Training on Data Acquisition and Dataset Development

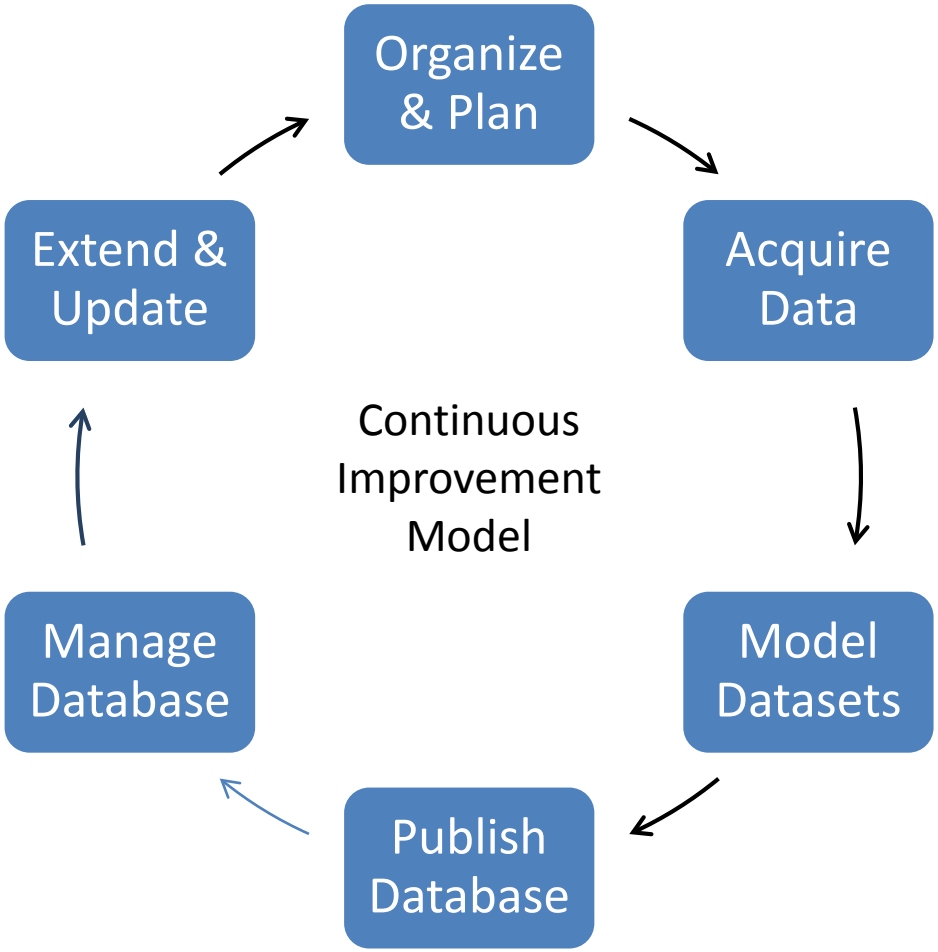
Introduction and Course Structure

August 2017
Version

Managed by SETAC

Training Course Rationale

- LCA databases are going global, including networking through GLAD*
- Public database offerings are increasing at the national level
- Developers of national databases want training on the procedures to create, publish, and manage their databases
- The National LCA Databases Helpdesk has created a validated training package to assist with these needs.



*Global LCA Access to Data

Training Course - Components

- Course consists of technical modules and learning exercises
- Designed to be delivered by a live trainer over a two day period, but
- Can be used in a self-taught, online mode
- Helpdesk support can augment the online materials

- Can be adjusted or tailored for trainee needs
- Can be preceded by general LCA training if needed
- Can be followed by specific training on LCA software if desired

- Relationship of data to databases and LCA software
- Raw data acquisition – primary and secondary
- Process modeling
- Alternative dataset development methods
- Datasets in databases, including documentation
- Linking databases in software

Training Course - Contents

Part 1 - Relationship of data to (national) databases and LCA software

Part 2 - Creating datasets from original sources
- Defining the process inputs and outputs
- Suggestion: Identifying and prioritizing the required data
- Identifying data sources and sample stratification
- Creating/tailoring a data collection strategy and provider coordination approach
- Confidentiality protection
- Creating/tailoring a data collection template
- Classroom exercise (mango production/processing)

Part 2a - Creating datasets from secondary data and quality assurance

Part 2b - Meta-indicators and GLAD (in development)

Part 3 - Creating datasets from public sources
- Defining the process inputs and outputs
- Identifying data sources and sample stratification
- Creating/tailoring a data collection strategy and coordination approach
- Creating/tailoring a data collection template
- Classroom exercise (electricity production)

Part 4 - Process modelling for LCI Datasets
- Concepts and practicalities
- Allocated versus unallocated processes
- Classroom exercise

Part 5 - Modelling processes for national dataset development, case study from Sri Lanka and Brazil

Part 6 - Regionalizing datasets of processes

Part 7 - Storing datasets into databases

Part 8 - Linking databases and LCA software, and alternative dataset development methods

Training on Data Dataset Acquisition and Database Development

Excel Sheets in Annex

Part 2 - Exercise 1 (Mango Production)

"data_collection_template_generic"

Part 3 - Exercise 2 (Sri Lanka Energy Mix)

"2nd exercise_processes"

Part 3 - Exercise 3 (Sustainable Apparel Coalition)

"data_collection_template_example_customized_may_2013"

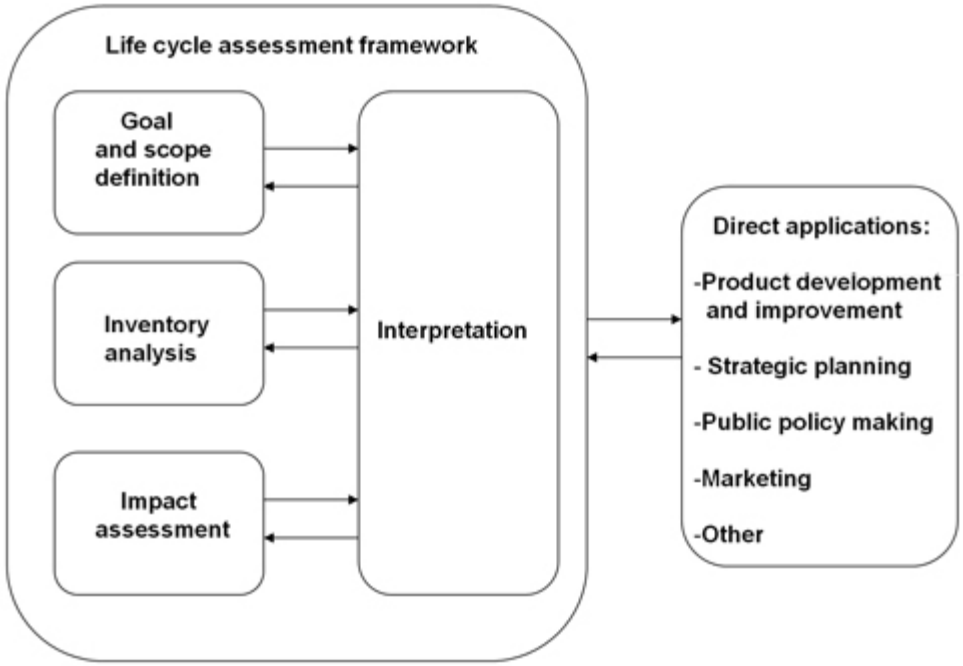
Part 1 - Relationship of data to (national) databases and LCA software

Content from: Andreas Ciroth, GreenDelta

Managed by
SETAC

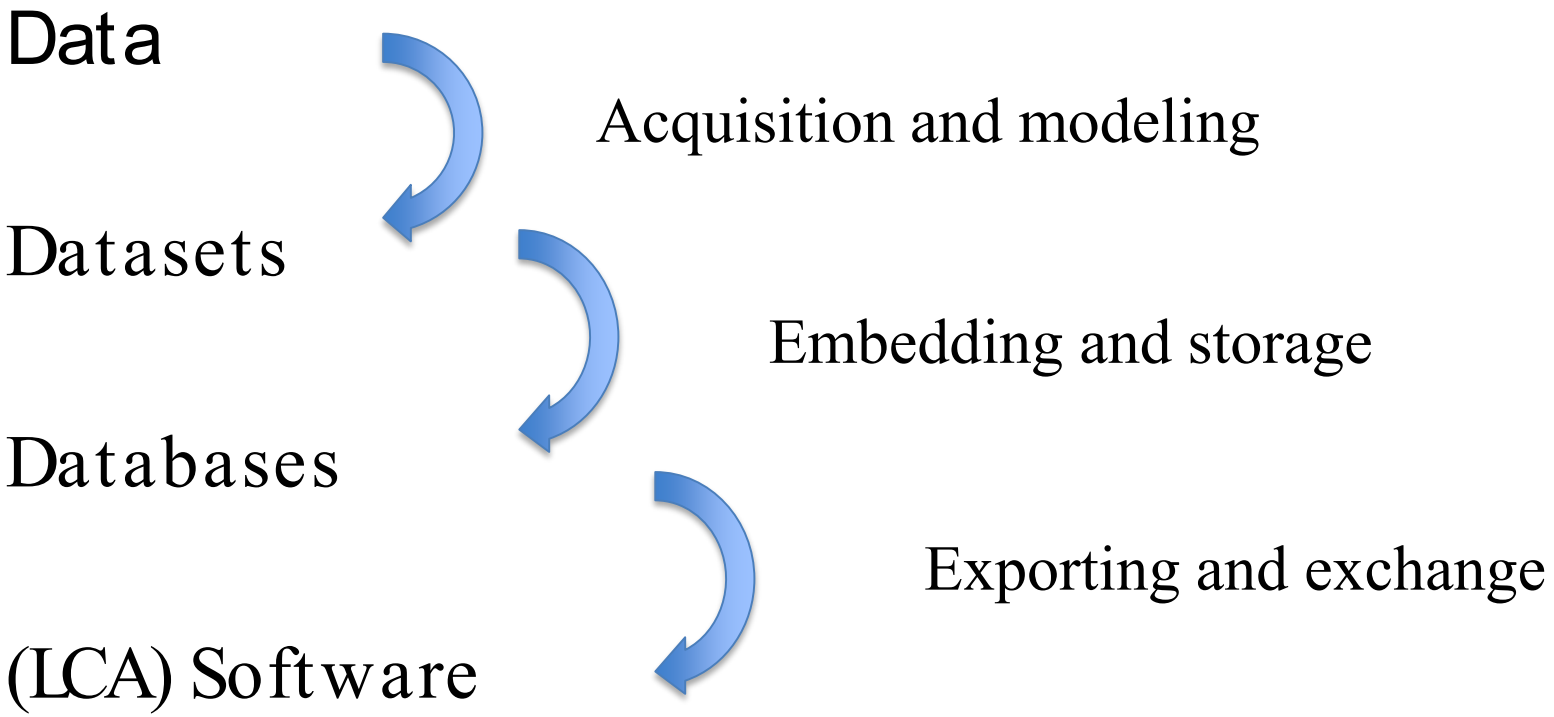
Relationship of data to (national) databases and LCA software

(Life Cycle Assessment)



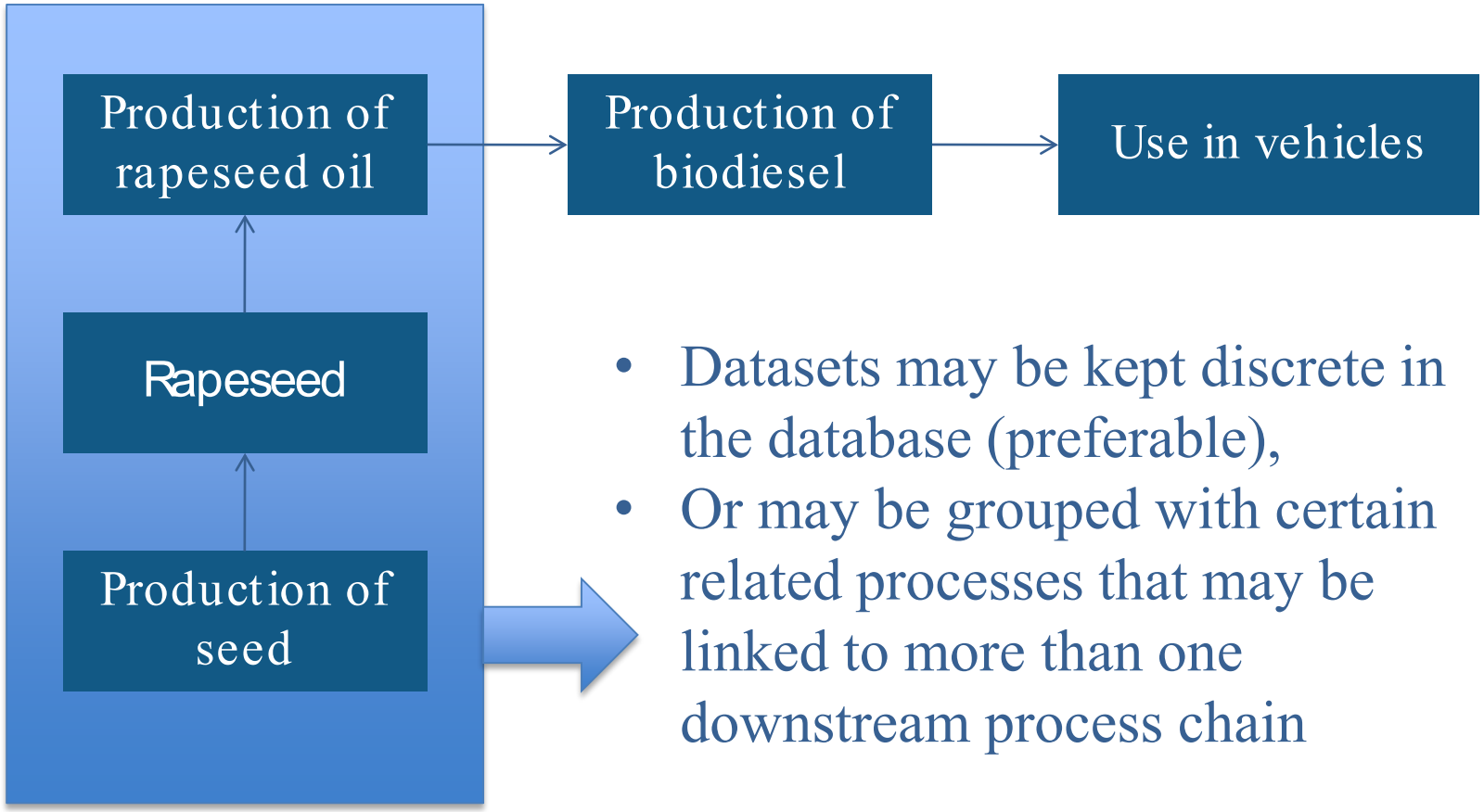
ISO 14040, 2006

Relationship of data to (national) databases and LCA software



Relationship of data to (national) databases and LCA software:

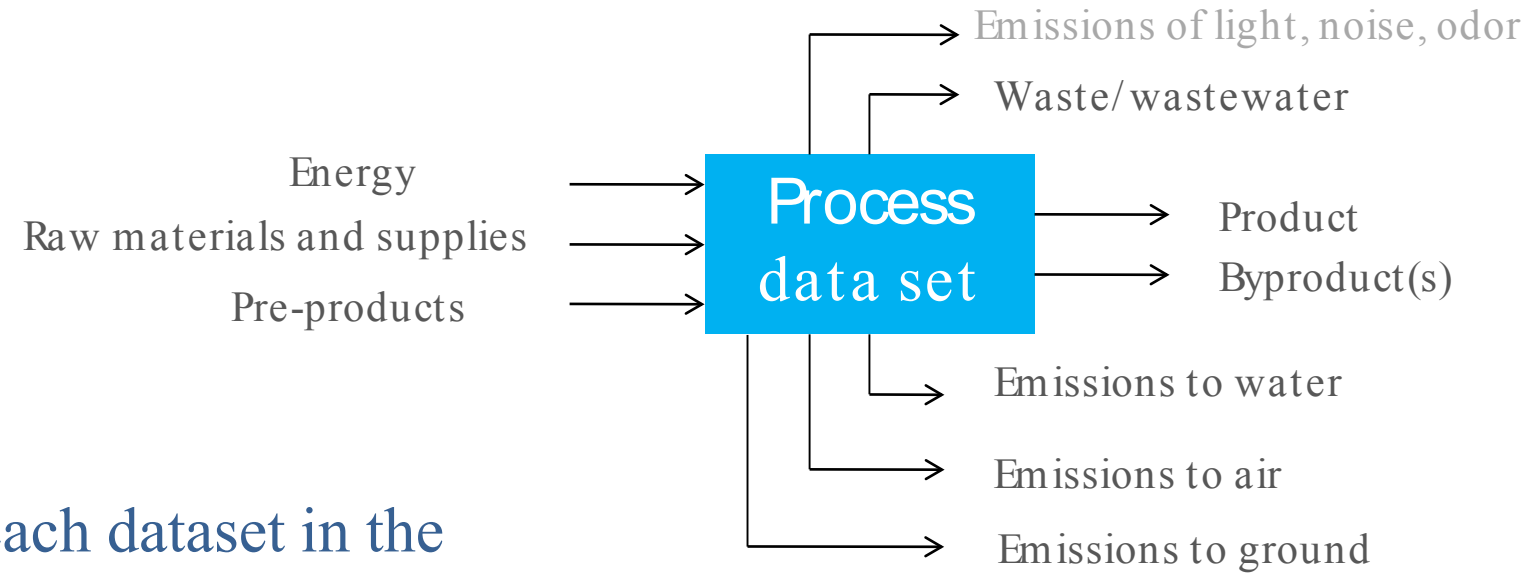
datasets in a supply chain



- Datasets may be kept discrete in the database (preferable),
- Or may be grouped with certain related processes that may be linked to more than one downstream process chain

Relationship of data to (national) databases and LCA software:

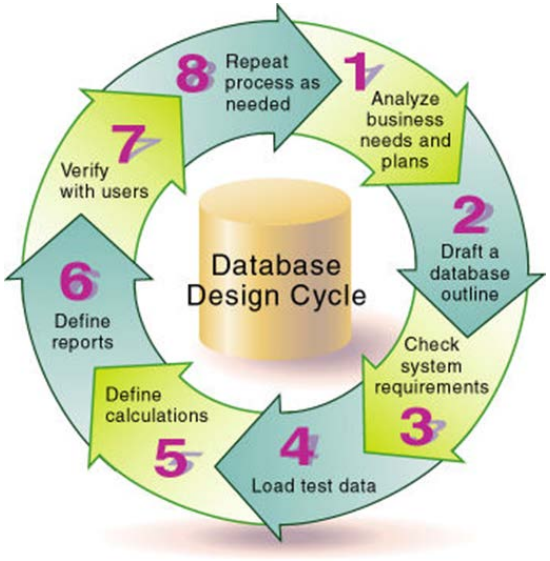
datasets



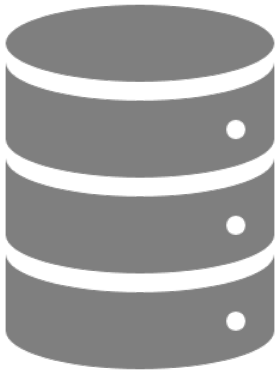
Each dataset in the database should be complete (inputs and outputs) and balanced (material and energy)

Relationship of data to (national) databases and LCA software:

database



- Databases are storage locations for data
- Traditional databases store data in one of two ways –
  - ✓ Flat file – all information, including meta-data, about a database object is contained in one place, termed a "record", or
  - ✓ Relational – individual attributes are grouped into tables and "related" through a common parameter such as a process name.
- Newer approaches are called "data repositories, curated collections of files that can be downloaded by users

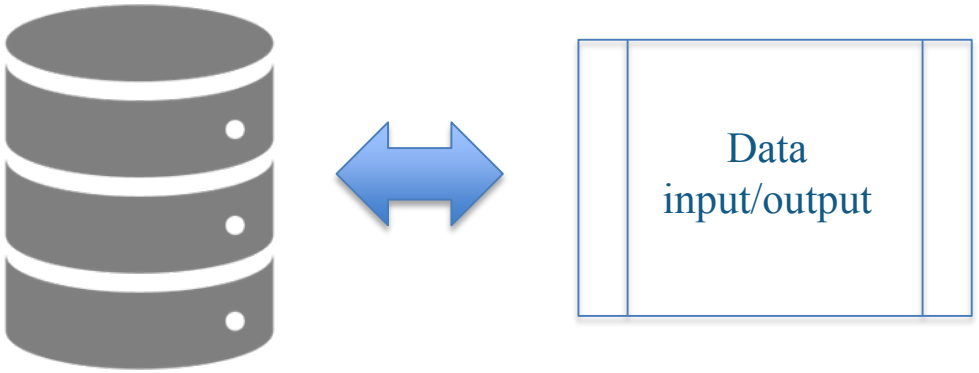


The most convenient databases from a user perspective are web-based.

Relationship of data to (national) databases and LCA software:

LCA / LCI database characteristics

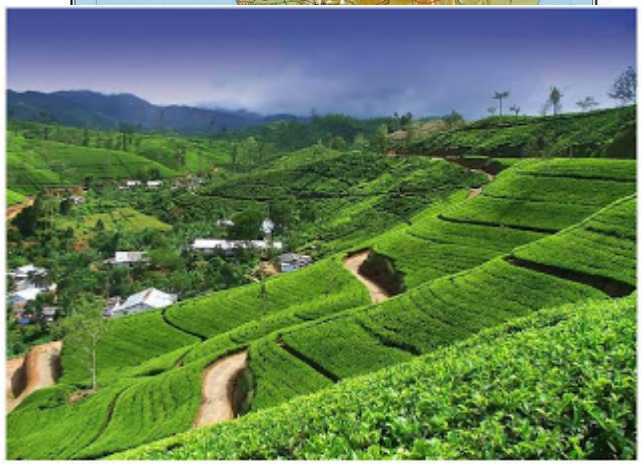
- Comprehensive input/output flows for datasets
- Consistent modeling approach and flow nomenclature
- Complete and consistent dataset documentation
- Often, capability beyond storing data, e.g. data collection and submittal support, review, aggregated process creation, …



Relationship of data to (national) databases and LCA software:

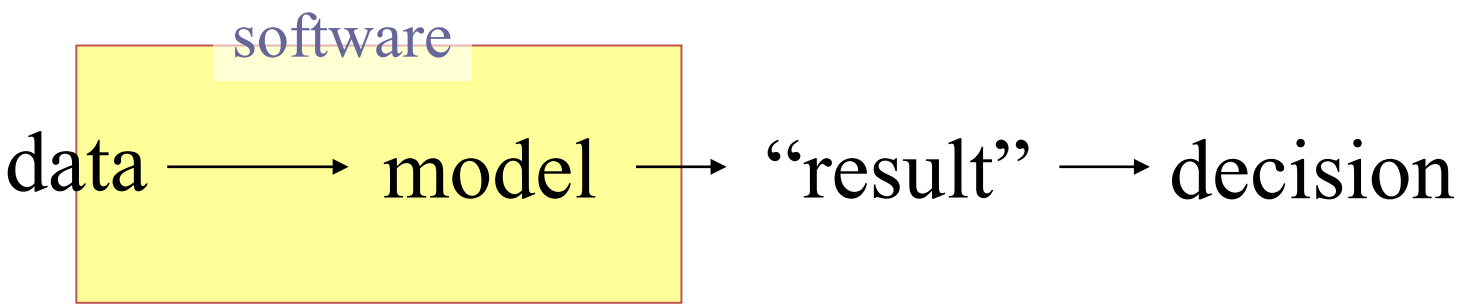
LCA / LCI database characteristics

- Comprehensive input/output flows for datasets
- Consistent modeling approach and flow nomenclature
- Complete and consistent dataset documentation
- Often, capability beyond storing data, e.g. data collection and submittal support, review, aggregated process creation, …
- National database: (often) contains reference datasets, representative for a country



Relationship of data to (national) databases and LCA software:

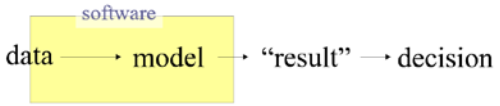
LCA software



Ciroth, A.: A new open source LCA software , EcoBalance conference, presentation, Tsukuba, 2006

Relationship of data to (national) databases and LCA software:

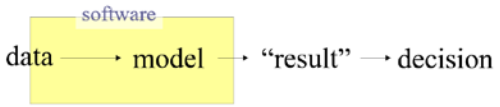
LCA software



- LCA Software works with databases, and makes the database much more useful for modeling and decision support and application
- There are many different LCA software packages; four are mainly important:
  - SimaPro
  - GaBi
  - openLCA
  - (Umberto)
- When developing a database, always consider its use in software!

Relationship of data to (national) databases and LCA software:

LCA software

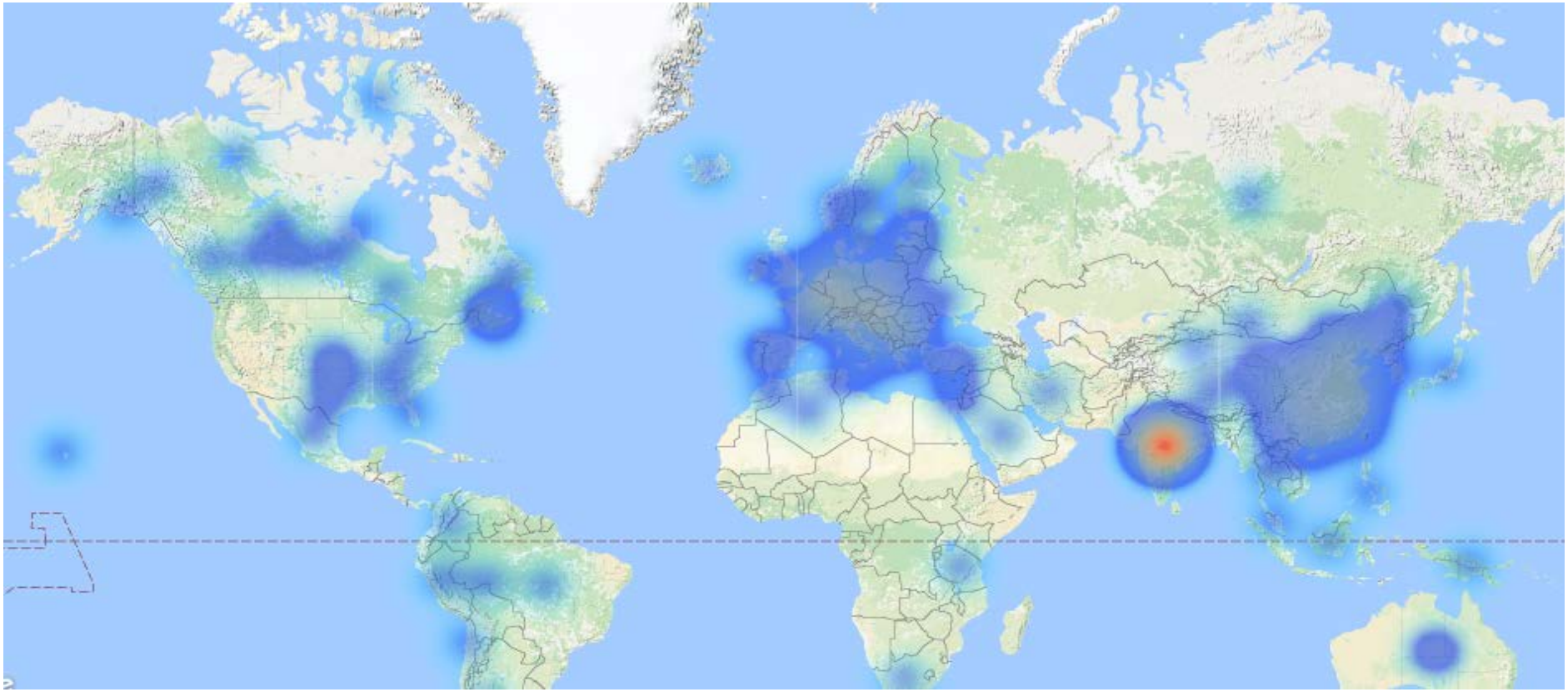


- LCA Software works with databases, and makes the database much more useful for modeling and decision support and application
  - Identify the main drivers along the supply chain
  - Make comparisons
  - Perform sensitivity analyses
  - Create and add user-specific datasets
  - …

Relationship of data to (national) databases and LCA software:

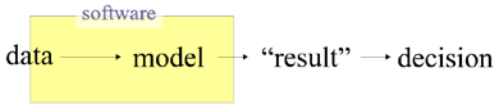
LCA software

- Rice production India, ecoinvent 3.3.cut-off, ReCiPe LCIA method, human health – human toxicity

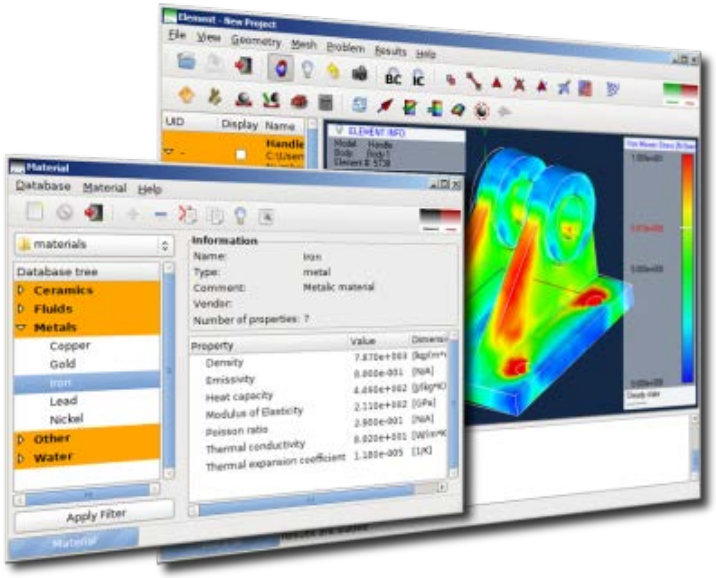


Relationship of data to (national) databases and LCA software:

LCA software



- When developing a database, consider how it would be used in software!
- Ideally, not only one but several broadly used LCA software systems should be considered



Training Workshop on Data Acquisition and Dataset Development for Life Cycle Inventory

Exercise 1 - Primary raw data acquisition and modeling

Data for producers:[1]

The dataset covers all relevant steps involved with a Mango production from cradle-to-gate, i.e. all processes from raw materials extraction till Mango harvesting are taken into account.

