Jul 1st, 12:00 AM

A Review of Approaches to Treat Uncertainty in LCA

Reinout Heijungs
Mark A. J. Huijbregts

Follow this and additional works at: https://scholarsarchive.byu.edu/iemssconference


This Event is brought to you for free and open access by the Civil and Environmental Engineering at BYU ScholarsArchive. It has been accepted for inclusion in International Congress on Environmental Modelling and Software by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.
A Review of Approaches to Treat Uncertainty in LCA

Reinout Heijungs\textsuperscript{a} and Mark A.J. Huijbregts\textsuperscript{b}

\textsuperscript{a} Institute of Environmental Sciences, Leiden University, Leiden, The Netherlands
e-mail: heijungs@cml.leidenuniv.nl

\textsuperscript{b} Department of Environmental Studies, University of Nijmegen, Nijmegen, The Netherlands

Abstract: It is important to know to what extent the outcome of an LCA is affected by various types of uncertainty, such as parameter, scenario and model uncertainty. These types may occur in the goal and scope definition, the inventory analysis and the impact assessment of an LCA. Information on the uncertainty of the model outcomes provides useful information to assess the reliability of LCA-based decisions and to guide future research towards reducing uncertainty. This paper reviews several approaches to treat different types of uncertainty in LCA. It will discuss the typology of uncertainty that may be encountered in LCA, the qualitative and quantitative techniques that are available to address these uncertainties, the inclusion of these techniques in LCA software tools, the (graphical) possibilities to show uncertainty in LCA outcomes, ways to simplify the uncertainty analysis, the inclusion of uncertainty analyses in case studies and (the difficulties in) the interpretation of uncertainty information.

Keywords: Uncertainty, Life cycle assessment; LCA

1. INTRODUCTION

If LCA is supposed to play a role in environmental decision-making, the quality of the decision-support should be made clear. It is natural that there is an interest of decision-makers and LCA-experts in the credibility of the results of LCA. In fact, it is amazing that this interest has not been natural since the development of LCA and the rise of its use. Although concerns about the quality of LCA-results have been raised at an early stage of LCA-development, assessment of this quality is still not a standard feature, and a systematic and comprehensive treatment is still lacking in most guidebooks, databases and software for LCA. Witness of this is the several guidebooks for LCA where the assessment of uncertainty is postponed and presented as some sort of additional feature. One still sees LCA case studies with bar charts showing that product A is 0.3\% better than product B, without any indication of the significance or robustness of this difference. A decision-maker then has to figure out whether or not the difference of 0.3\% is in any sense significant. The situation is even more complex due to the fact that writers of LCA-reports sometimes use the technical term “significant” when they have the intuitive term “large” in mind; in fact, some writers seem to think that “significant” is the scientific-sounding equivalent of “large”. The idea of statistical analysis is not always part of the standard vocabulary of practitioners and users of LCA.

There is, however, also good news. There have been quite a few initiatives and developments towards including uncertainty in LCA. Statistical uncertainty information is to an increasing extent percolating into methods, databases and software, and are increasingly being applied in case studies. Decision-makers increasingly recognize that uncertainties are important and should be made explicit.

The importance of including uncertainty in LCA has been long recognized. Already in 1992, a SETAC-workshop focussed exclusively on this topic (Fava et al., 1993). During these years, the discussion was largely restricted to acknowledging the possible prohibitive effects of uncertainty and the setting-up of schemes for data quality indicators for LCI data. Approaches towards analyzing the uncertainty in final results were published (Heijungs, 1992; 1994), but remained unused for two main reasons: lack of knowledge of input uncertainties, and lack of appropriate software. For almost a decade, the two lacks seemed to be trapped in a vicious circle: as long as there is no software that deals with uncertain data, there is no need to collect uncertainty information for the
data, and as long as there is no uncertainty information for the data, there is no need to develop software that deals with uncertain data. But the last few years, software and data providers are freeing themselves from this trap. Software that in one form or another supports Monte Carlo analysis is becoming standard, and one of the most widely used data sources, the Swiss ecoinvent, has started to include information on distribution and data quality indicators.

From the theoretical side, we mention the publication of a number of PhD theses in which uncertainty in LCA played a dominant role (Roï, 1998; Pohl, 1999; Huijbregts, 2001; Ciroth, 2001). Other important contributions are summarized by Heijungs & Suh (2002). Besides these structural developments, developments in case studies have shown some ingenuity in dealing with uncertainty issues. More and more case studies have used statistical methods to address uncertainties.

