A protocol for approaching process data uncertainties in Life Cycle Inventory Analysis

53rd LCA Discussion Forum, Uncertainty in life cycle assessment: state of the art and practical challenges", September 13th 2013, Zürich

Patrik Henriksson & Jeroen Guinée Institute of Environmental Sciences (CML), Dept. Industrial Ecology



Universiteit Leiden

The Netherlands

Contents:

- Introduction
- Purpose
 - Proposed protocol
- Practical implementation
 - Discussion
- Conclusions

LCI data sourcing

- Most LCAs today present point values as results



- LCA practitioners (we all!!) do a lot of data picking without knowing its accuracy

- For our studies we continuously use "averages" without really knowing that we do ... and/or how they have been developed

LCA uncertainty analysis developing fast

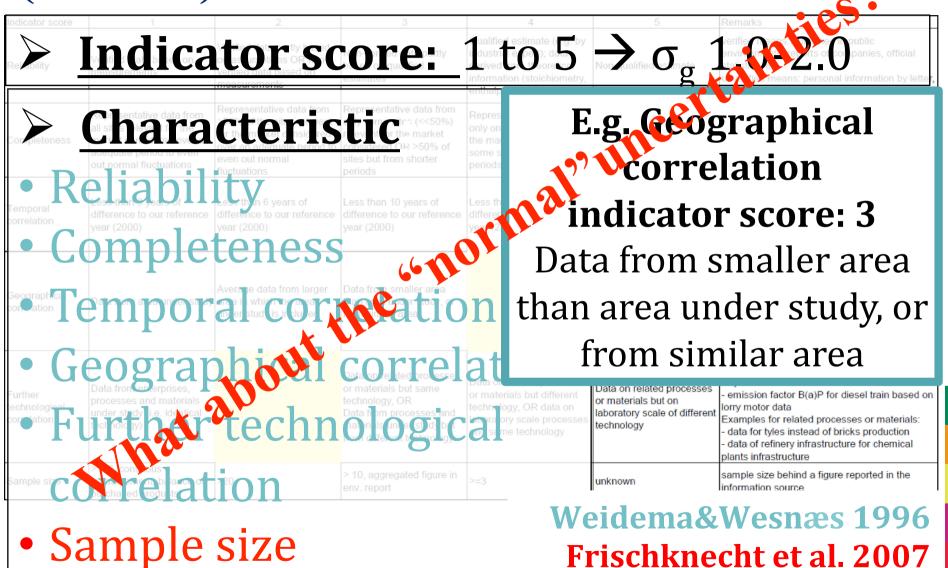
- Most attention up till now to propagation methods:
 - Monte-Carlo analysis
 - analytical error propagation
 - fuzzy logic
 - ...
- Producing uncertainty estimates requires:
 - input parameters
 - propagation method
 - definitions

Introduction

LCI uncertainty analysis largely disregarded

- Particularly input parameters for each Unit Process Data (UPD) point
 - basically ecoinvent pedigree approach ("post-normal" uncertainties)

Numeral Unit Spread Assessment Pedigree (NUSAP)

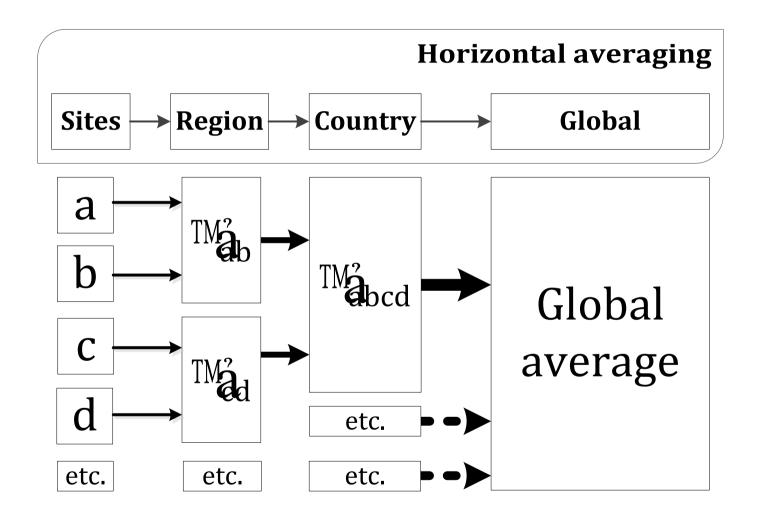


Purpose of this work

- To propose a method for <u>horizontal averaging of data</u> where dispersion from both <u>inherent uncertainty</u>, <u>spread</u> and <u>unrepresentativeness</u> are incorporated in the input parameters
 - not LCIA data
 - not methodological choices