PART 1.1 – Primary data collection

You are a successful producer that started your 20ha Mango plantation about 7 years ago, for that you needed to start from scratch and buy a piece of land that still had a forest on it, thus, first of all it was necessary to clean the "fields".

Preparing the field for the plantation demanded machines to cut the trees and prepare the fields, electricity to power some machinery, diesel and land.

The machines can clean and prepare 263 ha of field during its lifetime which means that for your Mango plantation only 0.076 Machine units were necessary.

In total 13360 kW of electricity to power machines working 4 hours per hectare was needed.

13360/4 = 3340 kWh per 20 ha

3340/20 = 167 kWh /ha

Necessary materials and resources		
..		
..		
Diesel	1.61	MJ
..		

In the End of this process you had 20ha of clean and prepared field, 1000m³ of Biomass coming from the cut trees that need to be disposed and emissions of pollutants to the air in a low populated area.

Products / by-products / waste		
..		
..		

Emissions to air per hectare		
CO2 (biogenic)	21.14	kg
N2O	2.6	kg
CO2 (fossil)	78.58	kg
PM (particulate matter)	15	kg

---

[1] This fact sheet describes a simplified Mango production farm based on data from EcoInvent and Agribalyse databases.

PART 1.2 - Primary data collection

After preparing the field you needed to start your production, for this step you needed to consider again all the necessary resources and materials.

The plantation needs irrigation through the whole year, each plant consumes 100 liters of water per week and there is a total of 140 plants per hectare.

You expect you can produce Mangos on this field for at least 25 years.

The irrigation system uses a 2.5kW pump per hectare that works every day for 12h.

The land needs to be fertilized every year, therefore 111.9 kg of fertilizer is applied to each hectare of cultivated land per year.

DATA:

Mango production area: 20 ha

Annual yield: 15 t/ha/year

Necessary materials and resources		
Pesticides	8.82E-05	kg
Transport of fertilizers and pesticides to plant	200	km

Products / by-products / waste		

Emissions per kg of Mango		
Phosphate emissions to groundwater	2.09E-06	kg
Pesticides emissions to soil	8.82E-05	kg

PART 2 – Allocation

In the end of the harvest you have 2 products:

- Mangos for selling at the market
- Mangos that are too ripe and can only be consumed after being processed (jams, dry mango snacks, Mango pickles, etc.)

The main product is the Market mango and the Ripe mango is the by product which is 15% of the production and is responsible for 5% of the farm income.

Because 2 products are generated the environmental burdens need to be correctly split between them. There are many ways of doing it but we will check the results for mass allocation and economic allocation.

Mass allocation:

1 ton of Mango = 850 kg of Market mango + 150 kg of Ripe mango, but emissions are for 1 kg of Mango

Products / by-products / waste		
Market mango	0.85	kg
Ripe mango	0.15	kg

Emissions per kg of Mango		
Phosphate emissions to groundwater	2.09E-06	kg
Pesticides emissions to soil	8.82E-05	kg

Mass distribution		

Emissions per product		
Phosphate emissions to groundwater		kg
Pesticides emissions to soil		kg
Phosphate emissions to groundwater		kg
Pesticides emissions to soil		kg

Economic allocation:

Market mango is worth Rs. 500/kg while Ripe mango is worth Rs. 25/kg

Products / by-products / waste		
Market mango	500	Rs
Ripe mango	25	Rs

Emissions per kg of Mango		
Phosphate emissions to groundwater	2.09E-06	kg
Pesticides emissions to soil	8.82E-05	kg

Mass distribution		

Emissions per product		
Phosphate emissions to groundwater		kg
Pesticides emissions to soil		kg
Phosphate emissions to groundwater		kg
Pesticides emissions to soil		kg

Part 2 - Creating datasets from original sources

Content from: Andreas Ciroth, GreenDelta and Amir Safaei, ecoinvent

Managed by SETAC

Primary raw data acquisition

Objective

- In this first practical exercise, we will create datasets from original sources, i.e. from "raw data".
- Raw data = data that is not yet modeling in an inventory dataset with inputs and outputs matched to a reference flow*

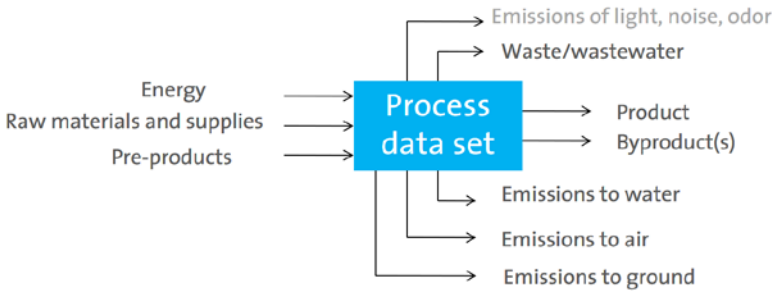
A formal definition of raw data is as follows:

- "Data used in unit process inventory modeling to deliver inventory data at the end, which are extracted from various data sources, such as bookkeeping of a plant, national statistics, or journal literature."

Source: Sonnemann, G. and B. Vigon, 2011. Global Guidance Principles for Life Cycle Assessment Databases, UN Environment Programme, ISBN: 978-92-807-3174-3, DTI/1410/PA

Primary raw data acquisition

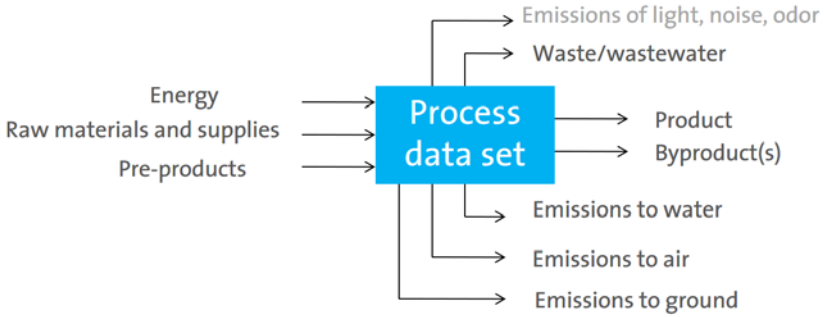
Background



- Mango production with two processes

Primary raw data acquisition

Background



- Mango production with two processes
- Two groups:

  1) LCA modeler, data collectors

  2) Mango producer, data owners

Primary raw data acquisition

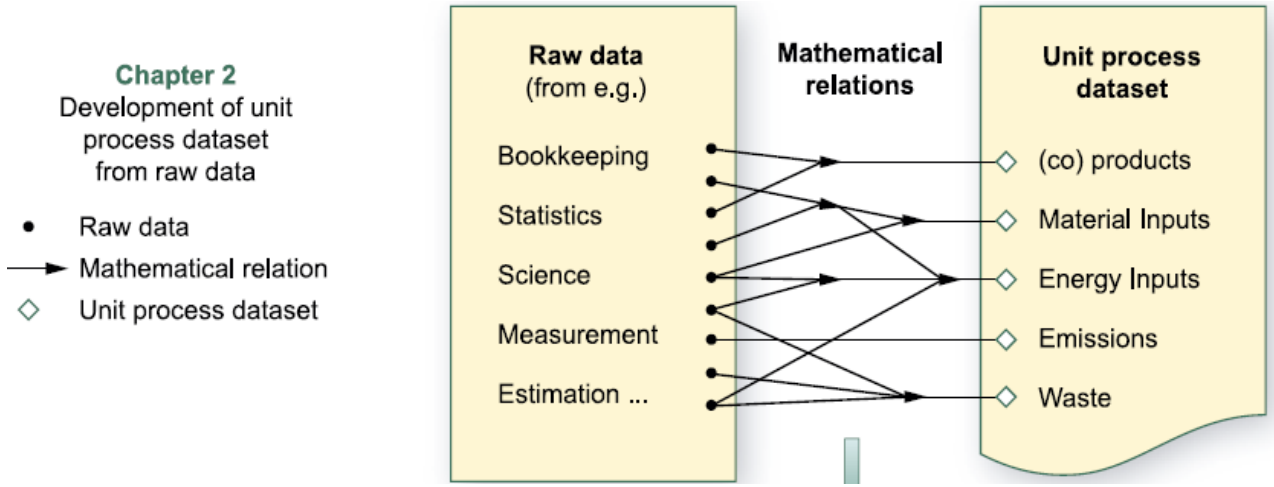
Exercise 1-- Mango Production

- Producers - Own the data, exercise guidance filename (MSWord):
  Mango_data_producer_1st_exercise_ver8-17
- Modelers - Collect the data, exercise guidance filename (MSWord):
  Mango_data_collector_modeler_1st_exercise_ver8-17
- Data collection template (generic), filename (MS Excel):
  Data_collection_template_generic_ver8-17

Primary raw data acquisition

Background

- Using "raw data" to create unit process datasets

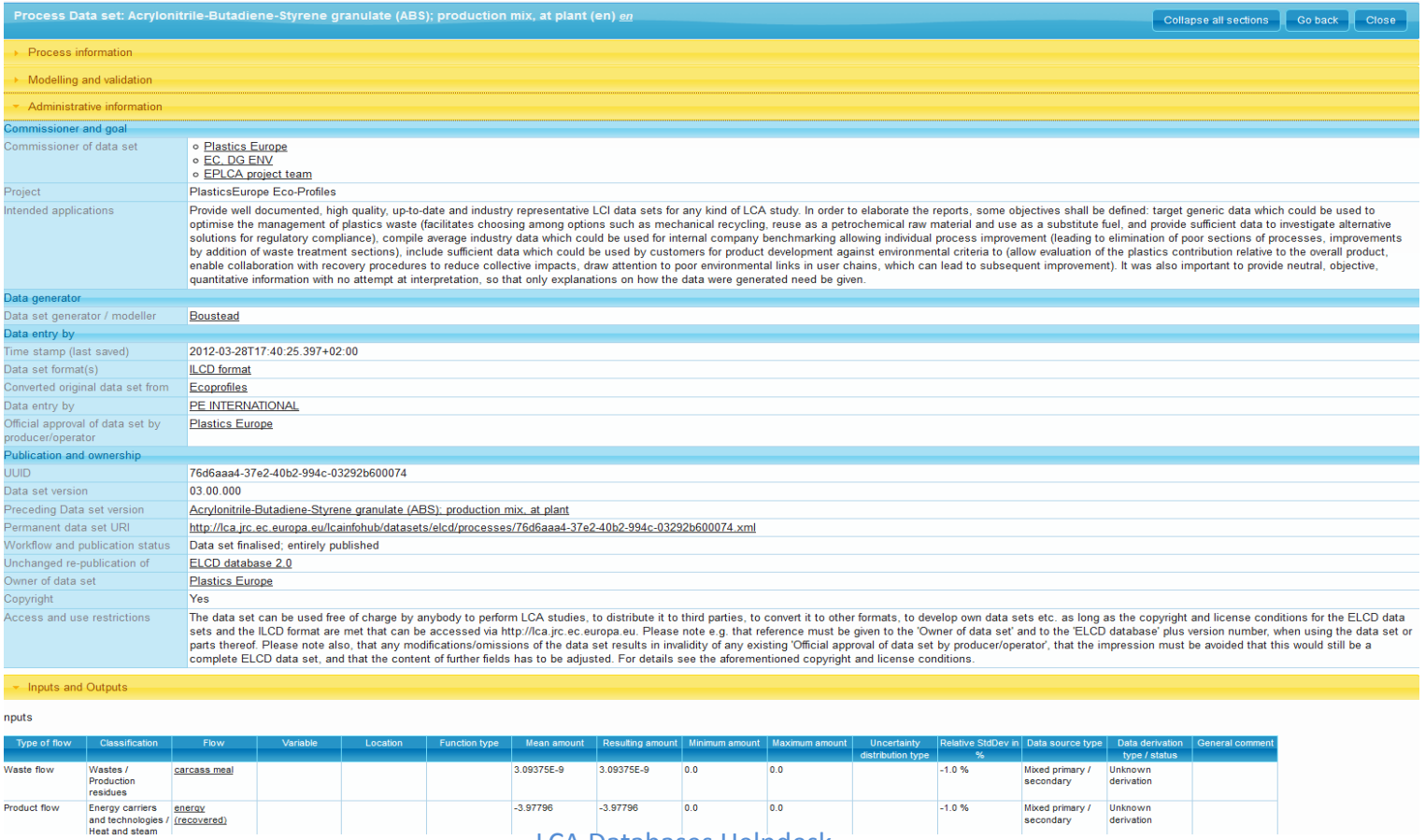


Source: Wang, H., et al., Development of Unit Process Datasets, Chapter 2 in Sonnemann and Vigon (eds.), 2011. Global Guidance Principles for Life Cycle Assessment Databases, UN Environment Programme, ISBN: 978-92-807-3174-3, DTI/1410/PA.

Primary raw data acquisition

Background

- What does a process data set in principle look like?



Process Data set: Acrylonitrile-Butadiene-Styrene granulate (ABS); production mix, at plant (en) *en*

Collapse all sections | Go back | Close

- Process information
- Modelling and validation
- Administrative information

Commissioner and goal	
Commissioner of data set	○ Plastics Europe ○ EC, DG ENV ○ EPLCA project team
Project	PlasticsEurope Eco-Profiles
Intended applications	Provide well documented, high quality, up-to-date and industry representative LCI data sets for any kind of LCA study. In order to elaborate the reports, some objectives shall be defined: target generic data which could be used to optimise the management of plastics waste (facilitates choosing among options such as mechanical recycling, reuse as a petrochemical raw material and use as a substitute fuel, and provide sufficient data to investigate alternative solutions for regulatory compliance), compile average industry data which could be used for internal company benchmarking allowing individual process improvement (leading to elimination of poor sections of processes, improvements by addition of waste treatment sections), include sufficient data which could be used by customers for product development against environmental criteria to (allow evaluation of the plastics contribution relative to the overall product, enable collaboration with recovery procedures to reduce collective impacts, draw attention to poor environmental links in user chains, which can lead to subsequent improvement). It was also important to provide neutral, objective, quantitative information with no attempt at interpretation, so that only explanations on how the data were generated need be given.
Data generator	
Data set generator / modeller	Boustead
Data entry by	
Time stamp (last saved)	2012-03-28T17:40:25.397+02:00
Data set format(s)	ILCD format
Converted original data set from	Ecoprofiles
Data entry by	PE INTERNATIONAL
Official approval of data set by producer/operator	Plastics Europe
Publication and ownership	
UUID	76d6aaa4-37e2-40b2-994c-03292b600074
Data set version	03.00.000
Preceding Data set version	Acrylonitrile-Butadiene-Styrene granulate (ABS); production mix, at plant
Permanent data set URI	http://lca.jrc.ec.europa.eu/lcainfohub/datasets/elcd/processes/76d6aaa4-37e2-40b2-994c-03292b600074.xml
Workflow and publication status	Data set finalised; entirely published
Unchanged re-publication of	ELCD database 2.0
Owner of data set	Plastics Europe
Copyright	Yes
Access and use restrictions	The data set can be used free of charge by anybody to perform LCA studies, to distribute it to third parties, to convert it to other formats, to develop own data sets etc. as long as the copyright and license conditions for the ELCD data sets and the ILCD format are met that can be accessed via http://lca.jrc.ec.europa.eu. Please note e.g. that reference must be given to the 'Owner of data set' and to the 'ELCD database' plus version number, when using the data set or parts thereof. Please note also, that any modifications/omissions of the data set results in invalidity of any existing 'Official approval of data set by producer/operator', that the impression must be avoided that this would still be a complete ELCD data set, and that the content of further fields has to be adjusted. For details see the aforementioned copyright and license conditions.

- Inputs and Outputs

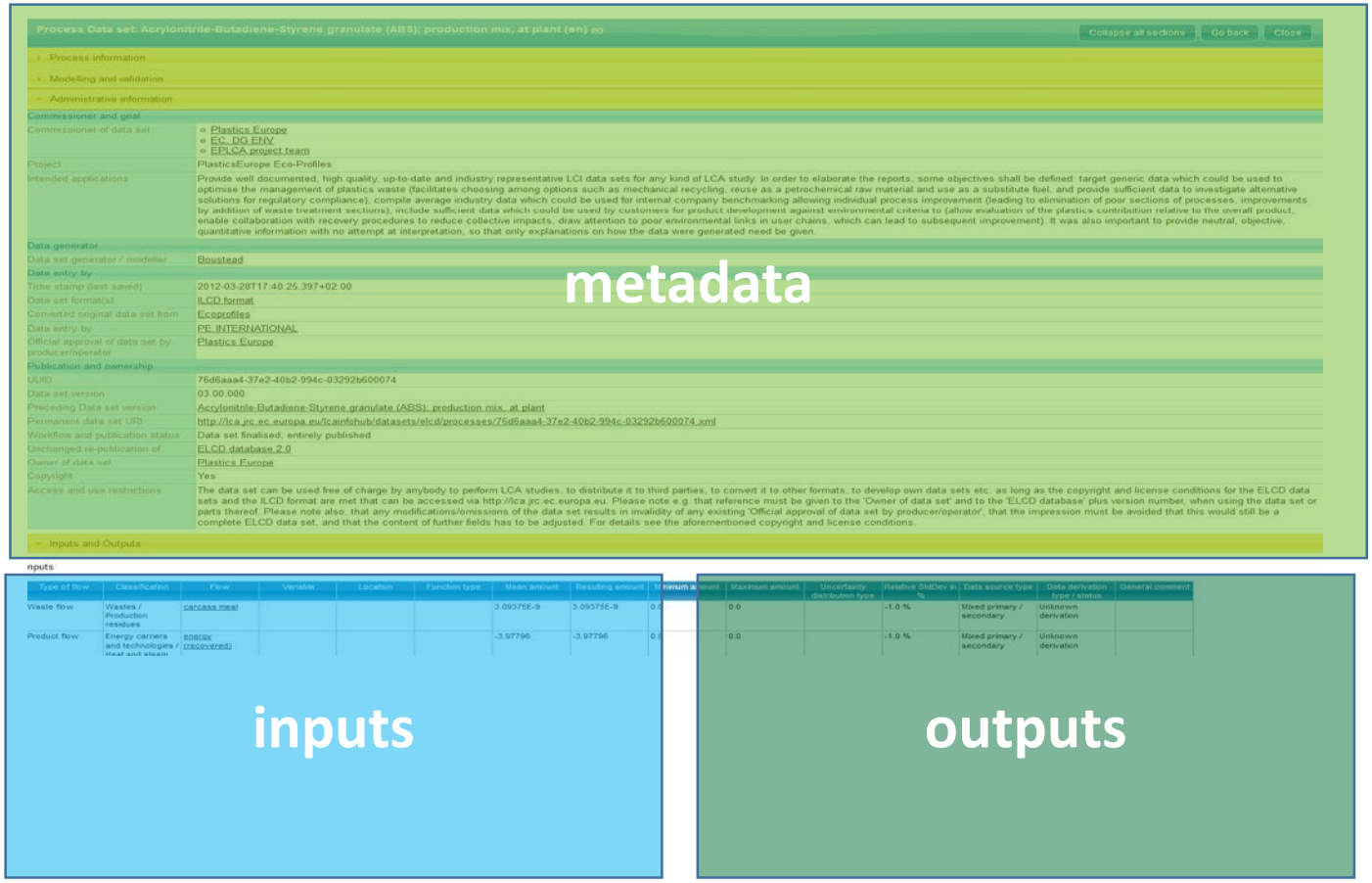
nputs

Type of flow	Classification	Flow	Variable	Location	Function type	Mean amount	Resulting amount	Minimum amount	Maximum amount	Uncertainty distribution type	Relative StdDev in %	Data source type	Data derivation type / status	General comment
Waste flow	Wastes / Production residues	carcass meal				3.09375E-9	3.09375E-9	0.0	0.0		-1.0 %	Mixed primary / secondary	Unknown derivation	
Product flow	Energy carriers and technologies / Heat and steam	energy (recovered)				-3.97796	-3.97796	0.0	0.0		-1.0 %	Mixed primary / secondary	Unknown derivation	

Primary raw data acquisition

Background

- What does a process dataset in principle look like?

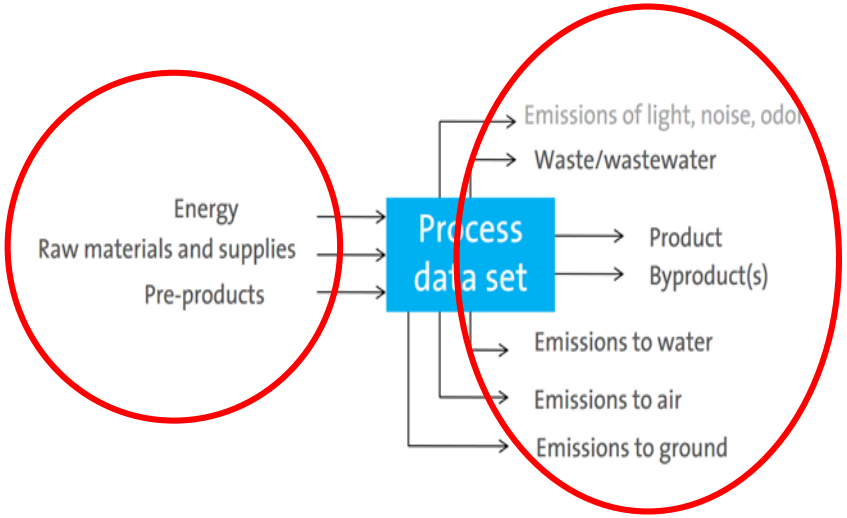


Primary raw data acquisition

The exercise

→What kind of information do we need to collect and store for a data set?

- ✓ Metadata?
- ✓ Inputs?
- ✓ Outputs?



Primary raw data acquisition

The exercise

Training Workshop on Data Acquisition and Dataset Development for Life Cycle Inventory

Exercise 1 - Primary raw data acquisition and modeling

Data for producers:[1]

The dataset covers all relevant steps involved with a Mango production from cradle-to-gate, i.e. all processes from raw materials extraction till Mango harvesting are taken into account.

PART 1.1 – Primary data collection

You are a successful producer that started your 20ha Mango plantation about 7 years ago, for that you needed to start from scratch and buy a piece of land that still had a forest on it, thus, first of all it was necessary to clean the "fields".

Preparing the field for the plantation demanded machines to cut the trees and prepare the fields, electricity to power some machinery, diesel and land.

The machines can clean and prepare 263 ha of field during its lifetime which means that for your Mango plantation only 0.076 Machine units were necessary.

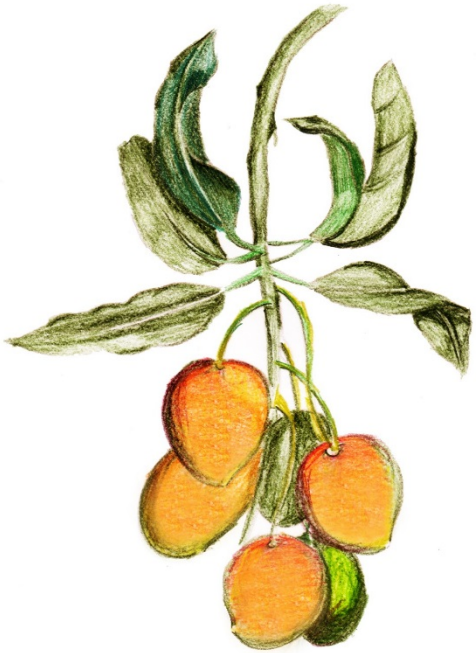
Primary raw data acquisition

The exercise*

Mango producer:

The information you have is in the exercise sheet.

You have of course an interest that your production appears environmental friendly.



* This exercise is more effective if the roles of data owner/provider and collector/modeler are played separately.

Primary raw data acquisition

The exercise

Data modeler: The mango producer is your main data source for what is happening in this specific production, but you have of course a basic understanding of what is happening when growing Mangos.

You want to model two processes:

- *Field preparation*, output is the prepared area
- *Mango production on the field*, output is the Mango produced.

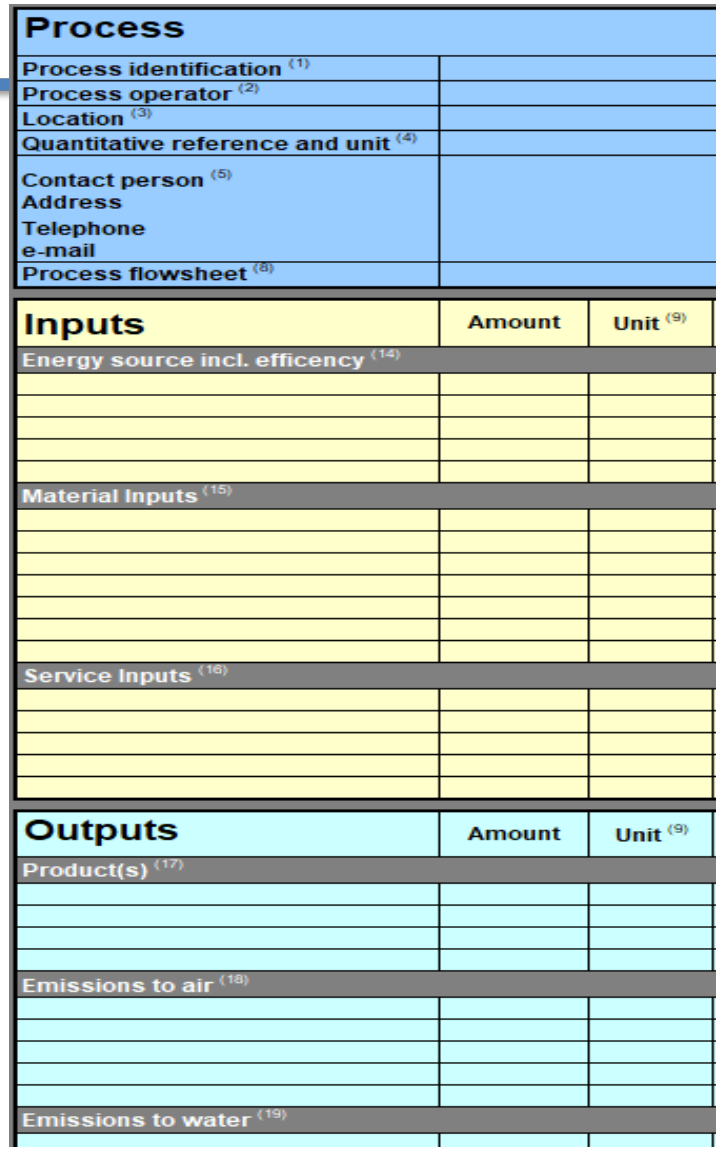
Simply create the processes in MS Excel, with inputs and outputs

Primary raw data acquisition

The exercise: example of an MS Excel sheet for data collection

In reality, two activities may be additionally needed in order to maximize the likelihood of getting complete and high quality data:

- Data collectors may need to establish a relationship with the data owners, possibly through a trade association, to help them understand the data needs better and to build credibility and
- The data collection template may need to be customized for the processes of interest. This minimizes the chances of data input errors and relieves data providers of the requirement to understand all the technicalities of the process.



Process	
Process identification [1]	
Process operator [2]	
Location [3]	
Quantitative reference and unit [4]	
Contact person [5] Address Telephone e-mail	
Process flowsheet [8]	

Inputs	Amount	Unit [9]
Energy source incl. efficency [14]		
Material Inputs [15]		
Service Inputs [16]		

Outputs	Amount	Unit [9]
Product(s) [17]		
Emissions to air [18]		
Emissions to water [19]		

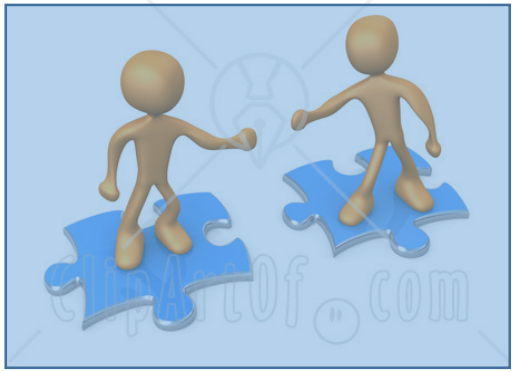
Primary raw data acquisition

The exercise - results

→The data collector groups are to present results

→The data owner groups are to describe their experience with supporting the collection effort*

* Oftentimes, data owners are given their own data (after modeling) to compare with the dataset average as a type of compensation for their cooperation.



Primary raw data acquisition

Elements in a data set, detail

- **General documentation considerations**
  - Name and classification – NACE or ISIC
  - Scope of the dataset
  - Functional unit and reference flow
  - Allocation information (if relevant)
  - Data quality information
  - Hints on interpretation
- **Specific requirements for documentation of unit process datasets**
  - Data source: raw data, mathematical relations, unit process dataset, supportive information
  - References and boundaries: description of goal and scope definition
  - Calculation models and other conventions:
  - Validation results

Primary raw data acquisition

Collecting information efficiently

- Data collection template*
- (Data collection software, with template)

*While it is always possible to use a generic data template, such as the supplied file for the trainee exercise, developing a process-specific template can have advantages. Such templates are more time-consuming to create, but once developed are more likely to make data providers more comfortable with the data request and are much less error prone. An example of such a customized template is:

Data collection template_example_customized_May 2013

Evaluate the pros and cons of a generic vs. custom template for your situation

Primary raw data acquisition

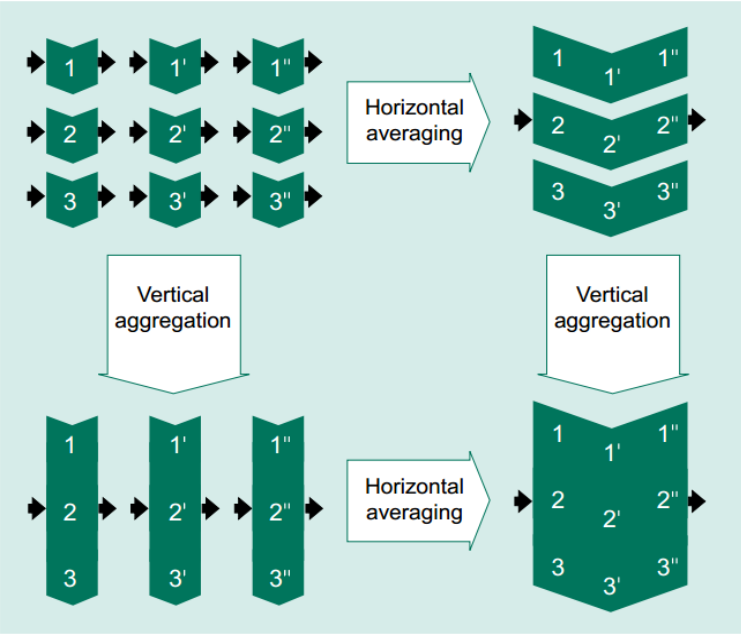
Protecting sensitive information

- Why is information collected for LCA potentially sensitive?
- Protecting sensitive information

Primary raw data acquisition

Protecting sensitive information

- Why is information collected for LCA potentially sensitive?
- Protecting sensitive information - approaches
  - Averages and aggregation



Source: Broadbent, C., et al.: Aggregated data development, Chapter 3 in Sonnemann and Vigon (eds.), 2011. Global Guidance Principles for LCA Databases, UN Environment Programme, ISBN: 978-92-807-3174-3, DTI/1410/PA.

