Supporting Information for:

How to conduct a proper sensitivity analysis in life cycle assessment: Taking into account correlations within LCI data and interactions within the LCA calculation model

W. Wei[1], ,P. Larrey-Lassalle[2] T. Faure[1], N. Dumoulin[1], P. Roux[2], J.D. Mathias[1]

[1] Irstea, UR LISC, 9 avenue Blaise Pascal, F-63178 Aubiere, France
[2] Irstea, UMR-Itap, 361 rue Jean-François Breton, F-34196 Montpellier, France

This document includes five supplementary materials and three supplementary tables. Terms, abbreviations and symbols are the same as in the main text.

S1. Supplementary matrix-based LCA

In the 1990s, Heijungs et al.[1] proposed a matrix method to solve the inventory problem in LCA (see Heijungs and Suh[2] for a brief history of the matrix method). It also provides explicit algebraic equations to calculate environmental impact though several intermediate steps. In the first step, they define economic flows (also known as technology flows) and environmental flows (also known as intervention flows) as follows:

- a_{ij}: economic flow i of unit process j.
- b_{ij}: environmental flow i of unit process j

Each elementary process is represented as a vector in a base of economic flows and a base of environmental flows giving the economic matrix $\boldsymbol{A}$ and the environmental matrix $\boldsymbol{B}$. The negative value corresponds to consumption, and the positive value corresponds to emissions.

We denote $\boldsymbol{S}$, the scaling matrix. We obtain the final demand matrix $\boldsymbol{f}$ as follows:

$$\boldsymbol{f} = \boldsymbol{A}\boldsymbol{S} \tag{1}$$

so the scaling matrix $\boldsymbol{S}$ is equal to :

$$\boldsymbol{S} = \boldsymbol{A}^{-1}\boldsymbol{f} \tag{2}$$

where final demand matrix $\boldsymbol{f}$ is a vector that is defined arbitrarily. The result $\boldsymbol{g}$ is then defined with $\boldsymbol{S}$ and $\boldsymbol{B}$:

$$\boldsymbol{g} = \boldsymbol{B}\boldsymbol{S} \tag{3}$$

Lastly,[3] went on from the inventory flows to the aggregation of environmental impacts $\boldsymbol{h}$ using the characterization matrix $\boldsymbol{Q}$:

$$\boldsymbol{h} = \boldsymbol{Q}\boldsymbol{g} \tag{4}$$

Matrix $\boldsymbol{Q}$ represents factors of characterization for each impact category. The goal is to reduce all the inventory flows participating into an impact category in a single unit quantifying the associated potential impact. For example, for the global warming impact category (greenhouse effect), many substances can lead to this effect, such as water vapor, CO_2, CO_3, CH_4, N_2O, etc. By the matrix $\boldsymbol{Q}$, everything is reduced to kg CO_2 equivalent. There are several methods of characterization (ReCiPe, UseTox, ILCD, IMPACT2002+, etc.); each method defines different impact categories, so each method has its own characterization matrix. In summary, the formula for the environmental impacts $\boldsymbol{h}$ is described by the following equation:

$$\boldsymbol{h} = \boldsymbol{Q}\boldsymbol{B}\boldsymbol{A}^{-1}\boldsymbol{f} \tag{5}$$

where $\Lambda = BA^{-1}$ is called intensity matrix. A overview of all matrix in the matrix-based LCA is given in the table S1.

Table S1: *Overview of the matrix in matrix-based LCA*

Symbol	***Name***	***Dimension (rows × columns)***
$\boldsymbol{A}$ (a_{ij})	Economic (technology) matrix	Economic flows×processes
$\boldsymbol{B}$ (a_{ij})	Environmental (intervention) matrix	Environmental flows×processes
$\boldsymbol{Q}$ (q_{ij})	Characterization matrix	Categories×environmental flows
$\boldsymbol{S}$ (s_j)	Scaling matrix	Processes×1
$\boldsymbol{g}$ (g_i)	inventory matrix	Environmental flows×1
$\boldsymbol{h}$ (h_k)	Environmental impacts	Categories×1
Λ (λ_{ij})	Intensity matrix	Environmental flows×economic flows
$\boldsymbol{f}$ (f_i)	Final demand matrix	Economic flows×1

