



# Identification of key issues for further investigation in improving the reliability of life-cycle assessments\*

Reinout Heijungs

Centre of Environmental Science, Leiden University, P.O. Box 9518, 2300 RA Leiden, The Netherlands

Received 13 October 1995; revised 10 October 1996; accepted 14 November 1996

The iterative nature of life-cycle assessment (LCA) means that more details are looked for until a certain level of reliability has been achieved. This paper is concerned with the identification of key issues for further investigation in such an iterative procedure. Key issues in this context are defined as those aspects of an LCA which need more detailed research to arrive at a solid conclusion. The main concept in the context of finding key issues is the study of the propagation of uncertainties in underlying data. The structured procedure of LCA can be described in mathematical terms, so that standard mathematical techniques for the study of the propagation of uncertainties can be employed. The influence of uncertainties in input data on uncertainties in output data can be calculated, and the main source of the resulting uncertainties can be identified. The result of the analysis is a list of prioritized key issues for more detailed research and more accurate data. © 1997 Elsevier Science Ltd

**Keywords:** life-cycle assessment; screening; uncertainty analysis; statistics

## Introduction

Many textbooks define life-cycle assessment as an iterative procedure to quantify and interpret the environmental repercussions of a product system from the cradle to the grave related to a functional unit of service. The word 'iterative' is explained by mere reference to the fact that a quick scan is performed before turning to a subsequent refined or even detailed analysis. Usually, guidelines for deciding whether the level of detail suffices, and, if not, at which points to improve on data quality, lack.

The SETAC-Europe Working Group on Screening and Streamlining has identified one of its main tasks as providing a number of initial answers to this problem. It has defined life-cycle screening as a procedure to identify key issues for investigation in a more detailed analysis<sup>1</sup>. It is therefore important to have operational methods and criteria which tell the LCA-

practitioner (or the LCA-commissioner) where to invest in further research with high priority.

The term 'key issues' also has another meaning: it sometimes means those areas where product and process improvement leads to the highest environmental gain. Obviously, the two meanings are related. Knowledge of areas where a small change has large consequences ('hot spots') serves both functions:

1. these areas represent highly sensitive parameters, in which a small deviation has large influences, and which must be accurately known before drawing conclusions;
2. these areas represent highly sensitive parameters, in which a small change has large influences, and which might be affected by alternative product or process design.

Observe that the words 'deviation' and 'change' occur in these lines. It appears that a central element in the procedure of identifying key issues is the study of the propagation of unintentional deviations or intentional changes.

A third meaning of the term 'key issue' refers to the life-cycle stages or emissions/extractions that contribute heavily to the total environmental impacts of a product system. It should be noticed that this meaning is not implied to be covered under 'key issue' in the present paper, because those 'hot spots' are

\*A first version of this paper was prepared for the SETAC-Europe Working Group on Screening and Streamlining under the title *Screening, key issues and sensitivity analyses*, dated 25 January 1995. This version contained the same ideas, but consisted essentially of only the first three sections. Comments from the Working Group, and in particular from Evert Nieuwlaar (Utrecht University), Elmar Heinzele (ETH Zürich) and Nico van den Berg (Leiden University), are acknowledged, as is the anonymous referee of this journal.

interesting in themselves, and not so much for a more detailed iteration of the LCA process in the context of life-cycle screening. The distinction between key issues as areas to concentrate on in a more detailed LCI (life-cycle inventory), and key issues as areas that make up a large contribution to the total burden is neglected in most literature,<sup>1,2</sup> whereas it is, according to this paper, of seminal importance.

In the present paper, the emphasis is on uncertainties, statistical issues, and life-cycle screening as a procedure to prioritize areas of further research. This means that the discussion will concentrate on unintentional deviations ('uncertainties', 'measurement errors'). Extension of the ideas developed to intentional changes ('improvements') is, although not discussed here, straightforward. Furthermore, the paper is restricted to screening on the basis of the inventory analysis, so its main emphasis is LCI rather than on full LCA. Extension to life-cycle impact analysis is trivial.

The next two sections give a general discussion on the propagation of uncertainties and the search for key issues *sensu stricto*. Next, a small example on how to find key issues in an LCI is given. The final section is devoted to discussions. Technical details are left to the appendix.

### Propagation of uncertainties

The study of the propagation of unintentional deviations is quite well known as 'uncertainty analysis', 'error analysis', or 'perturbation analysis'. It can be performed in various degrees of sophistication. Below, the general ideas will be illustrated using a simple geometric example.

Consider the following problem: given a rectangle of length ( $l$ ) 20 (neglecting units) and height ( $h$ ) 10, what is the area ( $a$ ), and how sensitive is this area for uncertainties in input data? Using the well-known formula  $a = l \times h$ , it is easily found that the area is 200. But now the sensitivity: how to determine it? Below, two approaches are described, one based on supplying extreme values, the other on statistical considerations.