Definitions



cf. UNEP-SETAC 'Shonan Guidance Principles'

Proposed protocol

Definitions

Dispersion	Any form of range around a variable, resulting from inherent uncertainty, spread or unrepresentativeness
Inherent uncertainty	Uncertainties related to the inaccuracies of measurements or model at no level of horizontal averaging
Spread	Variability around an average resulting from horizontal averaging
Unrepresentativeness	Uncertainty resulting from the level of representativeness

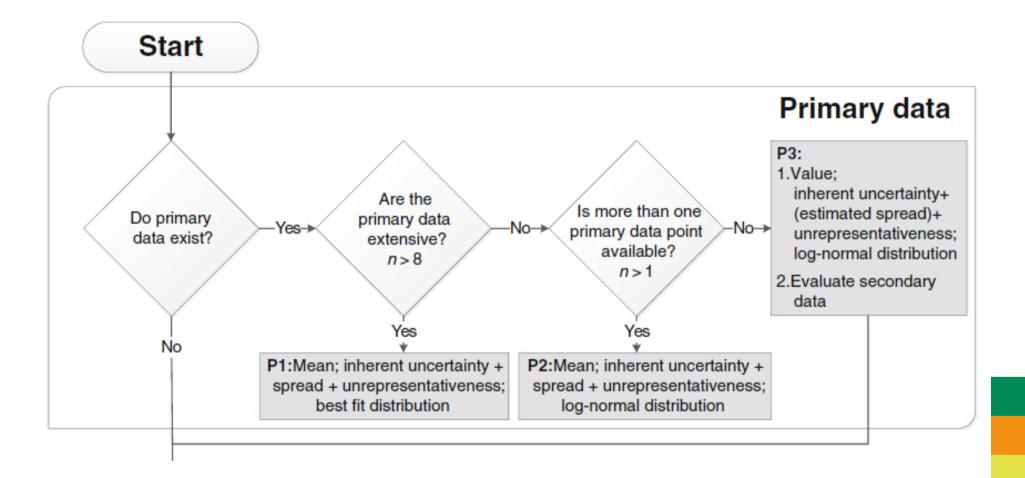
Proposed method

- Dispersion is the sum of:
 - Inherent uncertainty
 - Spread due to horizontal averaging
 - Unrepresentativeness **Post-normal (AP)**
- Combining the three into one quantitative measure of dispersion for each UPD point
- Propagation into LCI results

Leiden University. The university to discover.

Normal (NUS)

Decision tree



EXAMPLE SOYBEANS: PRACTICAL APPLICATION OF PROPOSED METHOD

Inherent uncertainty:

e.g. secondary data on Brazilian soybeans

TABLE 11.1 Default emission factors to estimate direct N2O emissions from managed soils									
Emission factor	Default value	Uncertainty range							
EF ₁ for N additions from mineral fertilisers, organic amendments and crop residues, and N mineralised from mineral soil as a result of loss of soil carbon [kg N ₂ O–N (kg N) ⁻¹]	0.01	0.003 - 0.03							
EF _{1FR} for flooded rice fields [kg N ₂ O–N (kg N) ⁻¹]	0.003	0.000 - 0.006							
$EF_{2\ CG,\ Temp}$ for temperate organic crop and grassland soils (kg $N_2ON\ ha^{-1})$	8	2 - 24							
$EF_{2CG,Trop}$ for tropical organic crop and grassland soils (kg $N_2ON\ ha^{-1})$	16	5 - 48							
$EF_{2F, Temp, Org, R}$ for temperate and boreal organic nutrient rich forest soils (kg N ₂ O–N ha ⁻¹)	0.6	0.16 - 2.4							
$EF_{2F, Temp, Org, P}$ for temperate and boreal organic nutrient poor forest soils (kg N ₂ O–N ha ⁻¹)	0.1	0.02 - 0.3							
EF _{2F, Trop} for tropical organic forest soils (kg N ₂ O–N ha ⁻¹)	8	0 - 24							

