GreenDelta
sustainability consulting + software

A procedure for LCI data gaps management applied for an aerospace case

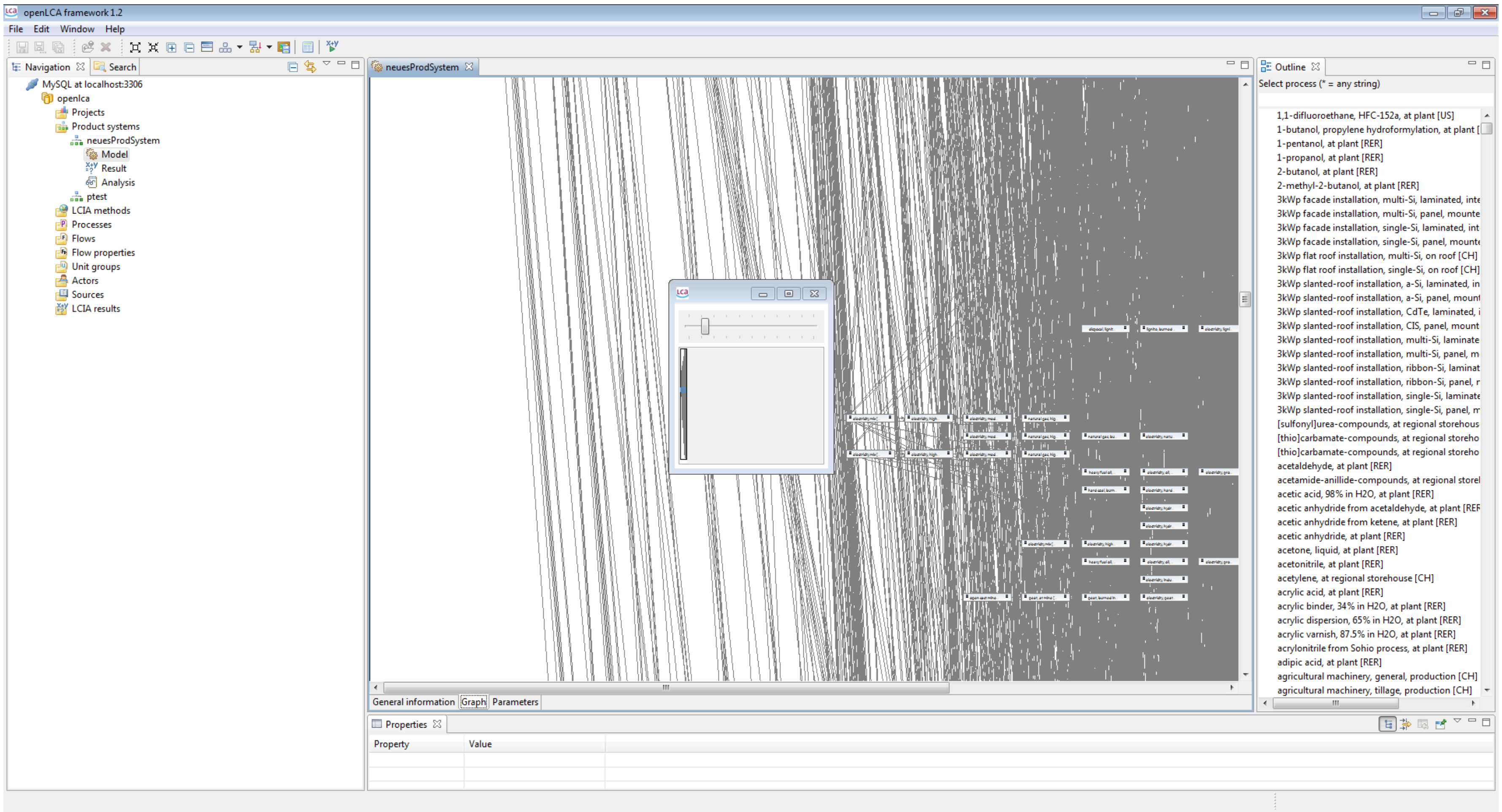
Andreas Ciroth, Cristina Rodríguez
GreenDelta GmbH