#### Calculating extreme values

One approach is to look for lower and upper values of every parameter, and to combine these to find a lower and upper value of the area. Suppose that various measurements have been performed, and that values for the lowest and the highest outcomes have been established. For instance, assume that  $l_{\text{lower}} = 18$  and  $l_{\text{upper}} = 23$ , and that  $h_{\text{lower}} = 8$  and  $h_{\text{upper}} = 12$ . There are four combinations:  $l_{\text{lower}} \times h_{\text{lower}}$ ,  $l_{\text{lower}} \times h_{\text{upper}}$ ,  $l_{\text{upper}} \times h_{\text{lower}}$ , and  $l_{\text{upper}} \times h_{\text{upper}}$ . It is evident that the lower value of  $a$  will be achieved by taking the lower values of  $l$  and  $h$ , so  $a_{\text{lower}} = 144$ . Similarly,  $a_{\text{upper}} = 276$ .

It should be observed that we used intuition to know that  $a_{\text{lower}}$  would be produced by combining  $l_{\text{lower}}$  and  $h_{\text{lower}}$ , and not by, e.g., combining  $a_{\text{lower}}$  and  $h_{\text{upper}}$ . It

can easily be shown, however, that in less straightforward examples, e.g., in LCIs where feedback loops due to recycling occur, the lowest value of the result can not be predicted by intuition. This would imply that all combinations of upper and lower values must be tried in order to find the upper and lower values of the result. If there are 10 000 figures used as input data—a typical number for a mediocre LCI—the number of combinations is  $2^{10000}$ , a number which amounts to a 1 with 3000 zeroes. A hopeless task: even for a computer that performs  $10^9$  operations per second, this would mean a calculation time that is considerably longer than the estimated age of the universe!

An alternative is to use Monte Carlo simulations to combine sets of data that deviate randomly within the specified margins. The usefulness of this method hinges on the assumption that a 'sufficient' number of trials has been carried out. It is unclear whether doing so for, say, one hour satisfies the expectations. An interesting third alternative that was published during the review of this paper is presented by Chevalier and Le Teno<sup>3</sup>.

#### Using statistical methods

The second approach involves the establishment of margins of uncertainty (or margins of confidence). Under the assumption of a particular distribution of measurement values, certain rules for calculating absolute errors, standard deviations and variances apply.

Suppose that the length of the rectangle is  $l = 20 \pm 3$  and the height is  $h = 10 \pm 2$ . It is not difficult to establish a formula to express the margin of uncertainty of  $a$ :  $\Delta a = l\Delta h + h\Delta l$ . This results in  $a = 200 \pm 70$ . The 'real' lower and upper values with these margins of uncertainty are 136 and 276. This formula presents a first-order approximation; the validity is restricted to small changes. Nevertheless, it also often gives for larger changes a fairly good approximation.

Mathematicians have established a general formula to express the propagation of uncertainties<sup>4-6</sup>. Let a variable  $y$  depend on a number of variables  $x_1, x_2, \dots$ , and let the functional relationship be denoted by  $f$ :

$$y = f(x_1, x_2, \dots) \quad (1)$$

The propagation of absolute errors in the variables  $x_1, x_2, \dots$ , indicated by  $\Delta x_1, \Delta x_2, \dots$ , into the dependent variable  $y$  is then given by

$$\Delta y = \left| \frac{\partial f}{\partial x_1} \right| \Delta x_1 + \left| \frac{\partial f}{\partial x_2} \right| \Delta x_2 + \dots \quad (2)$$

This equation can be used to derive a similar equation for the propagation of the variance  $\text{var}(x_1), \text{var}(x_2), \dots$ :

$$\text{var}(y) = \left( \frac{\partial f}{\partial x_1} \right)^2 \text{var}(x_1) + \left( \frac{\partial f}{\partial x_2} \right)^2 \text{var}(x_2) + \dots \quad (3)$$

The connection with the standard deviations  $\sigma(x_1)$ ,  $\sigma(x_2)$ , ... can then be made by using the fact that the variance is the square of the standard deviation:

$$\sigma(y) = \sqrt{\left(\frac{\partial f}{\partial x_1}\right)^2 (\sigma(x_1))^2 + \left(\frac{\partial f}{\partial x_2}\right)^2 (\sigma(x_2))^2 + \dots} \quad (4)$$

We see that there is some choice with respect to the exact definition of uncertainty. For now, let it be sufficient to mention that the remaining text is based on the absolute error (equation (2)) while Rønning *et al.*<sup>7</sup> use standard deviations (equation (4)). The argument concerning the choice for this paper is postponed to the discussion.

A formal expression of the uncertainty of the result as a function of the uncertainty of the data is thus possible, but requires that the result is not expressed as a number but as a formula (expressed by  $f(x_1, x_2, \dots)$ ). Fortunately, this is possible in LCI. The details are worked out in one of the next sections as well as in two earlier publications<sup>8,9</sup>.

