An implemented approach for estimating uncertainties for toxicological impact characterisation

R. Rosenbaum[a], D.W. Pennington[b], O. Jolliet[a]

[a] *Industrial Ecology & Life Cycle Systems, Swiss Federal Institute of Technology Lausanne (EPFL), 1015 Lausanne, Switzerland*

[b] *Soil and Waste Unit, Institute for Environment and Sustainability (IES), Joint Research Centre (JRC), European Commission (EC)*

Abstract: One approach accounting for parameter and model uncertainty is implemented in the LCIA (life cycle impact assessment) method IMPACT 2002. The uncertainty is estimated for intermediate results from the chemical fate, human intake fraction, and two toxicological effect modules. Overall uncertainty estimates are then arithmetically calculated. Results are presented for impact contributions in the contexts of aquatic ecosystems and human health. The approach of Hofstetter (1998) was adapted for estimating the uncertainty related to chemical fate and human intake fractions. A fundamental problem when estimating uncertainties for 1000's of substances consists of the lack of uncertainty distributions for all of the input data and the need to have a practical approach to assign distributions to each chemical. Hofstetter (1998) proposed the use of fixed factors for clusters of substances. The choice of a factor is then dependent on the emission medium, exposure route, and the robustness of the model relative to the chemical being considered. The factors are initially determined for representative substances for each category using evaluation data, expert judgement, or approaches such as Monte Carlo. There is then no need to repeat the Monte Carlo calculations. Multiplying and dividing the geometric mean estimate by a factor provides an estimate of the upper and lower 95th percentile confidence interval bounds. The human health effect factor uncertainty is similarly defined and readily combined through addition with that of the intake fraction. Using expert judgement, three uncertainty classes were proposed to estimate uncertainty related to the human effects input data. These effects data account for both the risk of an effect, as well as the potential consequences of population-based exposures. The uncertainty for ecotoxicological effects is currently related to the number of species tested for aquatic species in the water column. The more species test results available, the more robust the estimate of the ecotoxicological factor is assumed to be. For estimating the ecotoxicological effect factor uncertainty, the combined use of two distinct approaches was suggested, – the higher uncertainty estimate being adopted. The combination of both guaranteed more robust results compared to applying either method – both being based on differing assumptions related to the sample versus the population distribution. The presented approach proved to be very transparent, robust but while reflecting our current level of knowledge, quick to use, and is easily applied in practice to combine the uncertainty of the emissions inventory with those of the impact assessment phase in a life cycle assessment study.

Keywords: Uncertainty; LCIA; Toxicity; Multimedia Modelling

1. INTRODUCTION

Accounting for uncertainty is vital in comparative assessments, although this has remained neglected in most life cycle assessment (LCA) studies and in related decision making. In LCA, uncertainty is, in part, due to the uncertainty associated with the model input data, at least those that are significant in the calculations of a characterisation factor (the impact per unit emission) for a particular chemical. This is the parameter, or data, uncertainty. Additionally, the uncertainty is related to the inherent uncertainty of the overall models and the underlying correlations themselves (model uncertainty). In combination, these model and parameter uncertainties could be considered the accuracy of the model.

IMPACT 2002 (Pennington et al., 2003) provides estimates for the uncertainty accompanying every

characterisation factor (CF). By proposing a straightforward way of combining uncertainties of intermediate results to derive the final overall characterisation factor uncertainty, which can then be combined with LCA inventory uncertainties, the IMPACT 2002 methodology facilitates the calculation of LCA results with a related overall uncertainty estimate. In this way, judgement of LCA results is improved as their reliability can be taken into consideration by decision makers.

The general framework for calculating a human health effect characterisation factor is established as:

$$CF_h = FF \cdot XF \cdot EF_h = iF \cdot EF_h \qquad (1)$$

where FF denotes the fate factor, XF the exposure factor, EF_h the human health effect factor, and iF the human intake fraction which is the product of FF and XF.

Assuming that $XF = 1$, in IMPACT 2002 the parallel concept for (aquatic) ecotoxicological effect characterisation is:

$$CF_{aqu} = FF \cdot EF_{aqu} \qquad (2)$$

As an example we used the non-spatial version of IMPACT 2002 for an emission of 1,1,2,2-Tetrachloroethane into European surface waters. Based on the framework presented in Equations 1 and 2 we obtained the following illustrative results:

Table 1: Illustrative geometric mean estimates for an emission of 1,1,2,2-Tetrachloroethane into European surface waters.