Primary raw data acquisition

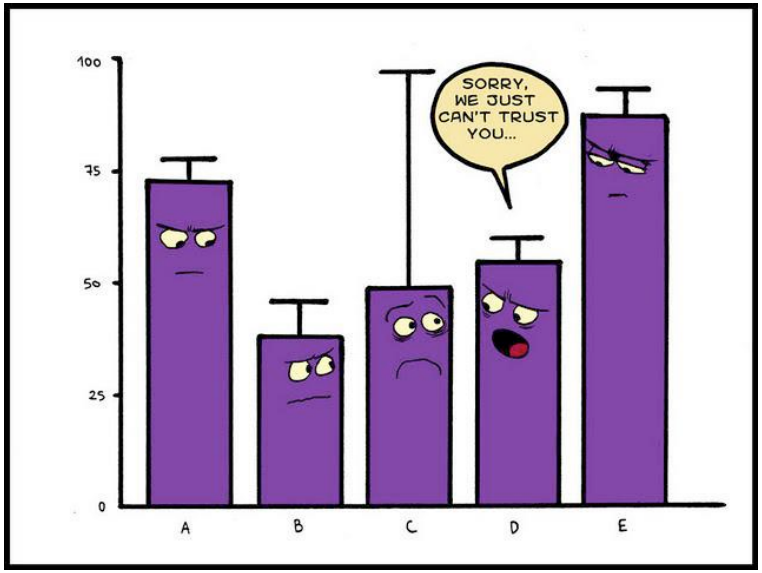
Protecting sensitive information

- Why is information collected for LCA potentially sensitive?
- Protecting sensitive information
  - Averages and aggregation
  - Remove specificity
  - Is the information really sensitive?
- Protection by trusted "man in the middle"
- Self-Protection by data owner

Primary raw data acquisition

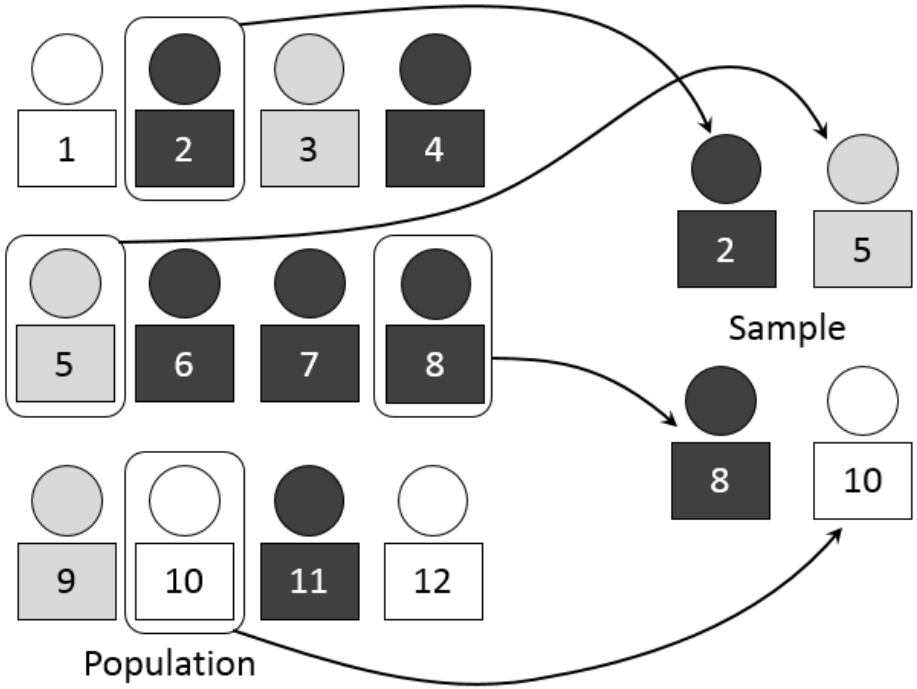
How to create representative data sets

- Common practice in LCA: market share
- This is not corresponding to scientific practice!
- Scientific: Statistical sampling



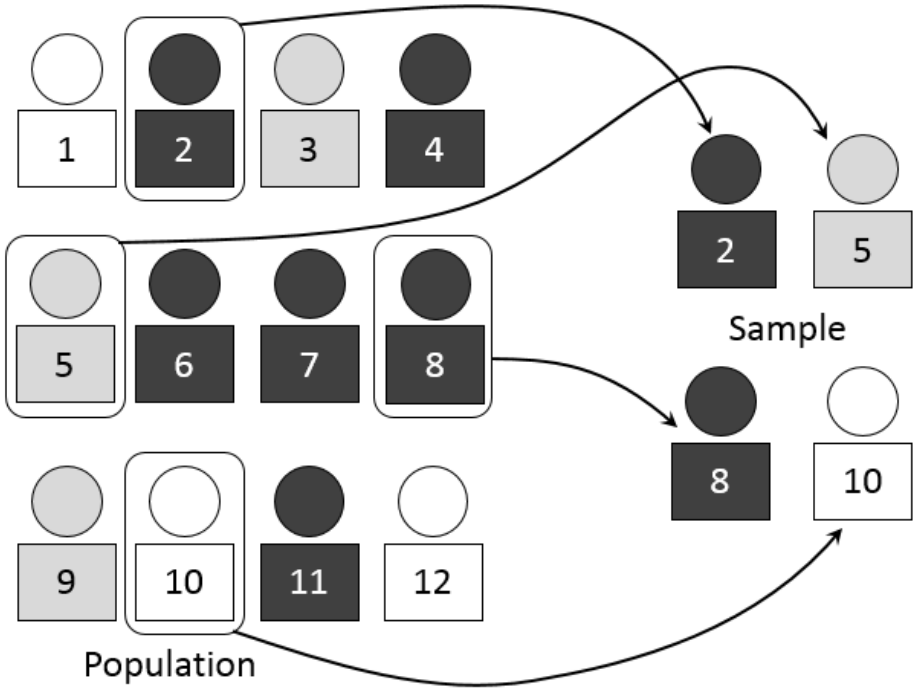
Primary raw data acquisition

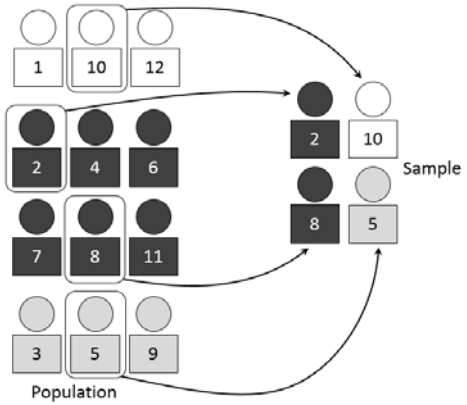
How to create representative data sets: statistical sampling



Primary raw data acquisition

How to create representative data sets:
statistical sampling

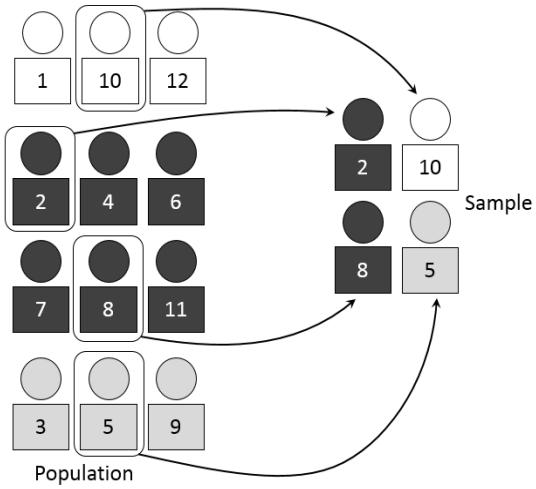


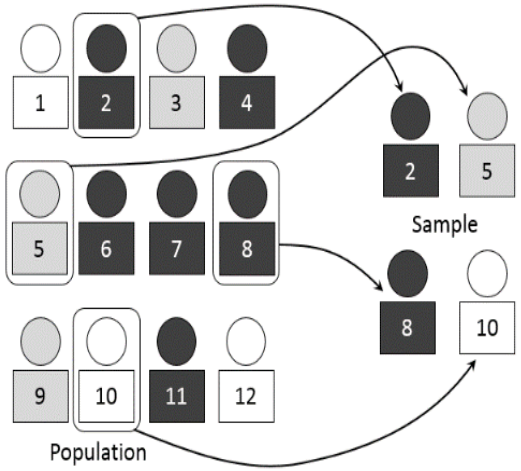


Primary raw data acquisition

How to create representative data sets: statistical sampling

- Would you see any issues in applying this technique for LCA data?





Training Workshop on Data Acquisition and Dataset Development for Life Cycle Inventory

Exercise 1 - Primary raw data acquisition and modelling
Exemplary Answer

Data for producers:[1]

The dataset covers all relevant steps involved with a Mango production from cradle-to-gate, i.e. all processes from raw materials extraction till Mango harvesting are taken into account.

PART 1.1 – Primary data collection

You are a successful producer that started your 20ha Mango plantation about 7 years ago, for that you needed to start from scratch and buy a piece of land that still had a forest on it, thus, first of all it was necessary to clean the "fields".

Preparing the field for the plantation demanded machines to cut the trees and prepare the fields, electricity to power some machinery, diesel and land.

The machines can clean and prepare 263 ha of field during its lifetime which means that for your Mango plantation only 0.076 Machine units were necessary.

In total 13360 kW of electricity to power machines working 4 hours per hectare was

needed. 13360/4 = 3340 kWh per 20 ha

3340/20 = 167 kWh /ha

Necessary materials and resources		
Electricity	167	kWh
Tree cutting machine	0.0038	Units
Diesel	1.61	MJ
Land	1	ha

In the End of this process you had 20ha of clean and prepared field, 1000m³ of Biomass coming from the cut trees that need to be disposed and emissions of pollutants to the air in a low populated area.

Products / by-products / waste		
land, prepared for tree growing	1	ha
Biomass	50	m³

Emissions to air per hectare		
CO2 (biogenic)	21.14	Kg
N2O	2.6	Kg
CO2 (fossil)	78.58	Kg
PM (particulate matter)	15	Kg

PART 1.2 - Primary data collection

[1] This fact sheet describes a simplified Mango production farm based on data from EcoInvent and Agribalyse databases.

After preparing the field you needed to start your production, for this step you needed to consider again all the necessary resources and materials.

The plantation needs irrigation through the whole year, each plant consumes 100 liters of water per week and there is a total of 140 plants per hectare.

You expect you can produce Mangos on this field for at least 25 years.

WATER:

Assumed water consumption: 100 liters per plant per week (140 plants per hectare), irrigation through the whole year:

100*140*52 = 728,000 l = 728 m3/ha/year

Water consumption: 728/15 = 48.53 m3/t = 0.04853 m3/kg

The irrigation system uses a 2.5kW pump per hectare that works every day for 12h.

ELECTRICITY:

One engine/pumping station that powers the irrigation machine, with a power of 2.5 kW; the irrigation machine works every day, but only half day (=8760/2=4380h):

4380 * 2.5 = 10,950 kWh/year

Electricity use: 10,950/15 = 730 kWh/t = 0.73 kWh/kg

The land needs to be fertilized every year, therefore 111.9 kg of fertilizer is applied to each hectare of cultivated land per year.

FERTILIZER:

111.9/15 = 7.46 kg of fertilizer per 1000kg Mango produced

Fertilizer use: 0.00746 kg/kg (Mango)

Land use change: 10000m²/15000kg Mangos/25years = 0.0267 m² land use change per kg Mango

DATA:

Mango production area: 20 ha

Annual yield: 15 t/ha/year

Necessary materials and resources		
Electricity for irrigation	0.73	kWh
Fertilizer (Phosphate P2O5)	0.00746	kg
Pesticides	8.82E-05	kg
Water for irrigation	0.04853	m³
Transport of fertilizers and pesticides to plant	200	km
Land use change	0.0267	m²

Products / by-products / waste		
Mango	1	kg

Emissions per kg of Mango		
Phosphate emissions to groundwater	2.09E-06	kg
Pesticides emissions to soil	8.82E-05	kg

PART 2 – Allocation

In the end of the harvest you have 2 products:

- Mangos for selling at the market
- Mangos that are too ripe and can only be consumed after being processed (jams, dry mango snacks, Mango pickles, etc.)

The main product is the Market mango and the Ripe mango is the by product which is 15% of the production and is responsible for 5% of the farm income.

Because 2 products are generated the environmental burdens need to be correctly split between them. There are many ways of doing it but we will check the results for mass allocation and economic allocation.

Mass allocation:

1 ton of Mango = 850 kg of Market mango + 150 kg of Ripe mango, but emissions are for 1 kg of Mango

Products / by-products / waste		
Market mango	0.85	kg
Ripe mango	0.15	kg

Emissions per kg of Mango		
Phosphate emissions to groundwater	2.09E-06	kg
Pesticides emissions to soil	8.82E-05	kg

Mass distribution		
Market mango	85	%
Ripe mango	15	%

Emissions per product		
Phosphate emissions to groundwater	1.78E-06	kg
Pesticides emissions to soil	7.50E-05	kg
Phosphate emissions to groundwater	0.31E-06	kg
Pesticides emissions to soil	1.32E-05	kg

Economic allocation:

Market mango is worth Rs. 500/kg while Ripe mango is worth Rs. 25/kg

Products / by-products / waste		
Market mango	500	Rs
Ripe mango	25	Rs

Emissions per kg of Mango		
Phosphate emissions to groundwater	2.09E-06	kg
Pesticides emissions to soil	8.82E-05	kg

Mass distribution		
Market mango	95	%
Ripe mango	5	%

Emissions per product		
Phosphate emissions to groundwater	1.99E-06	kg
Pesticides emissions to soil	8.38E-05	kg
Phosphate emissions to groundwater	0.10E-06	kg
Pesticides emissions to soil	0.44E-05	kg

Data for dataset producers:

You were recently hired to create datasets for mango producers that are interested in modeling their production chain from cradle to gate.

To create the datasets you will need all the information about the elements involved in Mango production since raw materials extraction and preprocessing until the harvest or storage of Mangos.

There are many ways of obtaining such information but the best way, in order to create a reliable dataset, is to go to the source i.e. interview the mango producers and find out what they need for their production.

It is worth noticing that the dataset may consist of more than one process and although for this exercise all producers cultivate mangos the same way, in reality there are several ways of producing the fruit, manure the soil, irrigate and so on.

To create your dataset, take the empty data collection formulary and interview one of the mango producers that should provide you a high quality primary data.

Define a quantitative reference and mind units and amounts while collecting the data.

(Handout: empty formulary from excel table)

Training on Data Acquisition and Dataset Development for Life Cycle Inventory

Exercise 1 - Primary raw data acquisition and modeling

Data for data collectors and dataset modelers:

You were recently hired to create datasets for mango producers that are interested in modeling their production chain from cradle to gate.

To create the datasets you will need all the information about the elements involved in Mango production from raw materials extraction and preprocessing until the harvest or storage of Mangos.

There are many ways of obtaining such information but the best way, in order to create a reliable dataset, is to go to the source i.e. interview the mango producers and find out what they use for their production.

It is worth noticing that the dataset may consist of more than one process and although for this exercise all producers cultivate mangos the same way, in reality there are several ways of producing the fruit, manure the soil, irrigate and so on. This might require what is called sample stratification. Stratification is a way to define relevant differences among producers and then refining your data collection strategy to ensure that the eventual dataset, e.g. a national production-weighted average, or datasets are representative

To create your dataset, take the empty data collection template and interview one of the mango producers that should provide you with primary data.

Define a quantitative reference flow and keep in mind the correct units and reasonableness of amounts while collecting the data.

Part 2a - Creating datasets from secondary data and quality assurance

Content from: Andreas Ciroth, GreenDelta
and Amir Safaei, ecoinvent

Managed by
SETAC

Primary raw data acquisition, 2a

Now we have collected data..

- What's next:
  - Input of data into a common form
  - Quality assurance (QA)
  - Documentation

Primary raw data acquisition, 2a

Data input and compilation

- MS Excel (generic or custom)
- Specific solutions
  - Ecoeditor: www.ecoinvent.org/data-provider/data-provider-toolkit/ecoeditor/ecoeditor.html
  - limesurvey query tool, configured (open source survey tool, web-based)
  - Excel template for openLCA
  - …
- LCA software:
  - ..

Primary raw data acquisition, 2a

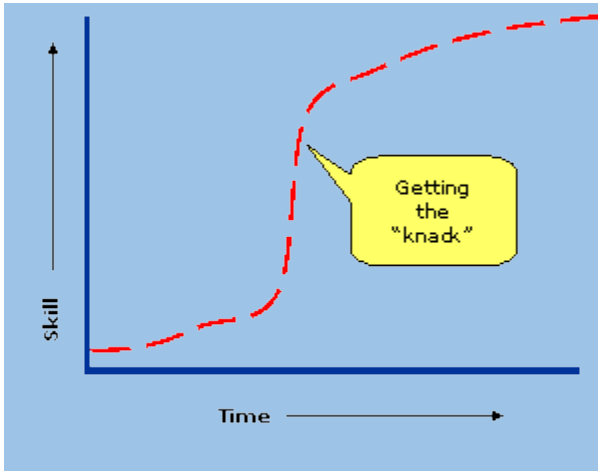
Data input and compilation

- LCA software:
  - openLCA
  - openLCA, with Collaboration Server (www.openlca.org/collaboration-server/)
  - openLCA, with soda4LCA hub
  - SimaPro
  - SimaPro, with share&collect
  - GaBi
  - GaBi, with LCA hub
  - …

Primary raw data acquisition, 2a

Data input and compilation

- Considerations in selecting :
  - ✓ Consistency
  - ✓ Costs and effort
  - ✓ Ease of use
  - ✓ Learning curves



Primary raw data acquisition, 2a

Data verification and quality assurance

- Did you collect or did you receive "correct" datasets?

Primary raw data acquisition, 2a

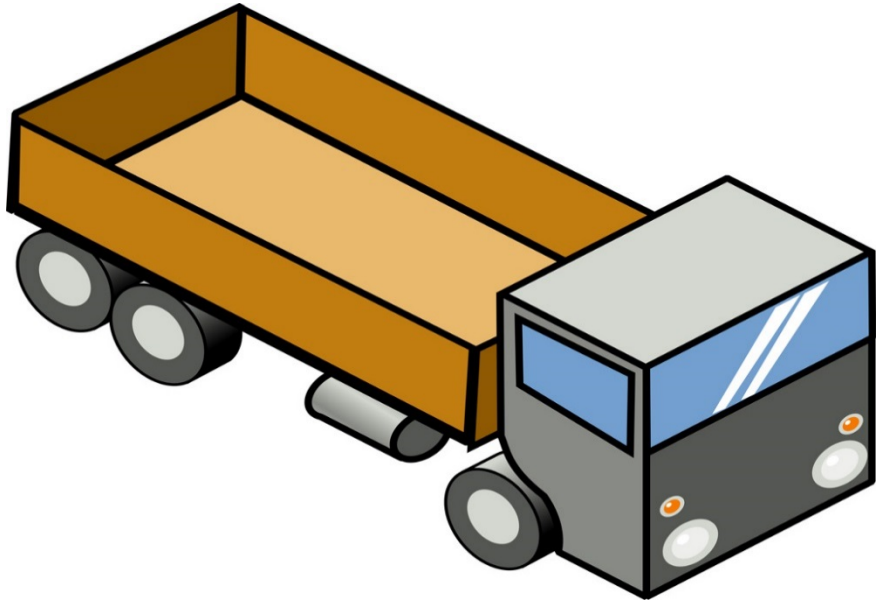
Data verification and quality assurance

- Did you collect, did you receive "correct data sets"?
  - What would you expect?
    - Missing flows
    - Unnecessary flows
    - Units plausible
  - Mass balance, energy balance
  - Relations between flows
    - water in, water out
    - SO2 output by fuel S-content

Primary raw data acquisition, 2a

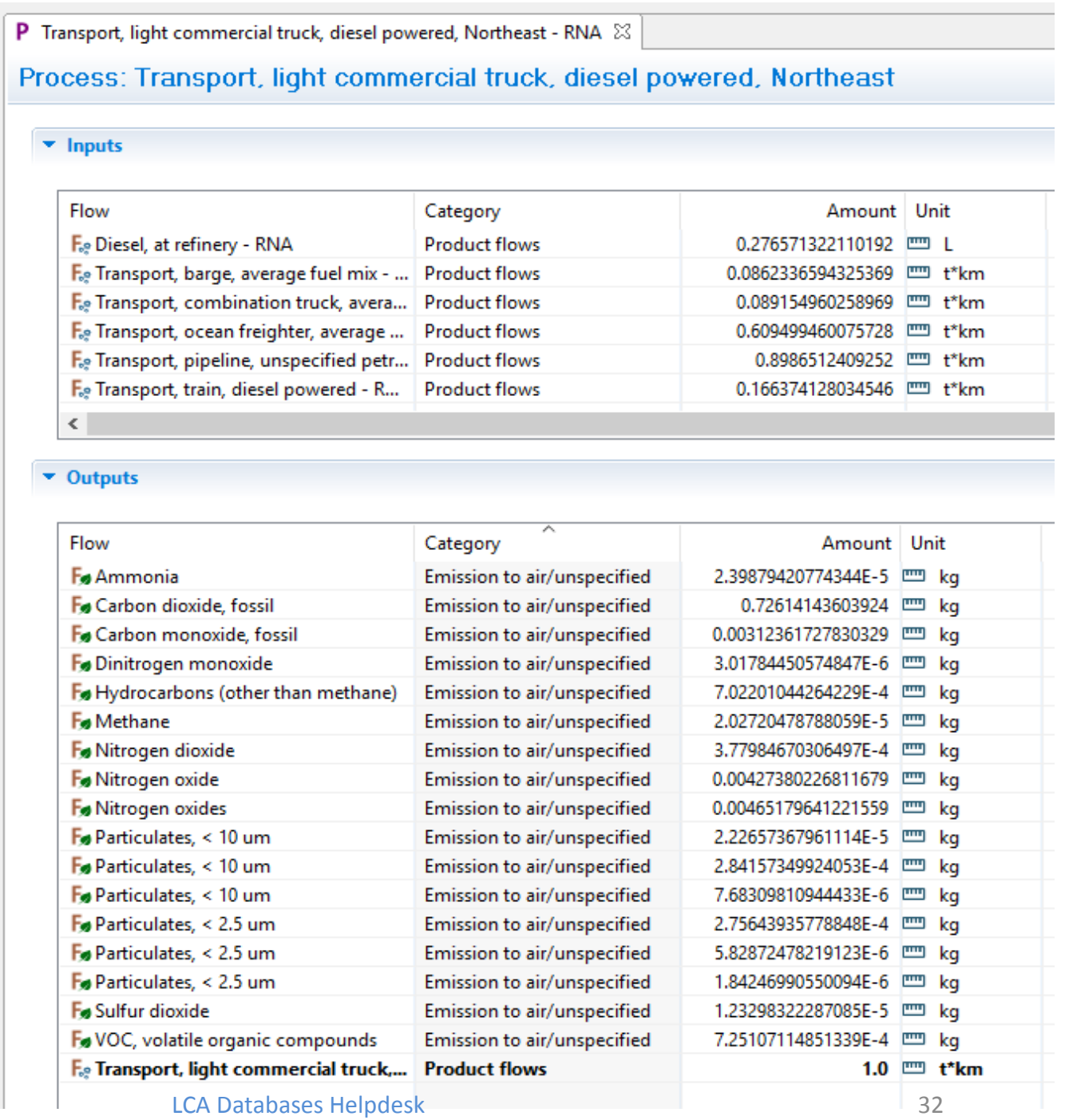
QA, real processes (already in databases)
Light commercial truck transport, USA

- What would you expect?



QA, real processes

From USLCA Commons database, 2004


Transport, light commercial truck, diesel powered, Northeast - RNA

Process: Transport, light commercial truck, diesel powered, Northeast

Inputs

Flow	Category	Amount	Unit
Diesel, at refinery - RNA	Product flows	0.276571322110192	L
Transport, barge, average fuel mix - ...	Product flows	0.0862336594325369	t*km
Transport, combination truck, avera...	Product flows	0.089154960258969	t*km
Transport, ocean freighter, average ...	Product flows	0.609499460075728	t*km
Transport, pipeline, unspecified petr...	Product flows	0.8986512409252	t*km
Transport, train, diesel powered - R...	Product flows	0.166374128034546	t*km

Outputs

Flow	Category	Amount	Unit
Ammonia	Emission to air/unspecified	2.39879420774344E-5	kg
Carbon dioxide, fossil	Emission to air/unspecified	0.72614143603924	kg
Carbon monoxide, fossil	Emission to air/unspecified	0.00312361727830329	kg
Dinitrogen monoxide	Emission to air/unspecified	3.01784450574847E-6	kg
Hydrocarbons (other than methane)	Emission to air/unspecified	7.02201044264229E-4	kg
Methane	Emission to air/unspecified	2.02720478788059E-5	kg
Nitrogen dioxide	Emission to air/unspecified	3.77984670306497E-4	kg
Nitrogen oxide	Emission to air/unspecified	0.00427380226811679	kg
Nitrogen oxides	Emission to air/unspecified	0.00465179641221559	kg
Particulates, < 10 um	Emission to air/unspecified	2.22657367961114E-5	kg
Particulates, < 10 um	Emission to air/unspecified	2.84157349924053E-4	kg
Particulates, < 10 um	Emission to air/unspecified	7.68309810944433E-6	kg
Particulates, < 2.5 um	Emission to air/unspecified	2.75643935778848E-4	kg
Particulates, < 2.5 um	Emission to air/unspecified	5.82872478219123E-6	kg
Particulates, < 2.5 um	Emission to air/unspecified	1.84246990550094E-6	kg
Sulfur dioxide	Emission to air/unspecified	1.23298322287085E-5	kg
VOC, volatile organic compounds	Emission to air/unspecified	7.25107114851339E-4	kg
Transport, light commercial truck,...	**Product flows**	**1.0**	**t*km**

QA, real processes

1. l Diesel / 100km?
2. S content in Diesel?
3. CO2 emissions in relation to Diesel consumption?
4. Mass balance?
5. Inputs?
6. Other aspects?

Transport, light commercial truck, diesel powered, Northeast - RNA

Process: Transport, light commercial truck, diesel powered, Northeast

Inputs

Flow	Category	Amount	Unit
Diesel, at refinery - RNA	Product flows	0.276571322110192	L
Transport, barge, average fuel mix - ...	Product flows	0.0862336594325369	t*km
Transport, combination truck, avera...	Product flows	0.089154960258969	t*km
Transport, ocean freighter, average ...	Product flows	0.609499460075728	t*km
Transport, pipeline, unspecified petr...	Product flows	0.8986512409252	t*km
Transport, train, diesel powered - R...	Product flows	0.166374128034546	t*km

Outputs

Flow	Category	Amount	Unit
Ammonia	Emission to air/unspecified	2.39879420774344E-5	kg
Carbon dioxide, fossil	Emission to air/unspecified	0.72614143603924	kg
Carbon monoxide, fossil	Emission to air/unspecified	0.00312361727830329	kg
Dinitrogen monoxide	Emission to air/unspecified	3.01784450574847E-6	kg
Hydrocarbons (other than methane)	Emission to air/unspecified	7.02201044264229E-4	kg
Methane	Emission to air/unspecified	2.02720478788059E-5	kg
Nitrogen dioxide	Emission to air/unspecified	3.77984670306497E-4	kg
Nitrogen oxide	Emission to air/unspecified	0.00427380226811679	kg
Nitrogen oxides	Emission to air/unspecified	0.00465179641221559	kg
Particulates, < 10 um	Emission to air/unspecified	2.22657367961114E-5	kg
Particulates, < 10 um	Emission to air/unspecified	2.84157349924053E-4	kg
Particulates, < 10 um	Emission to air/unspecified	7.68309810944433E-6	kg
Particulates, < 2.5 um	Emission to air/unspecified	2.75643935778848E-4	kg
Particulates, < 2.5 um	Emission to air/unspecified	5.82872478219123E-6	kg
Particulates, < 2.5 um	Emission to air/unspecified	1.84246990550094E-6	kg
Sulfur dioxide	Emission to air/unspecified	1.23298322287085E-5	kg
VOC, volatile organic compounds	Emission to air/unspecified	7.25107114851339E-4	kg
Transport, light commercial truck,...	**Product flows**	**1.0**	**t*km**

Primary raw data acquisition, 2a

Data documentation

- Author
- Date
- Aim of the data set
- Sources
- Potential modifications of the data set
- Limitations
- …

Primary raw data acquisition, 2a

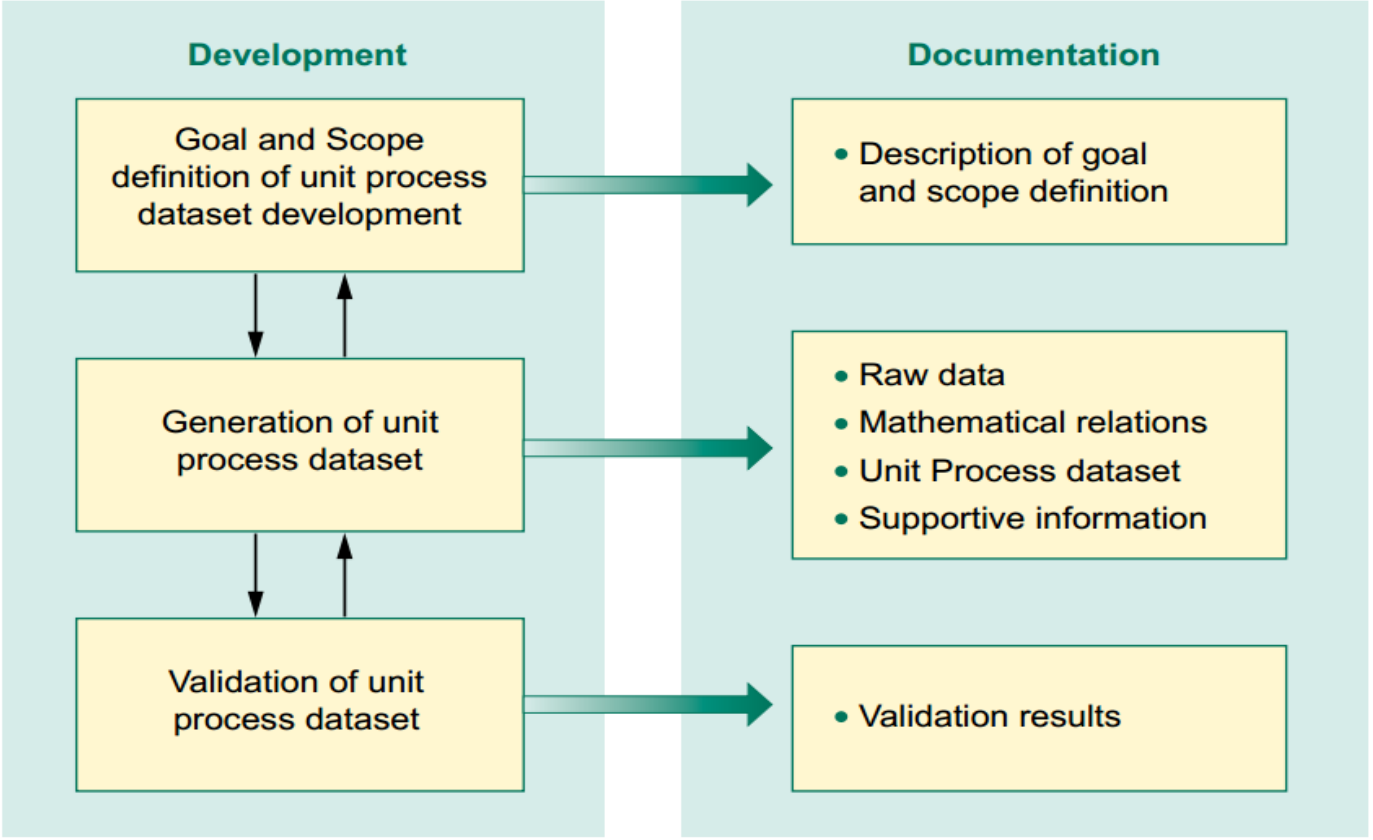
Data documentation

- Author
- Date
- Aim of the data set
- Sources
- Potential modifications of the data set
- Limitations
- …

- At which level?
  - Dataset
  - Individual input/output?