From: IPCC 2006 - N₂O Emissions from managed soils

Practical implementation

Spread due to horizontal averaging: e.g. *secondary data* on Brazilian soybeans

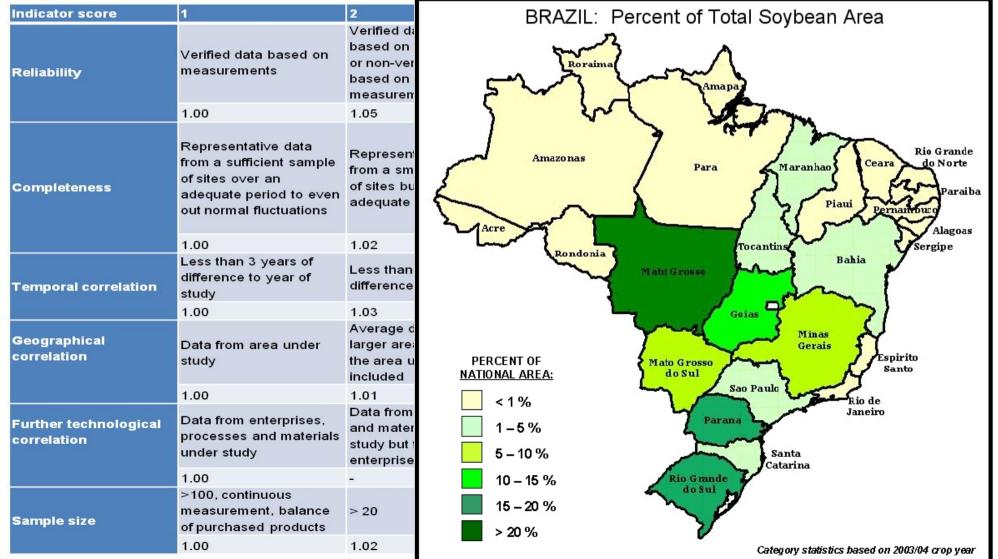
Average: Cederberg 1998	ecoinvent 2.2 (2007!)	Cederberg 1998	FAO 2002 & 2012	0		Castanheira 0 et al. 2013	
FNP 2000 quota – in Ostermayer 20		Database	Report	FAO	Report	Article	Article
Economical inputs		L					·
Diesel	kg	55.25	51.00	-	55.25	55.25	42.8
Nitrogen (kg N)	kg	3.1	0.00	3.7	8.0	0.00	0,00
Phosphorus (kg P)	kg	26.2	17.5	28.8	31.0	33.8	26.2
Potassium (kg K)	kg	24.9	33.2	51.5	57.0	65.4	49.8
Economic outputs							
Soybeans, kg	kg	2544	2200	2867	2500	2830	2940