(intermediate) Result	Value
Human health characterisation factor, CF_h [*number of cases /kg$_{emitted}$*]	1.66E-06
Ecotoxicological characterisation factor, CF_{aqu} [PAF $m^3 \cdot day^2/kg$]	2.93E+01
Human health effect factor, EF_h [*number of cases/kg$_{intake}$*]	8.74E-02
Ecotoxicological aquatic effect factor, EF_{aqu} [PAF $m^3 \cdot day/kg$]	2.4E+02
Human Intake Fraction, iF [$kg_{intake}/kg_{emitted}$]	1.90E-05
Human exposure factor, XF [1/*day*]	1.56E-04
Fate factor, FF [*day*]	1.22E-01

Based on these geometric mean estimates their 95th percentile confidence intervals are calculated in the following subsections.

2. UNCERTAINTY MODULES

2.1. Fate and exposure uncertainty

For estimating the overall uncertainty related to fate and exposure the approach of Hofstetter (1998) was adapted due to its straightforward and pragmatic, yet robust nature. By considering the emission medium, exposure route, and the appropriateness of the model relative to the chemical, this approach provides estimates for clusters of chemicals of the combined model, scenario, and parameter uncertainty – i.e. the likely overall accuracy of the results.

A fundamental problem when calculating the uncertainties of results for many substances is that not every input parameter has a "known" uncertainty distribution that can be used to estimate the final uncertainty. Hofstetter (1998) therefore proposed a set of fixed factors, based on representative distributions for clusters of chemicals with similar attributes. These factors are the square geometric standard deviations (SD_g^2) associated with log-normal distributions. The factors depend on the exposure route, emission medium, and certainty criteria classifying the robustness of the model in the context of the substance into one of three certainty classes (h – high, m – medium, l – low certainty).

The representative uncertainty distributions for each cluster of chemicals can be estimated using evaluation data for representative substances and expert judgement, or other approaches such as Monte Carlo. Caution is advocated when using approaches such as Monte Carlo as only input data uncertainty is often taken into account. The uncertainty of also the correlations in a model, etc., are often not addressed in practice and may be more significant. For example, estimates of contaminant uptake into meat and vegetation can be a major source of the uncertainty in human exposure estimates for most persistent organic chemicals that is often associated with the correlations adopted and not the data input into the model. The factors in the Hofstetter approach are the square geometric standard deviation values, SD_g^2. These factors are used to calculate the upper and the lower bound of the 95th percentile confidence interval by assuming a log-normal distribution of the uncertainty. For the intake fraction the calculations are:

$$iF_{97.5} = iF \cdot SD_g^2 \qquad (3)$$

$$iF_{2.5} = \frac{iF}{SD_g^2} \qquad (4)$$

Tables 2 and 3 list the values originally proposed by Hofstetter (1998) for emissions to air and water/soil, respectively. It should be noted that Hofstetter did not provide factors for soil emissions; hence the working proposal in Table 3 was adopted. Thus, the values for emissions to soil are the same as for water emissions. This might not be correct for soil emissions where exposure via drinking water or fish consumption are the main concern, but this may be rarely the case at a European population level for example (e.g. Bennett et al. 2002).

Table 2: Estimated fixed square geometric standard deviations for emissions to air (Hofstetter, 1998)

Certainty criteria	Certainty class	Exposure via:		
		air	water	food
- >95% of intake is inhaled - chem. properties are well known - substance is well suited for model	h	2	4	8
- substance partitions - chem. properties are well known - substance is well suited for model	m	3	6	12
- substance properties are uncertain or unknown - substance is not well suited for model	l	20	40	80

Table 3: Estimated fixed geometric standard deviations for emissions to water and soil (Hofstetter, 1998)

Certainty criteria	Certainty class	Exposure via:		
		air	water	food
- >95% of intake by drinking water - chem. properties are well known - substance is well suited for model	h	4	2	8
- substance partitions - chem. properties are well known - substance is well suited for model	m	6	3	12
- substance properties are uncertain or unknown - substance is not well suited for model	l	40	20	80

Hofstetter (1998) also provided a list of example substances that fit into each of the established clusters.

For the 1,1,2,2-Tetrachloroethane example introduced above we chose the certainty class *m* as inhalation is below 95% of intake, the properties

are peer reviewed high quality data and the model was considered well suited for this organic non dissociating compound.

For an emission to water we find the SD_g^2 estimates for different exposure routes in Table 3. The detailed results suggest drinking water as the exposure pathway associated with the highest risk, hence $SD_g^2 = 3$. The 95th percentile confidence interval is calculated according to Equations 3 and 4, resulting in:

$iF_{97.5} = 5.70E\text{-}05$

$iF_{2.5} = 6.33E\text{-}06$

Hence, it is assumed that there is a 95% confidence that the true intake fraction is within a factor of 3 of the geometric mean estimate.