Primary raw data acquisition, 2a

Data documentation



Source: Wang, H., et al., Development of Unit Process Datasets, Chapter 2 in Sonnemann and Vigon (eds.), 2011. Global Guidance Principles for Life Cycle Assessment Databases, UN Environment Programme, ISBN: 978-92-807-3174-3, DTI/1410/PA.

Primary raw data acquisition, 2a

Finally, "metaindicators"

Metaindicators = metadata descriptors

→ everything that describes a process to understand "fitness for purpose" (data quality) for a process data set.

Metaindicators

UNEGLAD WG 3 metadata descriptors, task 3, final draft report, June 2017

See also Part 2b on Meta-indicators and GLAD

		Goal	Value & representation	Conformance
ID	Process name		0a	
	Process type		0b	
gvc Descriptors	Time	Ia	IVa	IIa
	Geography	Ib	IVb	IIb
	Technology	Ic	IVc	IIc
	Model completeness	Id	IVf	IId
	Sample representativeness	Ie	IVg	IIe
	LCA nomenclature systems		IVd	
	LCIA methods	Ig	IVe	
Modeling	LCI modeling type		IVh	
	System boundaries		IVi	
	Multifunctional processes		IVj	
	Biogenic carbon		IVk	
	Land use		IVl	
	Wastes and end-of-life		IVm	
	Water		IVn	
	Infrastructure/capital goods		IVo	
	Long-term emissions		IVp	
	Temporary carbon storage		IVq	
Sampling	Sample approach	If		
	Reliability of the sources used		IIIa	
Calculation	Aggregation type if any		VIa	
QA	Data set review performed		Va	
	Type of data set review		Vb	
	Quality assurance performed		Vc	
	Reviewing person(s)		Vd	
Administrative	Copyright protected data set?		VIIa	
	Copyright holder		VIIb	
	Free data set or for purchase?		VIIc	
	Data set license		VIId	
	Data set contact		VIIe	

Descriptor element supported in / provided by

GLAD	ILCD	ILCD & EcoSpold02	EcoSpold02	not applicable	not foreseen	(ILCD)	(ILCD & EcoSpold02)	(EcoSpold02)

Primary raw data acquisition, 2a

Exercise, continued:

Do you need to add any metadata to your process, as an LCA modeler?

As producer, can you tell if additional metadata that are needed?

Part 3 - Creating datasets from public sources

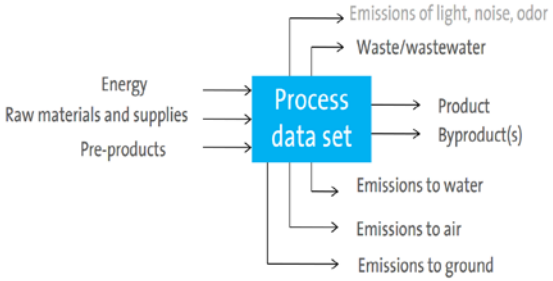
Content from Andreas Ciroth, GreenDelta
and Amir Safaei, ecoinvent

August 2017 Version

Managed by SETAC

Secondary data acquisition

Objective and Background

- In the second practical exercise, we will create datasets from secondary data, i.e. from data already published.
- 

- What are secondary sources that can be used here?

Secondary data acquisition

Where are secondary data available?

- Public, official statistics
- Wikipedia
- Company websites
- Journal papers
- Chemical engineering books
- Other LCA databases and datasets (!)
- …

Secondary data acquisition

Secondary data, example

Sri Lanka Energy Balance

Compiled by Sri Lanka Sustainable Energy Authority

Home | Electricity | Petroleum Products | Biomass | Energy Balance | Economic Indicators | Grid Emission | Conversion Factors | Coal

- Home
- Electricity Data
  - Overview
  - Electricity Generation
  - Own Use in Power Plants and Network Losses
  - Fuel Consumption in Power Plants
  - Electricity Sales by Sector
  - Capacities of Power Plants
  - Electricity Tariff
  - Customers, Employees and Finances
  - System Load Profile
- Petroleum Data
- Biomass Data
- Energy Balance
- Energy-Econ Indicators
- Grid Emission Factor
- Conversion Factors
- Coal

Gross Electricity Generation (GWh)

Select Year 2011 to 2015 | Update Table | Export to Excel

Links

Gross Generation of Thermal Power Stations
Gross Generation of Hydro Power Stations

CEB-Hydro (GWh)

	2011	2012	2013	2014	2015
Major Hydro	3,906.96	2,684.17	5,936.61	3,618.83	4,825.65
Small Hydro	65.72	42.55	73.49	30.89	78.76
Sub Total Hydro	3,972.67	2,726.72	6,010.10	3,649.72	4,904.41

CEB Non-Conventional (GWh)

	2011	2012	2013	2014	2015
CEB Wind	2.66	2.32	2.32	2.13	1.07
Sub Total Non Conventional	2.66	2.32	2.32	2.13	1.07

CEB-Thermal (GWh)

	2011	2012	2013	2014	2015
Steam - Fuel Oil	0.00	0.00	0.00	0.00	0.00
Steam - Coal	1,027.62	1,399.12	1,465.39	3,505.55	4,447.21
Steam - Diesel	10.49	4.61	3.98	19.40	9.97
Sub Total Steam	1,038.11	1,403.74	1,469.37	3,524.95	4,457.18
Diesel Engine - Residual Oil	903.37	921.55	559.81	647.10	271.93
Diesel Engine - Fuel Oil	0.00	0.00	111.48	95.69	87.94
Diesel Engine - Diesel Oil	14.19	9.42	27.26	9.31	22.48
Sub Total Diesel Engines	917.56	930.97	698.55	752.09	382.34
Gas Turbines - Diesel Oil	320.33	218.20	17.58	241.88	25.07
Gas Turbines - Naptha	0.00	0.00	0.00	0.00	0.00
Sub Total Gas Turbines	320.33	218.20	17.58	241.88	25.07
Combined Cycle - Diesel Oil	95.90	550.66	221.75	284.64	119.54
Combined Cycle - Naptha	159.78	329.05	388.53	465.54	540.26
Sub Total Combined Cycle	255.68	879.71	610.27	750.18	659.80

CEB - Sub Total (GWh)

	2011	2012	2013	2014	2015
Sub Total CEB Power plants	6,507.02	6,161.66	8,808.20	8,920.95	10,429.87

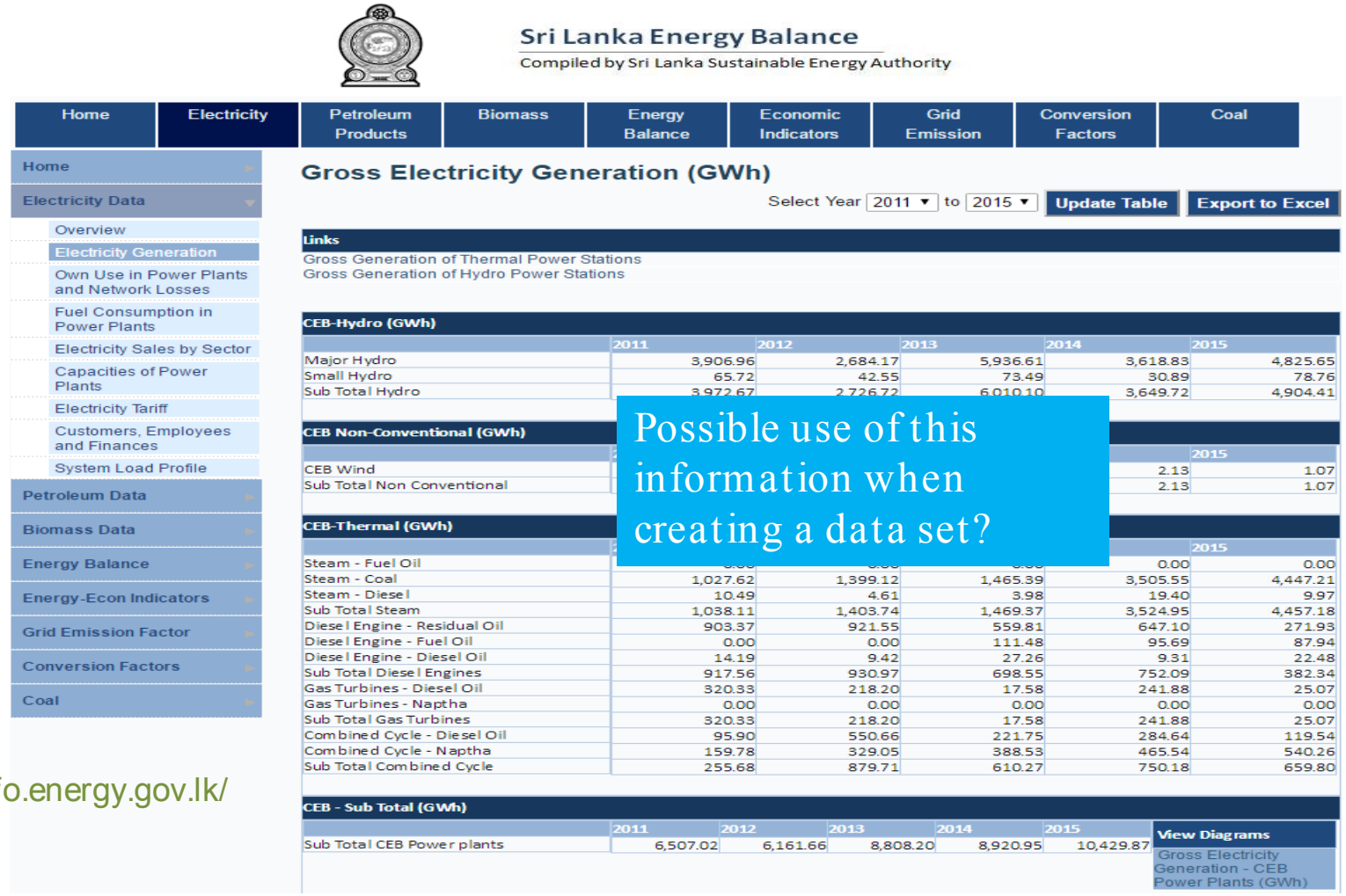
View Diagrams

Gross Electricity Generation - CEB Power Plants (GWh)

http://www.info.energy.gov.lk/

Secondary data acquisition

Secondary data, example



Sri Lanka Energy Balance

Compiled by Sri Lanka Sustainable Energy Authority

Home | Electricity | Petroleum Products | Biomass | Energy Balance | Economic Indicators | Grid Emission | Conversion Factors | Coal

- Home
- Electricity Data
  - Overview
  - Electricity Generation
  - Own Use in Power Plants and Network Losses
  - Fuel Consumption in Power Plants
  - Electricity Sales by Sector
  - Capacities of Power Plants
  - Electricity Tariff
  - Customers, Employees and Finances
  - System Load Profile
- Petroleum Data
- Biomass Data
- Energy Balance
- Energy-Econ Indicators
- Grid Emission Factor
- Conversion Factors
- Coal

Gross Electricity Generation (GWh)

Select Year 2011 to 2015 | Update Table | Export to Excel

Links

Gross Generation of Thermal Power Stations

Gross Generation of Hydro Power Stations

CEB-Hydro (GWh)

	2011	2012	2013	2014	2015
Major Hydro	3,906.96	2,684.17	5,936.61	3,618.83	4,825.65
Small Hydro	65.72	42.55	73.49	30.89	78.76
Sub Total Hydro	3,972.67	2,726.72	6,010.10	3,649.72	4,904.41

CEB Non-Conventional (GWh)

	2015
CEB Wind	1.07
Sub Total Non Conventional	1.07

CEB-Thermal (GWh)

	2015
Steam - Fuel Oil	0.00
Steam - Coal	4,447.21
Steam - Diesel	9.97
Sub Total Steam	4,457.18
Diesel Engine - Residual Oil	271.93
Diesel Engine - Fuel Oil	87.94
Diesel Engine - Diesel Oil	22.48
Sub Total Diesel Engines	382.34
Gas Turbines - Diesel Oil	25.07
Gas Turbines - Naptha	0.00
Sub Total Gas Turbines	25.07
Combined Cycle - Diesel Oil	119.54
Combined Cycle - Naptha	540.26
Sub Total Combined Cycle	659.80

CEB - Sub Total (GWh)

	2011	2012	2013	2014	2015
Sub Total CEB Power plants	6,507.02	6,161.66	8,808.20	8,920.95	10,429.87

View Diagrams

Gross Electricity Generation - CEB Power Plants (GWh)

Possible use of this information when creating a data set?

http://www.info.energy.gov.lk/

Secondary data acquisition

Secondary data – challenges, issues

- Quite often, data does not fully fit to what you need
  - Incomplete
    - Product missing
    - Only some emissions reported, or only composition of products
  - Not fully fitting (slightly different product, location, time than needed)
- Source reliability

Secondary data acquisition

Secondary data – challenges, issues

- Quite often, data does not fully fit to what you need
  - Incomplete
    - Product missing
    - Only some emissions reported, or only composition of products
  - Not fully fitting (slightly different product, location, time than needed)
- Source reliability

→ Need attention and care, but great potential.
Typically, results are "patchwork datasets"

Secondary data acquisition

Secondary data – challenges, issues

Exercise:

- Creating a national grid electricity data set for Sri Lanka
  - Files: Data_2nd_Exercise.docx, 2nd_exercise_processes.xlsx
- See MS Excel workbook and hand-outs
- Would you agree to the mix?
- Would you agree to the proposed processes?
- Is something essential lacking (yes..really..what is it)? How would you acquire the missing bits?

Exercise 2 – Secondary Data

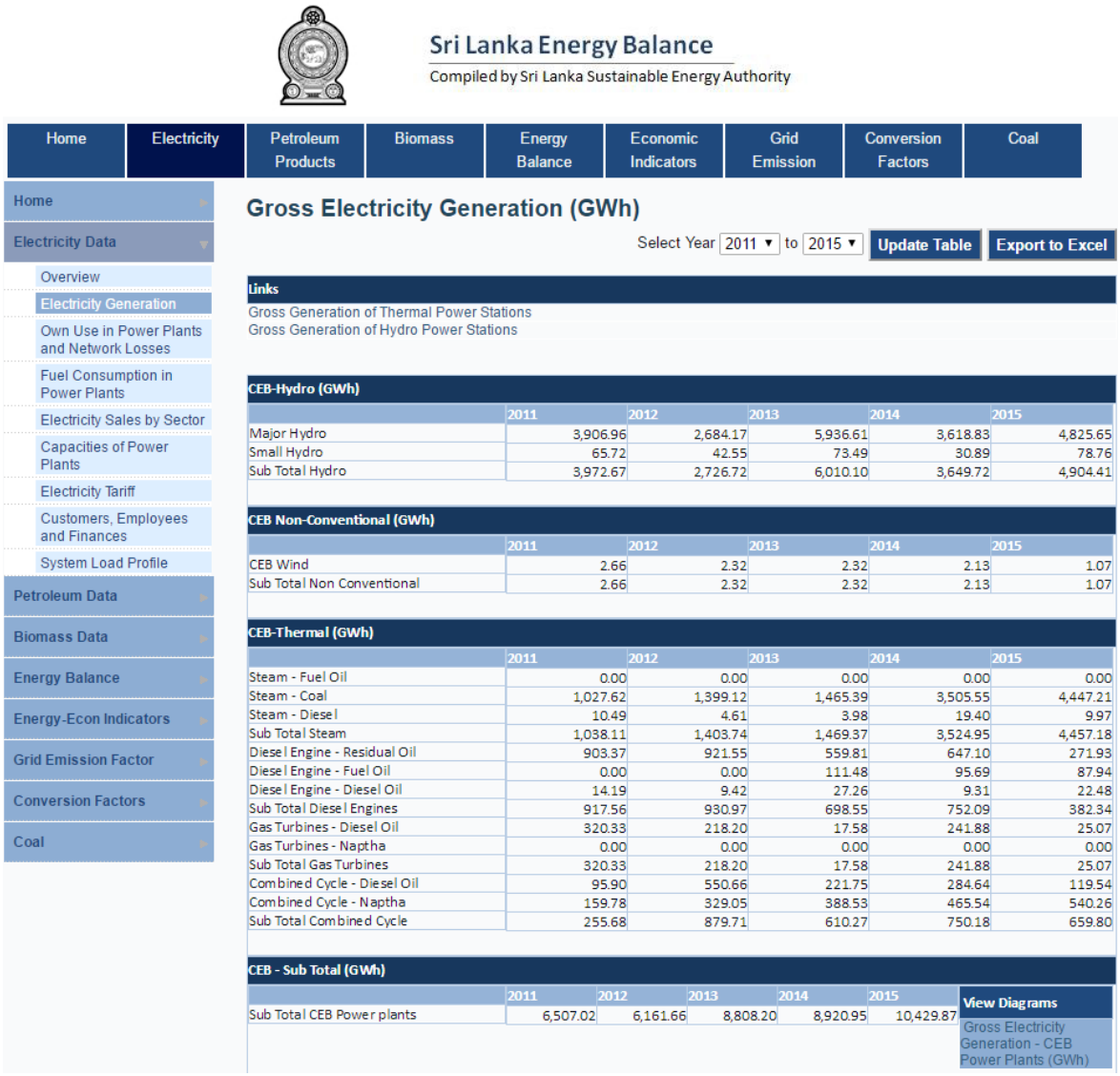
Apart from the primary data, secondary data is also necessary to model processes and attribute the necessary environmental burdens to them. Usually Electricity and Transport are secondary datasets that can be obtained from public databases and other sources, as this data is easily found on the internet, for instance on governmental websites or on scientific literature.

The figures below are screenshots from the Sri Lanka Sustainable Energy Authority[2], and the NATION newspaper, both are public data sources for the Sri Lanka energy mix dataset you want to create.

However, for the dataset you also need to account for the emissions from the energy production which are not provided and might require more research.

Take a look on the data provided and create an Energy Mix for Sri Lanka, it should be representative for the time period of 2015-2016.

Energy mix:



Sri Lanka Energy Balance

Compiled by Sri Lanka Sustainable Energy Authority

Home | Electricity | Petroleum Products | Biomass | Energy Balance | Economic Indicators | Grid Emission | Conversion Factors | Coal

- Home
- Electricity Data
  - Overview
  - Electricity Generation
  - Own Use in Power Plants and Network Losses
  - Fuel Consumption in Power Plants
  - Electricity Sales by Sector
  - Capacities of Power Plants
  - Electricity Tariff
  - Customers, Employees and Finances
  - System Load Profile
- Petroleum Data
- Biomass Data
- Energy Balance
- Energy-Econ Indicators
- Grid Emission Factor
- Conversion Factors
- Coal

Gross Electricity Generation (GWh)

Select Year 2011 to 2015 | Update Table | Export to Excel

Links

Gross Generation of Thermal Power Stations

Gross Generation of Hydro Power Stations

CEB-Hydro (GWh)

	2011	2012	2013	2014	2015
Major Hydro	3,906.96	2,684.17	5,936.61	3,618.83	4,825.65
Small Hydro	65.72	42.55	73.49	30.89	78.76
Sub Total Hydro	3,972.67	2,726.72	6,010.10	3,649.72	4,904.41

CEB Non-Conventional (GWh)

	2011	2012	2013	2014	2015
CEB Wind	2.66	2.32	2.32	2.13	1.07
Sub Total Non Conventional	2.66	2.32	2.32	2.13	1.07

CEB-Thermal (GWh)

	2011	2012	2013	2014	2015
Steam - Fuel Oil	0.00	0.00	0.00	0.00	0.00
Steam - Coal	1,027.62	1,399.12	1,465.39	3,505.55	4,447.21
Steam - Diesel	10.49	4.61	3.98	19.40	9.97
Sub Total Steam	1,038.11	1,403.74	1,469.37	3,524.95	4,457.18
Diesel Engine - Residual Oil	903.37	921.55	559.81	647.10	271.93
Diesel Engine - Fuel Oil	0.00	0.00	111.48	95.69	87.94
Diesel Engine - Diesel Oil	14.19	9.42	27.26	9.31	22.48
Sub Total Diesel Engines	917.56	930.97	698.55	752.09	382.34
Gas Turbines - Diesel Oil	320.33	218.20	17.58	241.88	25.07
Gas Turbines - Naptha	0.00	0.00	0.00	0.00	0.00
Sub Total Gas Turbines	320.33	218.20	17.58	241.88	25.07
Combined Cycle - Diesel Oil	95.90	550.66	221.75	284.64	119.54
Combined Cycle - Naptha	159.78	329.05	388.53	465.54	540.26
Sub Total Combined Cycle	255.68	879.71	610.27	750.18	659.80

CEB - Sub Total (GWh)

	2011	2012	2013	2014	2015
Sub Total CEB Power plants	6,507.02	6,161.66	8,808.20	8,920.95	10,429.87

View Diagrams

Gross Electricity Generation - CEB Power Plants (GWh)

[2] http://www.info.energy.gov.lk/

Self Generation By Customers	0.00	0.00	0.00	0.00	0.00
Off Grid Systems-Industrial	0.00	0.00	0.00	0.00	0.00
Off Grid Systems-Non Industrial	0.00	0.00	0.00	0.00	0.00
Sub Total Off Grid - Conventional	0.00	0.00	0.00	0.00	0.00

Gross Electricity Generation Off Grid Systems (GWh)

Off Grid Systems - Non Conventional (GWh)

	2011	2012	2013	2014	2015
Small Hydro, Industrial	7.07	7.07	7.07	7.07	7.07
Small Hydro, Household	3.58	3.69	3.69	3.69	3.69
Solar Photovoltaic, Household	7.60	8.01	8.01	8.01	8.01
Wind Energy, Household	0.01	0.01	0.01	0.01	0.01
Sub Total Off-Grid, Non-Conventional	18.26	18.77	18.77	18.77	18.77

Total Generation Sri Lanka

	2011	2012	2013	2014	2015
Total	11,599.87	11,897.62	12,024.32	12,848.88	13,225.55

Summary - 1 - CEB Grid

	2011	2012	2013	2014	2015
CEB Hydro	3,972.67	2,726.72	6,010.10	3,649.72	4,904.41
CEB Wind	2.66	2.32	2.32	2.13	1.07
CEB Thermal	2,531.68	3,432.62	2,795.78	5,269.10	5,524.39
IPP Thermal (Gross)	4,352.33	4,983.85	2,023.94	2,675.20	1,272.03
SPP Hydro	600.57	564.69	908.39	902.17	1,064.72
SPP Thermal	0.00	0.00	0.00	0.00	0.00
SPP Solar	1.11	2.00	1.68	1.47	1.87
SPP Biomass	31.63	22.17	26.39	41.39	57.31
SPP Wind	88.95	144.48	232.26	270.32	342.13
NmP Solar	0.00	0.00	4.69	18.60	38.84
Hired Thermal	0.00	0.00	0.00	0.00	0.00
Gross Generation to CEB Grid	11,581.61	11,878.85	12,005.55	12,830.11	13,206.77

Summary - 2 - CEB Grid

	2011	2012	2013	2014	2015
CEB Hydro	3,972.67	2,726.72	6,010.10	3,649.72	4,904.41
Thermal, CEB, IPP and Hired	6,884.01	8,416.47	4,819.72	7,944.30	6,796.42
CEB Wind	2.66	2.32	2.32	2.13	1.07
New Renewable Energy	722.26	733.34	1,168.72	1,215.36	1,466.04
Net-metered Projects	0.00	0.00	4.69	18.60	38.84
Gross Generation to CEB Grid	11,581.61	11,878.85	12,005.55	12,830.11	13,206.77
Generation Growth Rate: CEB Grid (%)	7.1	2.1	0.3	6.6	2.3

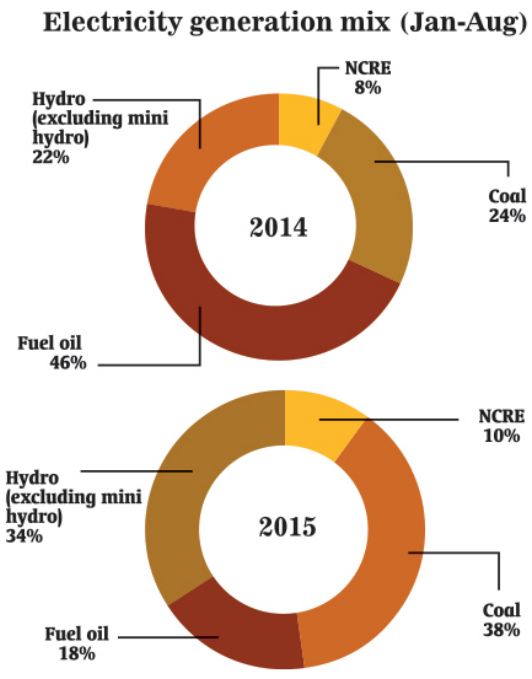
Summary- Sri Lanka

	2011	2012	2013	2014	2015
CEB Hydro	3,972.67	2,726.72	6,010.10	3,649.72	4,904.41
Thermal, CEB, IPP and Hired	6,884.01	8,416.47	4,819.72	7,944.30	6,796.42
CEB Wind	2.66	2.32	2.32	2.13	1.07
New Renewable Energy	722.26	733.34	1,168.72	1,215.36	1,466.04
Net-metered Projects	0.00	0.00	4.69	18.60	38.84
Self-Generation by Customers	0.00	0.00	0.00	0.00	0.00
Off-Grid, Conventional	0.00	0.00	0.00	0.00	0.00
Off-Grid, Non-Conventional	18.26	18.77	18.77	18.77	18.77
Gross Generation Sri Lanka	11,599.87	11,897.62	12,024.32	12,848.88	13,225.54
Generation Growth Rate Sri Lanka (%)	7.4	2.6	1.1	6.9	2.9

Sri Lanka's energy mix

By Azhar Razak - Mar 26, 2016 2460 0

SHARE Facebook Twitter



Solar power Grid-connected solar power has only recently been introduced. The only operational commercial-scale solar-powered facility is the Buruthakanda Solar Park of 1.2 MW, operated by the Sri Lanka Sustainable Energy Authority (SLSEA).

Sri Lanka's electricity demand per year is estimated to be at 2300 megawatts with the island's present generating capacity being around 3900 megawatts.

Hydroelectricity Currently, ten large hydroelectric power stations are in operation, with the single largest hydroelectric source being the Victoria Dam. Although a large portion of the country's hydroelectric resource is tapped, the government continues to issue small hydro development permits to the private sector, for projects up to a total installed capacity of 10 MW per project.

Thermal power Thermal power stations in Sri Lanka run on diesel, other fuel oils, naptha or coal. The Norochcholai Coal Power Station, the only coal-fired power station in the country, was commissioned in late 2011. The plant has since added 900 megawatts of electrical capacity to the grid.

Wind power The first commercial grid-connected wind farm is the 3 MW Hambantota Wind Farm, northwest of Hambantota. The government policy limit of 10 MW per wind project significantly decreases economies-of-scale, further straining such developments.

Geothermal power Geothermal power is under research, although no power stations of this type are operational.

Nuclear power The CEB has included a 600MW nuclear power plant as an option in its plans for 2031.

Source: http://nation.lk/online/2016/03/26/sri-lankas-energy-mix.html

Emissions:

After exhaustively researching you came to the conclusion that the detailed inventory of Sri Lanka energy production isn't available, therefore you decided to adapt datasets from other countries.