Practical implementation

Unrepresentativeness

e.g. secondary data on Brazilian soybeans

- Each data point is scored to the ecoinvent v2.2 pedigree



Combining the three

	A B	С	D	E	F	G	Н	- I	J	K	L	М	N	0	
1	Online resource: Henrik Int J LCA	sson et al. <mark>(</mark> 2013)	R	Reference:		Halberg et al. 2012		Perez 2009		IPCC modelled		Basset-Mens et al. (GAP scenario) 2007		Imbeah 1998	
2					Value 1		Value 2		Value 3		Value 4		Value 5		
3	White cells need in	out	Unit process flow		Value	Uncertainty	Value	Uncertainty	Value	Uncertainty	Value	Uncertainty	Value	Uncertainty	
4			Uncertainty factors	Unit	Indicator score	factor	Indicator score	factor	Indicator score	factor	Indicator score	factor	Indicator score	factor	
5			Economic inputs	Economic inputs											
6	Weighted mean, x _{a(w}	: 2,82E+0	Feed	kg	3,10E+03	0,0051476	2,87E+03	2,855E-05			2,70E+03	0,0044094			
7	X ₀₌	2,89E+03	Inherent uncertainty as $\sigma_8^{\ i}$:		1,05		1,05		1,05		1,05		1,05		
8	X g=	2,89E+03	Reliability:		2. Verified data partly	1,05	Verified data partly	1,05	1. Verified data basea	1,00	3. Non-verified data p	1,10	1. Verified data basea	1,00	
9	Inherent uncertainty o	: 1,050	Completeness:		4. Representative dat	1,10	3. Representative dat	1,05	1. Representative dat	1,00	 Representative dat 	1,10	1. Representative dat		
10	Inherent uncertainty C	/: 0,052	Temporal correlation:		1. Less than 3 years of	1,00	3. Less than 10 years o	1,10	1. Less than 3 years of	1,00	2. Less than 6 years oj	1,03	1. Less than 3 years of	1,00	
11	Representativeness CV	: 0,078	Geographic correlation:		3. Data from smaller a	1,02	3. Data from smaller o	1,02	1. Data from area und	1,00	3. Data from smaller o	1,02	1. Data from area und	1,00	
12	Spread σ_s^s	1,072	Further techonological correlation:		3. Data on related pro	1,20	1. Data from enterpris	1,00	1. Data from enterpris	1,00	1. Data from enterpris	1,00	1. Data from enterpris	1,00	
13	Spread, CV ⁱ	0,0694631	(Sample size):		5. Unknown	1,20	5. Unknown	1,20	1. n = >100, continous	1,00	3. n = > 10, aggregate	1,05	1. n = >100, continous	1,00	
14	Overall uncertainty I	°hi 0,116	Representativeness factor as σ _g ^r :		1,150	1,150 1,115			1,000		1,077		1,000		
15	Overall uncertainty	D95 1,119		W .=	182,11	564528,86	280,70	805611,13	*		510,59	1378591,64	*		

194			Economic outputs											
195	Weighted mean, x _{a(wt)} :	1,00E+03	Pigs, 85-100 kg each	kg	1,00E+03	0	1,00E+03	0			1,00E+03	0	1,00E+03	0
196	X ₀₌	1,00E+03	Inherent uncertainty as $\sigma_{g}^{\ i}$:		1,05		1,05		1,05		1,05		1,05	
197	$X_{g=}$	1,00E+03	Reliability:		Verified data partly	1,05	2. Verified data partly	1,05	1. Verified data based	1,00	3. Non-verified data p	1,10	1. Verified data basea	1,00
198	Inherent uncertainty σ_{r}^{i} :	1,050	Completeness:		4. Representative dat		3. Representative date		1. Representative dat	1,00	4. Representative dat	1,10	1. Representative dat	
199	Inherent uncertainty CV ¹ :	0,041	Temporal correlation:		1. Less than 3 years of	1,00	3. Less than 10 years d	1,10	1. Less than 3 years of	1,00	2. Less than 6 years of	1,03	1. Less than 3 years of	1,00
200	Representativeness CV ¹ :	0,078	Geographic correlation:		3. Data from smaller c	1,02	3. Data from smaller d	1,02	1. Data from area und	1,00	3. Data from smaller d	1,02	1. Data from area unc	1,00
201	Spread σ _p ^s	1,000	Further techonological correlation:		3. Data on related pro		1. Data from enterpris		1. Data from enterpris		1. Data from enterpris		1. Data from enterpris	
202	Spread, CV'	0,000	(Sample size):		5. Unknown	1,20	5. Unknown	1,20	1. n = >100, continous	1,00	3. n = > 10, aggregate	1,05	1. n = >100, continous	1,00
203	Overall uncertainty sigma	8,80E+01	Representativeness factor as $\sigma_8^{\ r}$:		1,150		1,115		1,000		1,077		1,000	
204	Overall uncertainty SD95	1,093		Wi=		182,11 182106,08 280,70 280700,74			510,59 510589,50			1680,33	1680333,29	
205														
206	Weighted mean, x _{a(wt)} :	3,89E+03	Solid manure	kg							4,00E+03	0,0027752	3,60E+03	0,0027752
207	<i>X</i> _{<i>a</i>=}	3,80E+03	Inherent uncertainty as $\sigma_{g}^{\ i}$:		1,05		1,05		1,05		1,05		1,05	
208	$X_{g=}$	3,79E+03	Reliability:		1. Verified data basea		1. Verified data based		1. Verified data based		3. Non-verified data p		1. Verified data basea	
209	Inherent uncertainty σ_{s}^{i} :	1,050	Completeness:		1. Representative date	1,00	1. Representative dat	1,00	1. Representative dat	1,00	4. Representative dat	1,10	4. Representative dat	
210	Inherent uncertainty CV ¹ :	0,052	Temporal correlation:		1. Less than 3 years of	1,00	1. Less than 3 years of	1,00	1. Less than 3 years of	1,00	2. Less than 6 years of	1,03	4. Less than 15 years c	1,20
211	Representativeness CV ¹ :	0,078	Geographic correlation:		1. Data from area und	1,00	1. Data from area und	1,00	1. Data from area und	1,00	3. Data from smaller d		3. Data from smaller c	1,02
212	Spread σ _p ^s	1,077	Further techonological correlation:		1. Data from enterpris		1. Data from enterpris		1. Data from enterpris		1. Data from enterpris		1. Data from enterpris	
213	Spread, CV'	0,074	(Sample size):		1. n = >100, continous	1,00	1. n = >100, continous	1,00	1. n = >100, continous	1,00	3. n = > 10, aggregate	1,05	5. Unknown	1,20
214	Overall uncertainty Phi	0,119	Representativeness factor as σ _g ^r :		1,000		1,000		1,000		1,077		1,148	
215	Overall uncertainty SD95	1,123		W _{i=}							510,59	2042357,98	187,18	673838,72