Thus we are able to provide a method which expresses the influence of uncertainties in input data on uncertainties in output data. It is thus possible to quantify the accuracy of an LCI result (the inventory table) in terms of the accuracy of the input data (the process data) used. This might provide a criterion for deciding whether more accurate data is needed. It is in principle possible to establish application-dependent criteria, e.g., for design-for-the-environment a relative error of 25% is tolerable, for ecolabelling it is 5%.

So far, a review of previous results. The new element of this paper is that it establishes a quantitative connection between the above estimation of the uncertainty of the LCI-result and the processes and emissions/extractions that contribute to this uncertainty, thereby enabling an operational and generic method for life-cycle screening. This is the topic of the next section.

### Finding key issues for a more detailed LCI

Suppose that an LCA-practitioner arrives at a margin of uncertainty of, say, 30%, and is not satisfied. Then, following the rules of life-cycle screening, he or she goes into a more sophisticated LCI and iterates back to data collection to improve on the key issues. There is thus a need to know which uncertainty of input data is mainly responsible for the large margin of error. Of course, other data sources could be improved as well, reducing the 30% to, say, 28%. What one is hoping for is an indication of which uncertainty should be removed in order to reduce the 30% to a mere 5%. It is therefore important to establish facts like: 'The emission of SO<sub>2</sub> is 23 kg with an uncertainty of 12 kg; 7 kg of this uncertainty is due to the uncertainty in the electricity consumption'. This section will concentrate on this question.

It is important to observe that highly uncertain input figures sometimes have a negligible influence on output uncertainty. An uncertainty of 100% in the CO emis-

sion from the factory that produces plastic wrappers in which pesticides, that are used for food production, are packed, will hardly show up in the uncertainty of the inventory table of the food product. On the other hand, a not too large uncertainty of 5% in the amount of pesticide that goes to the groundwater can make the final result highly uncertain.

It is therefore necessary to distinguish data that is uncertain from data that contributes grossly to the final result and for which the final result is quite sensitive (Figure 1). Life-cycle screening should categorize data accordingly. Here we also see the two distinct meanings of the term 'key issues': the life-cycle stages or emissions/extractions that make a large contribution to the total are in the rightmost column of Figure 1. Some of these will not be a key issue according to the present paper. On the other hand, the other column may also contain key issues in the present sense.

Going back to the example of  $l = 20 \pm 3$ ,  $h = 10 \pm 2$  and  $a = 200 \pm 70$ , we can find out that the uncertainty  $\Delta a$  is composed of two uncertainties:

in formula:  $\Delta a = h \times \Delta l + l \times \Delta h$

in numbers:  $70 = 10 \times 3 + 20 \times 2$

or more simply:  $70 = 30 + 40$

and relatively:  $0.35 = 0.15 + 0.20$

or as percentages:  $35\% = 15\% + 20\%$  (5)

The total uncertainty in  $a$  is 35%. The uncertainty in  $l$  contributes 15%; that in  $h$  contributes 20%. It is thus more efficient to reduce the uncertainty in  $h$  than the uncertainty in  $l$ , although the difference in efficiency is not large in this example.

Returning to the more general formulation in terms of  $y = f(x_1, x_2, \dots)$ , we need to know the total level of uncertainty

$$\frac{\Delta y}{|y|} \quad (6)$$

together with the parts of which it is composed

$$\left| \frac{\partial f}{\partial x_1} \right| \frac{\Delta x_1}{|y|}, \left| \frac{\partial f}{\partial x_2} \right| \frac{\Delta x_2}{|y|}, \dots \quad (7)$$

The former expression corresponds to the total relative uncertainty, like the 35% in the example above.

uncertainty	↑ <i>high</i>	perhaps a key issue	key issue
	↓ <i>low</i>	not a key issue	perhaps a key issue
		<i>low</i>	<i>high</i>
		←	→
		contribution	

Figure 1 Data that is certain and hardly contributes to the inventory table ('not a key issue') must be separated from data that is uncertain and makes quite some contribution ('key issue'). Anything in between must be considered carefully ('perhaps a key issue')

Table 1 Process data for the example; the functional unit is using 10 sandwich packages<sup>9</sup>

Commodity	Process				Product system
	Production of electricity	Production of aluminium	Production of aluminium foil	Usage of aluminium foil	
MJ electricity	1	-50	-1	0	0
kg aluminium	-0.01	1	-1	0	0
kg aluminium foil	0	0	1	-1	0
100 sandwich packages	0	0	0	1	0.1
kg bauxite	0	-5	0	0	?
kg crude oil	-0.5	0	0	0	?
kg CO <sub>2</sub>	3	0	0	0	?
kg solid waste	2	10	0	1	?

The latter series of expressions gives the contributions of individual uncertainties, like the 20% and the 15% above.

It is possible to rearrange this series of numbers in decreasing order, so as to obtain a ranking of the parameters which have the highest influence on the uncertainty of the result. This list can conveniently be called 'list of key issues for further investigation'. In the example above, the first key issue would be *h*, as its current uncertainty means that the result can never be known better than up to 20% certainty.