On ecoinvent you found the following datasets on which you could base Sri Lanka energy mix emissions, choose the dataset you find more appropriate to adapt to Sri Lanka.

electricity production, hydro, reservoir, alpine region | electricity, high voltage | cut-off, U

Inputs			
Flow	Category	Amount	Unit
Energy, potential (in hydropower reservoir), converted	Resource/in water	3.79	MJ
hydropower plant, reservoir, alpine region	4220:Construction of utility projects/4220a: Construction of utility pro...	4.04E-13	Item(s)
lubricating oil	192:Manufacture of refined petroleum products/1920:Manufacture of refine...	7.56E-06	kg
Occupation, water bodies, artificial	Resource/land	0.00345	m2*a
Transformation, from unknown	Resource/land	2.30E-05	m2
Transformation, to industrial area, built up	Resource/land	2.30E-07	m2
Transformation, to water bodies, artificial	Resource/land	2.28E-05	m2
Volume occupied, reservoir	Resource/in water	0.15	m3*a
waste mineral oil	382:Waste treatment and disposal/3822:Treatment and disposal of hazardou...	-7.56E-06	kg
Water, turbine use, unspecified natural origin	Resource/in water	0.81	m3
Outputs			
Flow	Category	Amount	Unit
Dinitrogen monoxide	Emission to air/low population density	7.70E-08	kg
electricity, high voltage	351:Electric power generation, transmission and distribution/3510:Electr...	1	kWh
Methane, biogenic	Emission to air/low population density	1.40E-05	kg
Water	Emission to air/unspecified	29.22168	kg
Water	Emission to water/unspecified	0.78078	m3

electricity production, wind, 1-3MW turbine, onshore | electricity, high voltage | cut-off, U

Inputs			
Flow	Category	Amount	Unit
Energy, kinetic (in wind), converted	Resource/in air	3.87	MJ
lubricating oil	192:Manufacture of refined petroleum products/1920:Manufacture of refine...	5.83E-05	kg
transport, freight, lorry 7.5-16 metric ton, EURO3	492:Other land transport/4923:Freight transport by road	4.31E-12	t*km

waste mineral oil	382:Waste treatment and disposal/3822:Treatment and disposal of hazardou...	-5.83E-05	kg
wind turbine network connection, 2MW, onshore	4220:Construction of utility projects/4220a: Construction of utility pro...	1.85E-08	Item(s)
wind turbine, 2MW, onshore	4220:Construction of utility projects/4220a: Construction of utility pro...	1.85E-08	Item(s)
Outputs			
Flow	Category	Amount	Unit
electricity, high voltage	351:Electric power generation, transmission and distribution/3510:Electr...	1	kWh

electricity production, hard coal | electricity, high voltage | cut-off, U

Inputs			
Flow	Category	Amount	Unit
hard coal	051:Mining of hard coal/0510:Mining of hard coal	0.63581	kg
hard coal ash	239:Manufacture of non-metallic mineral products n.e.c./2394:Manufacture...	-0.10858	kg
hard coal power plant	4220:Construction of utility projects/4220a: Construction of utility pro...	1.33E-11	Item(s)
SOx retained, in hard coal flue gas desulfurisation	351:Electric power generation, transmission and distribution/3510:Electr...	0.00073	kg
water, completely softened, from decarbonised water, at user	360:Water collection, treatment and supply/3600:Water collection, treatm...	0.0871	kg
Water, cooling, unspecified natural origin	Resource/in water	0.06639	m3
water, decarbonised, at user	201:Manufacture of basic chemicals, fertilizers and nitrogen compounds, ...	2.90323	kg
Outputs			
Flow	Category	Amount	Unit
Acenaphthene	Emission to air/low population density	6.10E-11	kg
Acrolein	Emission to air/low population density	3.47E-08	kg
…			

heat and power co-generation, biogas, gas engine | electricity, high voltage | cut-off, U

Inputs			
Flow	Category	Amount	Unit

biogas	382:Waste treatment and disposal/3821:Treatment and disposal of non-haza...	0.34433	m3
heat and power co-generation unit, 160kW electrical, common components for heat+electricity	4220:Construction of utility projects/4220a: Construction of utility pro...	3.91E-08	Item(s)
heat and power co-generation unit, 160kW electrical, components for electricity only	271:Manufacture of electric motors, generators, transformers and electri...	3.91E-08	Item(s)
heat and power co-generation unit, 160kW electrical, components for heat only	281:Manufacture of general-purpose machinery/2811:Manufacture of engines...	3.91E-08	Item(s)
lubricating oil	192:Manufacture of refined petroleum products/1920:Manufacture of refine...	0.00023	kg
waste mineral oil	382:Waste treatment and disposal/3822:Treatment and disposal of hazardou...	-0.00023	kg
Outputs			
Flow	Category	Amount	Unit
Carbon dioxide, biogenic	Emission to air/low population density	0.65358	kg
Carbon monoxide, biogenic	Emission to air/low population density	0.00038	kg
Dinitrogen monoxide	Emission to air/low population density	1.96E-05	kg
electricity, high voltage	351:Electric power generation, transmission and distribution/3510:Electr...	1	kWh
Methane, biogenic	Emission to air/low population density	0.00018	kg
Nitrogen oxides	Emission to air/low population density	0.00012	kg
NMVOC, non-methane volatile organic compounds, unspecified origin	Emission to air/low population density	1.57E-05	kg
Platinum	Emission to air/low population density	5.48E-11	kg
Sulfur dioxide	Emission to air/low population density	0.0002	kg

Assume this composition of the electricity production mix in Sri Lanka and calculate an overall data set based on the data sets listed above.

Electricity, high voltage, production mix - Sri Lanka

Inputs

Flow	Amount	Unit	%
electricity production, hydro	0.3714	kWh	37.14%
electricity production, wind	0.0008	kWh	0.01%
electricity production, nuclear	0.0000	kWh	0.00%
electricity production, fossil fuels and coal	0.5146	kWh	51.46%
electricity production, others (biomass, solar, etc.)	0.1139	kWh	11.39%

Outputs

Flow	category	amount	unit

electricity, high voltage	351:Electric power generation, transmission and distribution	1	kWh

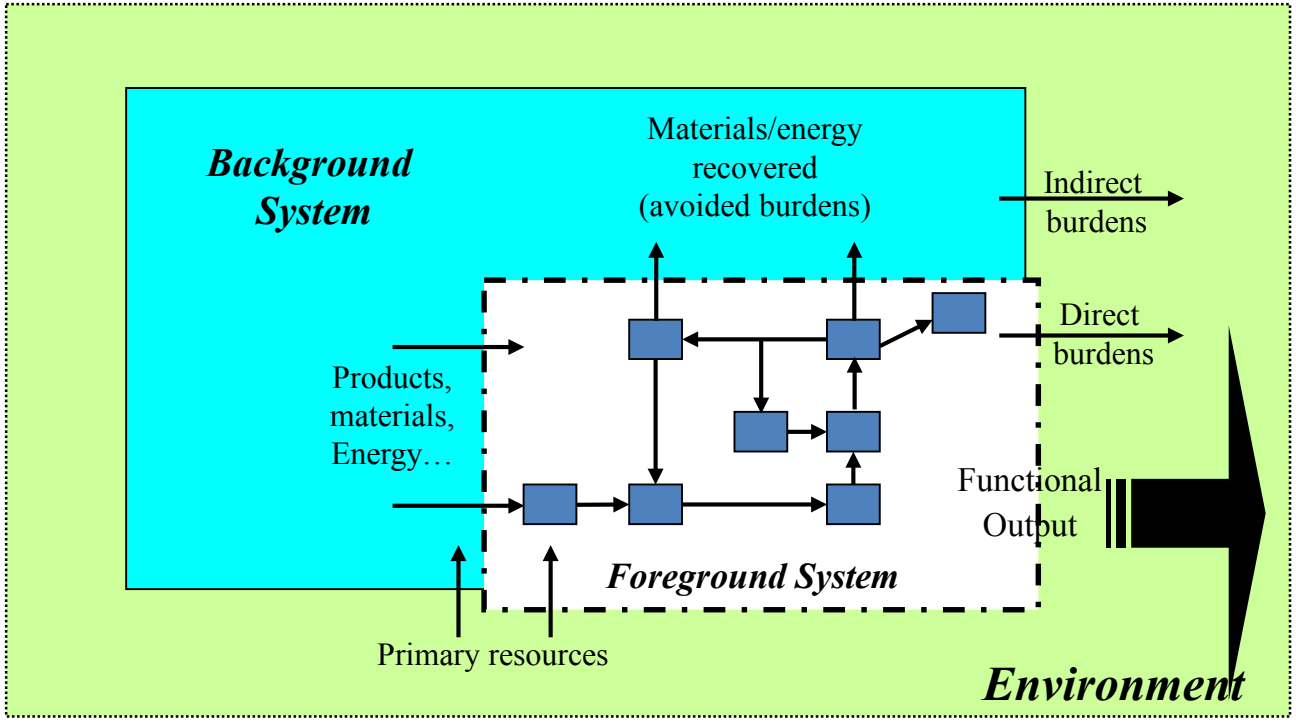
Part 4 - Process Modelling for LCI Datasets

Content from Amir Safaei,
ecoinvent

August 2017 Version

Managed by SETAC

Useful concepts
Foreground and Background system in LCA

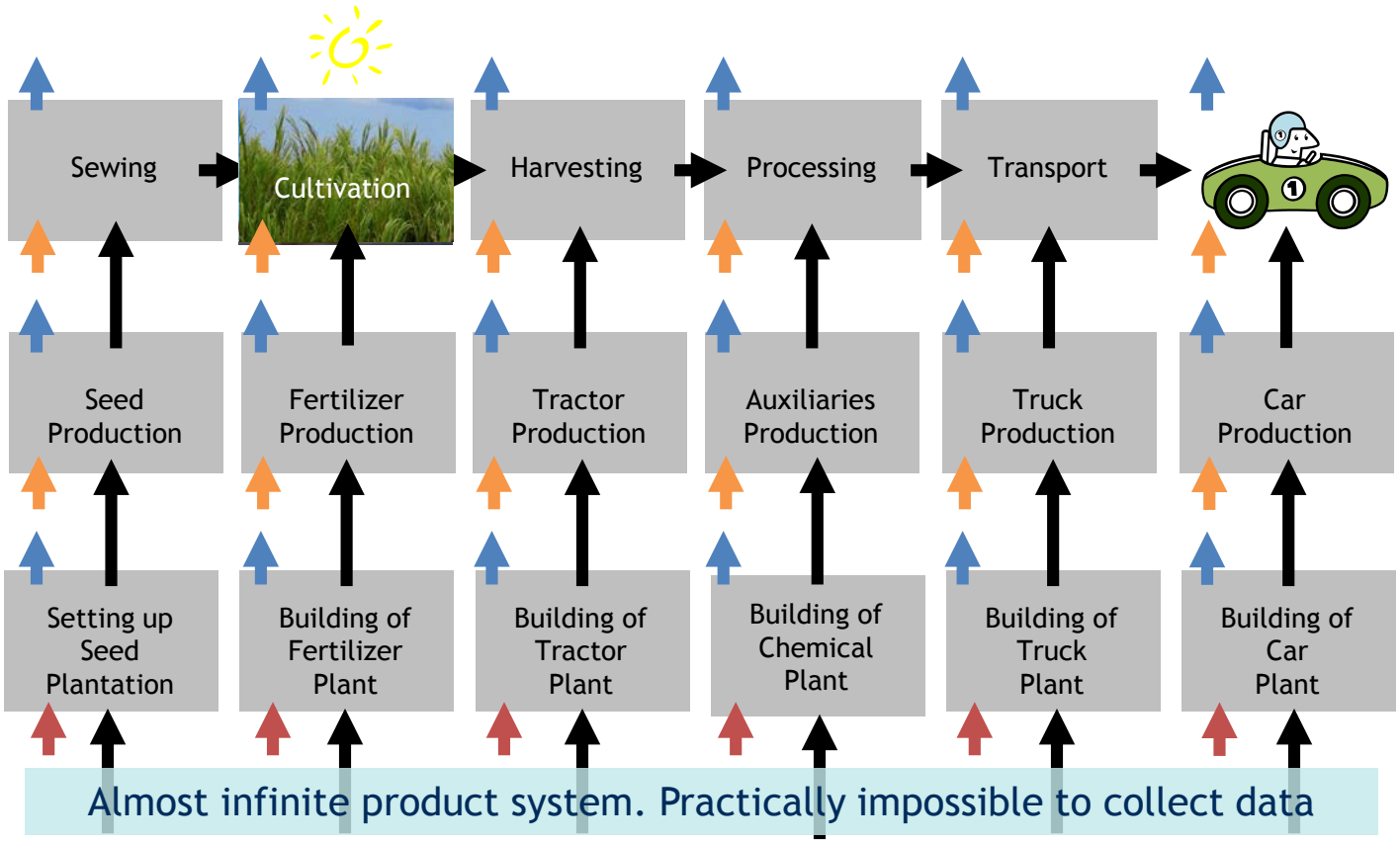


Foreground system: processes that actions can be directly taken wrt the results of the LCA, direct measurements can often be taken

Background system: processes that actions can not be directly taken wrt the results of the LCA, often, external secondary data used

The need for background data
Example: Biofuels for Transportation

ecoinvent

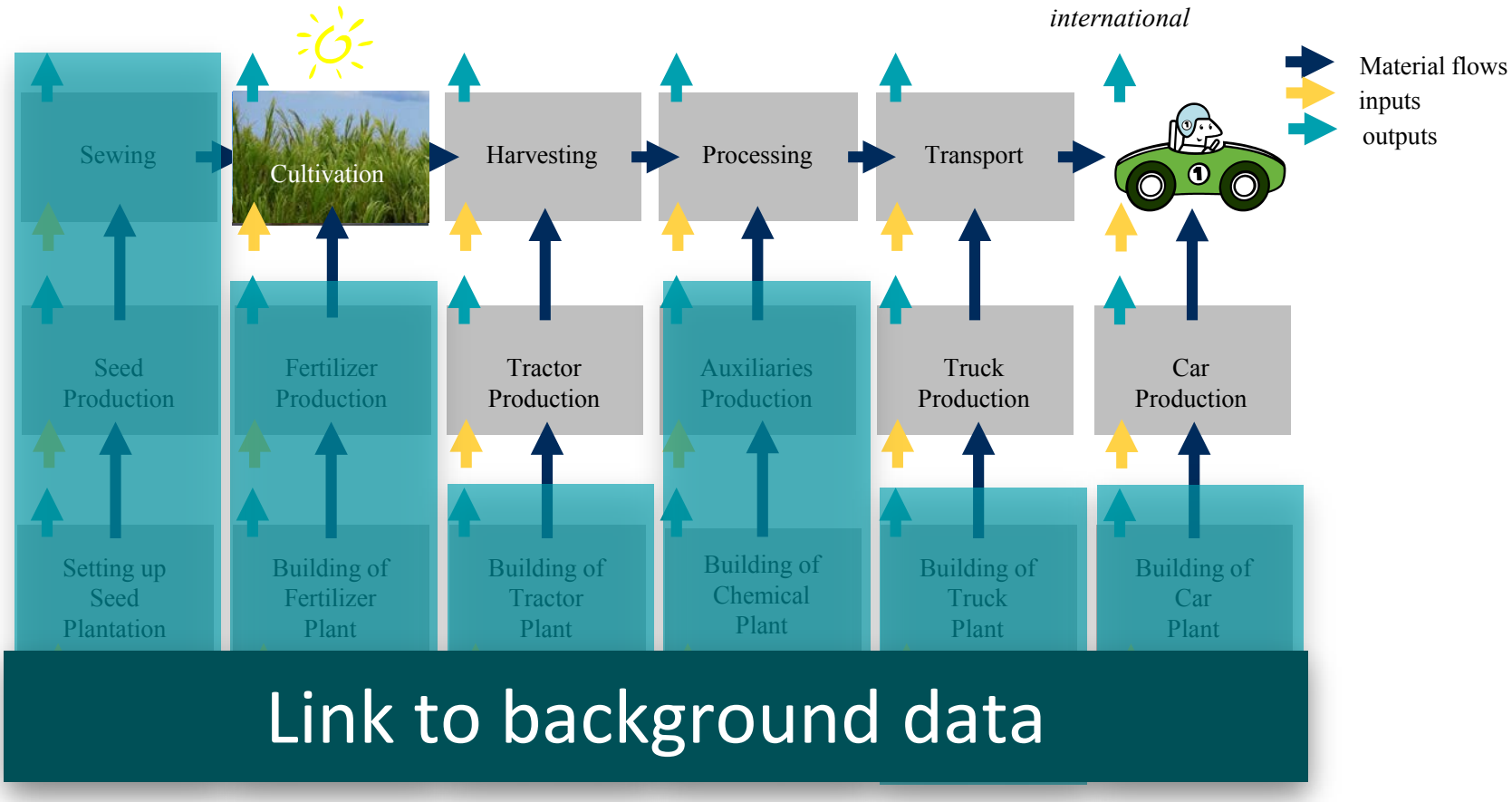


Acknowledgments:
Quantis international

The need for background data
Example: Fossil Energy Demand of Biofuels

ecoinvent

Acknowledgments:
Quantis
international



How to develop a dataset

ecoinvent

Step 0: Draw a basic technical flowchart of the unit process under development

Step 1: Prepare an inventory list of inputs and outputs

- ensure a complete list by referring to similar process/datasets and literatures

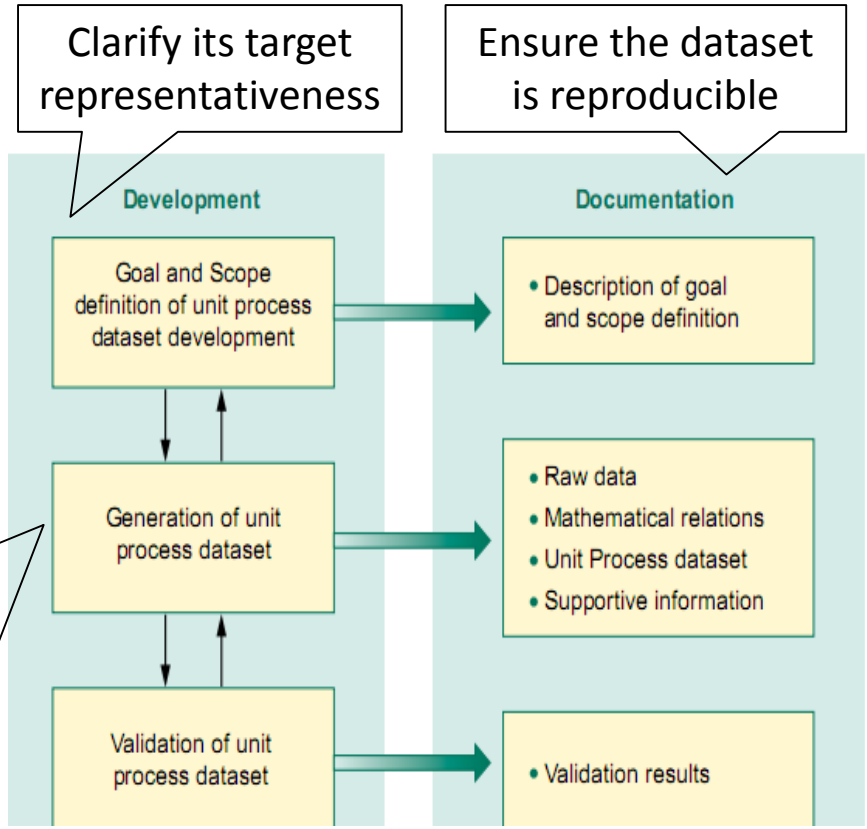
Step 2: Define the mathematical relationships

Step 3: Collect the raw data needed

- Step 2 & 3 are dependent to ensure data availability and data quality

Step 4: Perform relationship calculations to obtain a dataset

Step 5: Provide other supportive information, e.g. allocation, suggestions to users



Source: Sonnemann, G., & Vigon, B. (2011)

Dataset preparation steps

ecoinvent

- Process Modelling for datasets
  - Know the type of the DB (aggregated or unit process datasets)
  - Know the structure of the DS in the database
  - Know the guidelines of the DB
    - DB specific grammars (e.g. of type of datasets)
    - Naming conventions
  - Aggregation
  - Allocation
  - Documentation
  - Considerations for sectors with high modelling needs
  - Validation
  - Review process, how it works, plus time considerations
- Exercise

Know the structure of the database

ecoinvent

- Creating, modelling, and processing the datasets shall be done in accordance with the structure of the database you wish to work with
  - National database (SICV Brazil or a one-day-to-be Sri Lankan database)
  - ecoinvent
  - Other databases?

- Familiarize yourself with the structure of the database and its requirements (examples from DQG or simplified guidelines...)

Know the format of the database: ecoinvent

- Ecoinvent is a unit process database
- Data submitted to ecoinvent shall be in the form of unlinked unallocated multioutput Unit Processes

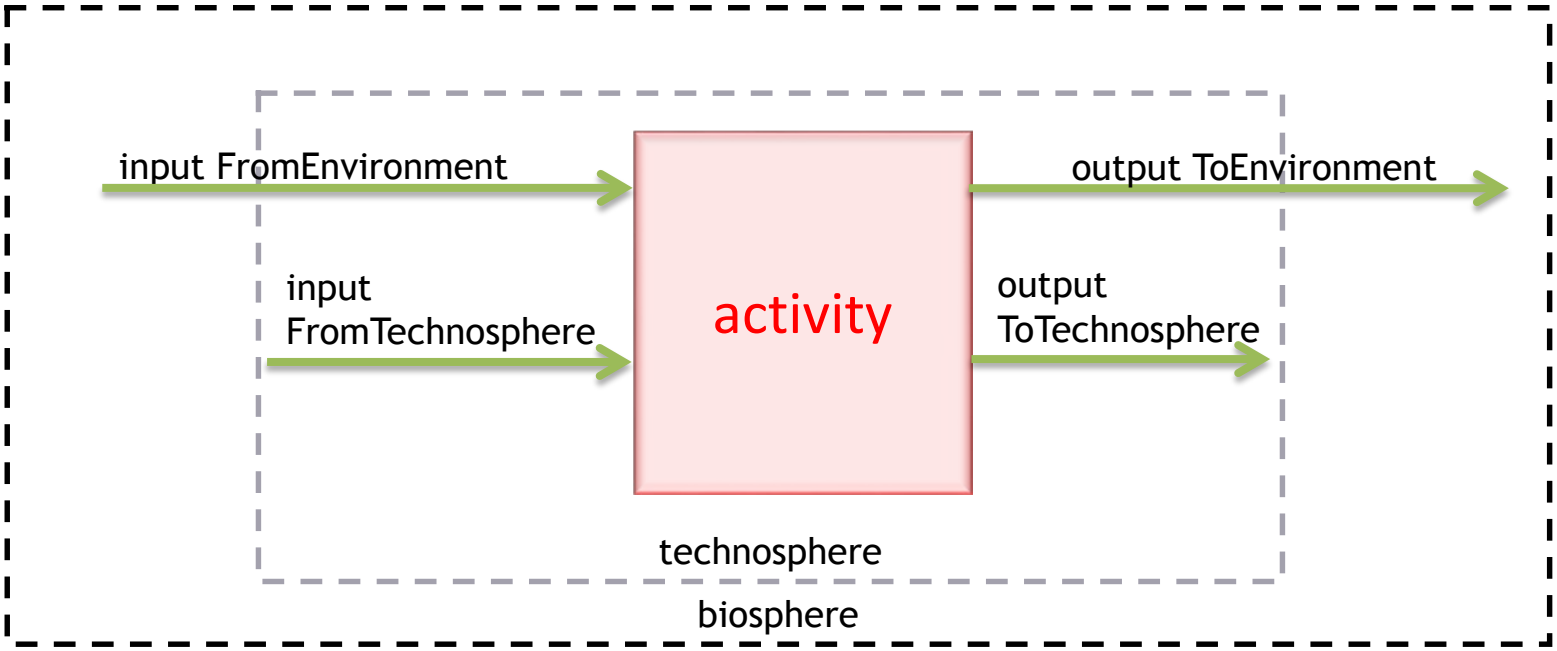
  unlinked: the inputs and outputs are *not* yet linked to the corresponding supply chains.

  unallocated multioutput: the activity models the physical reality of the process, *all co-products are listed.*

  Unit Process (UPR): the UPR represents all inputs and outputs to *each individual process*, from technosphere or environment.

This ensures transparency

UPR datasets in ecoinvent



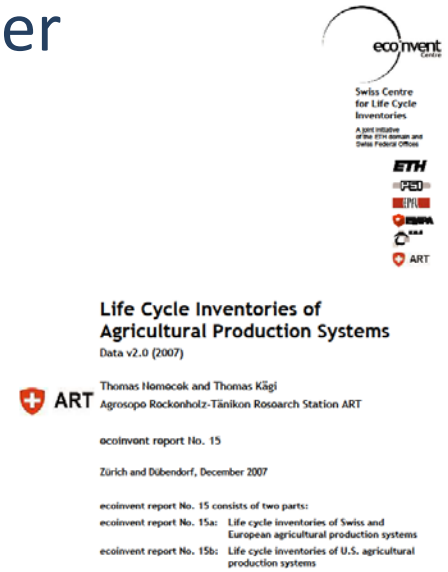
Explore similar datasets within the database

ecoinvent

- Check the DB for the same product (for other geographies)
- It exists? Cool! You are some steps ahead!
- Open the dataset
  - explore it
  - study the inputs/output
  - Glance over the references and guidelines
  - Any guidelines mentioned in the documentation?
- It does not exist? Check for similar products

  There is no rice, check wheat!

  There is no pear, check apple!


ecoinvent Centre
Swiss Centre for Life Cycle Inventories
A joint initiative of the ETH domain and Swiss Federal Offices
ETH
PSI
EMPA
ART

Life Cycle Inventories of Agricultural Production Systems
Data v2.0 (2007)

ART
Thomas Nemecek and Thomas Kägi
Agroscope Reckenholz-Tänikon Research Station ART

ecoinvent report No. 15

Zürich and Dübendorf, December 2007

ecoinvent report No. 15 consists of two parts:
ecoinvent report No. 15a: Life cycle inventories of Swiss and European agricultural production systems
ecoinvent report No. 15b: Life cycle inventories of U.S. agricultural production systems

Get to know the structure of the sector in the database

ecoinvent

- Find, explore and analyze the existing related datasets
  - How is the sector structured?
    - Study the sectorial guidelines
  - How are similar datasets structured?
    - What are the inputs and outputs to the datasets (flows)
  - Level of aggregation?
  - What are the related datasets? Do you have or do you require additional information for upstream/downstream processes of the supply chain?
  - How similar is your process to the existing one in the database?

Methods of assessment of direct field emissions for LCIs of agricultural production systems

Data v3.0 (2012)

ART Thomas Nemecek and Julian Schnetzer
Agroscope Reckenholz-Tänikon Research Station ART

Zurich, August 2011

This documentation file is mainly based on chapter 4 of ecoinvent report No. 15a (Nemecek et al. 2007) where the methods of assessment of direct field emissions for life cycle inventories (LCI) of agricultural crop production (ecoinvent data version 2) are described. Further elementary flows related to natural resources are described here, as well. This text represents an updated documentation of the methods and data sources used within the frame of updating agricultural LCIs for ecoinvent data version 3.

The issue of by-products and wastes

ecoinvent

- Always make sure you are looking at the whole process and considering all the inputs and outputs to the process
  - Does your process have multiple inputs/outputs?

  Most of the cases the answer is yes, and you might not have noticed.
  - Any by-product that you have to consider (and are not doing so)?

  Most of the fruits in the DB have "waste wood" as the by-product. Have you checked similar datasets in the DB?
  - Any waste? What about its treatment
  - Do all the products, by-products or waste types exist in the database? (or they should be modelled as from scratch?

Now that you have explored similar datasets....

ecoinvent

- How similar is your process to the existing one in the database?
  - Does it have a similar structure (input and output flows, technology, level of aggregation)?

  If no, how different it is?