Henriksson et al. 2013. A protocol for horizontal averaging of unit process data – including estimates for uncertainty. *Int J Life Cycle Assess (FirstOnLine)*. *DOI 10.1007/s11367-013-0647-4*

Practical implementation

Weighted factor – weighted means

- $x \downarrow wt = \sum i = 1 \uparrow n @w \downarrow i x \downarrow i / \sum i = 1 \uparrow n @w \downarrow i$

- wli=1/σliî î2 arithmetic standard deviation
- $w \downarrow i = 1/CV \downarrow i \uparrow \uparrow 2$ coefficient of variation
- $w li = 1/(\ln \sigma li) 12$ geometric standard deviation

http://cml.leiden.edu/software/software-quanlci.html

Online resources Sep03

A B	С	D	E	F	G	Н	l l	
1 Online resource: Henriksson et	al. (2013) Int J	Re	eference:					
2 <u>LCA</u>	_			Value 1		Value 2		
3 White cells need input		Unit process flow		Value	Uncertainty	Value	Uncertainty	
4	-	Uncertainty factors	Unit	Indicator score	factor	Indicator score	factor	
5		Economic inputs						
6 Weighted mean, X _{a(wt)} :	#DIV/0!							
7	#NUM!	Inherent uncertainty as $\sigma_{cv}{}^i$:		0.05		0.05		
Inherent uncertainty σ _e ⁱ :	5.86E+02	Reliability:		1. Verified data based	1.00	1. Verified data based	1.00	
Inherent uncertainty CV ⁱ :	1.00E+09	Completeness:		1. Representative data	1.00	1. Representative data	1.00	
0 Representativeness CV ⁱ :	4.04E+82	Temporal correlation:		1. Less than 3 years of	1.00	1. Less than 3 years of	1.00	
1 Spread σ _ε ⁵	1.00E+00	Geographic correlation:		1. Data from area unde	1.00	1. Data from area unde	1.00	
2 Spread, σ _O ⁱ	#DIV/0!	Further techonological correlation:		1. Data from enterprise	1.00	1. Data from enterprise	1.00	
3 Overall uncertainty CV	#DIV/0!	(Sample size):		1. n = >100, continous	1.00	1. n = >100, continous	1.00	
4 Overall uncertainty Phi	#DIV/0!	Representativeness factor as σ_{a} ':		1.000		1.000		
5 Overall uncertainty SD95	8.77E+02		₩.,					
6								
7 Weighted mean, X _{e(wt)} :	#DIV/0!							
.8 T _{g=}	#NUM!	Inherent uncertainty as $\sigma_{cv}{}^i$:		0.05		0.05		
9 Inherent uncertainty σ _e ⁱ :	5.86E+02	Reliability:		1. Verified data based	1.00	1. Verified data based	1.00	
0 Inherent uncertainty CV ⁱ :	1.00E+09	Completeness:		1. Representative data	1.00	1. Representative data	1.00	
1 Representativeness CV ⁱ :	4.04E+82	Temporal correlation:		1. Less than 3 years of	1.00	1. Less than 3 years of	1.00	
2 Spread σ _s ⁵	1.00E+00	Geographic correlation:		1. Data from area unde	1.00	1. Data from area unde	1.00	
²³ Spread, σ _O ¹	#DIV/0!	Further techonological correlation:		1. Data from enterprise	1.00	1. Data from enterprise	1.00	
0verall uncertainty CV	#DIV/0!	(Sample size):		1. n = >100, continous	1.00	1. n = >100, continous	1.00	
● ▶ ▶ Description / Index	🔬 Unit proce	ess template σa 🚽 Unit proces	s templa	te CV 🖉 Unit proce	ss template	ln(og) 📝 CV to og	σa ≈ σο	