San Francisco, LCAXIV, Oct 8 2014

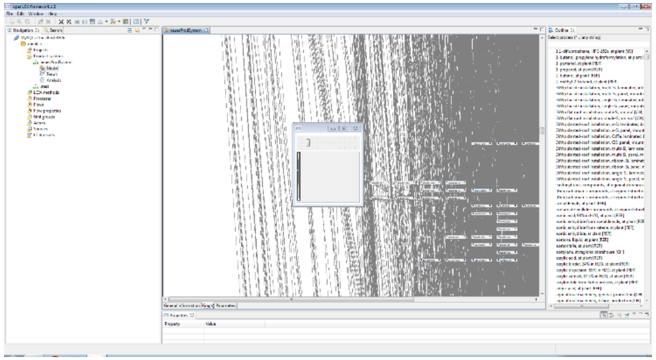
A procedure for LCI data gaps management
applied for an aerospace case

1 Problem setting

Problem setting



Problem setting



a) Larger and larger systems in Life Cycle Assessment studies today, with several thousand processes of mixed data quality

→ how is the overall quality, is it good enough, where to start to improve it

Problem setting

b) Data gaps in Life Cycle Assessment model (missing processes, relative to goal and scope system boundary) → where is it necessary to collect data

Problem setting

b) Data gaps in Life Cycle Assessment model (missing processes, relative to goal and scope system boundary) → where is it necessary to collect data

Since it is always possible to make a bad estimate, it is always possible to turn case b) into case a) (mixed data quality)

2 Concept and procedure for LCI data gaps management

A concept and procedure for LCI data gaps management

Basically, data gaps management in LCI is about managing data quality in LCI.

The procedure has two main elements:

a) Ensuring that in a life cycle model, minimum data quality requirements are met (optional)

b) Quantifying data quality and data quality contributions of all processes in the life cycle model to identify and manage where and if data quality improvements are necessary

A concept and procedure for LCI data gaps management

Basically, data gaps management in LCI is about managing data quality in LCI.

The procedure uses the pedigree matrix for data quality assessment

A concept and procedure for LCI data gaps management – pedigree matrix

Weidema and Wesnaes introduced the pedigree matrix into LCA in 1996, for quality assurance,

And it is in use in the ecoinvent database for generating uncertainty information for process flows

Indicator score	**1**	**2**	**3**	**4**	**5** (default)
Reliability	Verified[3] data based on measurements[4]	Verified data partly based on assumptions ***or*** non-verified data based on measurements	Non-verified data partly based on qualified estimates	Qualified estimate (e.g. by industrial expert)	Non-qualified estimate
Completeness	Representative data from all sites relevant for the market considered, over an adequate period to even out normal fluctuations	Representative data from >50% of the sites relevant for the market considered, over an adequate period to even out normal fluctuations	Representative data from only some sites (<<50%) relevant for the market considered ***or*** >50% of sites but from shorter periods	Representative data from only one site relevant for the market considered ***or*** some sites but from shorter periods	Representativeness unknown or data from a small number of sites ***and*** from shorter periods
Temporal correlation	Less than 3 years of difference to the time period of the dataset	Less than 6 years of difference to the time period of the dataset	Less than 10 years of difference to the time period of the dataset	Less than 15 years of difference to the time period of the dataset	Age of data unknown or more than 15 years of difference to the time period of the dataset
Geographical correlation	Data from area under study	Average data from larger area in which the area under study is included	Data from area with similar production conditions	Data from area with slightly similar production conditions	Data from unknown ***or*** distinctly different area (North America instead of Middle East, OECD-Europe instead of Russia)
Further technological correlation	Data from enterprises, processes and materials under study	Data from processes and materials under study (i.e. identical technology) but from different enterprises	Data from processes and materials under study but from different technology	Data on related processes or materials	Data on related processes on laboratory scale ***or*** from different technology

GreenDelta

A concept and procedure for LCI data gaps management – pedigree matrix

In ecoinvent 3: five data quality indicators, which are assessed in five scores, from 1 to 5 (bad)

Indicator score	**1**	**2**	**3**	**4**	**5** (default)
Reliability	Verified[3] data based on measurements[4]	Verified data partly based on assumptions ***or*** non-verified data based on measurements	Non-verified data partly based on qualified estimates	Qualified estimate (e.g. by industrial expert)	Non-qualified estimate
Completeness	Representative data from all sites relevant for the market considered, over an adequate period to even out normal fluctuations	Representative data from >50% of the sites relevant for the market considered, over an adequate period to even out normal fluctuations	Representative data from only some sites (<<50%) relevant for the market considered ***or*** >50% of sites but from shorter periods	Representative data from only one site relevant for the market considered ***or*** some sites but from shorter periods	Representativeness unknown or data from a small number of sites ***and*** from shorter periods
Temporal correlation	Less than 3 years of difference to the time period of the dataset	Less than 6 years of difference to the time period of the dataset	Less than 10 years of difference to the time period of the dataset	Less than 15 years of difference to the time period of the dataset	Age of data unknown or more than 15 years of difference to the time period of the dataset
Geographical correlation	Data from area under study	Average data from larger area in which the area under study is included	Data from area with similar production conditions	Data from area with slightly similar production conditions	Data from unknown ***or*** distinctly different area (North America instead of Middle East, OECD-Europe instead of Russia)
Further technological correlation	Data from enterprises, processes and materials under study	Data from processes and materials under study (i.e. identical technology) but from different enterprises	Data from processes and materials under study but from different technology	Data on related processes or materials	Data on related processes on laboratory scale ***or*** from different technology