S2. Supplementary relative sensitivity coefficients

Given a function (model) φ with output y and input parameters x_i and x_j:

$$y = \varphi(x_i, x_j) \tag{6}$$

The model φ can be a matrix-based LCA model (see Equation 5 for example). The sensitivity coefficients of x_i and x_j noted $\frac{\partial y}{\partial x_i}$ and $\frac{\partial y}{\partial x_j}$ can be calculated by the partial derivative of function φ:

$$\begin{cases} \dfrac{\partial y}{\partial x_i} = \dfrac{\partial \varphi(x_i, x_j)}{\partial x_i} \\ \dfrac{\partial y}{\partial x_j} = \dfrac{\partial \varphi(x_i, x_j)}{\partial x_j} \end{cases} \tag{7}$$

We then investigate dimensionless multipliers, such as:

$$\begin{cases} \dfrac{\partial y}{\partial x_i}\dfrac{x_i}{y} = \dfrac{\partial y / y}{\partial x_i / x_i} \\ \dfrac{\partial y}{\partial x_j}\dfrac{x_j}{y} = \dfrac{\partial y / y}{\partial x_j / x_j} \end{cases} \tag{8}$$

The results of these formulas are sensitivity ratios, and are termed "relative". The objective is to observe a small change in input parameter that leads to a change in the result. The partial derivatives for the matrix $\boldsymbol{A}$ and $\boldsymbol{B}$ are :

$$\frac{\partial h_k}{\partial a_{ij}} = -s_j \sum_l (q_{kl} \lambda_{li}) \tag{9}$$

$$\frac{\partial h_k}{\partial b_{ij}} = q_{ki} s_j \tag{10}$$

The relative sensitivity coefficient of the economic matrix $\boldsymbol{A}$ and the environmental matrix $\boldsymbol{B}$ are (see Table S1 for information on the parameters):

$$\frac{\partial h_k / h_k}{\partial a_{ij} / a_{ij}} = -\frac{a_{ij}}{h_k} s_j \sum_l (q_{kl} \lambda_{li}) \tag{11}$$

$$\frac{\partial h_k / h_k}{\partial b_{ij} / b_{ij}} = \frac{b_{ij}}{h_k} q_{ki} s_j \tag{12}$$

These values can be negative indicating that the change in the result is negative with respect to the varied parameter[4].

S3. Modeling uncertainty: univariate and multivariate random sampling

Log-normal distribution is defined by its mean μ and its variance σ^2; parameter σ is its standard deviation (noted SD). The distribution is often abbreviated $\ln N(\mu, \sigma)$. The probability density function is described so that :

$$f(x) \quad = \quad \frac{1}{x\sigma\sqrt{2\pi}} exp\left\{-\frac{(ln(x)-\mu)^2}{2\sigma^2}\right\}, \quad x > 0 \tag{13}$$

For the correlation case, a multivariate log-normal distribution must be used, it is definition is defined by:

$$f(X) \quad = \quad \frac{1}{\sqrt{(2\pi)^p|\Sigma|}} (\textstyle\prod_{i=1}^{p} \frac{1}{x_i}) exp\left\{-\frac{1}{2}(ln(X)-\mu)^T\Sigma^{-1}(ln(X)-\mu)\right\}, \; \forall x_i > 0 \tag{14}$$

where p represents the number of variables and μ is the means corresponding to vectors. $X = (x_1, x_2, \dots, x_p)$ and $ln(X) = (ln(x_1), ln(x_2), \dots, ln(x_p))$. $|\Sigma|$ is the determinant of the variance-covariance matrix Σ $(p \times p)$, and Σ^{-1} is the inverse covariance matrix. The variance-covariance matrix Σ $(p \times p)$ define as following:

$$\Sigma(p \times p) = \begin{pmatrix} var(x_1) & cov(x_1, x_2) & \cdots & cov(x_1, x_p) \\ cov(x_2, x_1) & \ddots & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ cov(x_p, x_1) & \cdots & \cdots & cov(x_p) \end{pmatrix} = \begin{pmatrix} \sigma^2_{x_1} & \sigma_{x_1x_2} & \cdots & \sigma_{x_1x_p} \\ \sigma_{x_2x_1} & \ddots & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{x_px_1} & \cdots & \cdots & \sigma^2_{x_p} \end{pmatrix} \tag{15}$$

where $cov(x_1, x_2) = \sigma_{x_1x_2} = \rho_{x_1x_2} * \sigma_{x_1} * \sigma_{x_2}$. We note that other probability distributions can be used. However, for the sake of simplicity, the study is limited to this probability distribution.

S4. Supplementary Sobol indices

TSI expresses all the effects of an uncertainty on the output result. The relationship between SI and TSI is described as follows :

$$TSI(x_i) = \underbrace{SI(x_i)}_{main\ effect} + \underbrace{\sum_{j \neq i} SI(x_{i,j}) + \sum_{j \neq i, k \neq i, j<k} SI(x_{i,j,k}) + \cdots}_{total\ interactions} = \sum_{l \in \# i} SI(x_l)$$

where $x_{i,j}$ (id. $x_{i,j,k}$) represents interactions between x_i and x_j (and x_k). $\# i$ represents all subsets which contain the index i. hence $\sum_{l \in \# i} SI(x_l)$ is the sum of all the sensitivity indices involving uncertainty x_i. $SI(x_i)$ is also known as the main effect, $SI(x_{i,j})$ (id. $SI(x_{i,j,k})$) is the interaction in the second degree (id. third degree), and the sum of all degrees is termed total interactions.

S5. Supplementary information for LCA practitioners

Figure S1 presents an example of an LCA model as done by LCA practitioners: a process tree. It shows that the different processes are split in foreground and background activities. Foreground activities concern processes of the system that are specific to it [5] and for which specific data is provided directly by LCA practitioner. Background activities are processes, which supply energy and materials to the foreground system, via a homogeneous market so that individual and operations cannot be identified [6]. Background LCI data is taken from LCI databases.