  Is it possible to do some modelling to adapt the data to the structure of the database, or

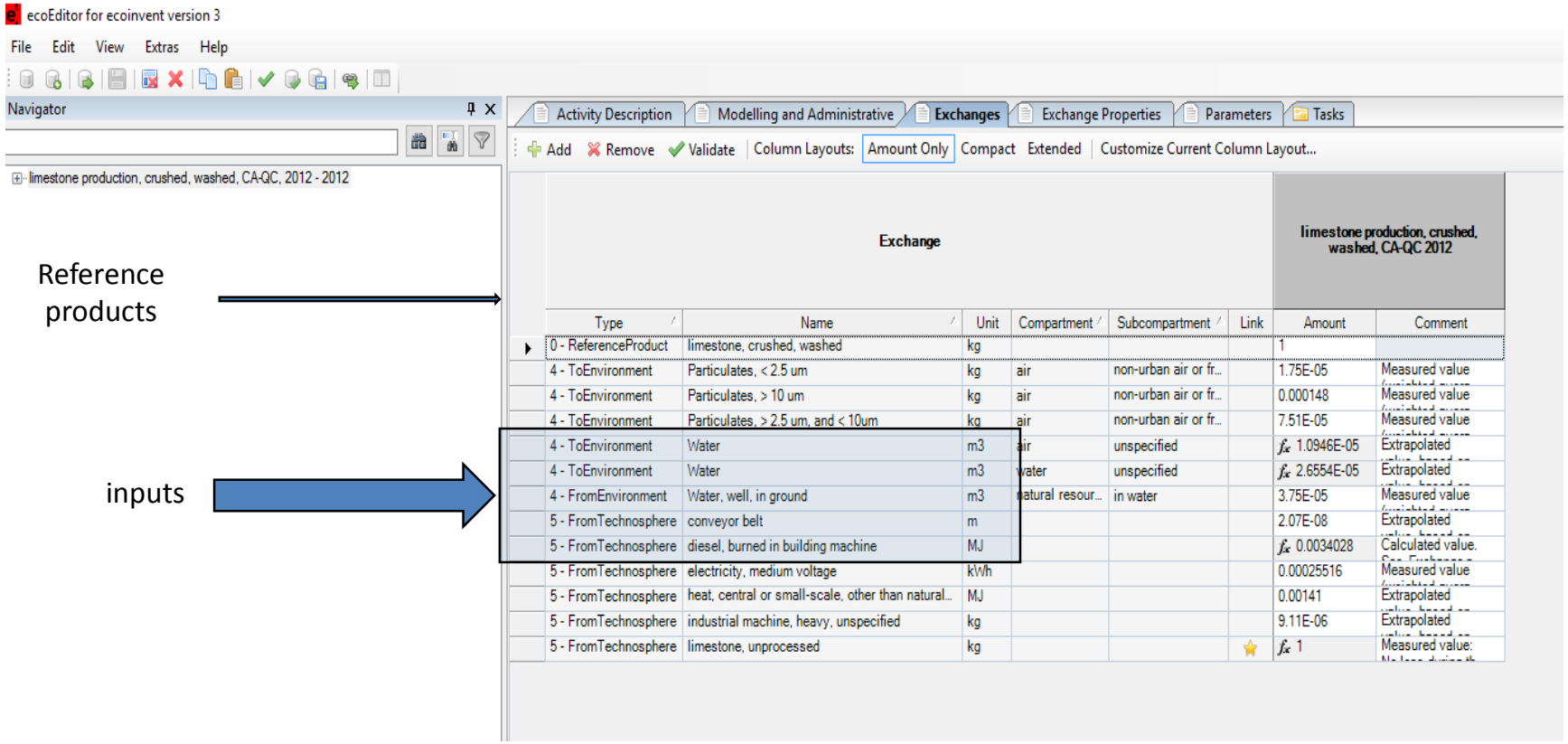
  is it so different (and prominent) that one should introduce a new structure for the DB type of activity in the dataset (*this is the less common case*)

Sensitivity analysis
Do not bark up the wrong tree...

ecoinvent

- When you have identified the similar processes in the DB, do a screening LCA or "process contribution" analysis
  - Choose a number of relevant LCIA categories
  - Do a "contribution analysis" to identify the most relevant flows and exchanges
  - You might be able get this info from the DB management

Example of process contribution



Aggregation

- Aggregation is about structuring of disparate data points that occur (naturally) in LCI data collection
- It refers to aggregation of data within and between system components
- It is performed to:
  - reduce variation and uncertainty
  - Obtain representative value according to the goal and scope of the dataset
  - Keep confidentiality
- Rice production in Sri Lanka
  - Collect and aggregate values from different national producers
- Rice production in Andhra Pradesh, IN
  - Collect data and aggregate the data point to obtain values representing state producers

Aggregation

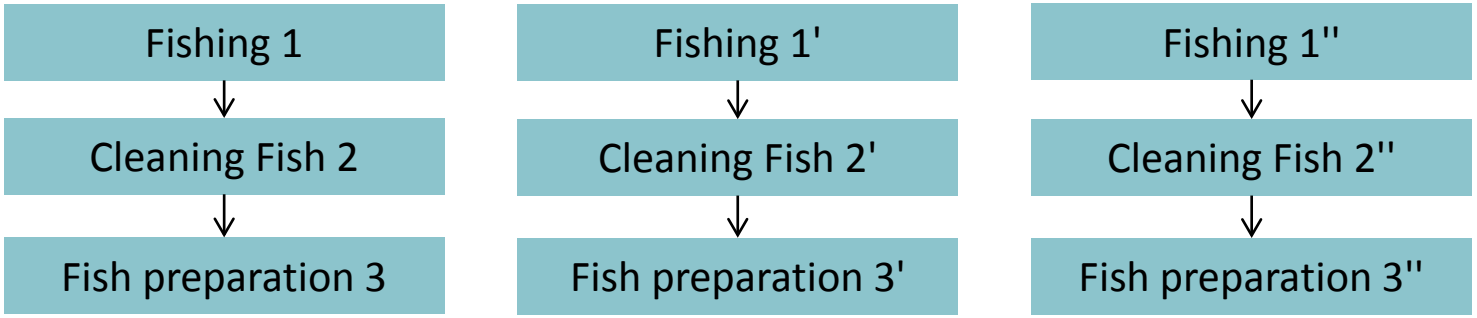
ecoinvent

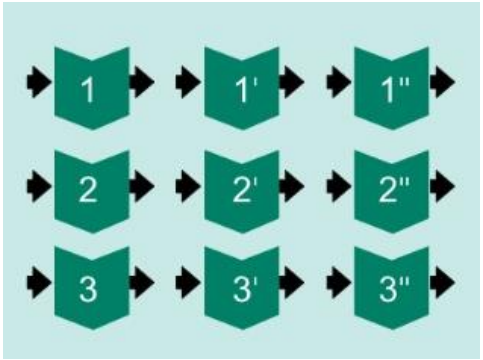
Typically LCI data vary based on ...

- The time-frame (over which the data was collected)
- The geographical representativeness (e.g. facility level vs regional level vs. national level)
- Product or process variation (multiple paths/processes to produce the same product, each with a distinct technology)
- Data value size (small vs large-scale measurements)
- Figures within the data (four digits decimals or one!)
- Data collection process (actual measurement vs expert judgement)
- Data manipulations (e.g. normalized per unit)

Example of unit process and aggregation

- Three supply chains. Each comprises three unit processes/producers.



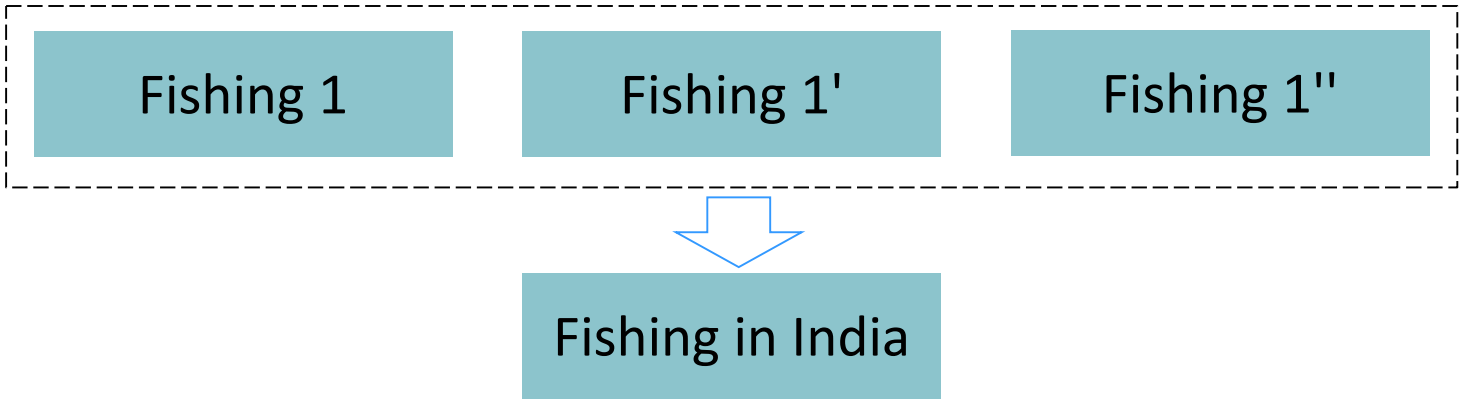


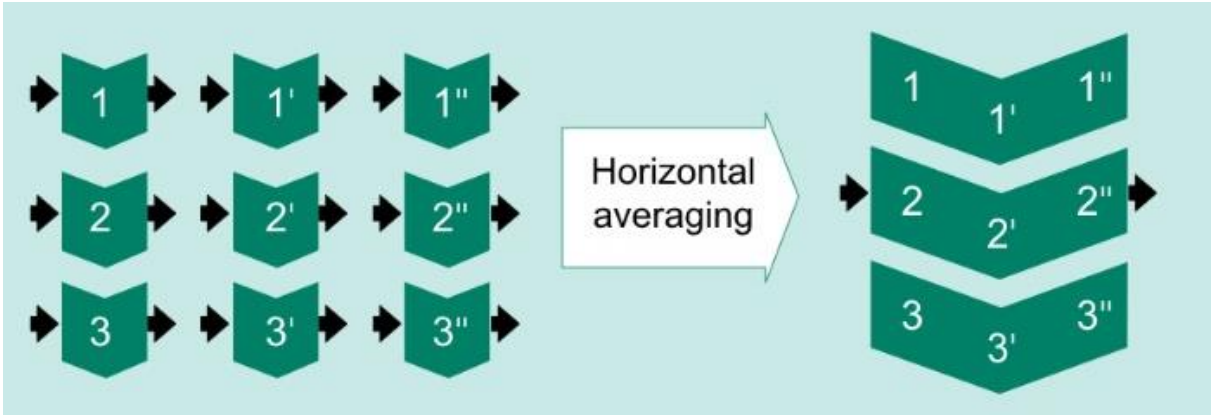
Source: Sonnemann, G., & Vigon, B. (2011).

Horizontal averaging

ecoinvent

- To represent, e.g. a regional average

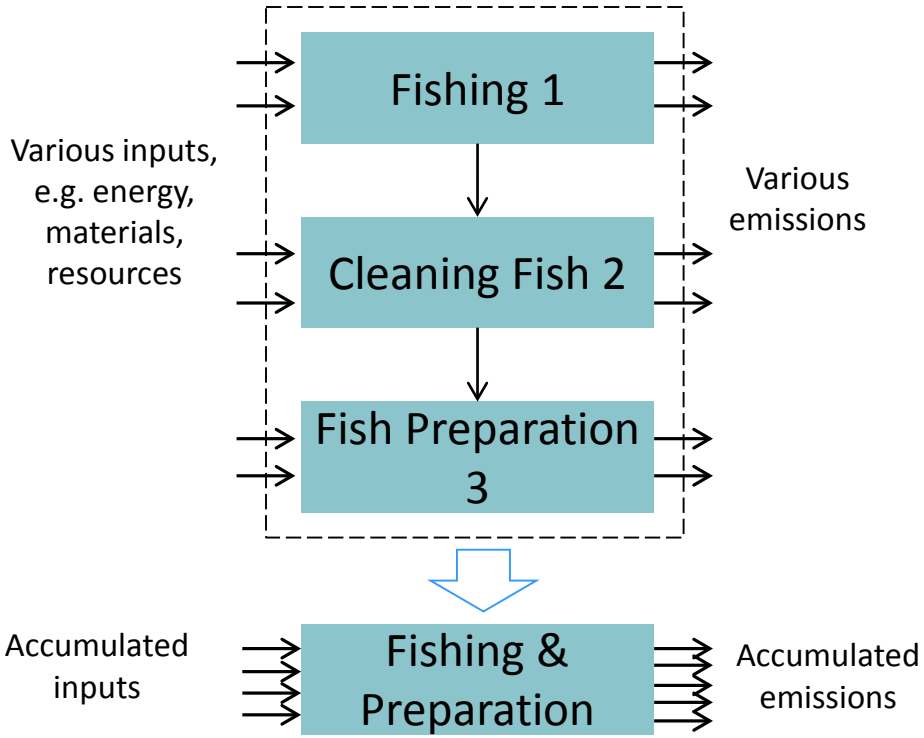


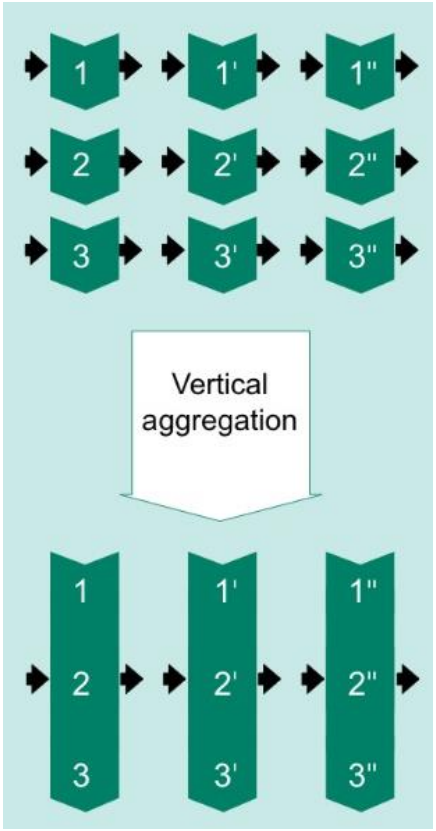


Source: Sonnemann, G., & Vigon, B. (2011)

Vertical aggregation

- To cumulate and represent multiple LC stages





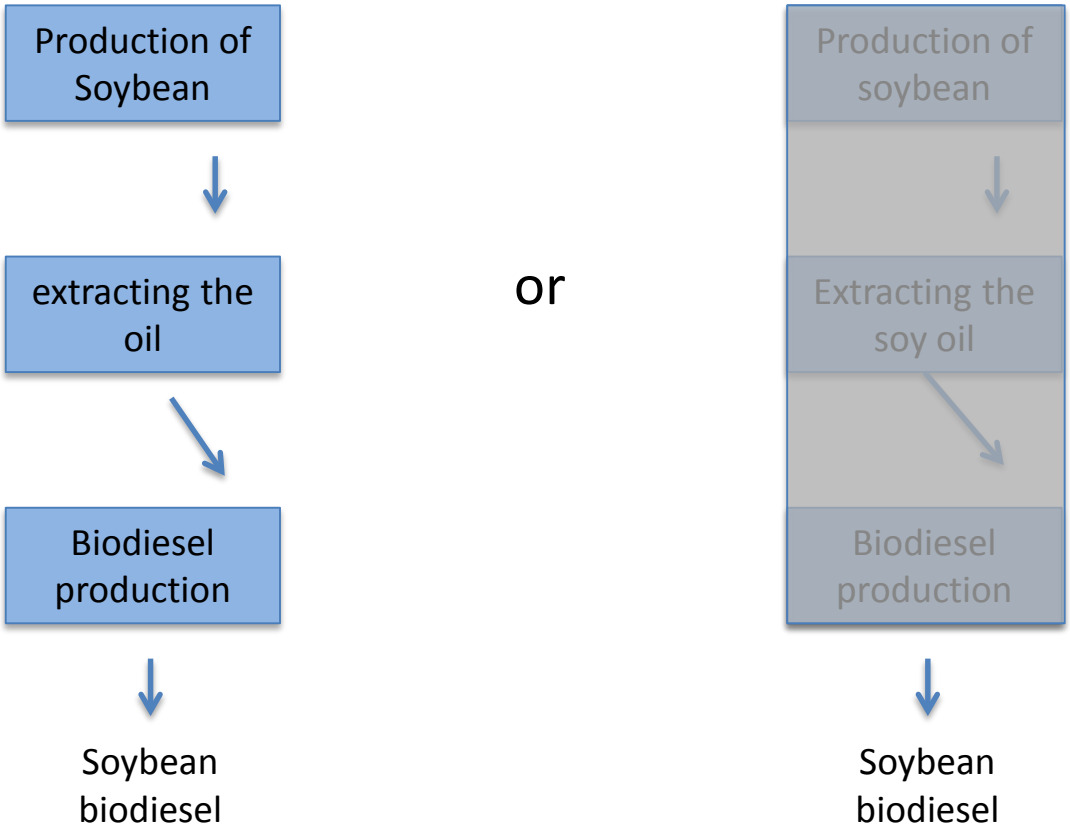
Source: Sonnemann, G., & Vigon, B. (2011)

Vertical aggregation

ecoinvent

- Vertical aggregation shall be performed considering
  - the structure of the sector (similar datasets) in the DB
  - level of confidentiality (more transparency, better; but also consider the effort and time required)
- It is desirable to have unit process for each activity (why?)
- You can aggregate the data, but dis-aggregation (for the user?) is another task.
- Documentation is key for vertically aggregated datasets
- This bounces back to the data collection; data collection strategy, if possible, **better be adjusted** to the unit process level and the aggregation level of the existing datasets

Vertical aggregation example



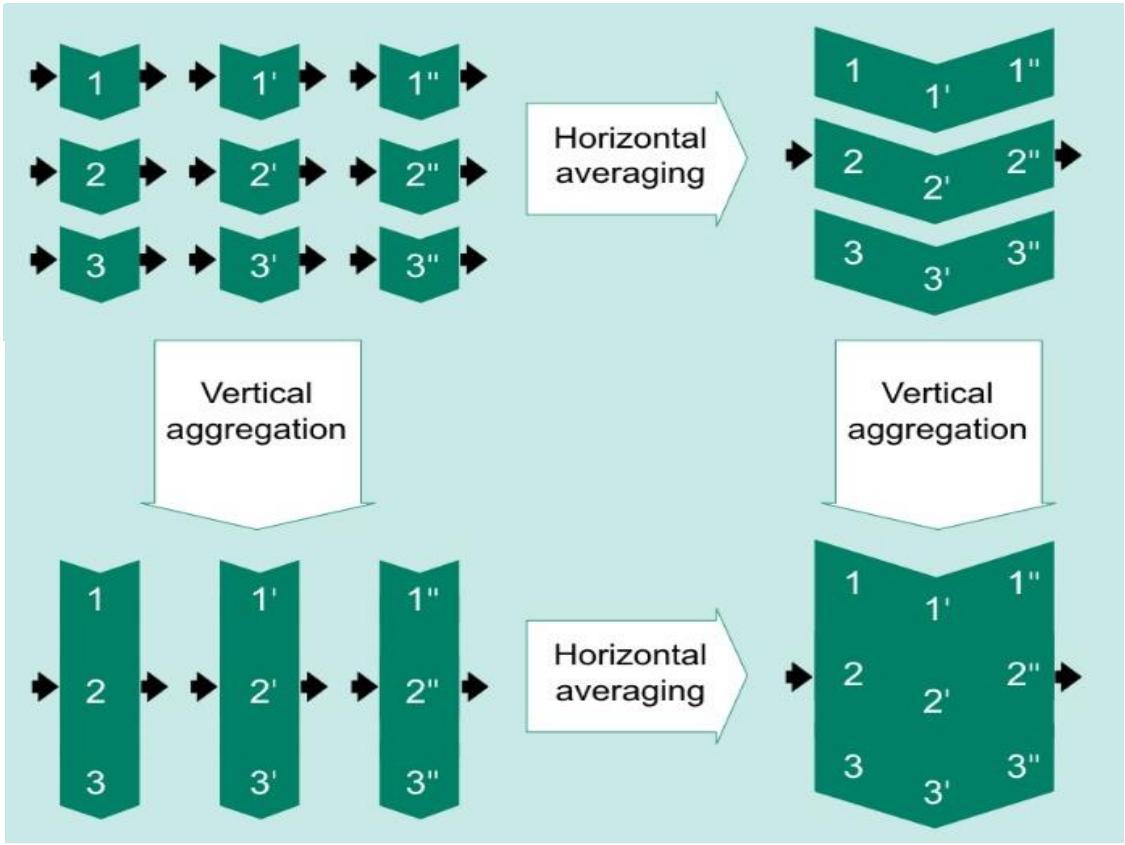
Which one is suitable for hot spot analysis?

Modelling your data according to the DB (vertical aggregation)

ecoinvent

- Check the system boundary(ies) of the similar dataset(s) in the DB
- Check what activities are included in each activity (unit process)
- Maybe read the sectorial guidelines of the DB
- You might be required to re-structure your data to fit with-in the structure of the DB
  - You might need to further divide your process
  - You might require to exclude some flows from your processes
  - You might be required to include some additional flows into your processes
  - You might need to gather additional information
  - You might need to shift some flows to other unit processes (transportation)

Total (bi-directional) averaging/aggregation



To retain

ecoinvent

- Aggregation shall be done considering the nature of the collected data, scope of the dataset, and the level of aggregation considered in the database and existing datasets
- It is about
  - modelling the data to suit (scale) the structure of the DB and existing datasets
  - reducing variety and uncertainty
  - better representation of the desired temporal/geographical/spatial etc. coverage of the dataset

Allocation

ecoinvent

- Allocation is partitioning the input and/or output flows of a process to the product system under study.
- Different criteria for allocation
  - Economic allocation
  - Physical allocation
  - Energy (exergy) allocation
  - Causal
  - Others.
- On the DB level, it is important to be able to use more than one allocation criteria (no hard allocation, but rather providing the possibility for using several allocation rules)
- Allocation principles and application may vary based on the database

Allocation in ecoinvent

ecoinvent

- In ecoinvent, datasets are in the form of unlinked, unallocated, multioutput Unit Processes
- As the data provider to the database, the data provider is NOT required to do any allocation
- You are, however, welcome to provide the additional info required for allocation purposes (i.e. price)
- The allocation is performed by the “system” on the database layer and according to the system model chosen

Additional considerations
Naming conventions

- Databases have defined nomenclature and naming rules
- The naming conventions shall be adjusted to that of the database
- This refers to how to “name your activity” as well as the exchange names.
  - *chapter 9 on naming rules from the Data Quality Guidelines of the ecoinvent v3*
  - *For chemicals, ecoinvent follows IUPAC rules*
  - *For products and services, ecoinvent follows International Standard Industrial Classification (ISIC)*
  - *and more…*

Documentation

ecoinvent

- Consider that your dataset has to be self-explanatory enough to be useful to the database users
  - Know the source(s) and age of the data;
  - Know how well the data represents an industry or process;
  - Understand how the underlying calculations were made;
  - Evaluate the appropriateness of the data for the user's intended application;
  - Validate the results through testing and cross-checking of data and modeling;
  - Make an informed determination concerning the extent to which they can rely on the data and conclusions drawn from it.
- Each DB can have its specific requirements for documentation

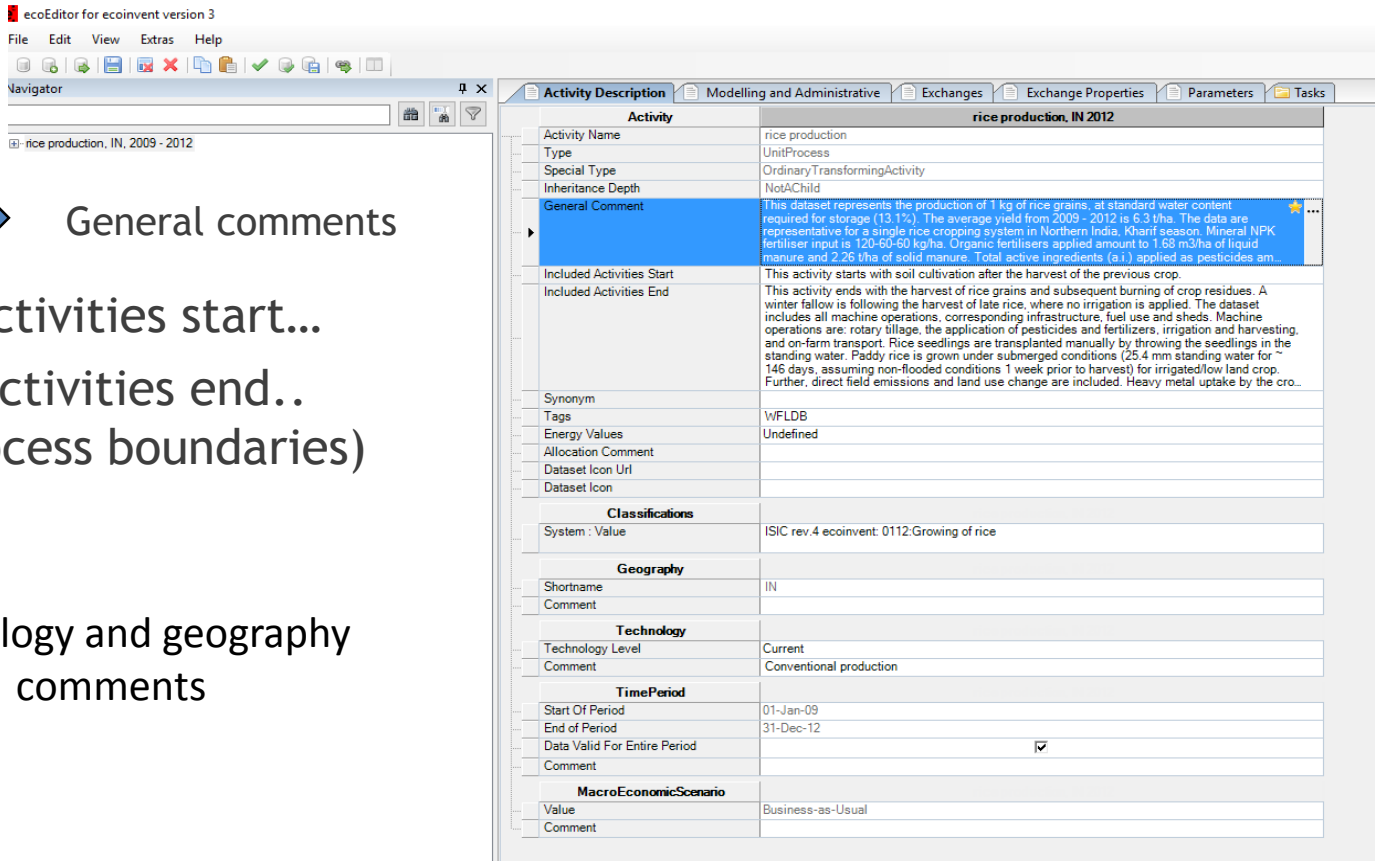
How is the documentation done

ecoinvent example

- In ecoinvent (ecospold 2), there are mandatory and optional fields for almost every value or exchange to be documented.
- Comments and references that are general to more than one entry are provided in the comment field most relevant for the nature of the value
- Comments and references to sources are given on the most detailed level possible, describing each individual value and their estimation i.e.
  - i.e. attributed to the particular exchanges of an activity
  - attributed to a particular property of an exchange, if possible and relevant)
  - …

Documentation, example

Rice production, IN



General comments

Activities start…

Activities end..
(process boundaries)

Technology and geography comments

Documentation, example

General comment for the dataset “Rice production, IN”

This dataset represents the production of 1 kg of rice grains, at standard water content required for storage (13.1%). The average yield from 2009 - 2012 is 6.3 t/ha. The data are representative for a single rice cropping system in Northern India, Kharif season. Mineral NPK fertilizer input is 120-60-60 kg/ha. Organic fertilizers applied amount to 1.68 m3/ha of liquid manure and 2.26 t/ha of solid manure. Total active ingredients (a.i.) applied as pesticides amount to 1.5 kg a.i./ha.

Documentation (cont'd)…

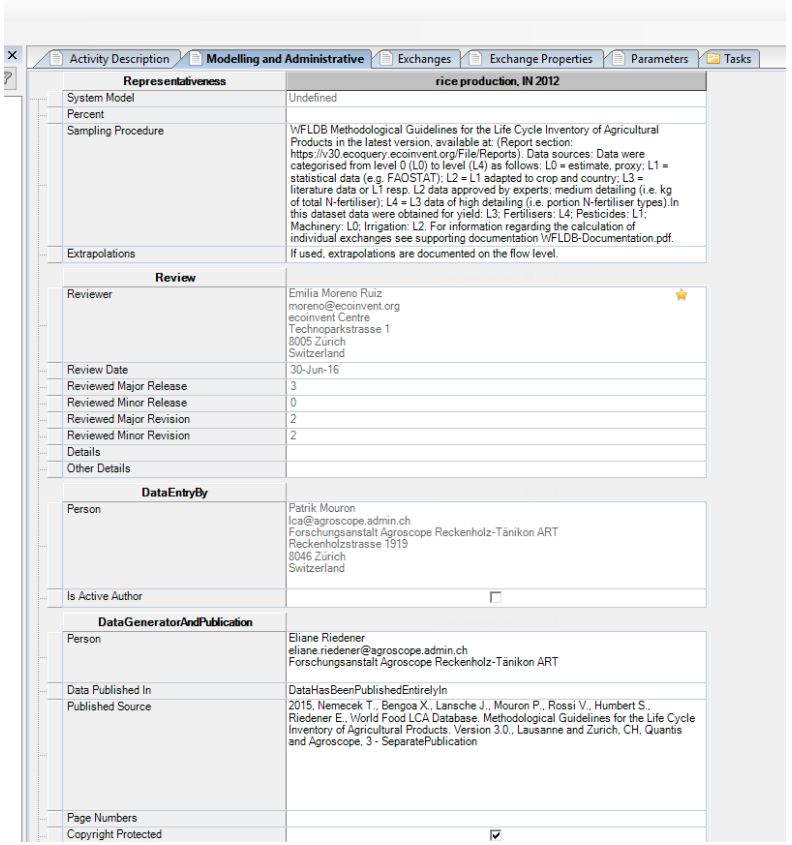
Sampling procedure

Extrapolations

Review process and reviewers

Data entry

Data generator

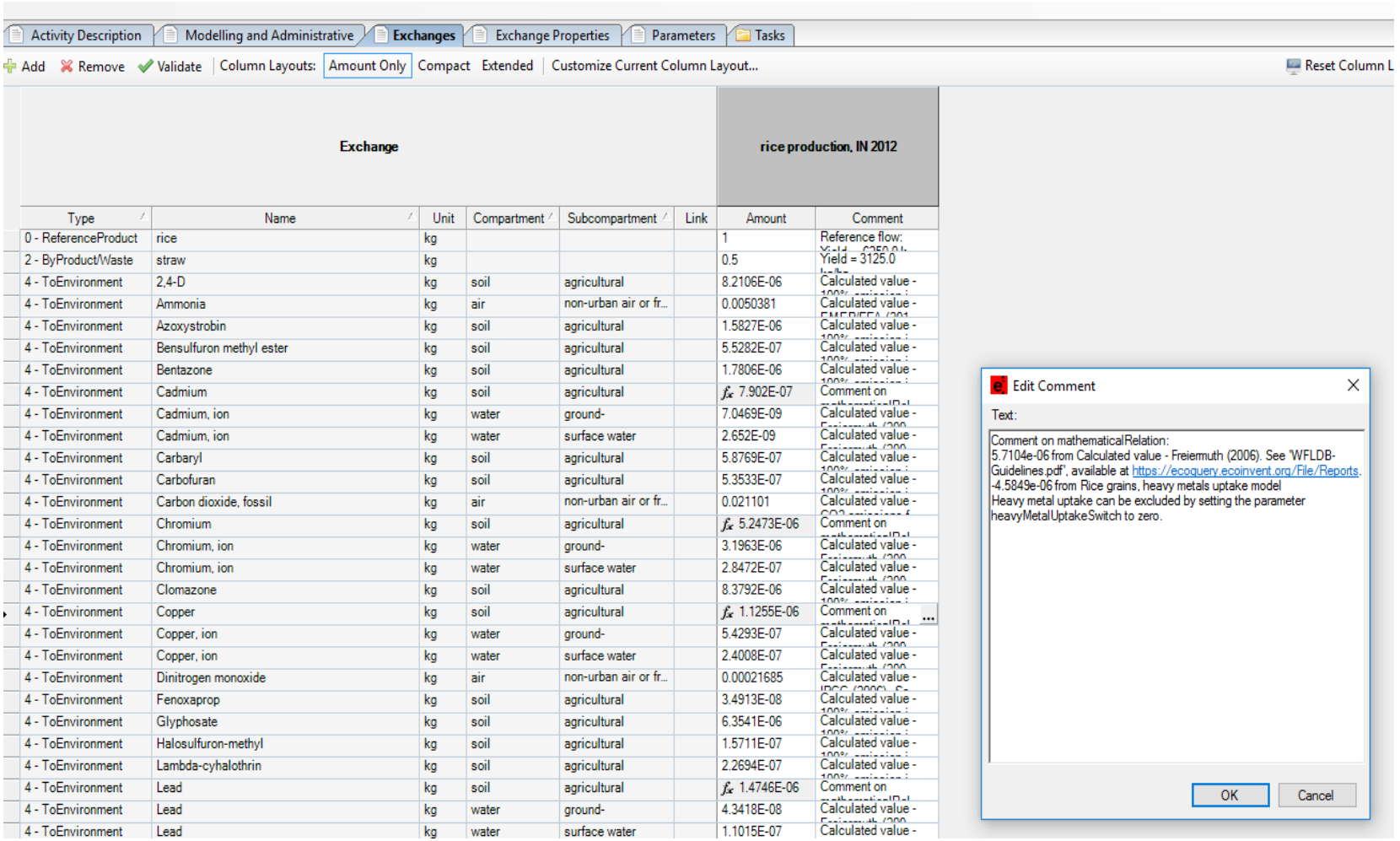


Documentation; system boundary

- *This activity **starts with** soil cultivation after the harvest of the previous crop.*

- *This activity **ends with** the harvest of rice grains and subsequent burning of crop residues. A winter fallow is following the harvest of late rice, where no irrigation is applied. The dataset includes all machine operations, corresponding infrastructure, fuel use and sheds. Machine operations are: rotary tillage, the application of pesticides and fertilizers, irrigation and harvesting, and on-farm transport. Rice seedlings are transplanted manually by throwing the seedlings in the standing water. Paddy rice is grown under submerged conditions (25.4 mm standing water for ~ 146 days, assuming non-flooded conditions 1 week prior to harvest) for irrigated/low land crop. Further, direct field emissions and land use change are included. Heavy metal uptake by the crop is considered.*

Comments on specific exchanges



To retain

ecoinvent

- Documentation is an integral part of dataset development and modelling
- Each DB can have specific requirements and fields for documentation
- Do not try to only “meet the minimum requirements”.
- Document your dataset in as much detail as possible, from general considerations to documenting specific values
- The documentation of your activity needs to be self-explanatory enough to be useful to the database users

Sources of missing data -- Step by step approach

ecoinvent

- If some data are missing, the data provider can do following
  - Contact the local producers with request to provide data
  - Check the same process in different geographies
  - Check the industry association documents – (IAI – International Aluminium Association, etc.)
  - Check the already published literature on the topic
  - Check the local statistical office documents or other public institutions (e.g. Eurostat, US EPA, USGS)
  - Check the international public ally available databases (e.g. FAOSTAT, AQUASTAT, IEA, other UN databases, etc.)
  - Etc. there are now boundaries to this process ;-)

Sources of missing data - Step by step approach

ecoinvent

- If some data are still missing…
  - Consider using models
    - rice production/CN – the use of pesticides and fertilizers during rice production in China can be modelled when knowing the average used pesticides and fertilizers in China and some distribution coefficient on how these chemicals behave once released to the environment)
    - production of chemicals – stoichiometric calculation
  - Think out of the box!
    - Seeking the electricity consumption of a facility? Check their electricity bill.

