GreenDelta

sustainability consulting + software

Refining the pedigree matrix approach in ecoinvent: Towards empirical uncertainty factors

Andreas Ciroth
GreenDelta

LCA Discussion Forum Zürich, September 13, 2013

Refining the pedigree matrix approach in ecoinvent: Towards empirical uncertainty factors

1 The pedigree matrix approach in ecoinvent

2 Uncertainty factors in the pedigree matrix: Current state

3 Empirically founded uncertainty factors:
- **Approach**
- **Results**
- **Next steps**

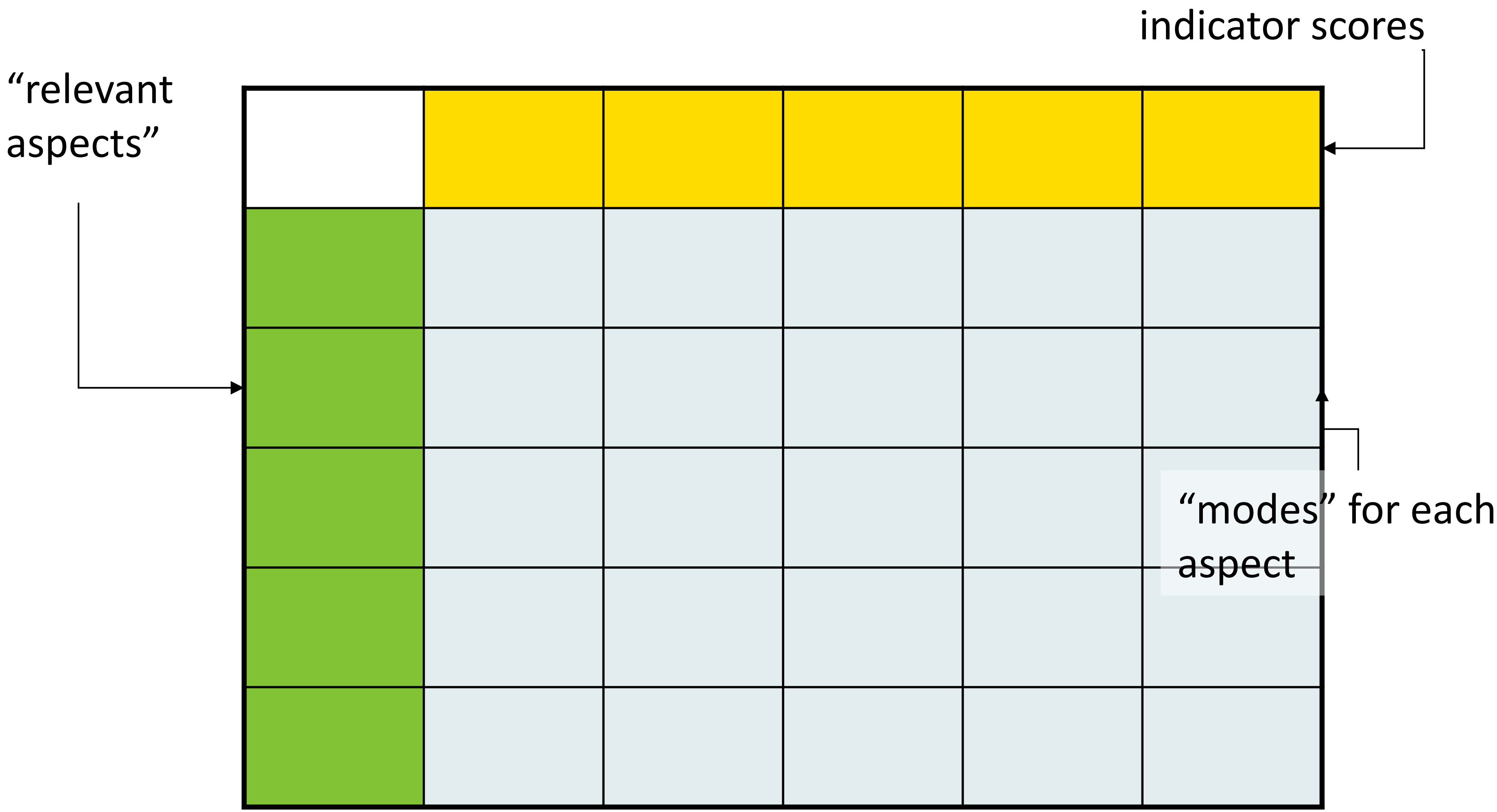
4 Discussion

1 The pedigree approach in ecoinvent

1 The pedigree matrix approach in ecoinvent

- Originating from Funtowicz & Ravetz (1990), as part of their NUSAP scheme for managing "all sorts of uncertainty"
- A pedigree expresses key components by means of a matrix. Its columns are basic aspects or "phases" and its lines qualitative "modes" of each aspect expressing different degrees of data quality or uncertainty
- Qualitative modes can be assigned to quantitative "codes" 1, 2, 3, .. . The lower the code the better.
- Pedigree matrix concept was transferred to environm. assessment by Weidema/Wesnaes in 1996

1 The pedigree matrix approach in ecoinvent



GreenDeLTa

The current matrix in ecoinvent 3

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

The current matrix in ecoinvent 3: Reliability (of the data source)

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

2 Uncertainty factors in the ecoinvent pedigree matrix: Current state

Uncertainty factors for the pedigree matrix 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
Sample size	1.00	1.02	1.05	1.10	1.20

Reliability: U1, Completeness: U2, asf.

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

GreenDeLTa

Basic uncertainty factors

input / output group	c	p	a
demand of:			
thermal energy, electricity, semi-finished products, working material, waste treatment services	1.05	1.05	1.05
transport services (tkm)	2.00	2.00	2.00
Infrastructure	3.00	3.00	3.00