http://cml.leiden.edu/software/software-quanlci.html

Online resource: Henriksson et al. (2013) Int J		R		<u>Ostermayer</u>	2002	Cederberg 1	998	Castanheira et al. 2013		
LCA					Value 1		Value 2		Value 3	
White cells need input		Unit process flow			Value	Uncertainty	Value	Uncertainty	Value	Uncertainty
		Uncertainty factors	Unit		dicator score	factor	Indicator score	factor	Indicator score	factor
		Economic inputs								
Weighted mean, X _{e(wt)} :	#DIV/0!	Diesel	<mark>∕ k</mark> g							
π _{g=}	#NUM!	Inhere t uncertainty as σ_{cv}^{i} :			0.05		0.05		0.05	
Inherent uncertainty σ_{e}^{i} :	5.86E+02	Reliał /ity:		1.	erified data based	1.00	1. Verified data based	1.00	1. Verified data based	1.00
Inherent uncertainty CV ⁱ :	1.00E+09	Com/ eteness:		1.	epresentative data	1.00	1. Representative data	1.00	1. Representative data	1.00
Representativeness CV ⁱ :	4.04E+82	Tem oral correlation:		1.	ess than 3 years of	1.00	1. Less than 3 years of	1.00	1. Less than 3 years of	1.00
Spread og ⁵	1.00E+00	Ge(;raphic correlation:		1.)	ata from area und	1.00	1. Data from area unde	1.00	1. Data from area unde	1.00
Spread, σ_{cv}^{i}	#DIV/0!	Fu her techonological correlative:		1)	ata from enterpris	1.00	1. Data from enterprise	1.00	1. Data from enterprise	1.00
Overall uncertainty CV	#DIV/0!	(S /nple size):		1 n	= >100, continous	1.00	1. n = >100, continous	1.00	1. n = >100, continous	1.00
Overall uncertainty Phi	#DIV/0!	Γ presentativeness factor as σ			1.000		1.000		1.000	
Overall uncertainty SD95	8.77E+02		¥.,							

- Inventory flow
- Unit
- Reference

http://cml.leiden.edu/software/software-quanlci.html

Online resource: Henriksson et o	Online resource: Henriksson et al. (2013) Int J		Reference:			Cederberg 1	998	Castanheira et al. 2013	
LCA			Value 1		Value 2		Value 3		
White cells need input		Unit process flow		Value	Uncertainty	Value	Uncertainty	Value	Uncertainty
		Uncertainty factors	Unit	Indicator score	factor	Indicator score	factor	Indicator score	factor
		Economic inputs							
Weighted mean, X _{s(wt)} :	4.97E+01	Diesel	kg	5.53E+01	0.0126324	5.10E+01	0.0009889	4.28E+01	0.0206902
π _{g=}	4.94E+01	Inherent uncertainty as σ_{cv}^{i} :		0.05		0.05		0.05	
Inherent uncertainty o _c ⁱ :	1.05E+00	Reliability:		1. Verified data bas	1.00	1. Verified data based	1.00	1. Verified data based	1.00
Inherent uncertainty CV ⁱ :	5.00E-02	Completeness:		1. Representative day	1.00	1. Representative data	1.00	1. Representative data	1.00
Representativeness CV ⁱ :	4.04E+82	Temporal correlation:		1. Less than 3 years of	1.00	1. Less than 3 years of	1.00	1. Less than 3 years of	1.00
Spread og ⁵	1.14E+00	Geographic correlation:		1. Data from area und	1.00	1. Data from area unde	1.00	1. Data from area unde	1.00
Spread, $\sigma_{\rm CV}^{\ \ i}$	1.28E-01	Further techonological correlatio		1. Data from enterpris	1.00	1. Data from enterprise	1.00	1. Data from enterprise	1.00
Overall uncertainty CV	4.04E+82	(Sample size):		1. n = >100, continous	1.00	1. n = >100, continous	1.00	1. n = >100, continous	1.00
Overall uncertainty Phi	1.95E+01	Representativeness factor ?		1.000		1.000		1.000	
Overall uncertainty SD95	1.00E+01		₩ _i .	400.00	22120 0	400.00	20400.00	400.00	17120.00