GreenDeLTa

A concept and procedure for LCI data gaps management – pedigree matrix

In ecoinvent 3: five data quality indicators, which are assessed in five scores, from 1 to 5 (bad, default)

Indicator score	**1**	**2**	**3**	**4**	**5** (default)
Reliability	Verified[3] data based on measurements[4]	Verified data partly based on assumptions *or* non-verified data based on measurements	Non-verified data partly based on qualified estimates	Qualified estimate (e.g. by industrial expert)	Non-qualified estimate

A concept and procedure for LCI data gaps management – pedigree matrix

In ecoinvent 3: five data quality indicators, which are assessed in five scores, from 1 to 5 (bad, default)

Indicator
Reliability
Completeness
Temporal correlation
Geographical correlation
Further technological correlation

A concept and procedure for LCI data gaps management – pedigree matrix

In ecoinvent 3: uncertainty information is provided for the flows, based on the assessment scores

Indicator score	1	2	3	4	5
Reliability	1.00	1.05	1.10	1.20	1.50
Completeness	1.00	1.02	1.05	1.10	1.20
Temporal correlation	1.00	1.03	1.10	1.20	1.50
Geographical correlation	1.00	1.01	1.02		1.10
Further technological correlation	1.00		1.20	1.50	2.00

"Default uncertainty factors (contributing to the square of the geometric standard deviation) applied together with the pedigree matrix", (Frischknecht, Jungbluth 2004 p 46)
Currently: expert judgement.

A concept and procedure for LCI data gaps management – pedigree matrix

In ecoinvent 3: uncertainty information is provided for the flows, based on the assessment scores

Indicator score	1	2	3	4	5
Reliability	1.00	1.05	1.10	1.20	1.50
Completeness	1.00	1.02	1.05	1.10	1.20
Temporal correlation	1.00	1.03	1.10	1.20	1.50
Geographical correlation	1.00	1.01	1.02		1.10
Further technological correlation	1.00		1.20	1.50	2.00

“Default uncertainty factors (contributing to the square of the geometric standard deviation) applied together with the pedigree matrix“, (Frischknecht, Jungbluth 2004 p 46)
Currently: expert judgement.

A concept and procedure for LCI data gaps management – pedigree matrix

In ecoinvent 3: uncertainty information is provided for the flows, based on the assessment scores

Indicator	1	2	3	4	5
Reliability	1	1.54*	1.61	1.69	(n.a.)
Completeness	1	1.03	1.04	1.08	(n.a.)
Temporal correlation	1	1.03	1.10	1.19	1.29
Geographical correlation	1	1.04	1.08	1.11	(n.a.)
Further technological correlation	1	1.18	1.65	2.08	2.80

*interim

Empirically based uncertainty factors (contributing to the square of the geometric standard deviation) for the pedigree matrix in ecoinvent, (Ciroth, Muller, Weidema Lesage 2013)
Based on empirical study

A concept and procedure for LCI data gaps management – pedigree matrix

In ecoinvent 3: uncertainty information is provided for the flows, based on the assessment scores

Indicator	1	2	3	4	5
Reliability	1	1.54*	1.61	1.69	(n.a.)
Completeness	1	1.03	1.04	1.08	(n.a.)
Temporal correlation	1	1.03	1.10	1.19	1.29
Geographical correlation	1	1.04	1.08	1.11	(n.a.)
Further technological correlation	1	1.18	1.65	2.08	2.80

*interim

Empirically based uncertainty factors (contributing to the square of the geometric standard deviation) for the pedigree matrix in ecoinvent, (Ciroth, Muller, Weidema Lesage 2013)
Based on empirical study

A concept and procedure for LCI data gaps management

Some first components of the procedure:

- Data gaps are interpreted as data quality gaps
- Data quality is assessed according to the data quality indicators in the pedigree matrix
- Data gaps are scored as 5, default
- Uncertainty from the pedigree matrix is used to estimate the relevance of the “imperfect” data quality
- Minimum requirements can be specified for each data quality indicator, for all indicators, or for the uncertainty. E.g.: no indicator >4.