2.2. Effects uncertainty

The uncertainty of the effect factors is exclusively linked here to the likely reliability of the related input data, e.g. the Toxic Dose causing tumours in 50% of species (TD_{50}), the No Observed Adverse Effect Level (*NOAEL*), etc. used to derive the effect factor (*EF*). Only parameter uncertainty is addressed.

Different approaches and values are proposed to estimate the uncertainties for human health and for ecosystem effects. These methods have been chosen with consideration of the need to assess thousands of substances and their related input data, as consistently as possible, while being appropriate for use in comparative assessments. Further research and development is strongly encouraged, particularly in the context of estimating the low dose-responses that will be associated with many of the emissions and toxicological impacts assessed using tools such as LCA.

2.3. Human health effects uncertainty

The human health effect factor (effect per unit intake) uncertainty is currently estimated based on a rough orders-of-magnitude approach, pending completion of reviews of the various uncertainties and propositions for more robust values and rules.

Using expert judgement and building on classical approaches, three uncertainty classes were proposed to estimate uncertainty related to the human health effects input data (see Table 4). This is typical of current practice for regulatory chemical risk screening, except the uncertainty factors are retained as a measure of the uncertainty and, unlike in screening, were not included in the estimate of the effect factor. This helps to avoid bias amongst factors with similar geometric means

in a relative comparison context, whilst retaining information on the different uncertainties.

Table 4: Expert judgement based uncertainty factors estimation of human health effects data uncertainty.

Uncertainty criteria	Uncertainty factor
Data from peer reviewed sources, such as the critical effect from US EPA's IRIS (IRIS, 2004) or Gold et al.'s carcinogenic data handbook (Gold & Zeiger, 1997) for example. These data should also be for chronic exposures.	10
Data from less peer reviewed sources that are based on chronic or sub-chronic exposure test results.	100
Data extrapolated from acute test results to humans (based on insights in e.g. Crettaz et al., 2002 and Pennington et al., 2002, as well as many other non-LCA sources).	1000

The calculation of the upper and the lower bound of the 95th percentile confidence interval of the effect factor (EF) is performed in a similar way to Equations 3 and 4.

For the 1,1,2,2-Tetrachloroethane example, peer reviewed high quality chronic effects data were used. This justifies an uncertainty factor of 10 (there is 95% confidence that the estimate is within a factor of 10 of the geometric mean). Hence, the upper and lower bounds of the effect factor can be estimated as:

$EF_{h, 97.5} = 8.74E\text{-}01$

$EF_{h, 2.5} = 8.74E\text{-}03$

2.4. Ecosystem effects uncertainty

Ecosystem effects uncertainty is currently related to the uncertainty of the underlying Hazardous Concentration data used to calculate the effect factor ($0.5/HC_{50}$) (Payet & Jolliet, 2004 Pennington et al. 2004). These HC_{50} data reflect the concentration at which 50% of species are likely to be affected.

In practice, the HC_{50} is estimated using the geometric mean of test results for different species in the relevant medium, currently in the water column for IMPACT 2002. The more species tested, the larger the sample size that is available to estimate the HC_{50} and the more robust it is considered to be. This robustness is expressed in terms of parameter uncertainty, the lower the sample size the higher the uncertainty.

As with the uncertainty calculations discussed above, the lower 95th percentile confidence interval HC_{50} is calculated by dividing by the square

geometric standard deviation (similar to Equation 4) and the upper bound by multiplying by SD_g^2 (similar to Equation 3).

For estimating uncertainties related to the ecotoxicological effect factors, the combined use of two distinct statistical approaches was suggested (pending further investigation) – the higher uncertainty estimate being adopted:

Payet & Jolliet (2004) proposed adoption of the Student approach to calculate the 95%ile confidence interval from the sample distribution of the logarithmic values to estimate the 2.5th and the 97.5th percentiles (assuming a log-normal distribution). Secondly, a fixed-factor approach, dependent only on the sample size (*n*), was adopted based on a review of methods, distributions, and available data to determine the 95th percentile confidence interval for the HC_{50} (Pennington, 2003).

As shown in Table 5 for the fixed factors, the uncertainty significantly decreases with sample sizes higher than three.