- Central value (arithmetic mean)
- Inherent uncertainty (coefficient of variation)

http://cml.leiden.edu/software/software-quanlci.html

Online resource: Henriksson et	al. (2013) Int J	Re	Ostermayer 2	2002	Cederberg 1	998	Castanheira et al. 2013		
LCA				Value 1	Value 1			Value 3	
White cells need input		Unit process flow		Value	Uncertainty	Value	Uncertainty	Value	Uncertainty
		Uncertainty factors	Unit	Indicator score	factor	Indicator score	factor	Indicator score	factor
		Economic inputs							
Weighted mean, X _{a(wt)} :	4.57E+01	Diesel	kg	5.53E+01	0.0126324	5.10E+01	0.0009889	4.28E+01	0.0206902
<i>Х_{g=}</i>	4.94E+U1	Inherent uncertainty as σ_{cv}^{i} :		0.05		0.05		0.05	
Inherent uncertainty o _e ⁱ :	1.05E+00	Reliability:		2. Verified data partly	1.05	3. Non-verified data po	1.10	2. Verified data partly	1.05
Inherent uncertainty CV ⁱ :	5.00E-02	Completeness:		3. Representative data	1.05	3. Representative data	1.05	3. Representative data	1.05
Representativeness CV ⁱ :	3.98E-02	Temporal correlation:		3. Less than 10 years o	1.10	4. Less than 15 years o	1.20	1. Less than 3 years of	1.00
Spread og ⁵	1.14E+00	Geographic correlation:		5. Data from unknown	1.10	5. Data from unknown	1.10	3. Data from smaller a	1.02
Spread, σ _{cv} ⁱ	1.28E-01	Further techonological correlation:		3. Data on related proc	1.20	3. Data on related proc	1.20	1. Data from enterprise	1.00
overall uncertainty cv	1.435-01	(Sample size):		3. n = > 10, aggregated	1.05	3. n = > 10, aggregated	1.05	2. n = >20	▼ 1.02
Overall uncertainty Phi	1.42E-01	Representativeness factor as σ_s' :		1.129		1.161		1. n = >100, continous meas (2. n = >20	urer
Overall uncertainty SD95	1.16E+00		¥.,	53.89	2980.24	37.16	10 5 0	3. n = > 10, aggregated figur	e in 10483.55
								4. n = >3 5. Unknown	

- Representativeness (dropdown menu)
- Weighted mean, (based upon inherent uncertainty + representativeness)
- Spread (calculated amongst values)

http://cml.leiden.edu/software/software-quanlci.html

- Weighted mean
- Inherent uncertainty
- Representativeness
- Spread
- Overall dispersion
 - Phi = CMLCA
 - SD95 = Simapro

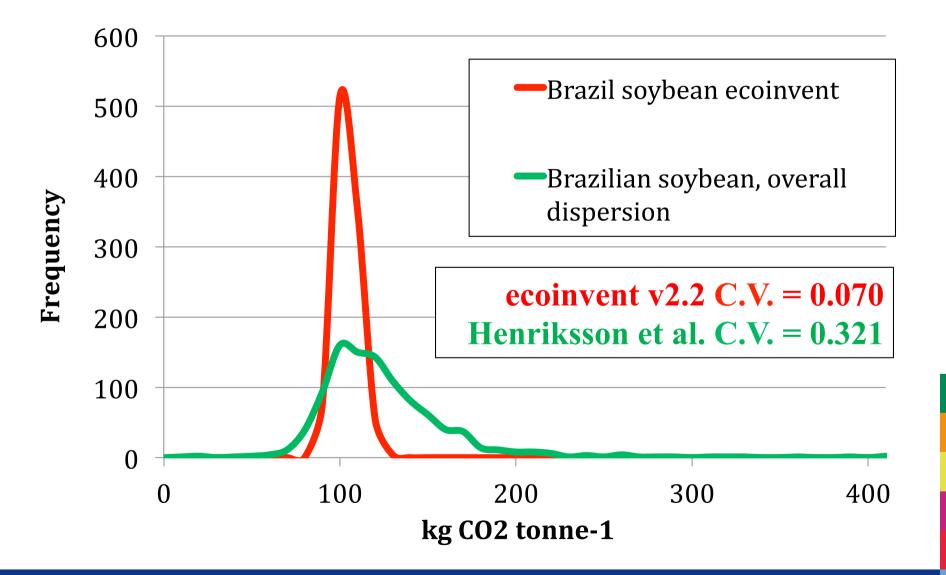
Online resource: Henriksson et al. (2013) Int