The increasing recognition of the role of uncertainty in LCA has also some darker sides. On the one hand, it may easily lead to pessimism or even cynicism. Results of LCA would be meaningless, as the uncertainties associated with these results would overshadow the results themselves. And carrying out LCA would become much more complicated, due to the additional data collection efforts and the more intricate calculations. Finally, interpretation of LCA-results would be more cumbersome, and involve a much more technical jargon: confidence intervals, significance levels, etc.

A practical problem with dealing with uncertainty is that the information is scattered and that terminology is confusingly non-standardized. This already applies to the definition of uncertainty. What is uncertainty, what is variability, what is sensitivity? And there are also rumors. Is Monte Carlo analysis the only possible method? Is it needed to take correlated variates into account?

This paper aims to provide an overview of the various aspects of uncertainty in relation to LCA, and of the practical approaches that have been proposed or employed. It partly builds on and supplements the survey of Björklund (2002). It starts with a theoretical part on the types of uncertainty and the techniques that are available to address these uncertainties. It then proceeds to give a survey of concrete proposals and implementation in guidebooks, databases, software and case studies. Finally, some proposals are made to arrive at a more uniform terminology of uncertainty issues in LCA.

2. TYPOLOGIES OF UNCERTAINTY

When speaking on uncertainties, one of the first things that could be defined is the very notion of uncertainty itself. Although a fully satisfying definition may be difficult to agree upon, we will here rely on a mere reference to the problem of using information that is unavailable, wrong, unreliable, or that show a certain degree of variability. The wording above suggests a division into three types:

- data for which no value is available;
- data for which an inappropriate value is available;
- data for which more than one value is available.

On top of that, one should acknowledge that LCA – and indeed any model – contains data, relationships and choices, so that the same division into three may be applied for relationships and choices as well, e.g., relationships for which no equation is available, or choices for which more than one option is available.

Before proceeding to study uncertainty in more detail, a contrast with variability should be made (US-EPA, 1989). Uncertainty relates to a lack of knowledge: no data is available, or the data that is available is wrong or ambiguous. Variability, in contrast, is a quality of data that is essentiality of a heterogeneous nature. The number of passengers in a specific train may be subject to uncertainty, while the number of passengers in a typical train may be subject to variability, because it differs from case to case. Likewise, the molecular weight of phenol may be uncertain, while the half-life time may be uncertain, because it depends on – variable – ambient conditions. Despite the different meaning and source of uncertainty and variability, the approaches for dealing with the two show a large overlap.

There are many ways of classifying uncertainty. Without going into the details of defining these categories, Table 1 lists a few typologies. Reviewing all these typologies, one might ask oneself whether a typology of uncertainties is useful at all. It appears that, no matter how you classify uncertainties, all uncertainties should be dealt with in the appropriate way. We believe, with Funtowicz & Ravetz (1990), that a typology is useful provided a distinction is made between sources and sorts of uncertainty. Moreover, we believe that it is the sorts of uncertainty that should be emphasized, because it ought to steer the approach taken to deal with uncertainty.

Another, perhaps underemphasized, aspect of uncertainty is that there are levels of uncertainty, relating to the role of the person that experiences the uncertainty. Thus, a scientist may feel uncer-
tain on the value of a certain parameter, while a
decision-maker may feel uncertain on the decision
to be taken. This distinction may be of critical
importance in the choice of methods to deal with
uncertainty. For instance, an ISO-standard may
settle the uncertainty problem for the decision-
maker, but not for the scientist, who will like to do
more research or to specify statistical distribu-
tions.

Table 1: Classification of uncertainties according to several authors.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>systematic errors</td>
<td>statistical variation</td>
<td>parameter uncertainty</td>
<td></td>
</tr>
<tr>
<td>random errors</td>
<td>subjective judgment</td>
<td>model uncertainty</td>
<td></td>
</tr>
<tr>
<td></td>
<td>linguistic imprecision</td>
<td>uncertainty due to choices</td>
<td></td>
</tr>
<tr>
<td></td>
<td>variability</td>
<td>spatial variability</td>
<td></td>
</tr>
<tr>
<td></td>
<td>inherent randomness</td>
<td>temporal variability</td>
<td></td>
</tr>
<tr>
<td></td>
<td>disagreement</td>
<td>variability between sources and objects</td>
<td></td>
</tr>
<tr>
<td></td>
<td>approximation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Funtowicz &amp; Ravetz (1990)</td>
<td>aleatory uncertainty</td>
<td></td>
<td></td>
</tr>
<tr>
<td>data uncertainty</td>
<td>epistemic uncertainty</td>
<td></td>
<td></td>
</tr>
<tr>
<td>model uncertainty</td>
<td>parameter uncertainty</td>
<td></td>
<td></td>
</tr>
<tr>
<td>completeness uncertainty</td>
<td>data uncertainty</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>model uncertainty</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ambiguity</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>volitional uncertainty</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. TECHNIQUES AND TOOLS TO ADDRESS UNCERTAINTY