Table 5: Median and upper 95th percentile confidence interval limits for extrapolations from the sample-based estimate to the population HC_{50} as a function of sample size – sample/population HC_{50} 97.5th percentile ratio (Pennington, 2003)

Sample size (*n*)	Log-logistic distributed	Log-triangular distributed	Log-normal distributed
1	-	-	-
2	231	266	231
3	134	242	168
4	53	88	68
5	44	87	69
6	27	46	38
7	25	45	36
8	18	29	26

The two approaches differ in terms of the assumed relationship between the distribution of the sample and of the actual population. The Student approach assumes that the distribution standard deviation is equivalent in the sample and in the population. Approaches considered in Pennington (2003) assumed that the standard distribution of the population was between likely limits observed from available toxicological insights, but not that it is equal to that of the sample.

So far, as both include necessary assumptions, the combination of the two estimates was considered to provide a more robust result compared to applying one or the other. In practice, this combination is applied by simply choosing the

higher uncertainty estimate from the two approaches.

From Table 5, the choice of model is less important for the estimation of the geometric mean. The log-normal distribution is consistent with the approach adopted in IMPACT 2002 and reflects the middle estimate for the uncertainty for each sample size from these different common distributions. When compared with the Students approach, the estimate may be higher or lower for a given sample size depending on the distribution of the actual sample of test results (noting again that this is then assumed to be equivalent to the actual, or population, distribution).

In practice, the tendency is that the Student's approach typically estimates lower uncertainty values and therefore reliance is more on the values in Table 5. The estimates in Table 5 were based on the assumption of a maximum plausible standard deviation from observations. It could be argued that smaller plausible deviations could be adopted for certain chemicals based on their likely modes of action, although this is not without complexity and may not be justifiable in a practical context in LCA.

These calculations do not currently account for acute-to-chronic uncertainty, which is likely to be negligible according to both Payet & Jolliet (2004) and Pennington (2003), nor was additional uncertainty from using QSARs taken into account (where such data are adopted). Further research is required to account for the scenario and model uncertainty associated with these effect factor estimates, particularly in the context of estimating toxicological effects at likely low concentrations in the context of complex mixtures at regional scales that are relevant in LCA.

Applying Student's approach to the 1,1,2,2-Tetrachloroethane example results in a SD_g^2 estimate of 4.7 for the lower bound and 0.75 for the upper bound HC_{50}. As the HC_{50} is based on 10 species choosing the log-normal distribution from Table 5 yields an SD_g^2 of 26. A higher uncertainty is indicated by the latter approach. Applying the principle of Equations 3 and 4 to the HC_{50} results in the following interval:

$HC_{50\,97.5}$ = 8.01E-05

$HC_{50\,2.5}$ = 5.42E-02

This will be a slight overestimation since the HC_{50} is based on 10 species, while Table 5 only provided SD_g^2 estimates for up to 8 species. The effect factor (0.5/HC_{50}) is:

$EF_{aqu,\,97.5}$ = 6.24E+03

$EF_{aqu,\,2.5}$ = 9.23E+00

3. COMBINING TWO INDEPENDENT UNCERTAINTIES

Combining uncertainties is a vital step throughout the process of deriving a final overall uncertainty estimate for the model result (the characterisation factor in this case). It also provides a basis to quickly identify what is causing the highest uncertainty.

Uncertainty estimates for the human intake fractions for the fate and exposure part and the effect factors for the effects side, are readily combined when both uncertainties are log-normally distributed (without the need for further assumptions). Table 6 presents the results for the 1,1,2,2-Tetrachloroethane.

The geometric mean fate and exposure estimate is multiplied by the geometric mean effect factor, to estimate the overall characterisation factor (Equations 1 and 2). The two square geometric-standard deviations of the uncertainty distributions can be simply summed to calculate the overall uncertainty distribution's standard deviation, as outlined for example in the appendices of Hofstetter (1998). This overall uncertainty estimate can similarly be combined in a quantitative manner with those of the inventory data if they are presented in the form of log-normal uncertainty distributions.

To derive the overall uncertainty in the 1,1,2,2-Tetrachloroethane example the previously calculated uncertainties have to be combined. For the human health CF:

$$SD_g^2(CF) = SD_g^2(iF) + SD_g^2(EF_h)$$
$$13 = 3 + 10$$

(6)

In this particular case and using the values presented, the uncertainty of the effect factor is clearly dominating the overall uncertainty. Similar to Equations 3 and 4, the 95th percentile confidence interval bounds on the CF_h are then given by multiplying and dividing the geometric mean estimate by a factor of 13. There is 95% confidence that the CF_h is within a factor of 13 of the geometric mean):

$CF_{h,\,97.5}$ = 2.16E-05

$CF_{h,\,2.5}$ = 1.28E-07

The uncertainty of EF_{aqu} is dominating the overall CF_{aqu} uncertainty (as the combined fate and

exposure uncertainty is a factor 3 compared to that of EF_{aqu} of 26).