3 Application in a satellite LCA case study

A satellite LCA case study

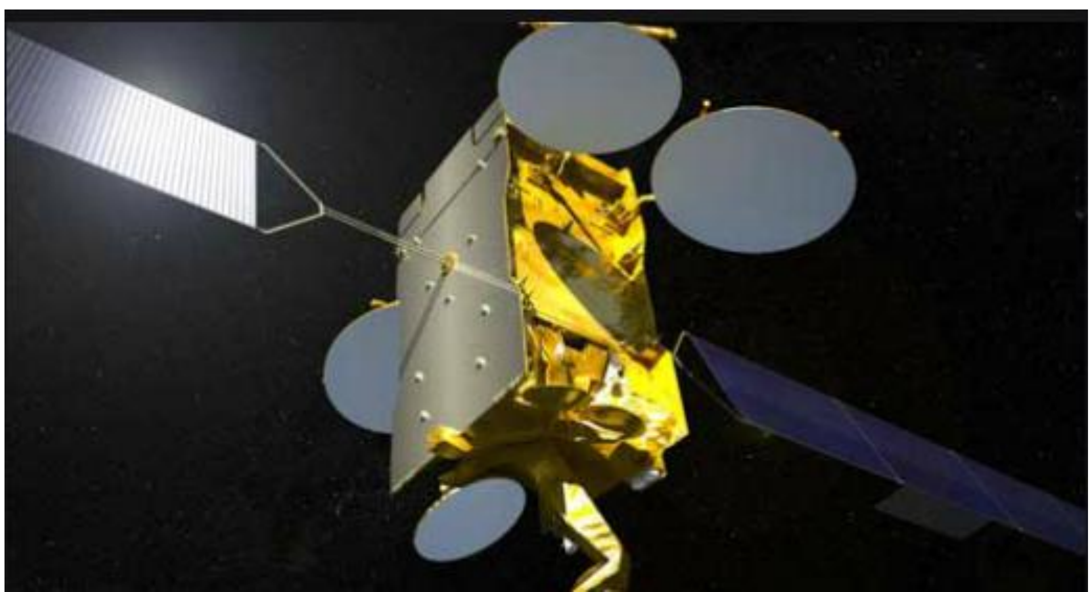
- Project commissioned by esa, European Space Agency, 2012
- Consortium:
  - **DAPP**: Italian engineering consulting company
  - **POLITO**: Italian technical university
  - **EMPA**: Swiss centre for material sciences and technology development
  - **GREENDELTA**: German sustainability consulting and software company

A satellite LCA case study

- Goal of the project: Develop a methodology and LCA model for two satellite types
  - → highly sensitive information
  - → few data available, many "uncommon" processes
  - → Idea for data gaps management in this project, by presenter; project applied an initial version