Approaches to deal with uncertainty exist in many
kinds. Consider a concrete example: an LCA-
practitioner runs across an uncertain data item,
say the characterization factor for human toxicity
of zinc. How could one proceed? Main lines are:
- the scientific approach (doing more research,
  like setting out laboratory tests to find out
  LC50s and other relevant parameters in the
  characterization model);
- the constructivist approach (involving stake-
  holders, discussing and finally deciding on or
  voting for a consensus characterization fac-
  tor);
- the legal approach (relying on what authorita-
  tive bodies, like ISO or US-EPA, have de-
  creed as the truth);
- the statistical approach (using methods from
  statistics, like Monte Carlo analysis or fuzzy
  set theory, to determine confidence intervals
  and other indicators of robustness).

It should be noted that the first three of these ap-
proaches aim to reduce uncertainty, while the last
approach aims to explicitly incorporate it. Reduc-
ing uncertainty – although in itself a noble aim –
will not further be discussed here; we merely refer
to Von Bahr & Steen (2004) as a recent example
in LCA. We will restrict the discussion to ap-
proaches to incorporate uncertainty. In doing so,
we will be close to practicing post-normal science
(Funtowicz & Ravetz, 1993), a form of applied
science which claims to deal with policy issues in
cases of large uncertainties and high decision
stakes. Nonetheless, the statistical approach is in
other respects more alien to post-normal scien-
ts. In general, post-normal science prefers con-
structivist approaches to statistical approaches,
thereby eventually doing away with the uncer-
tainty. One should note that even statisticians
eventually do away with the uncertainty, namely
in their process of null-hypothesis significance
testing, where all uncertainty information finally
condenses into a yes-no decision.

In dealing with uncertainty, one is faced with
problems at three places:
- the input side: where are the uncertainties,
  and how large are they?
- the processing side: how do we translate input
  uncertainties into output uncertainties?
- the output side: how can we visualize and
  communicate uncertain results?

Of course, answers within the three areas are
highly dependent. For instance, if we choose to
use Monte Carlo analysis for the processing side,
our possibilities for the input and output sides are
immediately restricted to a few options. One
might be tempted to confine the picture of input,
processing and output to those sources of uncer-
tainty that pertain to parameter uncertainty or data
uncertainty, the truly input-oriented elements of a
model. However, some of the other sources of un-
certainty can be captured in these terms as well. The choice between competing models, or the choice among the elements to be modeled can also be regarded as comprising an input uncertainty, for instance.

3.1 Processing uncertainties

As the method for processing are the pivot in this scheme, we start to discuss this aspect. Within the statistical approach, there are many possibilities. Some of these are:

- parameter variation/scenario analysis;
- sampling methods;
- analytical methods;
- non-traditional methods, such as the use of fuzzy set theory.

Each of these methods for processing uncertainties requires different forms of input uncertainties, and delivers different forms of output uncertainties. Below, we will discuss these methods and some applications in LCA.

With parameter variation/scenario analysis, a few different data sets and/or models and/or choices are investigated as to their consequences for the model results. For instance, the results are calculated for a data set with high emission values and a data set with low emission values. Good illustration of the use of this method within LCA are provided by Copius Peereboom et al. (1999) and Huijbregts (1998).

A sampling method is a method that employs the power of a computer for repeating calculations many times. If the input data for each parameter is drawn from some distribution, the results will differ from run to run and gradually give rise to a sample of results, of which the statistical properties may be investigated. Monte Carlo analysis, where a distribution of outcomes is calculated by running the model a number of times with randomly selected parameter representations, is the most well-known form, but there are more sophisticated ways, like Latin hypercube sampling (Morgan & Henrion, 1990). Monte Carlo analyses have been applied in LCA by a couple of authors, see, e.g., Meier (1997), Huijbregts (1998), Maurice et al. (2000) and Sonneman et al. (2003).