resources:			
primary energy carriers, metals, salts	1.05	1.05	1.05
land use, occupation	1.50	1.50	1.10
land use, transformation	2.00	2.00	1.20
pollutants emitted to water:			
BOD, COD, DOC, TOC, inorganic compounds (NH_4, PO_4, NO_3, Cl, Na etc.)		1.50	
individual hydrocarbons, PAH		3.00	
heavy metals		5.00	1.80
pesticides			1.50
NO_3, PO_4			1.50
pollutants emitted to soil:			
oil, hydrocarbon total		1.50	
heavy metals		1.50	1.50
pesticides			1.20

input / output group	c	p	a
pollutants emitted to air:			
CO_2	1.05	1.05	
SO_2	1.05		
NMVOC total	1.50		
NO_X, N_2O	1.50		1.40
CH_4, NH_3	1.50		1.20
individual hydrocarbons	1.50	2.00	
PM>10	1.50	1.50	
PM10	2.00	2.00	
PM2.5	3.00	3.00	
polycyclic aromatic hydrocarbons (PAH)	3.00		
CO, heavy metals	5.00		
inorganic emissions, others		1.50	
radionuclides (e.g., Radon-222)		3.00	

"Ub"

GreenDelta

Uncertainty factors contribute directly to quantitative uncertainty

$$SD_{g95} := \sigma_g{}^2 = \exp^{\sqrt{[\ln(U_1)]^2+[\ln(U_2)]^2+[\ln(U_3)]^2+[\ln(U_4)]^2+[\ln(U_5)]+[\ln(U_b)]^2}}$$

with :

U_1 : uncertainty factor of reliability

U_2 : uncertainty factor of completeness

U_3 : uncertainty factor of temporal correlation

U_4 : uncertainty factor of geographic correlation

U_5 : uncertainty factor of other technological correlation

U_b : basic uncertainty factor

Uncertainty factors contribute directly to quantitative uncertainty

$$SD_{g95} := \sigma_g{}^2 = \exp^{\sqrt{[\ln(U_1)]^2+[\ln(U_2)]^2+[\ln(U_3)]^2+[\ln(U_4)]^2+[\ln(U_5)]+[\ln(U_b)]^2}}$$

with :

U_1 : uncertainty factor of reliability

U_2 : uncertainty factor of completeness

U_3 : uncertainty factor of temporal correlation

U_4 : uncertainty factor of geographic correlation

U_5 : uncertainty factor of other technological correlation

U_b : basic uncertainty factor

Uncertainty factors contribute directly to quantitative uncertainty

$$SD_{g95} := \sigma_g{}^2 = \exp^{\sqrt{[\ln(U_1)]^2+[\ln(U_2)]}}$$

with :

Geometric standard deviation of a flow.