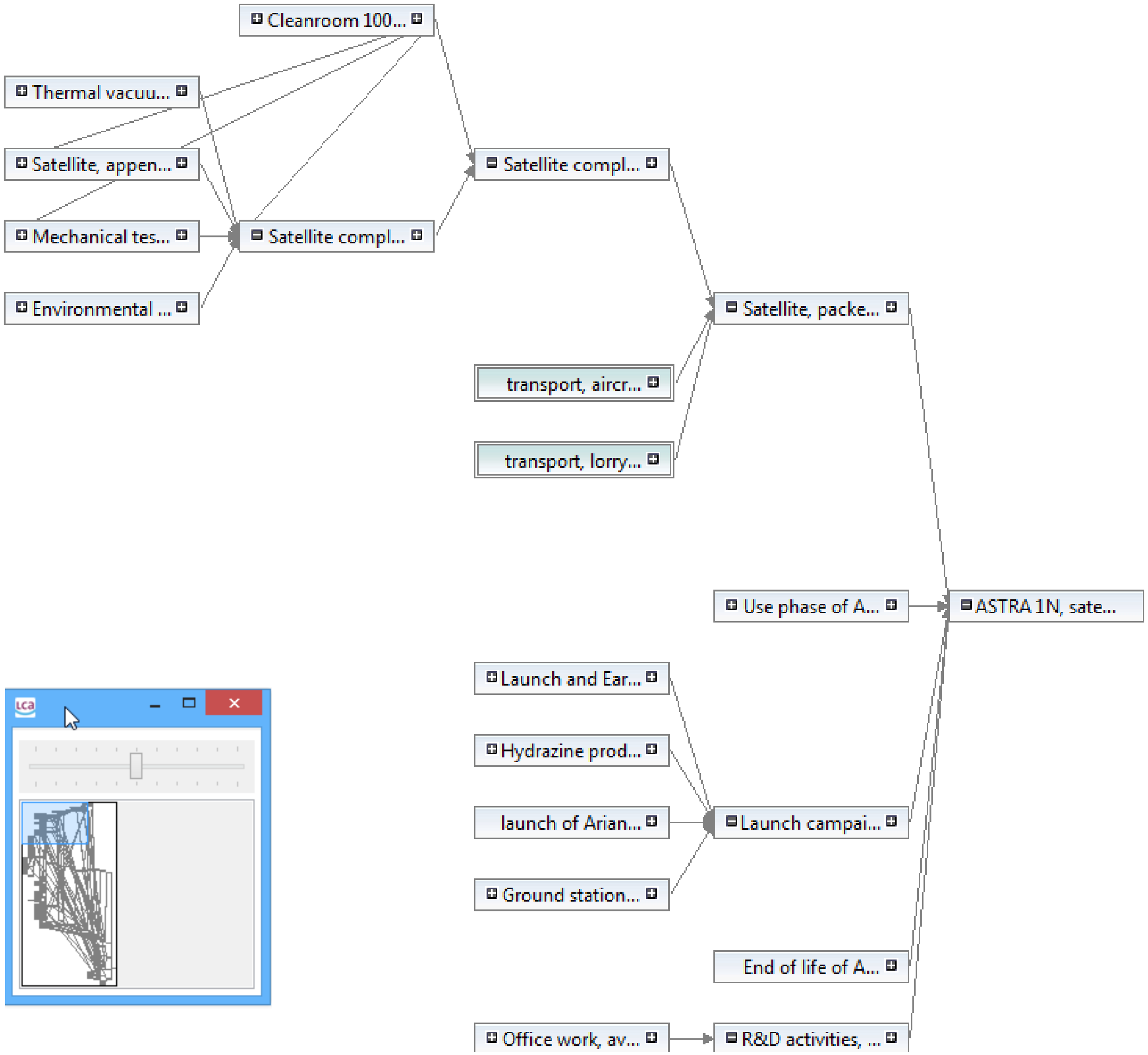
A satellite LCA case study: Astra 1N

- ASTRA 1N is a current satellite in use for digital and HD telecommunications
- ASTRA 1N is a Eurostar E3000 model with launch mass of 5400 kg and spacecraft power of more than 13 kW