- If data are missing, they cannot be simply excluded!!

Considerations for sectors with high modelling needs

ecoinvent

- Several sectors require high modelling needs to calculate the LCI and create datasets
- Examples include: agriculture
  - Agriculture sector and
- The development of agricultural LCI is time-consuming because of the complexity of the agricultural modelling.
- Consistency: Further, there is a risk that different models are used in developing the inventory, and so creating inconsistencies

Sectors with high modelling needs (example of agriculture)

- There are existing models to calculate the faith of several exchanges (emissions) due to agricultural practices, namely

  - Direct land use change
  - Irrigation
  - Soil erosion
  - Nitrate leaching
  - Phosphorus and phosphate to water
  - Ammonia
  - Heavy metal to agricultural soil, surface water and groundwater
  - Nitrous oxides
  - Nitrogen oxides
  - CO2 from urea or lime application
  - Methane from rice cultivation

Tools for easier generation of LCI data

ecoinvent

- Open ALCIG

https://alcig.quantis-software.com/

To retain

- Several sectors require further modeling considerations to be able to calculate and generate LCI datasets
- You should be aware of, and abide by, such modelling needs to generate **consistent datasets**
- Always consult the sectorial reports of the database you are working with or wish to emulate, and search for any tools which can facilitate the generation of such datasets

Uncertainty

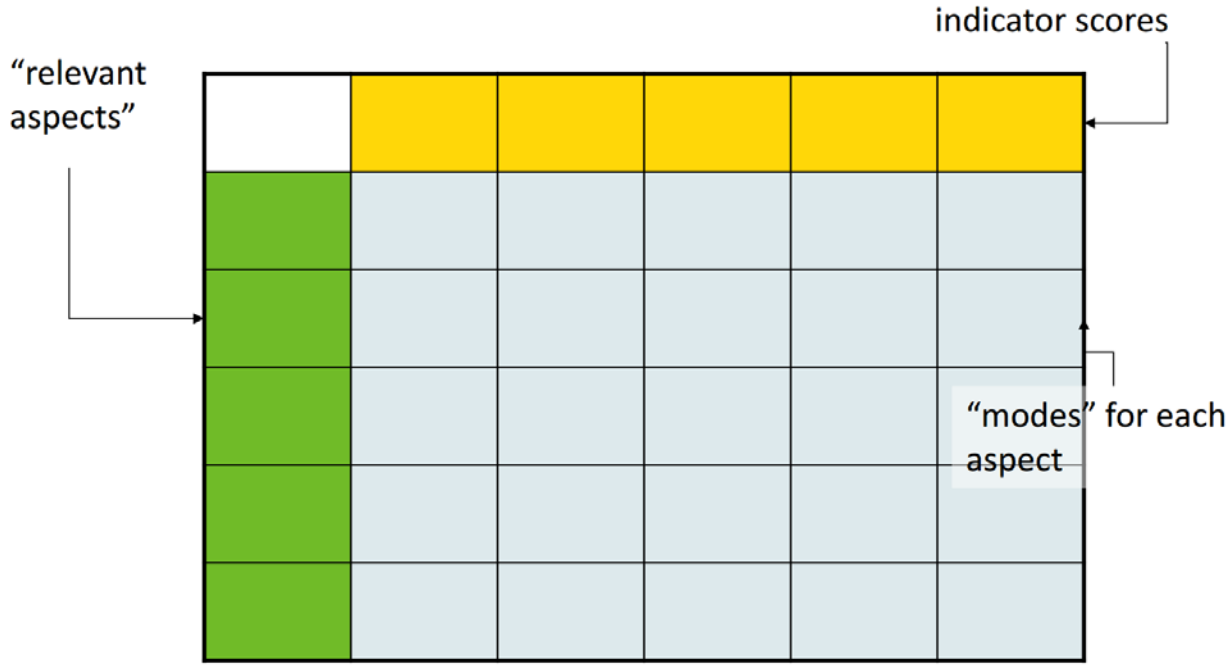
- Uncertainty is present in all phases of LCA
  - Goal and scope definition
  - LCI
  - LCIA
- Consequently, all the numerical values contain an associated uncertainty
- The approaches to calculate the uncertainty can be quantitative (Monte Carlo simulation e.g.) and qualitative
- Once again, each DB can have different requirements to deal with uncertainty
- A well-used qualitative method in LCA for estimation of uncertainty is **pedigree Matrix**

Pedigree Matrix

ecoinvent

- A method to transform qualitative modes of uncertainty into quantitative modes
- columns are basic aspects
- lines are 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

Pedigree Matrix concept



Uncertainty in ecoinvent

- All numerical fields have an associated uncertainty field.
- ecoinvent employs a hybrid between a quantitative (statistical analysis) and a qualitative approach (expert judgement)
- In case of uncertainty, one should follow the below steps:
  - select the probability function,
  - Provide the distribution parameters (e.g. variance of log-transformed data)
  - edit the Pedigree matrix using expert judgement.
- If you have enough data points to be able to obtain a “distribution” for your data, you can enter that in the DB
- Otherwise, you can use the «default values» according to the DQ guidelines

Pedigree matrix in ecoinvent

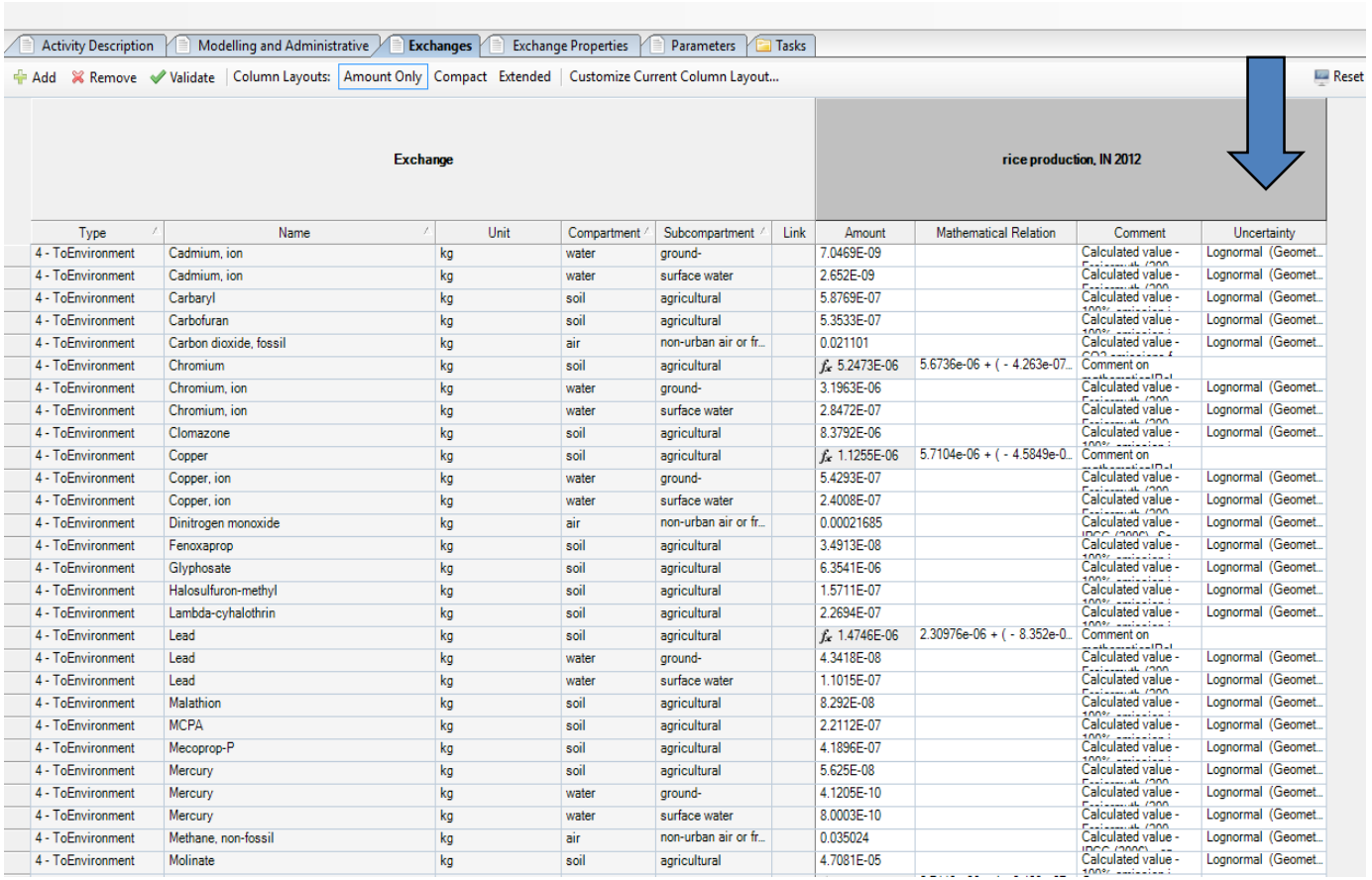
ecoinvent

Indicator score	1	2	3	4	5 (default)
Reliability	Verified[5] data based on measurements[6]	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	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

Uncertainty in ecoinvent

All numerical fields have an uncertainty field.

Uncertainty field



Exchange						rice production, IN 2012			
Type	Name	Unit	Compartment	Subcompartment	Link	Amount	Mathematical Relation	Comment	Uncertainty
4 - ToEnvironment	Cadmium, ion	kg	water	ground-		7.0469E-09		Calculated value -	Lognormal (Geomet...
4 - ToEnvironment	Cadmium, ion	kg	water	surface water		2.652E-09		Calculated value -	Lognormal (Geomet...
4 - ToEnvironment	Carbaryl	kg	soil	agricultural		5.8769E-07		Calculated value -	Lognormal (Geomet...
4 - ToEnvironment	Carbofuran	kg	soil	agricultural		5.3533E-07		Calculated value -	Lognormal (Geomet...
4 - ToEnvironment	Carbon dioxide, fossil	kg	air	non-urban air or fr...		0.021101		Calculated value -	Lognormal (Geomet...
4 - ToEnvironment	Chromium	kg	soil	agricultural		5.2473E-06	5.6736e-06 + (- 4.263e-07...	Comment on	
4 - ToEnvironment	Chromium, ion	kg	water	ground-		3.1963E-06		Calculated value -	Lognormal (Geomet...
4 - ToEnvironment	Chromium, ion	kg	water	surface water		2.8472E-07		Calculated value -	Lognormal (Geomet...
4 - ToEnvironment	Clomazone	kg	soil	agricultural		8.3792E-06		Calculated value -	Lognormal (Geomet...
4 - ToEnvironment	Copper	kg	soil	agricultural		1.1255E-06	5.7104e-06 + (- 4.5849e-0...	Comment on	
4 - ToEnvironment	Copper, ion	kg	water	ground-		5.4293E-07		Calculated value -	Lognormal (Geomet...
4 - ToEnvironment	Copper, ion	kg	water	surface water		2.4008E-07		Calculated value -	Lognormal (Geomet...
4 - ToEnvironment	Dinitrogen monoxide	kg	air	non-urban air or fr...		0.00021685		Calculated value -	Lognormal (Geomet...
4 - ToEnvironment	Fenoxaprop	kg	soil	agricultural		3.4913E-08		Calculated value -	Lognormal (Geomet...
4 - ToEnvironment	Glyphosate	kg	soil	agricultural		6.3541E-06		Calculated value -	Lognormal (Geomet...
4 - ToEnvironment	Halosulfuron-methyl	kg	soil	agricultural		1.5711E-07		Calculated value -	Lognormal (Geomet...
4 - ToEnvironment	Lambda-cyhalothrin	kg	soil	agricultural		2.2694E-07		Calculated value -	Lognormal (Geomet...
4 - ToEnvironment	Lead	kg	soil	agricultural		1.4746E-06	2.30976e-06 + (- 8.352e-0...	Comment on	
4 - ToEnvironment	Lead	kg	water	ground-		4.3418E-08		Calculated value -	Lognormal (Geomet...
4 - ToEnvironment	Lead	kg	water	surface water		1.1015E-07		Calculated value -	Lognormal (Geomet...
4 - ToEnvironment	Malathion	kg	soil	agricultural		8.292E-08		Calculated value -	Lognormal (Geomet...
4 - ToEnvironment	MCPA	kg	soil	agricultural		2.2112E-07		Calculated value -	Lognormal (Geomet...
4 - ToEnvironment	Mecoprop-P	kg	soil	agricultural		4.1896E-07		Calculated value -	Lognormal (Geomet...
4 - ToEnvironment	Mercury	kg	soil	agricultural		5.625E-08		Calculated value -	Lognormal (Geomet...
4 - ToEnvironment	Mercury	kg	water	ground-		4.1205E-10		Calculated value -	Lognormal (Geomet...
4 - ToEnvironment	Mercury	kg	water	surface water		8.0003E-10		Calculated value -	Lognormal (Geomet...
4 - ToEnvironment	Methane, non-fossil	kg	air	non-urban air or fr...		0.035024		Calculated value -	Lognormal (Geomet...
4 - ToEnvironment	Molinate	kg	soil	agricultural		4.7081E-05		Calculated value -	Lognormal (Geomet...

Dataset validation - first steps

ecoinvent

- Balance: is your dataset (mass, carbon, material) balanced?
  What goes IN, must go OUT!
  - Check the mass balance of your dataset
  - Check the carbon balance of your dataset
- Take a step back, and verify if data actually make sense when put together
  - Come out of the woods to see the woods!
- Validation is both on a DS level as well as on the LCIA level

LCI dataset review

ecoinvent

- General review criteria
  - **Dataset classification** follows database requirements
  - **Nomenclature** is correct and consistent with applied nomenclature and terminology
  - **Modeling method** is consistent with the requirements of the database
  - **Scope and boundary** is consistent with the requirements of the database
  - **Information regarding the data quality indicator DQIs (and if appropriate, how the aggregated DQI results were determined)** is necessary
  - **The appropriateness, correctness, extent of documentation, and the metadata information** in the dataset are consistent with the requirements of the database
- Specific requirements for review of aggregated process datasets
  - Mainly focus on modeling and documentation

Source: Sonnemann, G., & Vigon, B. (2011); see also Vigon et al. (2016

Checking your created datasets

ecoinvent

- **Completeness check**
  - Inventory list (inputs/outputs) completeness
    - compared with the general target representativeness and impact categories of database
    - check against process boundary (process activities included)
    - compare with similar datasets in other databases and technical literature (special effort if a never before created dataset)
  - Document completeness
    - Raw data, mathematical relations and literature cites or original bases of each input/output
    - Supportive information and else according to documentation format

Source: Sonnemann, G., & Vigon, B. (2011)

UP dataset checking (cont'd)

ecoinvent

- **Plausibility check**
  - Balance check: mass, element, water, energy balance
  - Magnitude checks: to prevent typos and unit conversion mistakes
  - Comparison with process data and LCIA results from alternative data sources or mathematical relationships
  - Expert information exchange
  - Statistical tool used to identify outliers - if sample data population exists and technical understanding is not sufficient
- **Consistency check**
  - Technological representativeness
  - Temporal representativeness
  - Geographical representativeness
  - Goal, scope, models and assumption
- **Sensitivity**
  - Can be analyzed when a life cycle model is established.

Source: Sonnemann, G., & Vigon, B. (2011)

Main steps of dataset submission and validation

ecoinvent example

data provider	ecoinvent Centre management	ecoinvent editor
contact the ecoinvent Centre (support@ecoinvent.org)	initial agreement of cooperation	
creates **list of datasets** (activity and reference product names) which he/she wants to submit	identify and contact the responsible editor	decide whether the dataset in question is **relevant for the database**
enter the data into ecoEditor -> creation of the spold file	guidance during the dataset creation in the ecoEditor	
run off-line and on-line **validation**, edit dataset accordingly	guidance during dataset's editing (based on validation errors and warning messages)	
submission of the dataset into review		start the **review** of the dataset, enter notes on what should be corrected
edits the dataset according to comments from the editor		**final approval** of the dataset

References

- Sonnemann, G. and B. Vigon, Eds. (2011). Global guidance principles for Life Cycle Assessment (LCA) databases: a basis for greener processes and products. ISBN: 978-92-807-3174-3, 2011, 156 pages.
- Vigon, B. G. Sonnemann, A. Asselin, D. Schrijvers, A. Ciroth, S. S. Chen, T. Braga, N. Poolsawad, J. Mungkalasiri, F. Boureima and L. Milà i Canals, Review of LCA datasets in three emerging economies: a summary of learnings, International Journal of Life Cycle Assessment, September 2016, DOI:10.1007/s11367-016-1198-2.

Part 5 - Modelling processes for national dataset development

Case Studies from Sri Lanka and Brazil

Content from Andreas Ciroth, GreenDelta and Amir Safaei, ecoinvent

August 2017 Version

Managed by SETAC

Process datasets for Sri Lanka (how to)

Objective

- We will now, with the information that we have received so far, create new data sets with secondary data information, for the following products:
  - Rice production and milling
  - Tea production and processing
  - Dairy production
- For information, use the internet and also the material provided in the exercise
  - Filename: exercise_3_datasources.docx
- Split in groups for exercise

Exercise 3 – National dataset development by adjustment

With the information obtained in the training, how would you create datasets that are representative for Sri Lanka for one of the following options, Rice production and milling, Tea production and processing or Dairy production?

Below you can find examples of sources for **Rice datasets.**

Source: http://irri.org/our-work/locations/sri-lanka

Sri Lanka and IRRI started collaborating in 1960 through training and exchange of rice varieties. In 1967, an agreement between Sri Lanka and the Ford Foundation led to a two-year program between IRRI and the country's Department of Agriculture where its scientists eventually trained at IRRI. Renewed in 1969, the program also included technology transfer activities.

Different projects were conducted with the collaborative efforts of IRRI and partner institutions in Sri Lanka, one of which is a collaboration with the country's Department of Agriculture (DOASL) and the United States Agency for International Aid (USAID). They undertook a Rice Research project where Sri Lankans were provided with nearly 25 graduate scholarships to different universities in the US, UK, and the Philippines, as well as 100 short-term trainings in rice production, and cropping systems. Moreover, the project also provided research equipment and infrastructure development in rice research stations, and included as a component the improvement of rice varieties.

In 2007, DOASL and IRRI held a collaborative workplan meeting where they identified some areas for mutual cooperation. This included improving existing rice varieties in terms of quality and yield, enhancing conservation of rice genetic resources, increasing labor productivity, and strengthening the delivery and impact of technology through good extension models. Similar with past agreements, DOASL continued to serve as the clearing house for IRRI's activities in Sri Lanka. Both institutions continue to find ways and means of improving collaboration including funding support.

RICE RESEARCH AND DEVELOPMENT INSTITUTE
Department of Agriculture - Sri Lanka

RICE CULTIVATION

RICE

Rice is the single most important crop occupying 34 percent (0.77 /million ha) of the total cultivated area in Sri Lanka. On average 560,000 ha are cultivated during maha and 310,000 ha during yala making the average annual extent sown with rice to about 870,000 ha. About 1.8 million farm families are engaged in paddy cultivation island-wide. Sri Lanka currently produces 2.7 million t of rough rice annually and satisfies around 95 percent of the domestic requirement. Rice provides 45% total calorie and 40% total protein requirement of an average Sri Lankan. The per capita consumption of rice fluctuates around 100 kg per year depending on the price of rice, bread and wheat flour.

The current cost of production of rough rice is Rs. 8.57 per kg. The cost of labor, farm power and tradable inputs constitutes 55%, 23% and 23% respectively.

Field water requirement for a rice crop depends mainly on the growth duration of the crop and its growing environment. It is calculated that about 30-40% of the total water supplied to an irrigated crop is often supplied before the establishment of the rice crop and the amount is dependent on the soil drainage class, weed density and time taken for land preparation. Time taken for land preparation could be minimised to about 2 weeks using total killing herbicides (e.g. Paraquat) which also would help to reduce one tillage operation and conserve irrigation water.

Recommended Herbicides for Rice Weeds
One shot herbicides

Trade Name	Common Name	Rate of application/ha
Sofit	Pretilachlor 300 g/lit.EC	1.6 lit.
Goal	Oxyflorfen 240 g/lit.EC	0.5 lit.

Broad leaves/Sedge Herbicide

Trade Name	Common name	Rate of application/ha
Hedonal	2.4 D 550 g/lit.SL	0.95 - 1.2 lit.
	MCPA 600 g/lit.SL/EC	0.8 - 1.1lit.

Source: http://www.doa.gov.lk/rrdi/index.php/en/rice

New Rathna Rice Mill
South Asia's Largest Rice Production Line

Head Office & Factory 01

New Rathna Rice (PVT) Ltd.
No. 54, Somawathiya Road,
Pulasthigama, Polonnaruwa,
Sri Lanka.
Tel - 027-2242627
Email - newrathna.rice@gmail.com

The Rice Grain

The rice grain has both physical and chemical characteristics.

Physical Structure

A rice grain is made up of an outside husk layer, a bran layer, and the endosperm, see Figure 1. The husk layer (lemma and palea) accounts for 20% of the weight of paddy and helps protect the grain kernel from insect and fungal attack. When the husk is removed, the rice is called brown rice. Brown rice contains the bran layer and the endosperm. The bran layer is made up of the pericarp and testa, the aluerone layer and the embryo. The degree to which this bran layer is removed is known as the milling degree. The desired amount of bran removed depends on the country. In Japan, the aluerone layer is often not removed however in many other countries all bran layers are removed to give very highly polished rice. The storage life of milled rice is improved when all of the bran layers are removed.

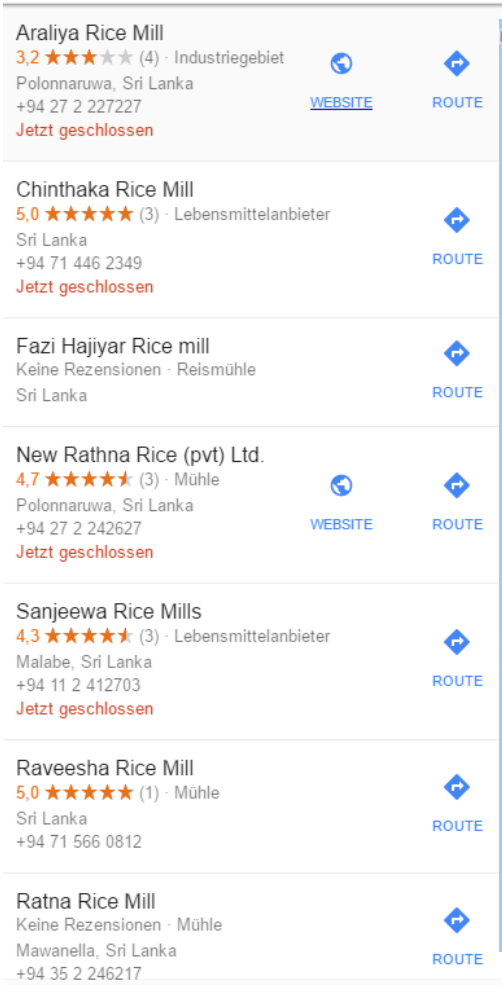
Physical characteristic	Percentage
Paddy	100
Husk	20
Brown rice	80
Meal Pericarp and testa (5-6%) Aluerone (1%) Embryo (3%)	8-10
White rice	70-72

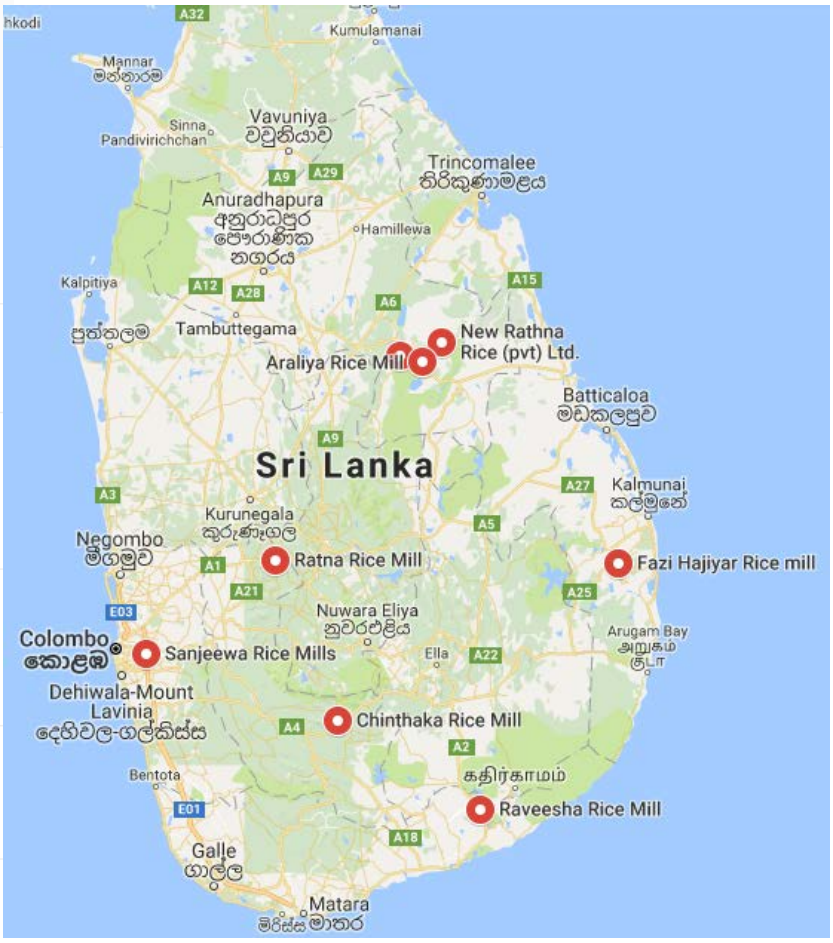
Chemical Composition of Milled Rice

Rice at 12% moisture contains approximately 80% starch and 7% protein. (Currey, 1984
Starch occurs in the endosperm as small many-sided granules while protein is present a
particles that lie between the starch granules. Rice grain also contains sugars, fat, dietar
fiber and minerals.

Component	Brown rice	White rice	Bran
Water (%)	13-14	13-14	13-14
Starch (%)	68-70	80	9
Amylose	28-30	33	6
Protein (%)	6-8	6-7	14
Fat	3	1	20
Fiber	2-3	0.5	25
Crude ash	1-1.5	0.5	9-10

Source: http://www.newrathnarice.com/





Source: https://www.google.de

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/agwat

Agricultural Water Management

Predicting the impacts of climate change—A case study of paddy irrigation water requirements in Sri Lanka

C.S. De Silva [a,*], E.K. Weatherhead [b], J.W. Knox [b], J.A. Rodriguez-Diaz [c]

[a] Department of Agricultural Engineering, Faculty of Engineering Technology, The Open University of Sri Lanka, Nawala, Nugegoda, Sri Lanka

[b] Cranfield University, Bedford MK43 0AL, UK

[c] IFAPA, Junta de Andalucia, CIFA Alameda del Obispo, 14080 Cordoba, Spain

ARTICLE INFO

Article history:
Received 3 March 2006
Accepted 13 June 2007
Published on line 23 July 2007

Keywords:
Food production
Geographical information system
Water resources
Spatial variability
Adaptation

ABSTRACT

Nearly 72% of paddy production, the staple food in Sri Lanka, is grown during the wet season in dry areas where water resources are already stressed. Climate change datasets for Sri Lanka were derived using outputs from the UK Hadley Centre for Climate Prediction and Research Model (HadCM3) for selected scenarios for the 2050s, chosen from the Intergovernmental Panel of Climate Change Special Emission Scenarios Report. Water balance modelling and a geographical information system were used to model and map the impacts on irrigation requirements for wet season paddy. We examined two scenarios. The A2 scenario represents a heterogeneous, regionalised, market-led world, with high population growth, leading to a rapid increase in atmospheric carbon dioxide levels. The B2 scenario follows a similar regionalised future but with more moderate population growth and more concern for the environment and local sustainability, and a slower rate of increase in atmospheric carbon dioxide.