In principle, this is the entire idea of using statistical methods for the identification of key issues. The appendix is devoted to the remaining problem of the exact form of *f* in  $y = f(x_1, x_2, \dots)$ , and to the sensitivity analysis that can be derived from that equation by means of the partial derivatives. The next section is simpler; the reader is encouraged to go through this example of the LCI of sandwich packagings.

### A simple example

A small numerical example is presented below. The example builds on the product system introduced by Heijungs<sup>9</sup>. Some of the words and symbols (such as 'technology matrix', **G**) are defined in the appendix. It is expected that the reader will be able to understand what is meant without going through the appendix.

The table of process data is reproduced in Table 1. This gives rise to the following technology matrix **G** and intervention matrix **H**:

$$\mathbf{G} = \begin{pmatrix} 1 & -50 & -1 & 0 \\ -0.01 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix};$$

$$\mathbf{H} = \begin{pmatrix} 0 & -5 & 0 & 0 \\ -0.5 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 2 & 10 & 0 & 1 \end{pmatrix} \quad (8)$$

The external supply vector **u** is

$$\mathbf{u} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0.1 \end{pmatrix} \quad (9)$$

hence, the functional unit is the use of 10 sandwich packages. Using the formalism described in the appendix, it is easily found that the inventory table is given by

$$\mathbf{y} = \begin{pmatrix} -1.01 \\ -5.1 \\ 30.6 \\ 22.52 \end{pmatrix} \quad (10)$$

or, in words, extraction of 1.01 kg bauxite, extraction of 5.1 kg crude oil, emission of 30.6 kg CO<sub>2</sub>, and generation of 22.52 kg solid waste (which is considered in this example as an ecologic commodity).

We introduce margins of uncertainty as

$$\Delta \mathbf{G} = \begin{pmatrix} 0.01 & 0.5 & 0.01 & 0 \\ 0.0001 & 0.01 & 0.01 & 0 \\ 0 & 0 & 0.01 & 0.01 \\ 0 & 0 & 0 & 0.01 \end{pmatrix} \text{ and}$$

$$\Delta \mathbf{H} = \begin{pmatrix} 0 & 0.05 & 0 & 0 \\ 0.005 & 0 & 0 & 0 \\ 0.03 & 0 & 0 & 0 \\ 0.02 & 0.1 & 0 & 0.01 \end{pmatrix} \quad (11)$$

i.e. we assume all process data to have a rather small uncertainty of 1%. Furthermore, we will assume that there is no uncertainty in the functional unit, i.e.  $\Delta \mathbf{u} = \mathbf{0}$ . We find that the inventory table **y** has margins of uncertainty of

**Table 2** List of key issues for further investigation, aiming at a reduction of the uncertainty in the extraction of crude oil, assuming that all process data of *Table 1* have an inaccuracy of 1%

Parameter	Contribution to uncertainty (total = 10.4%)
$g_{22}$ (output of aluminium by aluminium production process)	2.2%
$g_{11}$ (output of electricity by electricity production process)	2.0%
$g_{12}$ (input of electricity by aluminium production process)	2.0%
$g_{33}$ (output of aluminium foil by aluminium foil production process)	1.1%
$g_{44}$ (output of sandwich packages by aluminium foil usage process)	1.1%
$g_{21}$ (input of aluminium by electricity production process)	1.0%
$h_{21}$ (input of crude oil by electricity production process)	1.0%

$$\Delta y = \begin{pmatrix} 0.32 \\ 0.53 \\ 31.9 \\ 2.92 \end{pmatrix} \quad (12)$$

Concentrating on, say  $y_2$  (extraction of crude oil), a list that indicates which inaccuracies contribute most to the total inaccuracy of  $0.53/5.1 = 10.4\%$  can be compiled (see *Table 2*).

It is easily seen that the uncertainties in the data of aluminium production (process 2) contributes 4.2% to the total uncertainty of 10.4%. A similar remark applies to the data on electricity production (process 1). A first research priority could quite efficiently be aimed at reducing the uncertainties of these process data. The uncertainties of the data of the other two processes are less important. Observe that it can be helpful to aggregate the percentages of contributions per process (see *Table 3*).

A combination of these tables is perhaps the most useful form (*Table 4*). This table shows that a first priority is to consult the manager of the aluminium producer, and ask him or her in particular about their aluminium output and their electricity input. One could envisage facilitating the screening procedure by adding

**Table 3** List of processes that need further investigation; assumptions are summarized in the caption of *Table 2*

Process	Contribution to uncertainty (total = 10.4%)
2 (aluminium production)	4.2%
1 (electricity production)	4.0%
3 (aluminium foil production)	1.1%
4 (aluminium foil usage)	1.1%

**Table 4** Combined list of processes and parameters that need further investigation; assumptions are summarized in the caption of *Table 2*

Process/Parameter	Contribution to uncertainty (total = 10.4%)
2 (aluminium production)	4.2%
$g_{22}$ (output of aluminium)	2.2%
$g_{12}$ (input of electricity)	2.0%
1 (electricity production)	4.0%
$g_{11}$ (production of electricity)	2.0%
$g_{21}$ (input of aluminium)	1.0%
$h_{21}$ (input of crude oil)	1.0%
3 (aluminium foil production)	1.1%
$g_{33}$ (output of aluminium foil)	1.1%
4 (aluminium foil usage)	1.1%
$g_{44}$ (output of sandwich packages)	1.1%

a third column with the telephone number of the process engineer that provided the data.