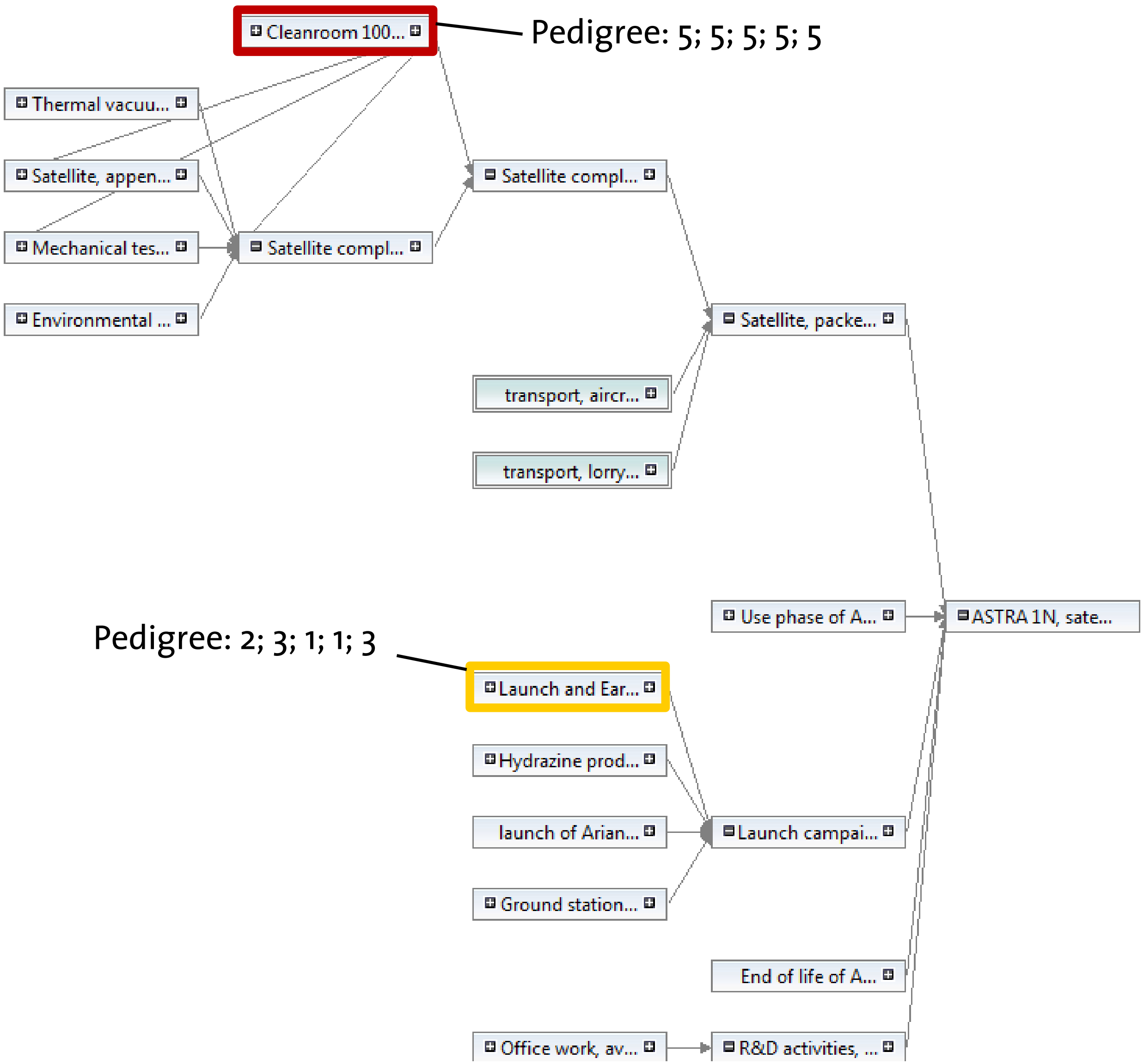


GreenDelta

A satellite LCA case study: Astra 1N

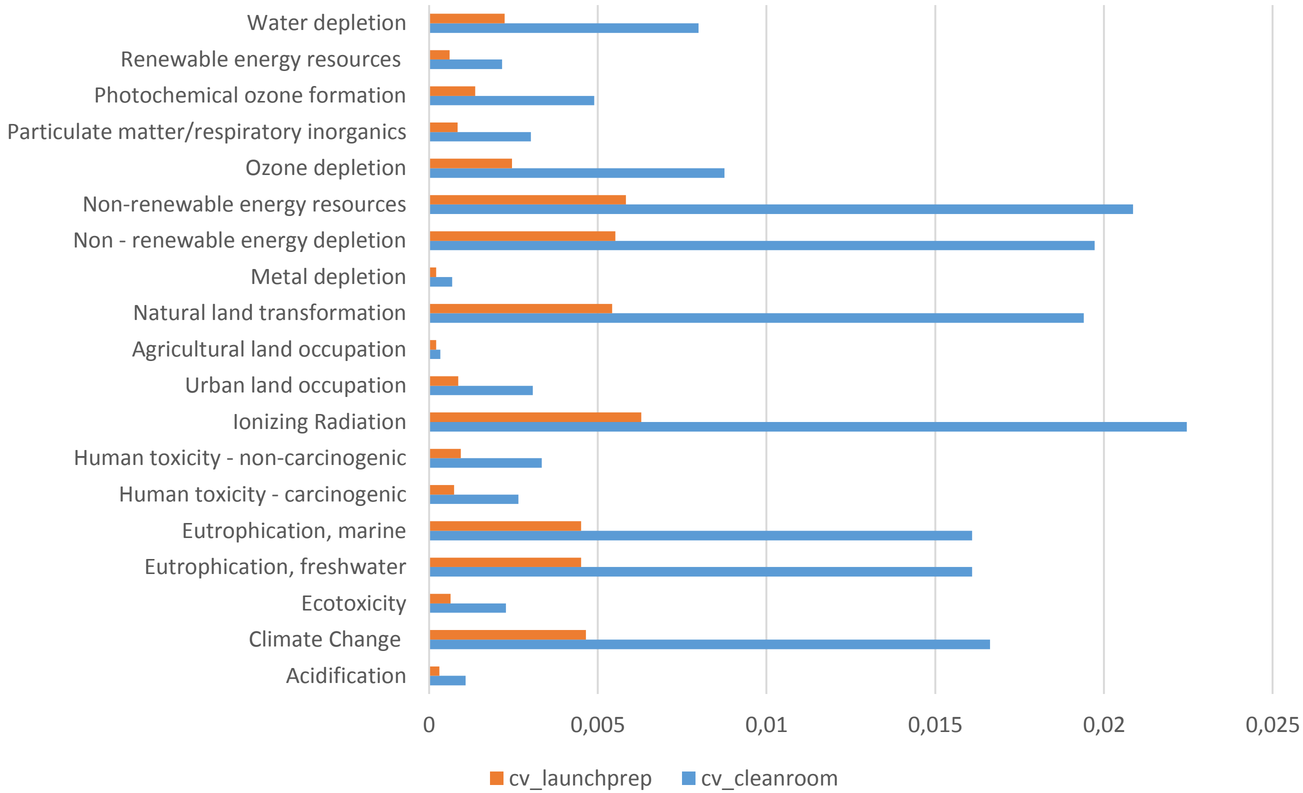


Astra 1N (data quality mock up values)



Indicator
Reliability
Completeness
Temporal correlation