Sampling methods can also be used to address scenarios. In that case, the sample consists of combinations of different decision scenarios and model formulations, with a subjective probability reflecting the preference of the decision-maker or the faith of the modeler in a particular model formulation for an alternative (Efron & Tibshirani, 1991). According to Huijbregts et al. (2003), the resulting output distribution reflects the uncertainty of the decision-maker regarding the normative choices involved (scenario uncertainty) or the uncertainty of the modeler regarding the alternative model formulations (model uncertainty).

Analytical methods are based upon explicit mathematical expressions for the distributions of the model results. Their use is based on a first-order approximation of the Taylor expansion of the underlying model (Heijungs & Suh, 2002; Heijungs, 2002). Distribution-free variances of input parameters can then be used to calculate variances of output variables. Their use in LCA has been limited so far, probably because the mathematics was too complicated to be implemented in software. Results have, however, been reported by Heijungs et al. (in press).

Under non-traditional methods, we will capture all methods of dealing with uncertainty that are not part of the traditional statistics curriculum. It comprises a variety of methods, for instance:

- fuzzy set methods;
- Bayesian methods;
- non-parametric statistics;
- robust statistics;
- neural networks and other methods from artificial intelligence.

Methods for uncertainty analysis based on fuzzy sets have been introduced into LCA by several authors, see, e.g., Weckemann & Schwan (2001), Chevalier & Le Téno (1996), Rong et al. (1998), Roš (1998). Bayesian statistics has hardly been mentioned in the context of LCA, although Shipworth (2002) provides an exception. The other mentioned methods are even less used within LCA, although the sign test and the Kruskall-Wallis test are briefly touched by Heijungs & Kleijn (2001). It should be noted that some of these methods are sometimes mentioned within LCA (see, e.g., Sangle et al., 1999), but that the emphasis is in those cases not so much on the processing side of uncertain information, but on the approaches that have been developed within decision theory for dealing with unclear preferences.

3.2 Input uncertainties

Parameter variation requires that a number of different values is available for one or more parameters. Treating all parameters individually may lead to an exceedingly large number of scenarios. Therefore, it is usual to vary one parameter and keep all other parameters fixed at some “most probable value”, and to repeat this procedure for all parameters in separate analysis. This type of analysis can be found in Copius Peereboom et al. (1999). An alternative is to define a limited number of scenarios with specific but consistent realizations of each parameter. Hofstetter (1998) and
Goedkoop & Spriensma (2000) employ these types of “perspectives” in their analyses. A more systematic treatment of the use of scenarios of the future in LCA is given by Pesonen et al. (2000) and Fukushima & Hirao (2002), and applied by Contadini et al. (2002).

Sampling methods are based on the random variation of uncertain parameters. They require the specification of a statistical distribution of every stochastic parameter. For instance, an emission may be specified as following a normal distribution with a mean of 12 kg and a standard deviation of 1 kg. Frequently encountered distributions are:

- the normal distribution;
- the lognormal distribution;
- the uniform distribution;
- the triangular distribution.

These distributions may or may not be correlated across parameters. In principle, correlations between parameters may be expressed by a correlation matrix or a covariance matrix (Heijungs & Suh, 2002). Huijbregts et al. (2003) showed how correlations between input parameters can be included in Monte Carlo analyses. Apart from correlations between input parameters, correlations between model outputs should be accounted for in comparative LCAs (Huijbregts et al., 2001; 2003). This can be done in the form of a comparison index (Huijbregts, 1998) for the case of two alternatives, or in a more general discernibility analysis (Heijungs & Kleijn, 2001; Heijungs & Suh, 2002).

Analytical methods are based on the estimation of the moments of the distributions (Morgan & Henrion, 1990). In particular the second moment, the variance, is used in a first order Taylor approximation. Thus, not the distribution, but only the variance (or standard deviation) of the parameter is needed here. Thus, less information is needed for analytical methods than for sampling methods. Like for Monte Carlo analysis, correlations between variates can in principle be included, although this is seldom seen in practice. Inclusion of correlations in the analytical case implies a broadening of the scope to second-order Taylor approximations (Heijungs & Suh, 2002).

Because methods for processing uncertainties on the basis of non-traditional methods have hardly been applied in LCA, it is not clear which types of input information would be needed.

3.3 Output uncertainties

At the output side there are fewer differences. In combination of parameter variation, one often sees the consecutive presentation of tables and/or graphs for the different sets of parameters or scenarios; see e.g. Copius Peereboom et al. (1999), Huijbregts (1998).