→ Project commissioned by ecoinvent:
(try to)
Provide an empirical basis
for the uncertainty factors

3 Empirically founded uncertainty factors

3 Definition of terms: Uncertainty

"Uncertainty means, basically, lack of certainty. "

3 Definition of terms: Uncertainty

A quantitative figure for the emission of a flow is not exactly known;

the correct allocation method for a multi output process is not exactly known;

it is unclear whether electric arc furnace steel should be included in a product system, or converter steel:

all these situations "contain" uncertainty

GreenDeLTa

3 Definition of terms: Uncertainty

The lack of certainty depends on the level of detail that is taken into account.

3 Definition of terms: Uncertainty

An example: The amount of fertiliser used by farmers.

With data sets for several farmers, over a certain time interval, the amount will vary, and the exact amount used in a specific farm will not be known precisely. The amount of fertiliser used is uncertain.

Image: James T M Towill, CC licence

GreenDelta

3 Definition of terms: Uncertainty, farmer example

This uncertainty will be lower, if we know in addition

- the time interval covered
- the size of the farms
- the type of farm, their products
- the geographical area where the farm is located
- the (micro-)climate where the farm is located
- the management type of the farm (organic farming e.g.)
- the farming background and expertise of the farmers
- asf.

Image: James T M Towill, CC licence

3 Definition of terms: Empirical

"Empirical: Derived from experiment and observation rather than theory and expert guesses."

3 Approach

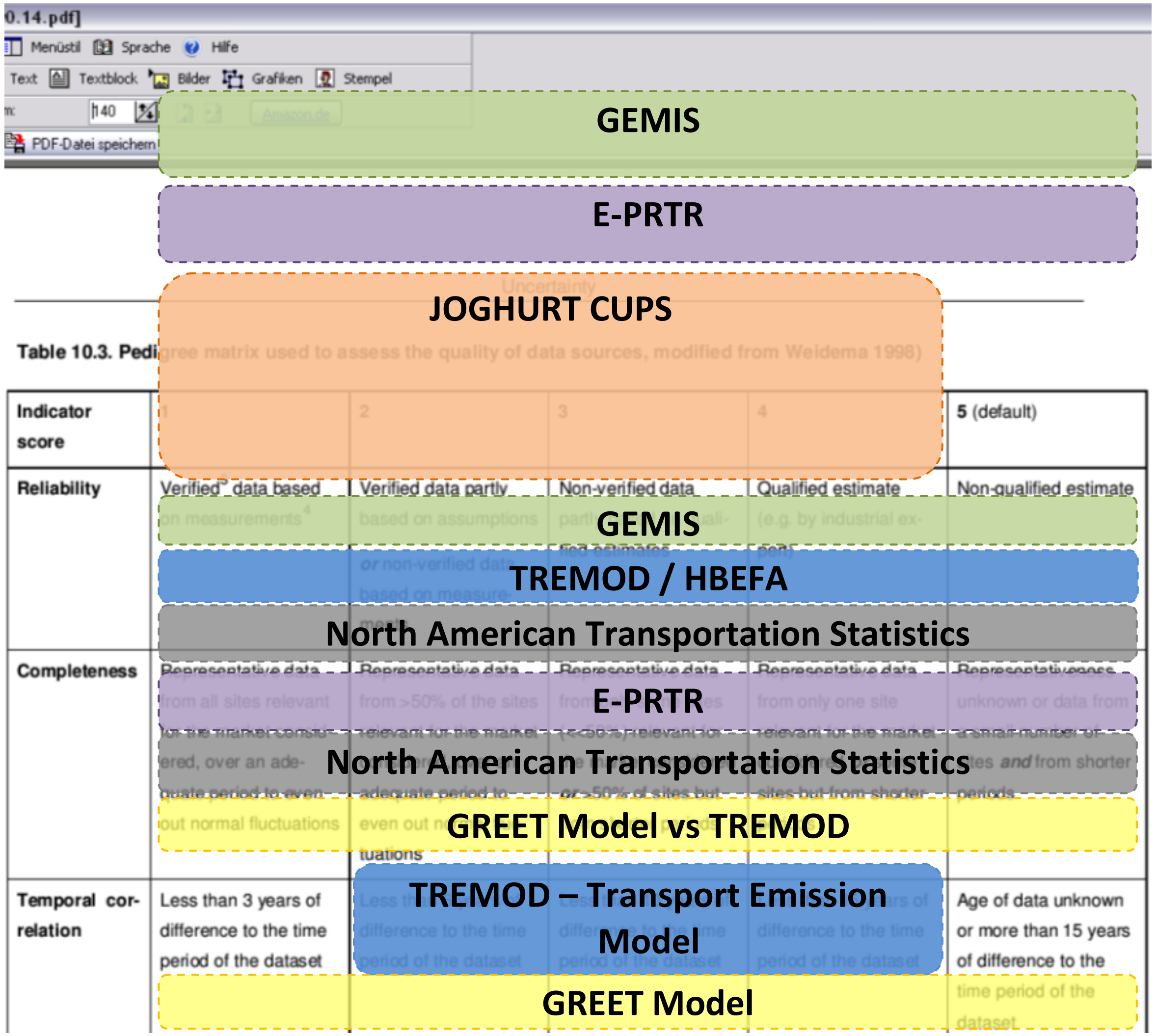
- Data from different sources analysed, from LCA and non-LCA sources. Data must not be related to the ecoinvent database.
- The pedigree parameters are then "relaxed", i.e. made less precise, and the resulting uncertainty in data is investigated
- Resulting uncertainty is the ratio of the geometric standard deviation (GSD) of relaxed to ideal sample