J LCA

White cells need input

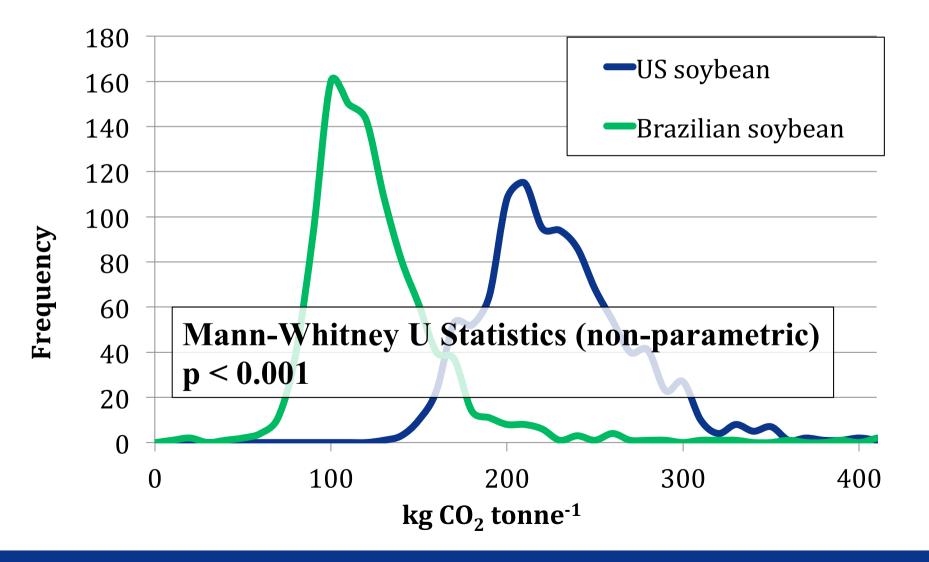
Weighted mean, X _{a(wt)} :	4.57E+01			
$x_{\overline{g}=}$	4.94E+01			
Inherent uncertainty σ_{g}^{i} :	1.05E+00			
Inherent uncertainty CV ⁱ :	5.00E-02			
Representativeness CV ⁱ :	3.98E-02			
Spread σ _g ^s	1.14E+00			
Spread, σ_{CV}^{i}	1.28E-01			
Overall dispersion CV	1.43E-01			
Overall dispersion Phi	1.42E-01			
Overall dispersion SD95	1.16E+00			

Propagation into LCI results - Histograms (= further work under construction)



Practical implementation

Propagation into LCI results - Histograms (= further work under construction)



Discussion

- Quantitative use of NUSAP is beyond original intent
- Averages or values behind the averages?
- Spread and inherent uncertainty often out-weight representativeness
- Absolute or comparative uncertainty

Conclusions

- Method is practical but intensive
- Better reporting of raw data
 - including inherent uncertainties, spread in background data & distributions instead of averaged point values
- Propagation and analysis work in progress...



Acknowledgement: This work is part of the Sustaining Ethical Aquaculture Trade (SEAT) project, which is co-funded by the European Commission within the Seventh Framework Programme—Sustainable Development Global Change and Ecosystem (project no. 222889). http://www.seatglobal.eu

Conclusions



Henriksson, Guinée, Heijungs, de Koning & Green (2013) A protocol for horizontal averaging of unit process data including estimates for uncertainty. Int J of LCA. DOI 10.1007/s11367-013-0647-4

<u>cml.leiden.edu/software/software-quanlci.html</u> <u>www.seatglobal.eu</u>