## Discussion

This paper has presented an operational and generic method for the identification of key issues for further analysis in a subsequent more detailed LCI. It is operational in the sense that the paper provides all necessary equations to perform the screening procedure, and it is generic in the sense that it can be applied regardless of system boundaries and functional unit.

An important point to note is that, although the mathematics are complicated, the fact that this necessarily involves automatized computation means that the method is very easy to operate in practice. One supplies, with the usual process data on inputs and outputs, estimated margins of error. By pressing one button and after waiting for some time, the computer gives the inventory table or environmental profile, their estimated margins of error, and lists of key issues, similar to *Table 2* and *Table 3* or, alternatively, *Table 4*.

Although the method itself is operational, we often lack knowledge of data on the uncertainties of the input data. This makes the application of the screening method described here problematic. A straightforward solution is the one applied in the simple example: give every figure an identical margin of error, e.g. 5%, and use the subsequent analysis to identify key issues, without reference to actual margins of error. Evidently, this procedure can be applied only once, not in an iterative procedure. It is more a sensitivity analysis than a screening procedure. We must be aware that the absence of error data could discourage the collection of these data. Why should one try to figure them out for one process, if they are absent for the other processes? It is clear that it is crucial to escape such negative spirals. The procedure clearly shows that it is better to have a rough estimate of the margins of uncertainty than to simply leave them out. Weidema & Wesnæs<sup>10</sup> describe some heuristical rules to estimate uncertainty figures.

A last point of discussion is related to the form of the propagation formula (the linear one of this paper, the Pythagorean one of Rønning *et al.*<sup>7</sup>, or still another one) in relation to the choice for representing absolute errors, standard deviations, or any other statistical measure of uncertainty. This paper is not so much devoted to representing the reliability of an LCI in the most accurate way, as emphasising the need to look for key issues. As equation (5) and Tables 2, 3 and 4 indicate, an important aspect in finding key issues is linearity. Suppose that the Pythagorean propagation rule (equation (4)) for standard deviations was chosen in the case of the rectangle. Assuming the errors now representing the standard deviations, equation (5) would change into

$$\begin{aligned} \text{in formula: } (\sigma(a))^2 &= h^2 \times (\sigma(l))^2 + l^2 \times (\sigma(h))^2 \\ \text{in numbers: } 2500 &= 100 \times 9 + 400 \times 4 \\ \text{or more simply: } 2500 &= 900 + 1600 \end{aligned} \quad (13)$$

$$\text{and taking the root: } 50 = 30 + 40$$

$$\text{and relatively: } 0.25 = 0.15 + 0.20$$

$$\text{or as percentages: } 25\% = 15\% + 20\%$$

with the serious implication that the total uncertainty of 25% is composed of two parts, 15% and 20%, that together do not make the total. In fact, it would disable the search for key issues, as it inhibits a sentence like 'The emission of SO<sub>2</sub> is 23 kg with an uncertainty of 12 kg; 7 kg of this uncertainty is due to the uncertainty in the electricity consumption'. Altogether, this makes the use of the linear propagation rule far more suitable in the context of finding key issues. The question of statistical soundness, however, could turn things upside down, in the sense that a careful statistical treatment would perhaps not allow the procedure to find key issues.

## References, also for the Appendix

- 1 Christiansen, K. (Ed.): SETAC Europe LCA Screening and Streamlining Working Group. Draft Report, 6 September 1996. Final report forthcoming.
- 2 Lindfors, L. G., Christiansen, K., Hoffman, L., Virtanen, Y., Juntilla, V., Hansen, O.J., Rønning, A., Ekvall, T. and Finnveden, G., in *LCA-Nordic. Technical reports No 1-9*, eds Lindfors, L.G., K. Christiansen, L. Hoffman, Y. Virtanen, V. Juntilla, A. Leskinen, O.J. Hansen, A. Ronning, T. Ekvall and G. Finnveden. Nordic Council of Ministers, Copenhagen, 1995.
- 3 Chevalier, J.L. and Le Téo, J.F., *International Journal of Life Cycle Assessment*, 1996, 1(2), 90-96.
- 4 Bevington, P.R. and Robinson, D.K., *Data reduction and error analysis for the physical sciences*. McGraw-Hill, New York, 1994, p. 328.
- 5 Hettelingh, J.P., *Uncertainty in modelling regional environmental systems. The generalization of a watershed acidification model for predicting broad scale effects*. Vrije Universiteit, Amsterdam, 1989, p. 224.
- 6 Abramowitz, M., in *Handbook of mathematical functions. With formulas, graphs, and mathematical tables*, eds Abramowitz, M. and I. Stegun. Dover Publications, New York, 1965, p. 1046.
- 7 Rønning, A., Hansen, O.J., Møller, H., Gade, A.L. and Haug, U.C., *Life cycle assessment of two paint products*. STØ, Frederikstad, 1993, p. 52.

