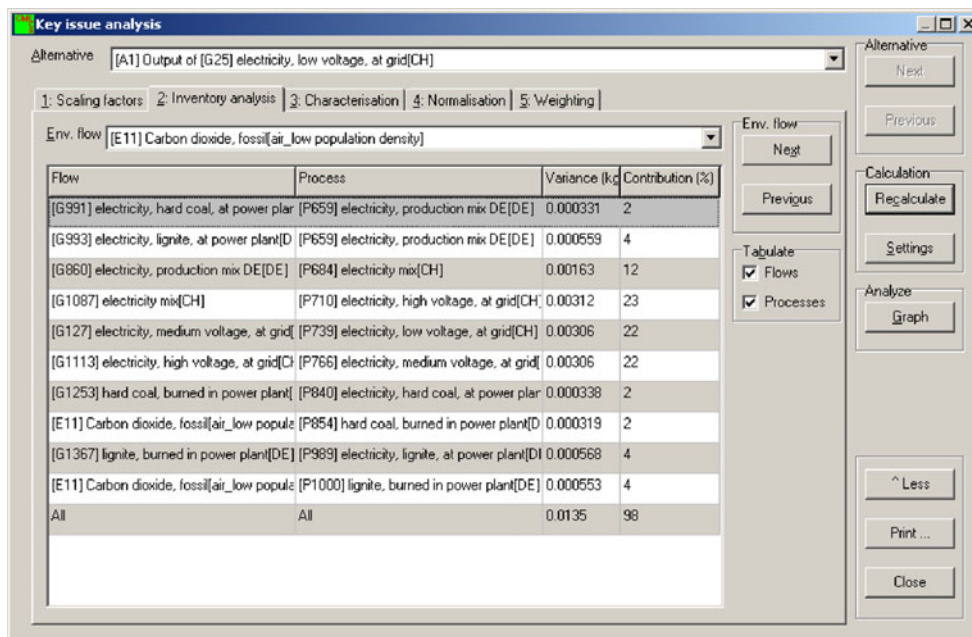


Fig. 3 Screenshot of CMLCA showing the results of a key issue analysis, the decomposition of the uncertainty of an output result in terms of contributions by the uncertainties of the input data. Contributions below 1% are not shown

these are extremely time-consuming for large LCA systems.

As noted above, Monte Carlo analyses for large LCA systems may be unfeasible. This paper develops sensitivity coefficients that serve to derive the formulas for analytical error propagation based on first (or higher)-order Taylor series approximation. This has other limitations, for instance, relating to a more restricted range of validity. Therefore, we welcome the simultaneous development of more sophisticated sampling methods such as Latin hypercube sampling and response surface methods, as well as the development of more efficient algorithms such as those based on a power series expansion. Yet, even when this would overtake the analytical error propagation, we still see a role for the sensitivity coefficients in perturbation analysis and in key issue analysis.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Bevington PR, Robinson DK (1994) Data reduction and error analysis for the physical sciences. McGraw-Hill, New York
- Ciroth A, Fleischer G, Steinbach J (2004) Uncertainty calculation in life cycle assessments: a combined model of simulation and approximation. *Int J Life Cycle Assess* 9(4):216–226
- Heijungs R, Guinée JB, Huppes G, Lankreijer RM, Udo de Haes HA, Wegener Sleeswijk A, Ansems AMM, Eggels PG, van Duin R, de Goede HP (1992) Environmental life cycle assessment of products. Backgrounds—October 1992. CML, Leiden
- Heijungs R (1994) A generic method for the identification of options for cleaner products. *Ecol Econ* 10(1):69–81
- Heijungs R (1996) Identification of key issues for further investigation in improving the reliability of life-cycle assessments. *J Clean Prod* 4(3–4):159–166
- Heijungs R, Frischknecht R (1998) A special view on the nature of the allocation problem. *Int J Life Cycle Assess* 3(6):321–332
- Heijungs R, Kleijn R (2001) Numerical approaches to life-cycle interpretation: five examples. *Int J Life Cycle Assess* 6(3):141–148
- Heijungs R, Suh S (2002) The computational structure of life cycle assessment. Kluwer, Dordrecht
- Heijungs R, Frischknecht R (2005) Representing statistical distributions for uncertain parameters in LCA. Relationships between mathematical forms, their representation in EcoSpold, and their representation in CMLCA. *Int J Life Cycle Assess* 10(4):248–254
- Heijungs R, Suh S, Kleijn R (2005) Numerical approaches to life cycle interpretation. The case of the Ecoinvent'96 database. *Int J Life Cycle Assess* 10(2):103–112
- Heijungs R, Suh S (2006) Reformulation of matrix-based LCI: from product balance to process balance. *Int J Life Cycle Assess* 14:47–51
- Heijungs R, de Koning A, Suh S, Huppes G (2006) Toward an information tool for integrated product policy: requirements for data and computation. *J Ind Ecol* 10(3):147–158
- Hong J, Rosenbaum R, Jolliet O (2008) Analysis of uncertainty propagation in life cycle inventory and impact assessment: application to an aluminum front panel. SETAC Europe 18th Annual Meeting, Warsaw, p 326
- Lloyd SM, Ries R (2007) Characterizing, propagating, and analyzing uncertainty in life-cycle assessment. *J Ind Ecol* 11(1):161–179
- Möller F-J. (1992) Ökobilanzen erstellen und anwenden. Entwicklung eines Untersuchungsmodells für die umweltverträglichkeit von Verpackungen. München: Ecobalance Applied Research
- Morgan MG, Henrion M (1990) Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge University Press, Cambridge
- Peters G (2007) Efficient algorithms for life cycle assessment, input–output analysis, and Monte-Carlo analysis. *Int J Life Cycle Assess* 12(6):373–380
- Sakai S, Yokoyama K (2002) Formulation of sensitivity analysis in life cycle assessment using a perturbation method. *Clean Technology and Environmental Policy* 4:72–78
- Saltelli A, Chan K, Scott EM (2000) Sensitivity analysis. Wiley, Chichester
- Sherman J, Morrison WJ (1950) Adjustment of an inverse matrix corresponding to a change in one element of a given matrix. *Ann Math Stat* 21(1):124–127
- Suh S, Huppes G (2005) Techniques for life cycle inventory of a product. *J Clean Prod* 13:687–697
- Suh S, Heijungs R (2007) Power series expansion and structural analysis for life cycle assessment. *Int J Life Cycle Assess* 12(6):381–390
- Tan RR (2008) Using fuzzy numbers to propagate uncertainty in matrix-based LCI. *Int J Life Cycle Assess* 13(7):585–592