Table 6: Illustrative IMPACT 2002 based geometric mean estimates accompanied by upper and lower bounds of the 95th percentile confidence interval for an emission of 1,1,2,2-Tetrachloroethane into European surface water.

(intermediate) Result	Lower bound	Geometric mean	Upper bound
Human health CF_h [*number of cases* /$kg_{emitted}$]	1.28E-07	1.66E-06	2.16E-05
Ecotoxicological CF_{aqu} [PAF $m^3{\cdot}day^2/kg$]	1.13E+00	2.93E+01	7.62E+02
Human health eff. fact., EF_h [*number of cases*/kg_{intake}]	8.74E-03	8.74E-02	8.74E-01
Ecotox. aquatic eff. f., EF_{aqu} [PAF $m^3{\cdot}day/kg$]	9.23E+00	2.4E+02	6.24E+03
Human Intake Fraction, iF [$kg_{intake}/kg_{emitted}$]	6.33E-06	1.90E-05	5.70E-05
Human exposure factor, XF [1/*day*]		1.56E-04	
Fate factor, FF [*day*]		1.22E-01	

While some insights have been provided in the literature related to uncertainty in LCA (Hofstetter, 1998; Hertwich et al., 1999; Hertwich et al., 2000; Huijbregts et al., 2000) and a working framework has been presented here for use in a relative comparison context, further research and development remain necessary for the quantification, but also to help take these uncertainties better into account in decision making.

4. REFERENCES

Bennett D.H. Margni M.D., McKone T.E., Jolliet O. (2002): Intake Fraction for Multimedia Pollutants: A Tool for Life Cycle Analysis and Comparative Risk Assessment. Risk Analysis, vol. 22 (5), 903-916.

Crettaz P., Pennington D.W., Rhomberg L., Brand K., Jolliet O. (2002): Assessing Human Health Response in Life Cycle Assessment using ED10s and DALYs - Part 1: Cancer Effects. Risk Analysis, vol. 22 (5), 931-946.

Gold L.S., Zeiger E. (1997): Handbook of carcinogenic potency and genotoxicity databases; ISBN 0-8493-2684-2, CRC Press.

Hertwich E.G., McKone T.E., Pease W.S. (2000): A systematic uncertainty analysis of an evaluative fate and exposure model. Risk Analysis 20: 439-454.

Hertwich E.G., McKone T.E., Pease W.S. (1999): Parameter uncertainty and variability in evaluative fate and exposure models. Risk Analysis 19: 1193-1204.

Hofstetter P. (1998): Perspectives in life cycle impact assessment: A structured approach to combine models of the technosphere, ecosphere and valuesphere. Kluwer Academic Publishers, Boston.

Huijbregts M.A.J., Thissen U., Jager T., van de Meent D., Ragas A.M.J. (2000): Priority assessment of toxic substances in life cycle assessment. Part II: assessing parameter uncertainty and human variability in the calculation of toxicity potentials. Chemosphere 41: 575-588.

IRIS (2004): US EPA's Integrated Risk Information System, available online at http://www.epa.gov/iris/.

Payet J. & Jolliet O. (2004): Comparative Assessment of the Toxic Impact of metals on aquatic ecosystems: the AMI method. In Life Cycle Assessment of Metals: Issues and research directions, A. Dubreuil Ed., SETAC Press, in Press.

Pennington, D.W.; Payet, J.; Hauschild, M. (2004): Aquatic ecotoxicological indicators in life cycle assessment. Environmental Toxicology & Chemistry, 23, in press.

Pennington D.W., Margni M., Payet J., Charles R., Jolliet O. (2003): Estimating cumulative toxicological risks and potential impacts for human health and ecosystems in LCA, submitted to Environmental Toxicology & Chemistry.

Pennington D.W. (2003): Extrapolating ecotoxicological measures from small data sets. Ecotoxicology and Environmental Safety, 56, 238-250.

Pennington D.W., Crettaz P., Tauxe A., Rhomberg L., Brand K., Jolliet O. (2002): Assessing Human Health Response in Life Cycle Assessment using ED10s and DALYs - Part 2: Non-Cancer Effects. Risk Analysis, vol. 22 (5), 947-963.

US EPA (1997): Exposure Factors Handbook, Vol. I-III. EPA/600/P-95/002Fa. Office of Research and Development, U.S. Environmental Agency, Washington.