3 Approach

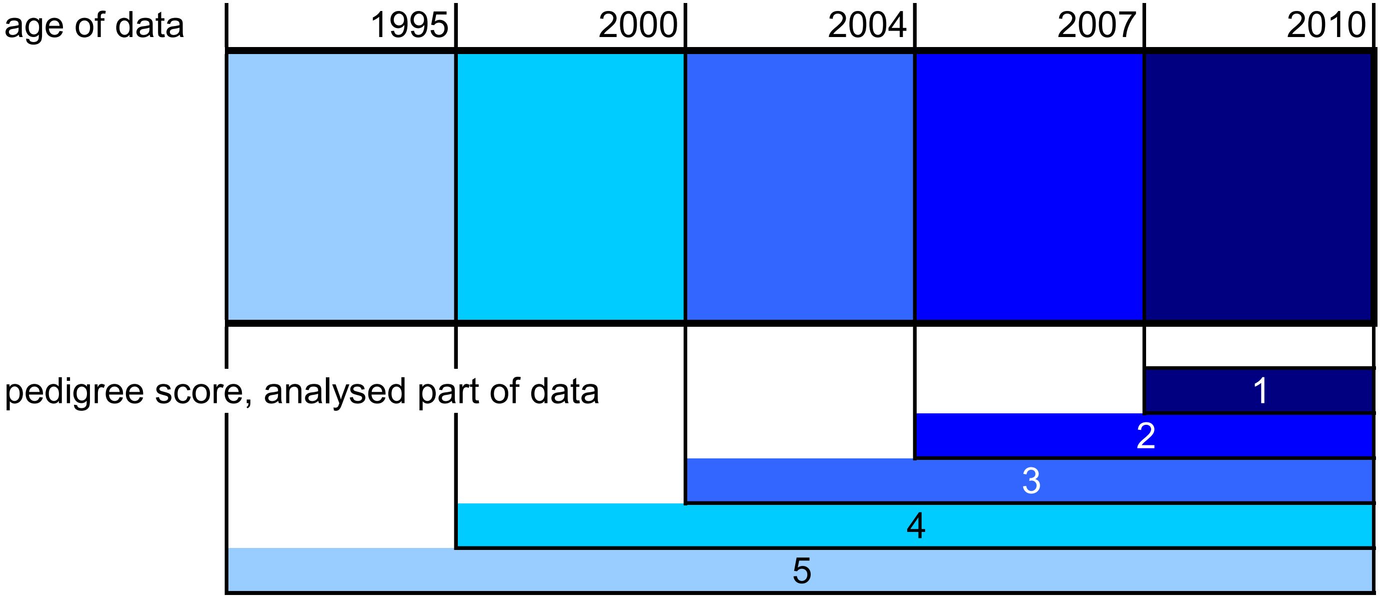
- Data from different sources analysed, from LCA and non-LCA sources. Data must not be related to the ecoinvent database.
- The pedigree parameters are then “relaxed”, i.e. made less precise, and the resulting uncertainty in data is investigated
- Resulting uncertainty is the ratio of the geometric standard deviation (GSD) of relaxed to ideal sample

(motivation: there is a true uncertainty in the sample, due to other things than the investigated score; the ratio of GSD expresses the “uncertainty difference” of the less ideal sample to the ideal sample)