Geographical correlation
Further technological correlatio

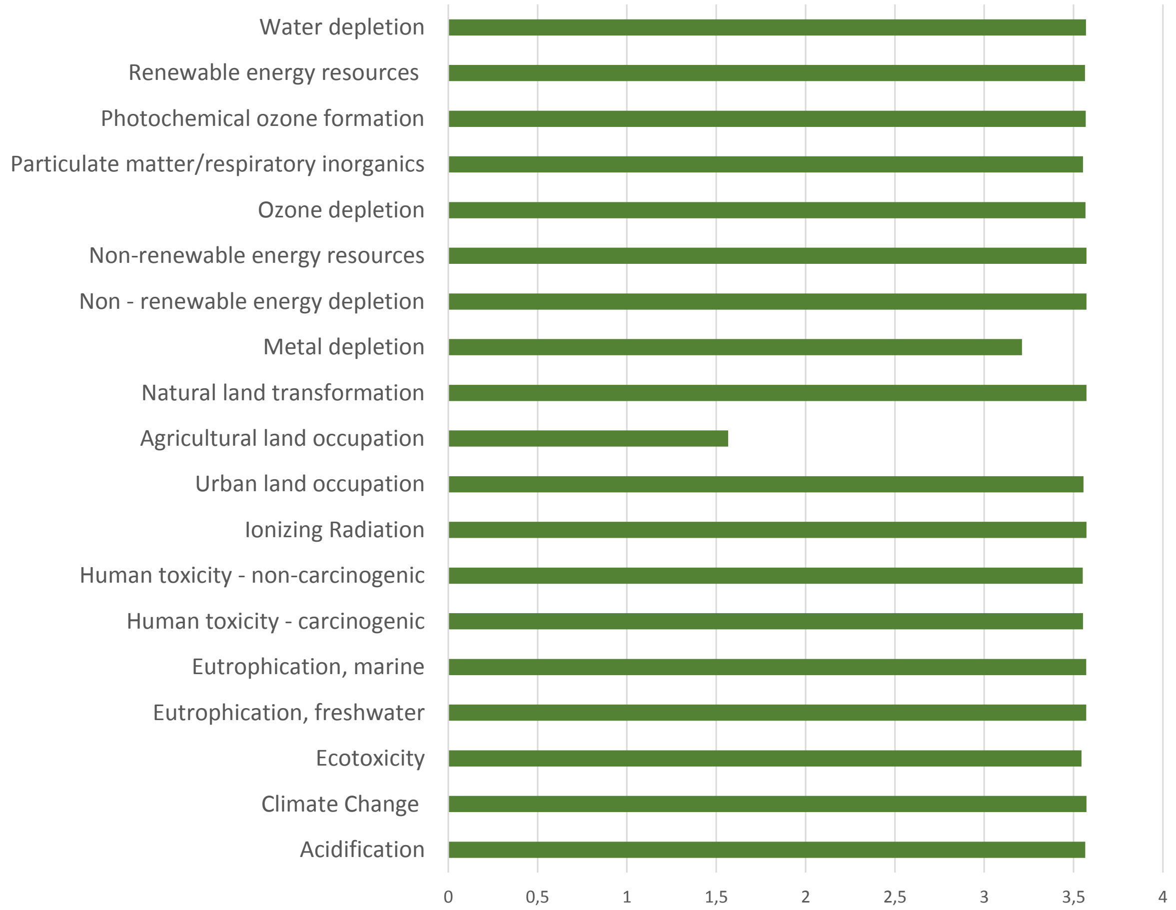
A satellite LCA case study: pedigree-induced variation coefficients in end-result for cleanroom process and for launch-preparation