- 8 Heijungs, R., Guinée, J.B., Huppes, G., Lankreijer, R.M., Udo de Haes, H.A., Wegener Sleswijk, A., Ansems, A.M.M., Eggels, P.G., van Duin, R. and de Goede, H.P., *Environmental life cycle assessment of products. I: Guide—October 1992. II: Backgrounds—October 1992*. CML, Leiden, 1992, 96-130.
- 9 Heijungs, R., *Ecological Economics*, 10(1) (1994), 69-81.
- 10 Weidema, B.P. and Wesnæs, M.S., *Journal of Cleaner Production*, 1996, 4(3-4), 165. (This issue.)
- 11 Hondo, H. and Uchiyama, Y., in *Proceedings of international conference in ecobalance. Life cycle assessment for development of materials and technologies*, ed. Anonymous. The Society for Non-traditional Technology, Tokyo, 1994, pp. 143-148.
- 12 Suga, M., Ishikawa, M. and Yoshioka, K., in *Proceedings of international conference in ecobalance. Life cycle assessment for development of materials and technologies*, ed. Anonymous. The Society for Non-traditional Technology, Tokyo, 1994, pp. 201-209.
- 13 Lave, L.B., Cobas-Flores, E., Hendrickson, C.T. and McMichael, F.C., *Environmental Science & Technology* 29(9), 1995, 420-426.
- 14 Leontief, W.W., *The Review of Economic Statistics*, 18(3), 1936, 105-125.
- 15 Leontief, W.W., *The structure of American Economy, 1919-1939. An empirical application of equilibrium analysis*. Oxford University Press, NY, 1953, p. 264.
- 16 Leontief, W.W., *Review of Economics and Statistics*, 52, 1970, 262-271.
- 17 Rosenbluth, G., *Statistische Hefte*, 9(4), 1968, 255-268.
- 18 Victor, P.A., *Pollution. Economy and environment*. George Allen & Unwin Ltd., London, 1972, p. 247.
- 19 Konijn, P.J.A., *The make and use of commodities by industries. On the compilation of input-output data from the national accounts*. Faculteit Bestuurskunde, Universiteit Twente, Enschede, 1994, p. 281.
- 20 Van Engelenburg, B.C.W., Van Rossum, T.F.M., Blok, K. and Vringer, K., *Energy Policy*, 1994, 22(8), 648-656.
- 21 Dohnomae, H., Tani, M., Shibata, K. and Okumura, N., in *Proceedings of international conference on ecobalance. Life cycle assessment for development of materials and technologies*, ed. Anonymous. The Society of Non-traditional Technology, Tokyo, 1994, pp. 137-142.
- 22 Koopmans, T.C., *Activity Analysis of Production and Allocation*, ed. T.C. Koopmans. John Wiley, NY, 1951, pp. 33-97.
- 23 von Neumann, J., *Review of Economic Studies*, 13, 1945-1946, 1-9.
- 24 Georgescu-Roegen, N., in *Activity analysis of production and allocation*, ed. T.C. Koopmans. John Wiley, NY, 1951, pp. 98-115.
- 25 Saxton, J.C. and Ayres, R.U., in *Mineral materials modelling. A state-of-the-art review*, ed. W.A. Vogely. John Hopkins University Press, Baltimore, 1975, pp. 178-244.
- 26 Georgescu-Roegen, N., *The entropy law and the economic process*. Harvard University Press, Cambridge, 1971, p. 457.
- 27 van Rijkeghem, W., *Review of Economics and Statistics*, 49, 1967, 607-608.
- 28 Heijungs, R., *Economic drama and the environmental stage. The derivation of a number of tools for environmental decision-support from a unified epistemological principle*. Rijksuniversiteit Leiden, Leiden, to appear in 1997.
- 29 Heijungs, R. and Frischknecht, R. *On the nature of the allocation problem*. In preparation.
- 30 Balestra, P., *La dérivation matricielle. Technique et résultats pour économistes*. Sirey, Paris, 1976, p. 115.
- 31 Stewart, G.W., *Introduction to matrix computations*. Academic Press, Orlando, 1973, p. 441.

## Appendix: what is *f*?

It was stated that the computational problem in LCI could be expressed in an explicit formula like  $y = f(x_1, x_2, \dots)$ . From that equation, the partial derivatives  $\partial f / \partial x_1, \partial f / \partial x_2, \dots$  needed to be calculated in order to express coefficients for the sensitivity analysis. This appendix is devoted to a discussion on the form of *f*. The description largely follows a previous analysis<sup>8</sup>,

but gives more credit to older literature on interindustry analysis or activity analysis. The reader will notice that some of the literature referred to is quite old. A general observation is that LCI is an extension of a quite well developed form of economic analysis, on which a rich literature is available. The reader should also observe that certain quantities have a symbol that differs from Heijungs<sup>9</sup>.