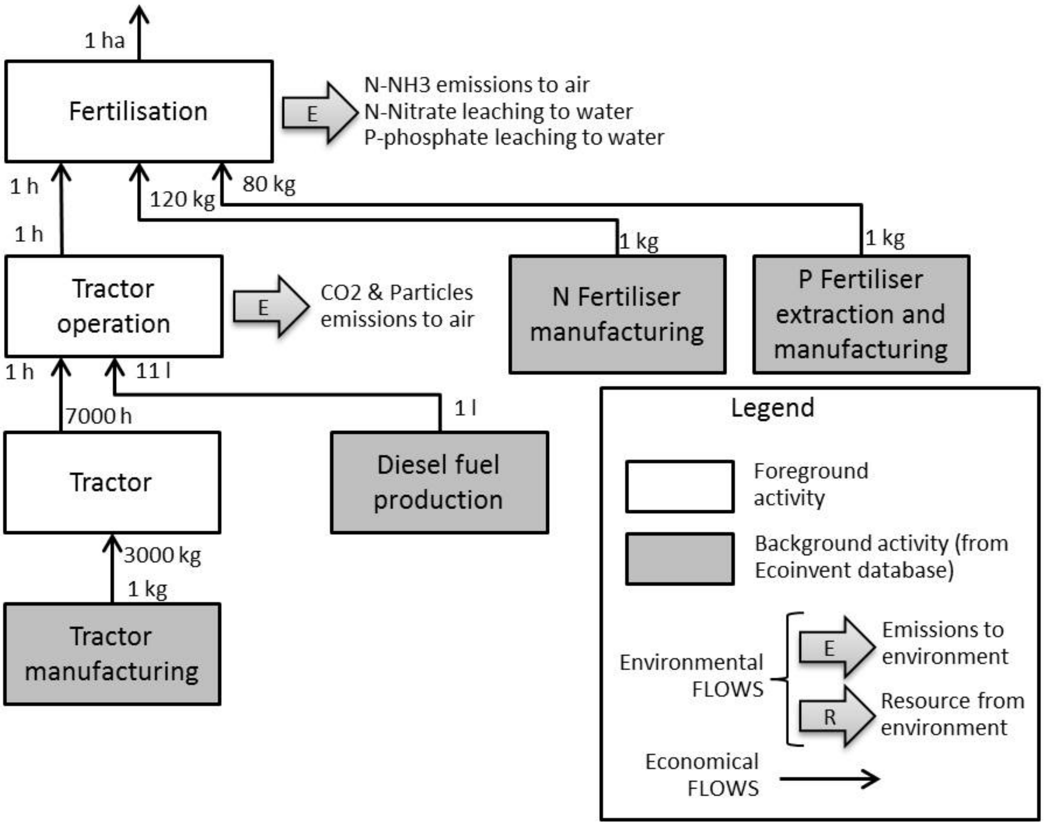

Figure S1: Example of LCA model (Process tree) illustrating the fertilisation of 1 ha (reference flow)

Interactions are produced by the model itself (process tree) and due to explicit causality relations. Since this information is embedded in $\boldsymbol{A}$ matrix for economic flows and $\boldsymbol{B}$ matrix for environmental flows, the LCA practitioner do not need to supply any additional data for interactions assessment.

Correlation presents the relationship among random variables. It quantify whether and how strongly pairs of random variable are related. The information related to correlation is generally not embedded in most of Life Cycle Inventories databases and models. Since it is a difficult task to identify all potential correlations within a LCA process tree, we propose to split it in two parts:

- For **foreground activities**, the practitioner has the expertise and should be able to complete the correlation matrix. Some methods [7;8;9;10] have been proposed in literature for assessing correlations; but their adaptation to the LCA specificities have not yet being done. For the illustration example hereunder, correlation coefficients have been estimated empirically, and thus are quite arbitrary and subjective. As such methods will become more used within the LCA community, guidelines should be provided in order to help practitioners assessing correlation between the variables of their modelled system, combining, if need both empirical and existing methods. We could imagine recommendations linked to families of correlated issues, in the same way that have been established Data Quality Indicators (DQI) to assess uncertainty in LCI databases.

- Until Databases include such information, the identification of correlated input parameters for **background activities** seems to be often not feasible for a common LCA practitioner (it should need expertise for all background processes and activities).

For the construction of the correlation matrix of the agricultural example presented in Figure S1, the first step is to identify stochastic external processes or phenomena, which can affect uncertainty associated to LCA model variables, such as:

- Climatic factors (e.g. rain).
- Field topography and soil conditions.
- Agricultural practices (e.g. human factor associated to the technology used for spraying).

These external stochastic processes can have a combined effect (i.e. a correlated effect) on the uncertainties of couples of LCA model variables. For instance, climatic factor such as rain can affect both field nitrates and phosphorus emissions (via leaching). Field topography and soil conditions have a coupled effect on the sprayed amount of N&P-fertilizers per hectare and on the diesel consumption: flow rate (i.e. dose per ha) is proportional to the sprayer advance (using electronic device measuring wheels rotations) and is correlated to soil conditions due to wheel spin which also affects the fuel consumption. Agricultural practices (e.g. spraying tractor speed and operator carefulness for avoiding overlapping) may lead to a heterogeneous distribution of fertilizers on the field and may affect as well fuel consumption per hectare.

On the contrary, CO_2 emissions from tractor may directly proportional to diesel consumption, given the explicit internal causality relation between these two variables (fuel combustion model). This is clearly involved by the model itself (process tree). We consider that no values of correlation factors are implemented in the correlation matrix.

Based on that, Table S2 shows how the correlation matrix can be determined empirically for the agricultural example (Figure S1).