Results of sampling methods can be presented in different forms. Sampled probability density plots, so-called histograms, are a typical example; see, e.g., Huijbregts (1998) and Sonneman et al. (2003). An alternative is the graphical representation of an average value with two boundary values. These boundary values may indicate the smallest and largest value obtained, or a more robust measure such as the 5 and 95 percentile values (Huijbregts et al., 2003).

Analytical methods do not provide a distribution of outcomes. Instead, they provide moments of the distributions, such as the standard deviation. These can be used to calculate and visualise 95% confidence intervals. As analytical methods have hardly been applied in LCA, we cannot give an example of its use.

This holds even more true for the non-traditional methods, like fuzzy sets methods and Bayesian methods.

4. PROSPECTS FOR INCLUSION OF UNCERTAINTY IN LCA

Reviewing the developments that have taken place in the last few years, it seems likely that discussion and inclusion of uncertainty issues in LCA will no longer be restricted to academic exploratory work, like PhD-theses, and will no longer be regarded as a curiosity in real practical work. Rather, we expect that inclusion of uncertainties will become a standard feature of case studies. The three requirements for becoming a standard procedure, availability of data of input uncertainties, availability of methods and software for processing uncertainties, and availability of methods for interpreting and visualizing output uncertainties, start to be satisfied.

There is perhaps one more aspect that can be seen as a requirement: standardization (cf. Björklund, 2002). The ISO-standards for LCA have canonized parts of the terminology used. On top of that, the format by SPOLD has provided a standard for data exchange. But especially for uncertainty, clear standards are lacking.

As to terminology, there is first of all the confusion between uncertainty and variability, and within uncertainty all the sorts and sources of uncertainty. It may be difficult (and unnecessary) to single out one single terminology. Then, there is a large number of types of approaches that are used interchangeably, or at least in a non-standardized way. We mention just a few:
- uncertainty analysis;
- sensitivity analysis;
- perturbation analysis;
- scenario analysis;
- error analysis;
- discernibility analysis.

It is a disturbing (or perhaps: consoling) fact that even outside LCA, within the uncertainty community itself, meanings and nomenclature give rise to disagreement. For instance, sensitivity analysis means to US-EPA (1989) the systematic changing of one parameter while keeping the other parameters constant, whereas to Saltelli et al. (2000) it means the apportioning of an output uncertainty to the various contributing input uncertainties, which is in turn referred to as key issue analysis by Heijungs (1996) and as uncertainty importance analysis by Björklund (2002).

But at least the way uncertain data is described can be standardized. In fact, this should be part of the normal data exchange process. In a small study, Heijungs & Frischknecht (in prep.) discussed the differences in representing a basic entity like the uniform distribution in just one database (ecoinvent), one LCA-program (CMLCA) and mathematical statistics. Clearly, one may choose in representing a uniform distribution between giving:
- the mean value and the width;
- the mean value and the half-width;
- the lowest and highest value;

The required transformations in going from one representation to another one is quite simple. For more complicated distributions, like the lognormal distribution, the expressions are much more involved. As long as these different options are not clearly defined and distinguished, one is likely to confuse a standard deviation with a variance, a width with a half-width, or worse.

Obviously, those aspects of uncertainty that pertain to parameter uncertainty or data uncertainty have received most attention so far, at least in practical cases. The model uncertainty is much less addressed. And the more profound forms of uncertainty, for instance epistemic uncertainty may fundamentally be difficult to deal with. In this, we agree with parts of the analysis of Funтовicz & Ravetz (1993), who promote the development of non-traditional modes of research (i.e. post-normal science) to deal with intrinsically uncertain policy questions. Their NUSAP-scheme (Funтовicz & Ravetz, 1990) has been brought into LCA by Weidema & Wenes (1996). We think that the separation with which this paper started, into data for which more than one value is available, data for which an inappropriate value is available, and data for which no value is available, can be connected to the SAP-part of the NUSAP-scheme:
- spread, for data for which more than one value is available;
- assessment, for data for which an inappropriate value is available;
- pedigree, for data for which no value is available.

It is especially the S-part which can be processed with sampling or analytical techniques, and the A-part by parameter variation/scenario analysis techniques. The P-part is supposed to reflect our “ignorance of ignorance”, for which the use of precision-suggesting numbers is by definition inappropriate.

5. REFERENCES


Fukushima, Y. and M. Hirao. A Structured Framework and Language for Scenario-Based


Shipworth, D. A Bayesian Hybrid LCA Model of Stochastic Embodied Greenhouse Gas Emissions...