Life-cycle inventory analysis is sometimes regarded or performed as a form of input-output analysis<sup>11,12</sup>. It is also sometimes proposed to be replaced by input-output analysis<sup>13</sup>. These ideas neglect the essential point that LCI is set up as a commodity-by-industry analysis, and not as a commodity-by-commodity analysis or an industry-by-industry analysis, as is the usual input-output approach<sup>14,15</sup> even when extended to include pollution<sup>16</sup>. In fact, we must embrace the point that LCI is a commodity-by-industry analysis, as this form of analysis is generally seen as superior to other forms of interindustry analysis<sup>17-19</sup>. Occasional proposals to supplement LCI by data from input-output analysis<sup>20,21</sup> on the other hand could be regarded as a solution to the data intensity of LCI.

Following Koopmans<sup>22</sup>, an economic process can be represented by a column vector  $\mathbf{p}$ , in which the number in the  $i$ th row represents the input or output of the  $i$ th commodity. Negative numbers denote inputs, positive numbers denote outputs.

Following Victor<sup>18</sup>, a distinction between economic commodities (interindustry flows, such as steel, electricity, products, waste-to-be-processed) and ecologic (or environmental) commodities (natural resources, emissions) can be made. Making this division corresponds to setting the system boundaries: if a dump site is included in the product system as an economic process, waste-to-be-dumped is treated as an economic commodity. *Mutatis mutandis*, if the dump site is outside the system, waste-to-be-dumped is an ecologic commodity. Thus, the process vector  $\mathbf{p}$  is split into two parts: an economic part  $\mathbf{a}$  and an ecologic part  $\mathbf{b}$ :

$$\mathbf{p} = \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix} \quad (14)$$

It is furthermore assumed<sup>22-24</sup> that the process proceeds linearly through time, and that the activity level<sup>25</sup> or—better in the present context—duration<sup>26</sup> of a process can be seen as a proportionality factor:

$$\mathbf{p} = \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix} = \begin{pmatrix} \mathbf{g} \\ \mathbf{h} \end{pmatrix} t \quad (15)$$

where  $\mathbf{g}$  could be interpreted as a vector of economic coefficients,  $\mathbf{h}$  as a vector of ecologic coefficients, and  $t$  is the duration (operating time) of the process. Similar to the technical coefficients in input-output analysis, these coefficients are assumed to be independent of production volumes, i.e. there are no economies of scale; they reflect solely the state of technology.

A number of different processes can be combined into a process tree. Processes will be indexed by a subscript  $j$ : the process vector  $(\mathbf{p})_j$  contains elements  $p_{ij}$  which denote the input (negative) or output (positive) of the  $i$ th commodity by the  $j$ th process. These coefficients  $p_{ij}$  can be put into a process matrix  $\mathbf{P}$ , which consists of a matrix of economic coefficients  $\mathbf{A}$  and a matrix of ecologic coefficients  $\mathbf{B}$ :

$$\mathbf{P} = \begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix} \quad (16)$$

Introducing  $t_j$  as the operating time of the  $j$ th process, and putting these  $t_j$ s into the vector  $\mathbf{t}$ , we can identify the technology matrix<sup>22,27</sup>  $\mathbf{G}$ , containing the elements  $g_{ij}$ :

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots \\ a_{21} & a_{22} & \dots \\ \dots & \dots & \dots \end{pmatrix} = \begin{pmatrix} g_{11}t_1 & g_{12}t_2 & \dots \\ g_{21}t_1 & g_{22}t_2 & \dots \\ \dots & \dots & \dots \end{pmatrix} = \mathbf{G} \cdot \mathbf{t} \quad (17)$$

and the intervention matrix  $\mathbf{H}$ <sup>28</sup>, containing the elements  $h_{ij}$ :

$$\mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & \dots \\ b_{21} & b_{22} & \dots \\ \dots & \dots & \dots \end{pmatrix} = \begin{pmatrix} h_{11}t_1 & h_{12}t_2 & \dots \\ h_{21}t_1 & h_{22}t_2 & \dots \\ \dots & \dots & \dots \end{pmatrix} = \mathbf{H} \cdot \mathbf{t} \quad (18)$$

Aggregation of all inputs and outputs of one commodity over all processes in the process tree

$$\sum_j p_{ij}$$

yields the external flow of that commodity. This quantity represents the product system's external flow of commodity  $i$ ; for economic commodities ( $u_i$ ) this corresponds to the functional unit, for ecologic commodities ( $y_i$ ) this corresponds to the associated inventory table. In vector notation:

$$\begin{pmatrix} \mathbf{u} \\ \mathbf{y} \end{pmatrix} = \begin{pmatrix} \mathbf{G} \cdot \mathbf{t} \\ \mathbf{H} \cdot \mathbf{t} \end{pmatrix} \quad (19)$$