3 Approach: Data sources



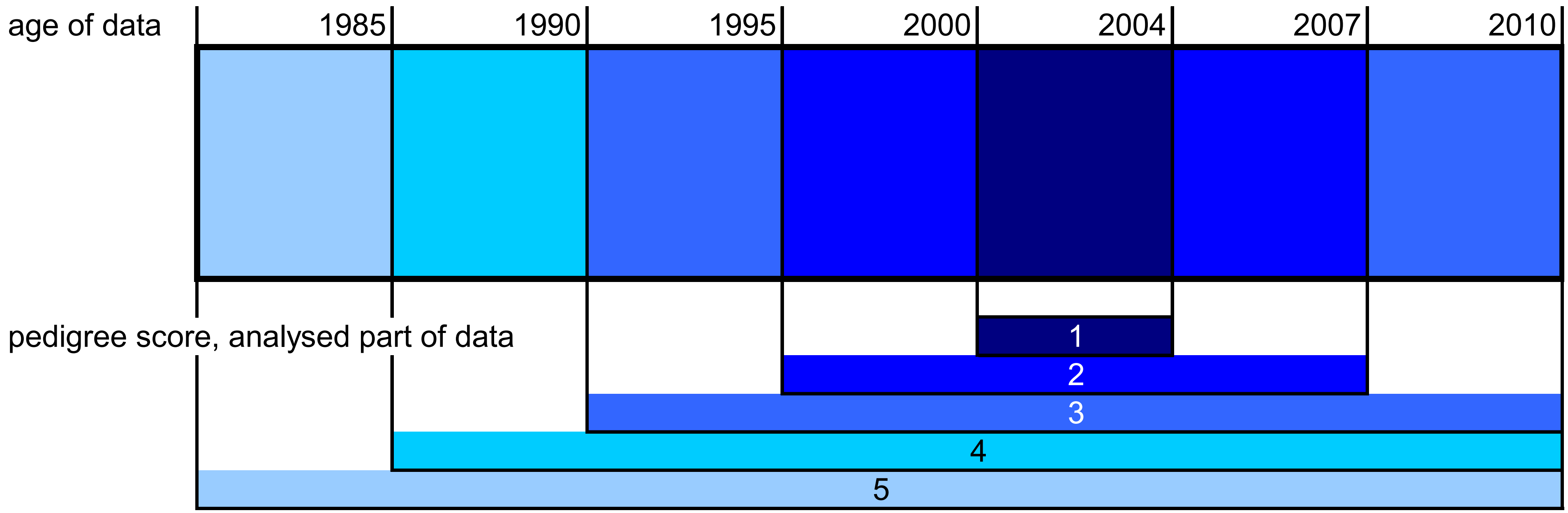
3 Approach: For example time



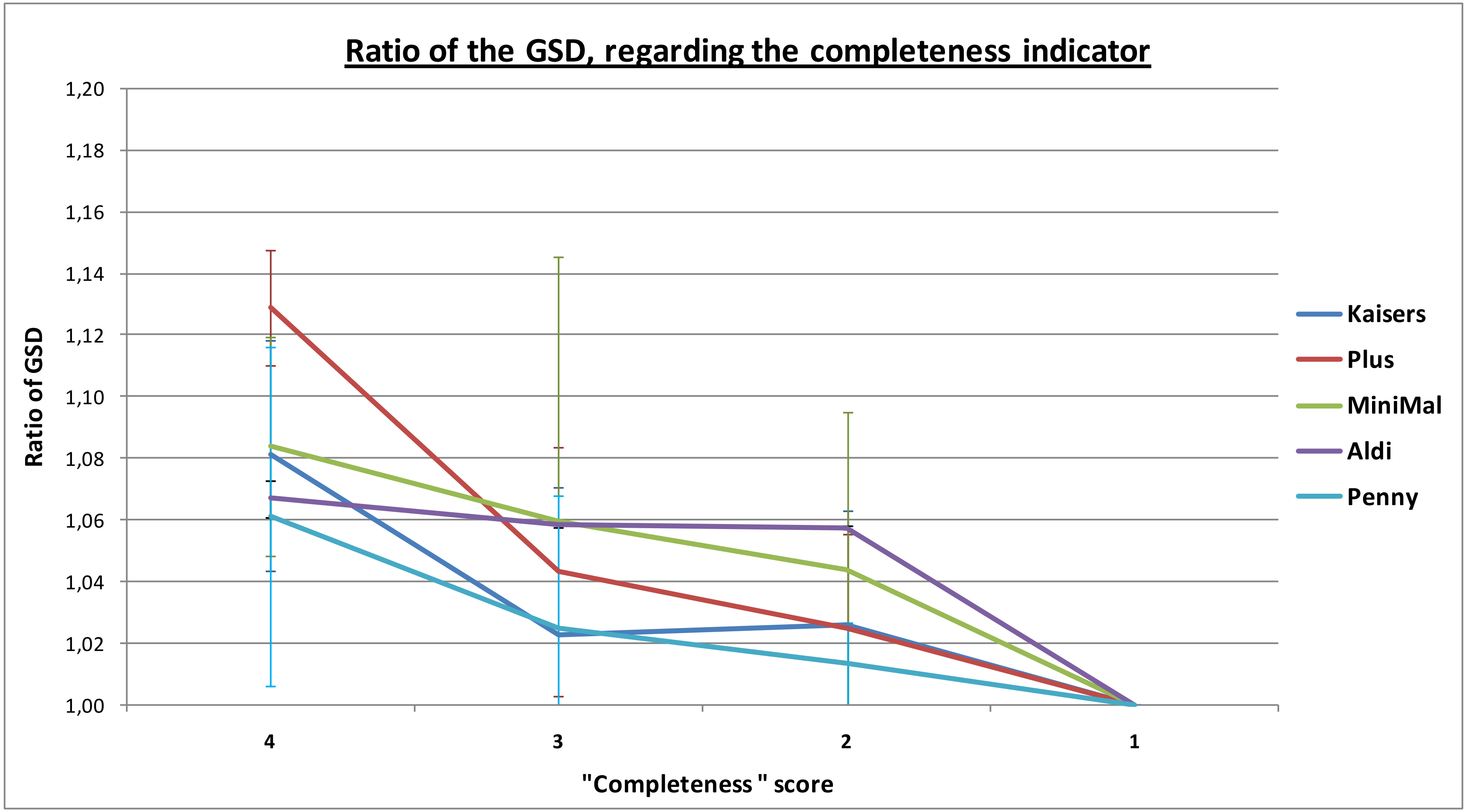
pedigree score definitions

1	data less than 3 years difference to the time period of the data set
2	data less than 6 years difference to the time period of the data set
3	data less than 10 years difference to the time period of the data set
4	data less than 15 years difference to the time period of the data set
5	age of data unknown or more than 15 years difference to the time period of the data set