Table S2: *Correlation matrix for example of Figure* S*1.*

	CO_2 emission in Tractor operation	N-Nitrate emissions to water in Fertilization	P-Phosophate emissions to water in Fertilization	Diesel consumption in Tractor operation	N-Fertilizer consumption in Fertilization	P-Fertilizer amount per hectare in Fertilization
CO_2 emission in Tractor operation	1	0	0	0	0	0
N-Nitrate emissions to water in Fertilization	0	1	0.9(a)	0	0	0
P-Phosophate emissions to water in Fertilization	0	0.9(a)	1	0	0	0
Diesel consumption in Tractor operation	0	0	0	1	0.4(b)(c)	0.4(b)(c)
N-Fertilizer consumption in Fertilization	0	0	0	0.4(b)(c)	1	0.8(b)(c)
P-Fertilizer amount per hectare in Fertilization	0	0	0	0.4(b)(c)	0.8(b)(c)	1

Legend:

(a) Due to uncertainty related to the climatic factor (e.g. rain), N-Nitrate emissions to water are correlated with P-Phosphate emissions.

(b) Due to uncertainty related to field topography and soil conditions, N&P-fertilizers amounts per hectare and Diesel consumption are correlated. N-fertilizers and P-fertilizers doses may be more correlated between themselves than with Diesel consumption, which explains the differences between the proposed correlation factors.

(c) Due to uncertainty related to agricultural practices, N&P-fertilizers amounts per hectare and Diesel consumption are also correlated.

Table S3. Correlation matrix. We consider that Phenol and Gas are correlated. Indeed, gas is used in the manufacture of methanol which is based on the production of formaldehyde itself. Formaldehyde and Phenol are both binding agents (fixing agents/binders) used to secure the glass fibers in the manufacture of glass wool[11].

Parameters	***CO2in Ammonia***	***CO2in LignitePower***	***Phenol***	***Gas***
CO2inAmmonia	1	0	0	0
CO2inLignitePower	0	1	0	0
Phenol	0	0	1	0.9
Gas	0	0	0.9	1

References

(1) R. Heijungs, J. Guinée, and C. voor Milieukunde, *Environmental life cycle assessment of products: backgrounds-October, 1992.* Centre of Environmental Science, 1992.

(2) R. Heijungs and S. Suh, "Reformulation of matrix-based lci: from product balance to process balance," *Journal of Cleaner Production*, vol. 14, no. 1, pp. 47–51, 2006.

(3) R. Heijungs and S. Suh, *The computational structure of life cycle assessment*, vol. 11. Springer, 2002.

(4) R. Hischier, P. Wäger, and J. Gauglhofer, "Does weee recycling make sense from an environmental perspective?: The environmental impacts of the swiss take-back and recycling systems for waste electrical and electronic equipment (weee)," *Environmental Impact Assessment Review*, vol. 25, no. 5, pp. 525–539, 2005.

(5) I. Handbook, "Ilcd handbook: General guide for life cycle assessment-detailed guidance. 417 pages," *Téléchargeable à l'adresse: http://lct. jrc. ec. europa. eu/pdf-directory/ILCD-Handbook-Generalguide-for-LCA-DETAIL-online-12March2010. pdf*, 2010.

(6) A. Azapagic, *Environmental System Analysis: The Application of Linear Programming to Life Cycle Assessment Volume 1.* PhD thesis, University of Surrey, 1996.

(7) K. Pearson, "On a new method of determining correlation between a measured character a, and a character b, of which only the percentage of cases wherein b exceeds (or falls short of) a given intensity is recorded for each grade of a," *Biometrika*, vol. 7, no. 1/2, pp. 96–105, 1909.

(8) S. Wright, "Correlation and causation," *Journal of agricultural research*, vol. 20, no. 7, pp. 557–585, 1921.

(9) F. Esscher, "On a method of determining correlation from the ranks of the variates," *Scandinavian Actuarial Journal*, vol. 1924, no. 1, pp. 201–219, 1924.

(10) H. Hotelling, "Relations between two sets of variates," *Biometrika*, pp. 321–377, 1936.

(11) J. Wang, M.-P. G. Laborie, and M. P. Wolcott, "Correlation of mechanical and chemical cure development for phenol–formaldehyde resin bonded wood joints," *Thermochimica Acta*, vol. 513, no. 1, pp. 20–25, 2011.