In the goal definition of an LCA, a functional unit  $\mathbf{u}$  is chosen. The LCI ends up with an inventory table  $\mathbf{y}$ . The relation between the two can now be expressed by the simultaneous set

$$\begin{cases} \mathbf{u} = \mathbf{G} \cdot \mathbf{t} \\ \mathbf{y} = \mathbf{H} \cdot \mathbf{t} \end{cases} \quad (20)$$

The vector  $\mathbf{t}$  containing the operating times of all processes in the system can be eliminated by inverting the technology matrix  $\mathbf{G}$ :

$$\mathbf{t} = \mathbf{G}^{-1} \cdot \mathbf{u} \quad (21)$$

and substituting this into the equation for  $\mathbf{y}$ :

$$\mathbf{y} = \mathbf{H} \cdot \mathbf{G}^{-1} \cdot \mathbf{u} \quad (22)$$

This presumes that  $\mathbf{G}$  indeed is invertible, which requires, amongst others, that it is square. I will not go in this paper into the details of the validity of this assumption, which touches at the heart of the allocation problem.<sup>29</sup>

Equation (22) is the fundamental equation of LCI: it expresses the inventory table  $\mathbf{y}$  as the product of the intervention matrix  $\mathbf{H}$ , the inverse of the technology matrix  $\mathbf{G}$ , and the functional unit  $\mathbf{u}$ . This is hence the explicit algebraic form which is being sought for: the form that we were looking for was

$$y = f(x_1, x_2, \dots) \quad (23)$$

It turns out to be a vector-valued function  $\mathbf{f}$ :

$$\mathbf{y} = \mathbf{f}(\mathbf{H}, \mathbf{G}, \mathbf{u}) \quad (24)$$

in other words, there is not one  $y$  but a number of them, and there are three different types of  $x$ : corresponding to the functional unit  $\mathbf{u}$ , corresponding to the technology matrix  $\mathbf{G}$ , and corresponding to the intervention matrix  $\mathbf{H}$ .

The next stage of the procedure for sensitivity analysis is to find  $\partial f / \partial x_i$ , etc. What we need is

$$\left| \frac{\partial f_k}{\partial h_{ij}} \right| \frac{\Delta h_{ij}}{|y_k|}, \left| \frac{\partial f_k}{\partial g_{ij}} \right| \frac{\Delta g_{ij}}{|y_k|}, \left| \frac{\partial f_k}{\partial u_i} \right| \frac{\Delta u_i}{|y_k|} \quad (\forall i, \forall j, \forall k) \quad (25)$$

as well as the total relative uncertainty

$$\frac{\Delta y_k}{|y_k|} = \sum_j \sum_i \left| \frac{\partial f_k}{\partial h_{ij}} \right| \frac{\Delta h_{ij}}{|y_k|} + \sum_j \sum_i \left| \frac{\partial f_k}{\partial g_{ij}} \right| \frac{\Delta g_{ij}}{|y_k|}$$

$$+ \sum_i \left| \frac{\partial f_k}{\partial u_i} \right| \frac{\Delta u_i}{|y_k|} \quad (\forall k) \quad (26)$$

The partial derivatives have to be worked out; most of them are straightforward, but the presence of the  $\mathbf{G}^{-1}$  in  $f_k$  complicates matters somewhat, see Heijungs<sup>9</sup> and Balestra<sup>30</sup> for elaborations. The final result can be shown to be

$$\Delta y_k = \sum_j |t_j| \Delta h_{kj} + \sum_j |h_{kj}| \Delta t_j$$

where

$$t_j = \frac{\det(\mathbf{G}_j(\mathbf{u}))}{\det(\mathbf{G})}$$

and

$$\Delta t_j = \sum_l \sum_m \left| \frac{-t_j \text{cof}(\mathbf{G})_{lm}}{\det(\mathbf{G})} \right| \Delta g_{lm} + \sum_l \sum_{m \neq j} \left| \frac{\text{cof}(\mathbf{G}_j(\mathbf{u}))_{lm}}{\det(\mathbf{G})} \right| \Delta g_{lm} + \sum_l \left| \frac{\text{cof}(\mathbf{G})_{lj}}{\det(\mathbf{G})} \right| \Delta u_j \quad (27)$$

where a number of notational conventions which were defined by Stewart<sup>31</sup> have been introduced. To be specific,  $\mathbf{G}_{lm}$  denotes the matrix  $\mathbf{G}$  in which the  $l$ th row and the  $m$ th column have been deleted,  $\text{cof}(\mathbf{G})_{lm}$  is the cofactor of  $\mathbf{G}$ , which is defined as

$$\text{cof}(\mathbf{G})_{lm} = (-1)^{l+m} \det(\mathbf{G})_{lm} \quad (28)$$

and  $\mathbf{G}_j(\mathbf{u})$  is defined as the matrix  $\mathbf{G}$  in which the  $j$ th column has been replaced by the vector  $\mathbf{u}$ .