3 Approach: For example time



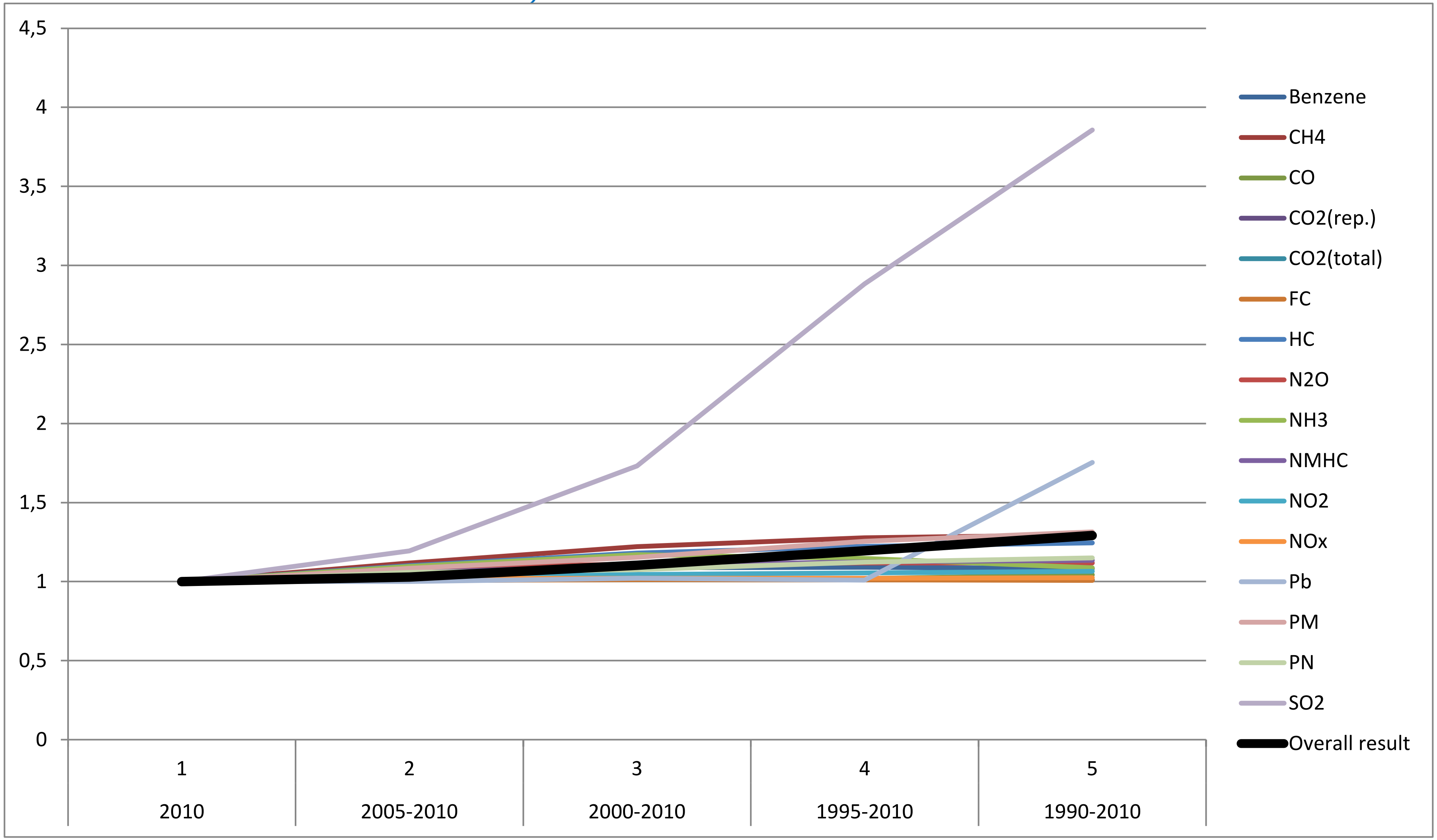
3 Results



Source: Refining the pedigree matrix approach in ecoinvent
Andreas Ciroth, With contributions from Stéphanie Muller and Bo Weidema
May 2012, final draft report

GreenDeLTa

3 Results: Time, Tremod database



Source: Refining the pedigree matrix approach in ecoinvent
Andreas Ciroth, With contributions from Stéphanie Muller and Bo Weidema
May 2012, final draft report

GreenDeLTa

3 Results: Geography, different sources

Indicator value	Tremod / GREET	North American Transport Statistics Database	PRTR
1	1	1	1
2	(n.a.)	1,159084043	1,043919013
3	1,020439873* / 1,032117664**	1,482781663	1,082233009
4	(n.a.)	(n.a.)	1,105217922
5	(n.a.)	(n.a.)	(n.a.)

Comparison of obtained GSD contributions for the indicator geographical correlation in the pedigree matrix
* with Tremod as reference
**with GREET as reference

Source: Refining the pedigree matrix approach in ecoinvent
Andreas Ciroth, With contributions from Stéphanie Muller and Bo Weidema
May 2012, final draft report

3 Results: Summary of uncertainty factors for ecoinvent

Indicator score	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

Source: Refining the pedigree matrix approach in ecoinvent
Andreas Ciroth, With contributions from Stéphanie Muller and Bo Weidema
May 2012, final draft report

GreenDeLTa

3 Results: Summary of uncertainty factors for ecoinvent

- Basic uncertainty: Previous factors are taken (not very large difference to new data; more investigations needed to really change the previous data sets)