GreenDelta

Pedigree-induced variation coefficients in end-result for cleanroom process and for launch-preparation:

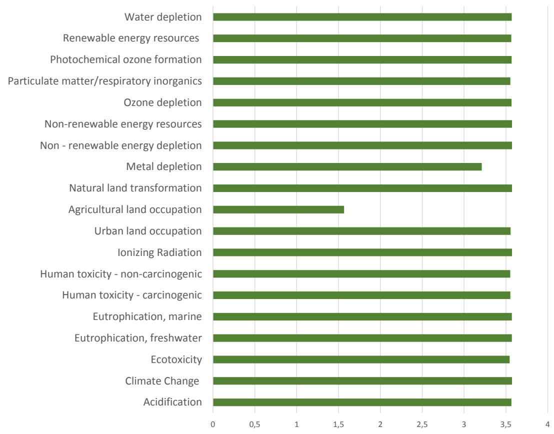
Ratio CV cleanroom / CV launch prep.



GreenDelta

A satellite LCA case study: Astra 1N

→ Uncertainty caused by the omission of cleanroom construction process in the system is about 3.5 times higher than uncertainty caused by an imperfectly modelled launch preparation.



4 Discussion and outlook

Discussion and outlook

An approach has been presented to estimate the relevance of data gaps and also of imperfect data quality in the inventory,

...by interpreting data gaps as presence of processes of bad data quality. Processes need to be roughly estimated for this.

Discussion and outlook

The approach uses the pedigree matrix uncertainties, as proposed by the ecoinvent database; the calculated uncertainties serve as a proxy for data quality.

The approach offers a structured procedure for focussing data collection effort on where it is relevant in an LCA study.

Discussion and outlook

Several aspects still can be refined:

Calculating the uncertainties: instead of MC simulation, variance contribution algorithms which are not yet available in LCA software

The pedigree uncertainty factors in ecoinvent: the traditional values are lower than those in an empirical study especially for medium data quality

GreenDelta

sustainability consulting + software

Thank you very much.

Contact: Dr. Andreas Ciroth

GreenDelta GmbH

Müllerstrasse 135, 13349 Berlin, Germany

ciroth@greendelta.com

www.greendelta.com

A concept and procedure for LCI data gaps management: "data quality propagation"

Since data quality is seen as uncertainty, the "propagation" of data quality to the end result of the Life Cycle Inventory is available.

The result can be the LCI amount per elementary flow, or the relative amount in a comparison

A concept and procedure for LCI data gaps management: "data quality propagation"

Since data quality is seen as uncertainty, the "propagation" of data quality to the end result of the Life Cycle Inventory is available.

→ You are able to see how relevant good data quality in a specific process is!

→ And you are also able to see where it makes sense to focus efforts.

A concept and procedure for LCI data gaps management:
"data quality propagation": tools

Monte Carlo simulation is able to calculate the uncertainty propagation perfectly

- available in LCA software tools
- after sufficient runs
- mainly the overall uncertainty is calculated, not individual contributions

A concept and procedure for LCI data gaps management:
"data quality propagation": tools

Taylor series and perturbation analysis: Ciroth 2004, Heijungs 2010

Variance contribution analysis: Sobol 2001, Saltelli 2010

Ciroth 2004: Ciroth, A., Fleischer, G., Steinbach, J. (2004): Uncertainty Calculation in Life Cycle Assessments - A Combined Model of Simulation and Approximation, in: The Int J of LCA 9 (4) S. 216 - 226 (2004).

Heijungs 2010: Heijungs, R.: Sensitivity coefficients for matrix-based LCA, Int J Life Cycle Assess (2010) 15:511–520

Sobol 2001: Sobol, I.: Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, Mathe. Comput. Simul., 55, 271–280, 2001.

Andrea Saltelli, Paola Annoni, Ivano Azzini, Francesca Campolongo, Marco Ratto, and Stefano Tarantola. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. Computer Physics Communications, 181(2):259-270, 2010

A concept and procedure for LCI data gaps management:
"data quality propagation": tools

Final remark: if the data quality has been changed in one process, the contributions to data quality of this process and of other processes can change

→ The application is iterative.