4 Discussion

4 Discussion

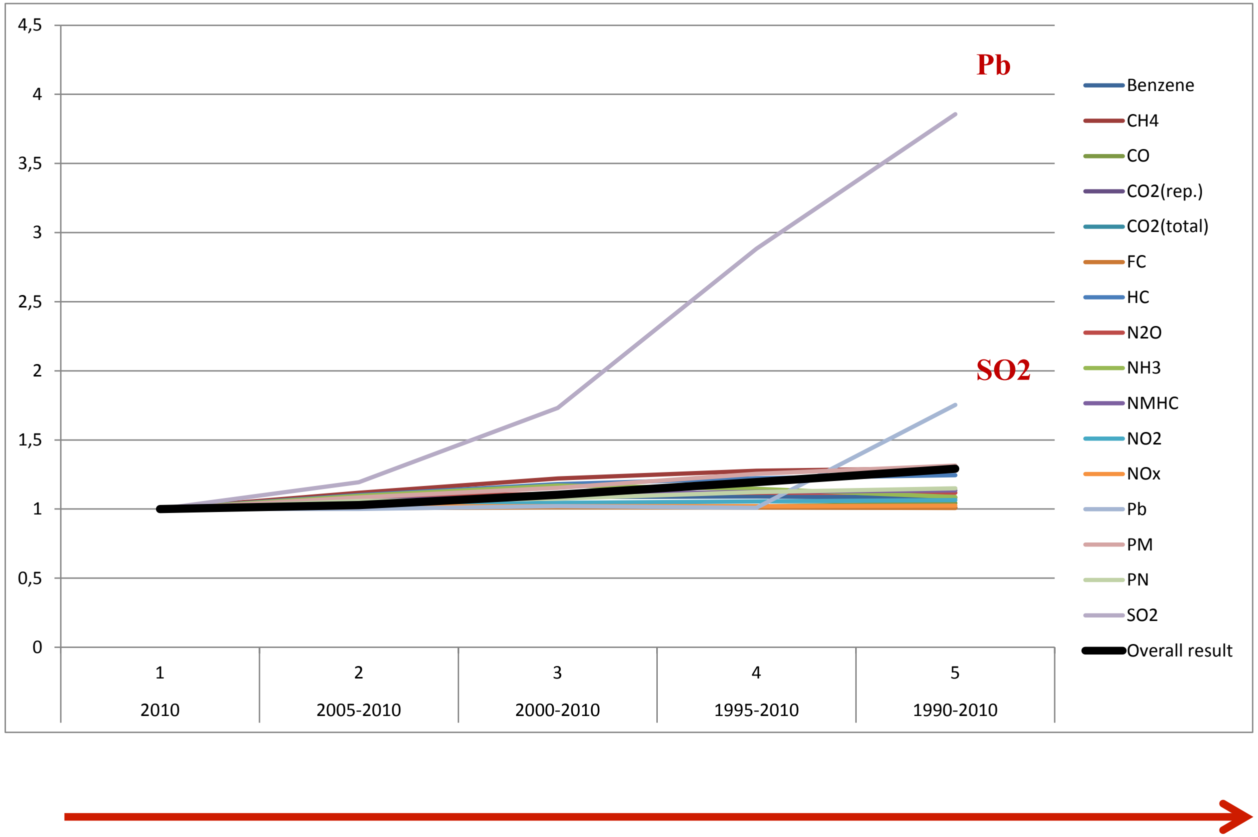
- It was indeed possible to obtain uncertainty factors based on empirical data
- With these factors, the whole generic uncertainty assessment in ecoinvent is put on a better founded basis
- The identified factors are different, but not *very* different, from previous ecoinvent factors

4 Discussion

- However, several aspects deserve further attention, e.g.:
  - Uncertainty distribution,
  - general limitations of the generic factor concept,
  - factor / indicator dependency,
  - basic uncertainty factors.

→ See also please the next presentation!

4 Discussion: General limitations of the generic factor concept



Temporal correlation score, transport database, Europe

4 Discussion

- Factor dependency:

e.g., time and technology: as technology evolves over time, changes in time also, most likely, relate to changes in technology (personal cars 1990 – personal cars 2010)

Needs to be considered in the uncertainty factor development.

(not in factor application – why)

4 Discussion

Application: (I think) guidance is needed, e.g. for ecoinvent:

- How to combine specific and generic factors
- How to obtain specific factors, when are they needed
- Perform the Pedigree approach twice? Once generic (data set against data set documentation, as now stored in the ecoinvent database), and then case study specific (database dataset against ideally required data set)

4b Conclusion

- Uncertainty factors now serve to provide better founded generic uncertainty information for ecoinvent flows;
- They should ideally be applied in combination with a case-specific uncertainty assessment.
- More experiences in practical application will be certainly useful.

GreenDelta

sustainability consulting + software

Thank you!

Contact: Dr. Andreas Ciroth
GreenDelta GmbH
Müllerstrasse 135, 13349 Berlin, Germany
ciroth@greendelta.com
www.greendelta.com