Results suggests that, during the wet season, average rainfall decreases by 17% (A2) and 9% (B2), with rains ending earlier, and potential evapotranspiration increasing by 3.5% (A2) and 3% (B2). Consequently, the average paddy irrigation water requirement increases by 23% (A2) and 13% (B2).

Mapping with GIS highlights the importance of considering spatial variation. Climate change impacts on wet season paddy production are positive in the extreme south, confirming results of a previous study. However, the impacts are negative across most of Sri Lanka. The adaptations needed are different in the two regions. Furthermore, spatial variation points to a further adaptation; the transfer of some paddy production to positively affected areas, which would not have been so clear if only point modelling had been used.

Source: http://www.sciencedirect.com/science/article/pii/S037837740700162X

Process datasets for Sri Lanka (how to)

Recapitulation

- What have we learned
  - Rice production and milling
  - Tea production and processing
  - Dairy production
- Do you feel yourself confident with the datasets, do you see need for improvement
  - Rice production and milling
  - Tea production and processing
  - Dairy production

Part 5 - Modelling processes for national dataset development

Case Study from Brazil

Content from Amir Sfaei, ecoinvent

August 2017 Version

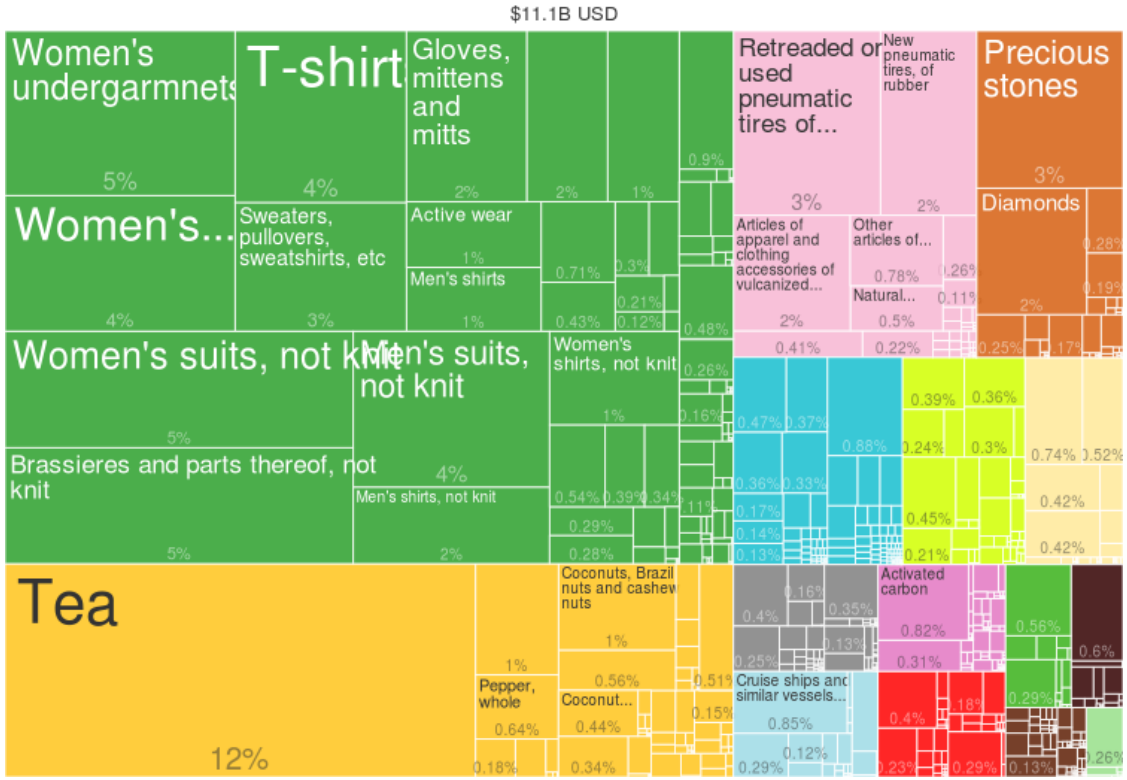
Managed by SETAC

Different approaches to prioritize data collection

ecoinvent

Economic Relevance

Use export statistics, production statistics to rank the relevance of products for the economy!



- PEF – Product Environmental Footprint
- Others?
  - Contribution to GDP

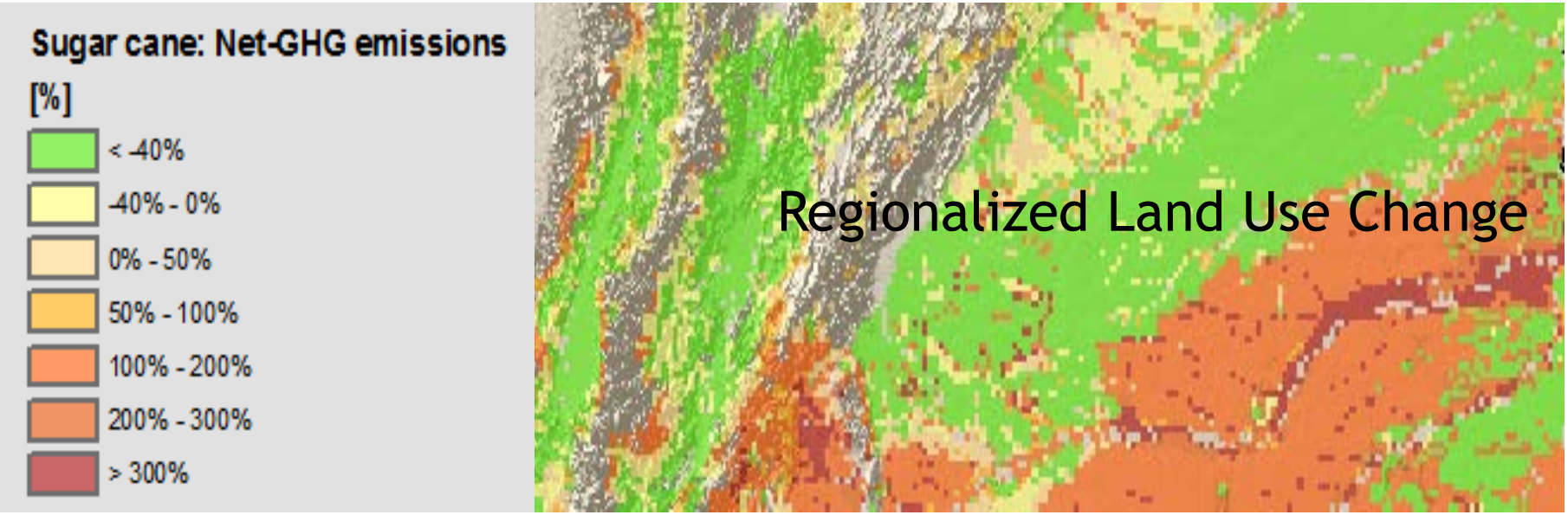
From:*http://atlas.cid.harvard.edu*

Sensitivity to the local context

ecoinvent

- Some datasets are more sensitive to local conditions than others

Diesel burned in building machine



Database Relevance
Case of Brazil

ecoinvent

Goal and Starting Point

ecoinvent

- Goal: Identification of the priority products for a national database based on expected contribution to environmental impacts;
- Starting point: The key products based on production and economic importance;
  - Source: The Brazilian Institute of Geography and Statistics (IBGE) report of production and sales of the top 100 products and/or industrial services in 2011, according to the national position in the value of sales;
  - Besides the report items, the electricity was also included in the assessment;

Chiumento e Ugaya, 2014

Method (1)

- After choosing the products: find their correspondence with database (ecoinvent 3.1);
  - 67 products of the report plus electricity were considered
  - Some data treatment were required to include relevant products
    - i.e. VHP sugar and crystal sugar were associated to the main correspondence: sugar.

Chiumento e Ugaya, 2014

Method (2)

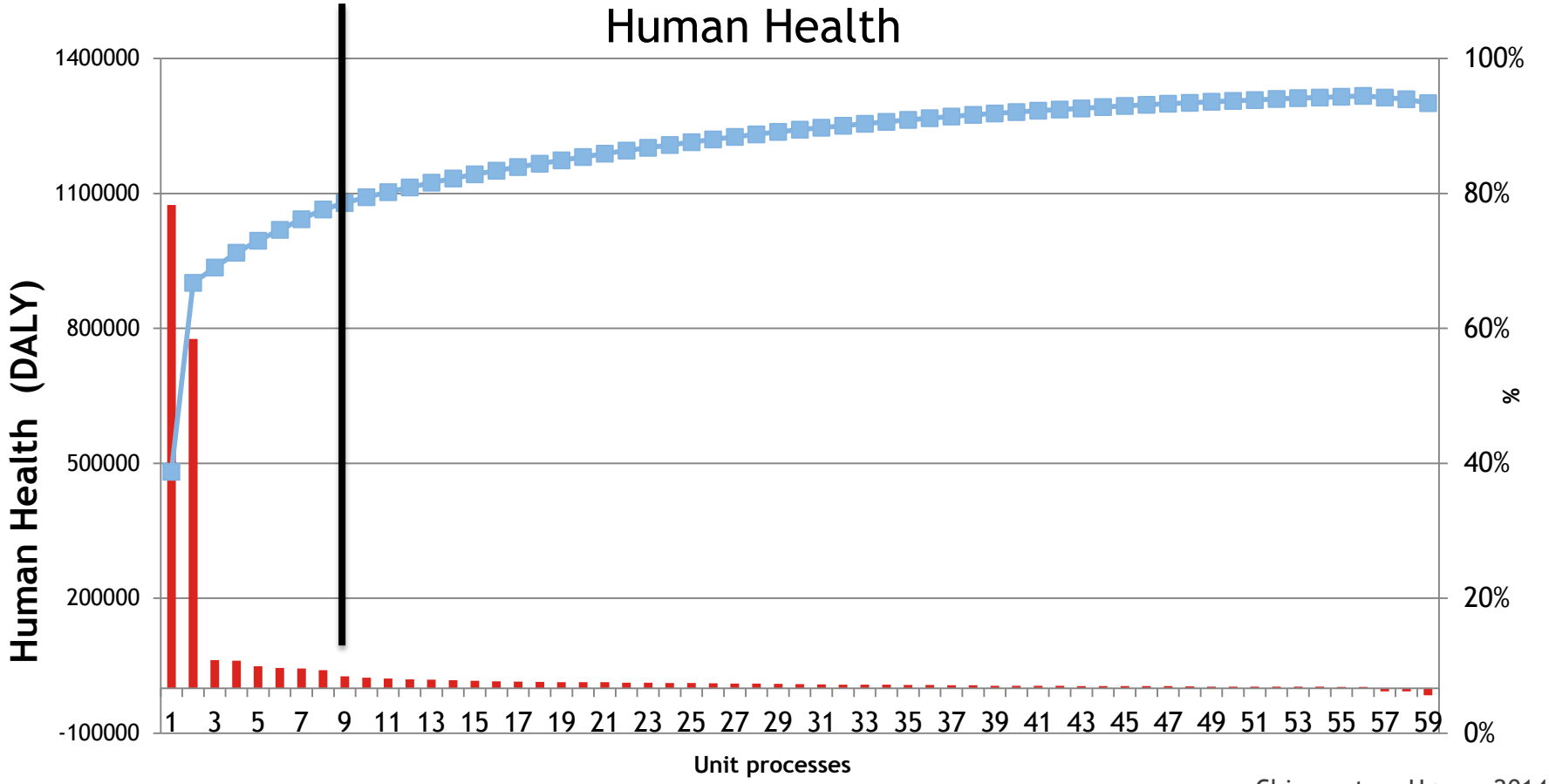
ecoinvent

- Identification of priority unit process: LCA technique;
  - Life Cycle Impact Assessment Method: Impact 2002+ endpoint;
  - After classifying the products in each impact category, Pareto's principle was used as a cut-off rule to select the most relevant results.
  - Products contributing up to 80% of the LCIA totals were selected.

Chiumento e Ugaya, 2014

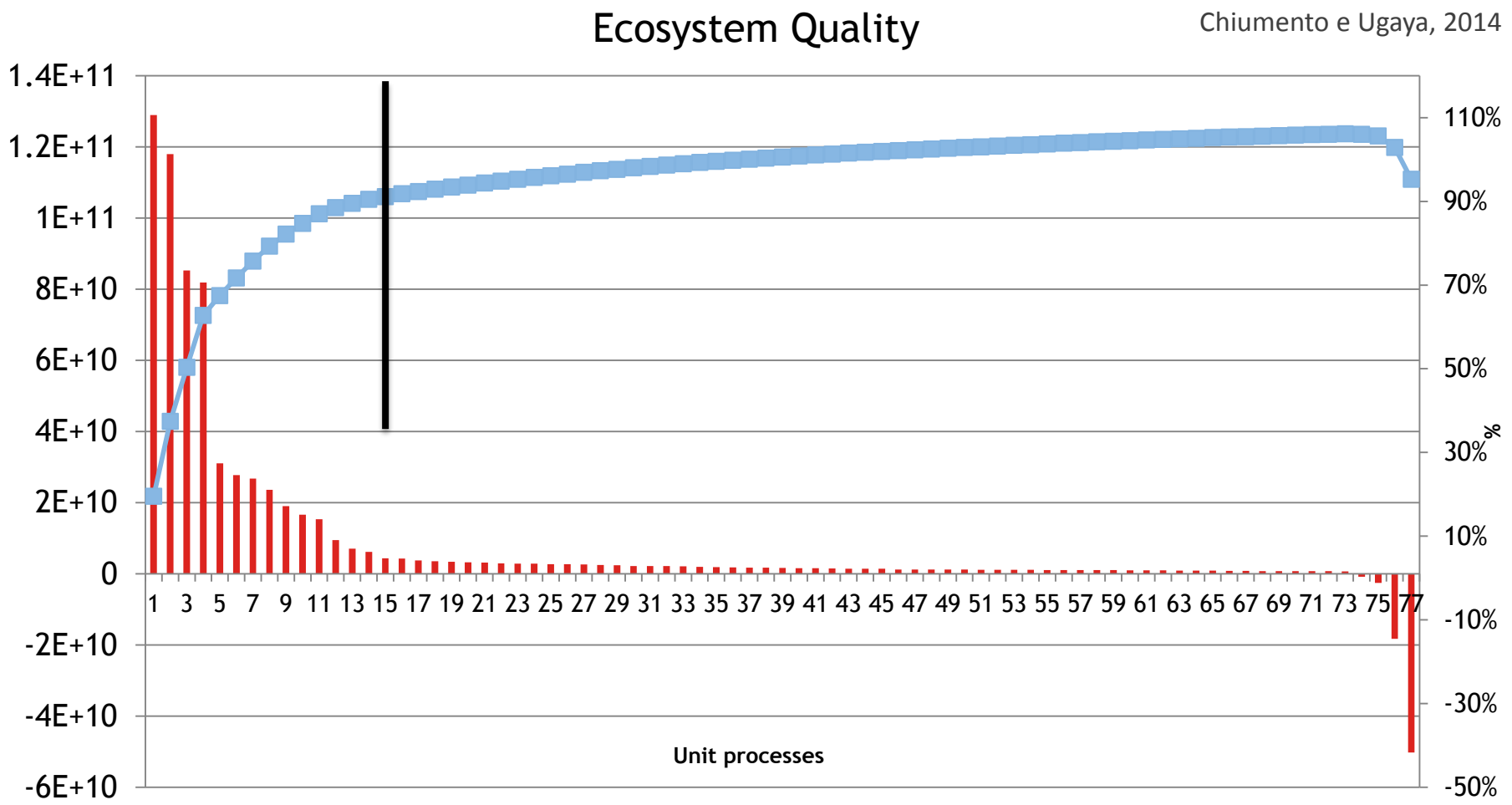
Results: Human Health

ecoinvent

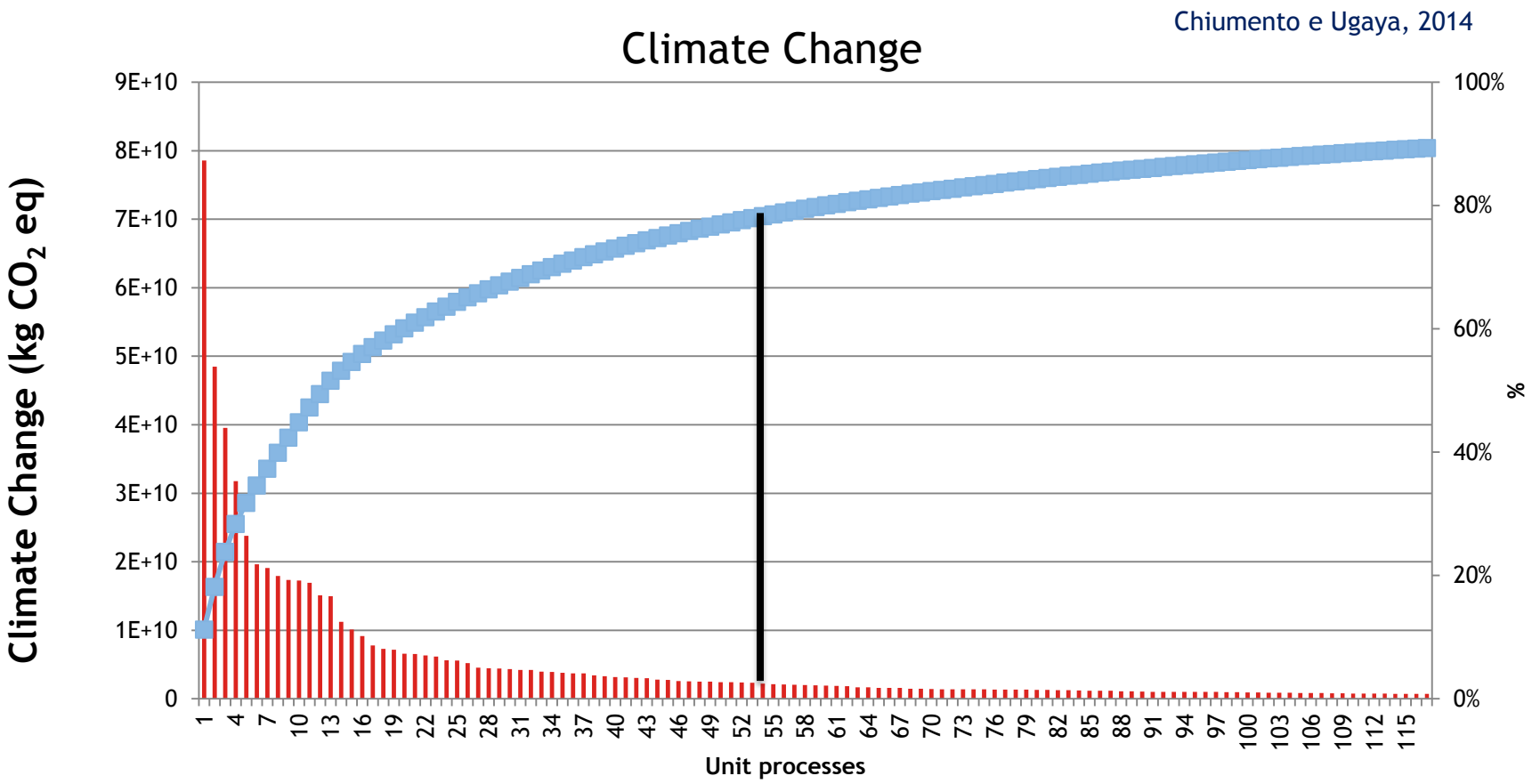


Chiumento e Ugaya, 2014

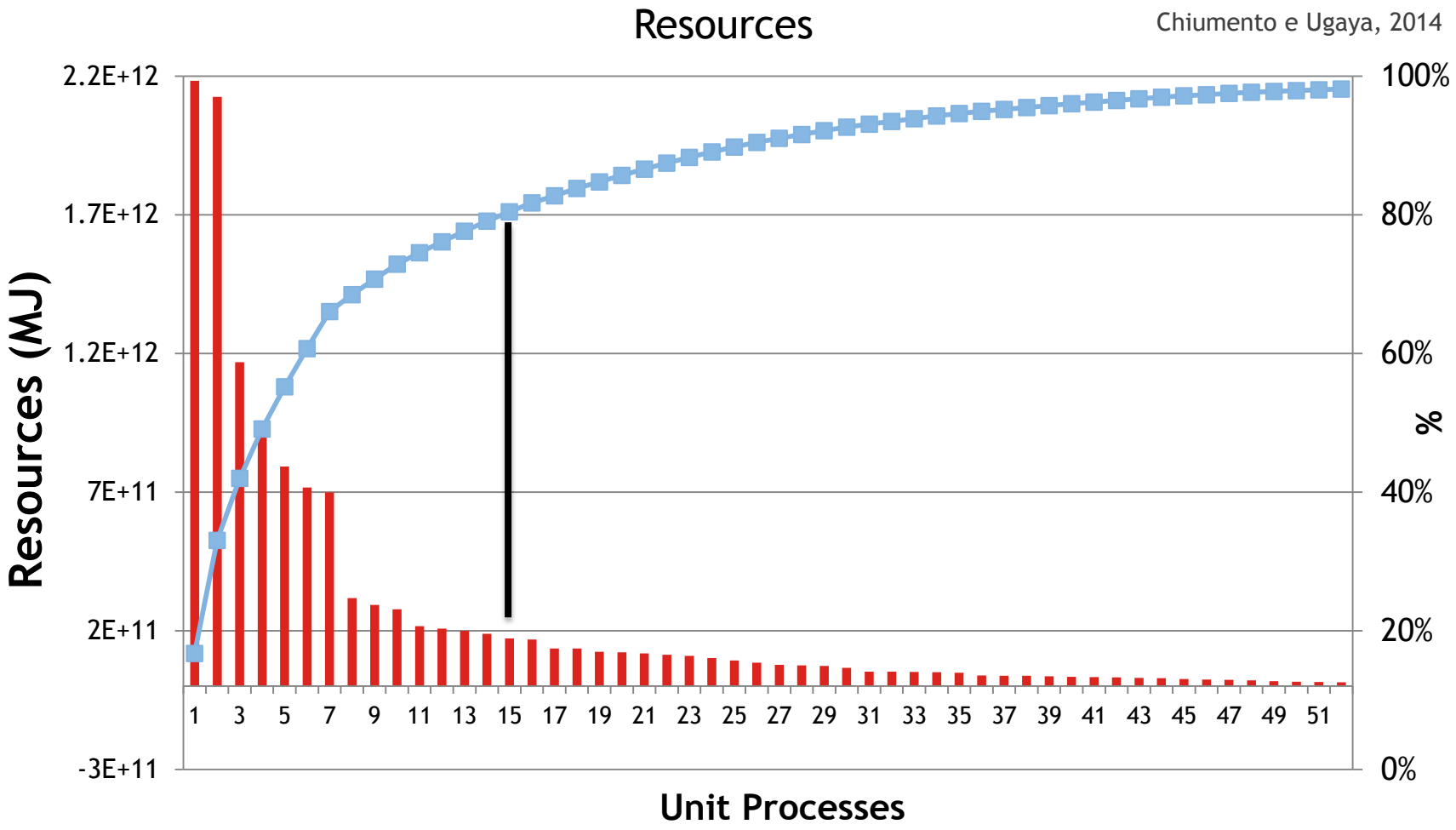
Results: Ecosystem Quality



Results: Climate Change



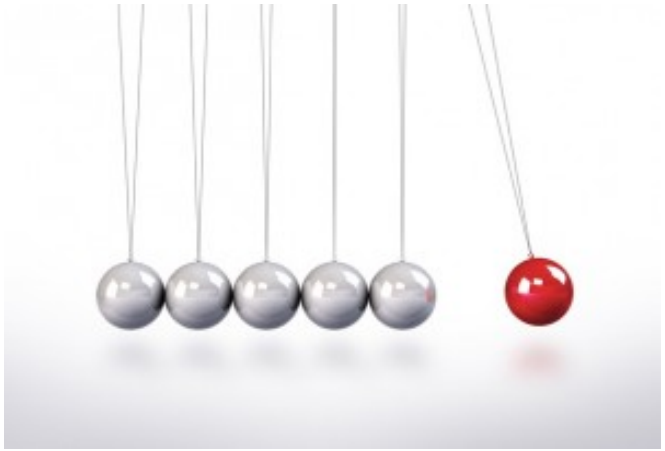
Results: Resources



Overall Results

ecoinvent

- Few products are responsible for the impact assessment results in each category evaluated (and are listed in the next slide)



Specific Results

ecoinvent

Top 15 unit processes
sugarcane//[BR] sugarcane production
sugarcane//[BR] sugarcane production, on land recently transformed
clinker//[RoW] clinker production
electricity, high voltage//[CN] electricity production, hard coal
electricity, high voltage//[BR] electricity production, hydro, reservoir, tropical region
pig iron//[GLO] pig iron production
cattle for slaughtering, live weight//[RoW] milk production, from cow
grass silage, Swiss integrated production//[CH] grass silage production, Swiss integrated production, intensive
land tenure, arable land, measured as carbon net primary productivity//[BR] clear-cutting, shrubland to arable land
land tenure, arable land, measured as carbon net primary productivity//[BR] clear-cutting, primary forest to arable land
digester sludge//[GLO] treatment of digester sludge, municipal incineration
heat, district or industrial, other than natural gas//[RoW] refinery gas, burned in furnace
heat, district or industrial, other than natural gas//[RoW] heat production, at hard coal industrial furnace 1-10MW
electricity, high voltage//[BR] electricity production, lignite
soybean//[BR] soybean production, on land recently transformed

Reference: Chiumento e Ugaya, Identificação de processos elementares prioritários para adaptação de Base de Dados de Inventário do Ciclo de Vida: estudo de caso para o Brasil. Anais do IV CBGCV. 2014. 10p.

Part 6 - Regionalizing datasets of processes

Content from Andreas Ciroth, GreenDelta and Amir Safaei, ecoinvent

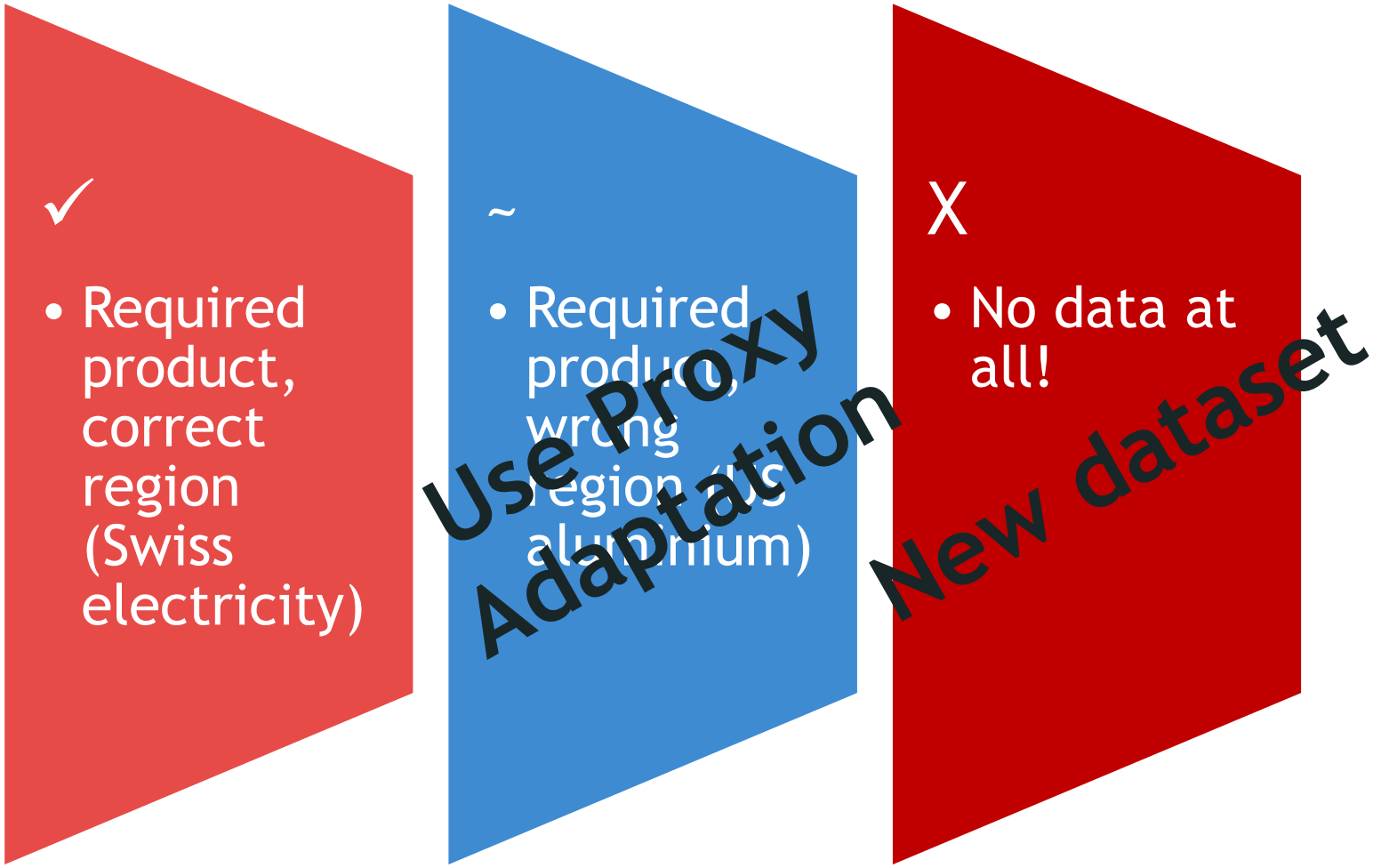
August 2017 Version

Managed by SETAC

August 2017 Version

How to regionalize: Start with existing!

ecoinvent



How to regionalize: Start with existing!

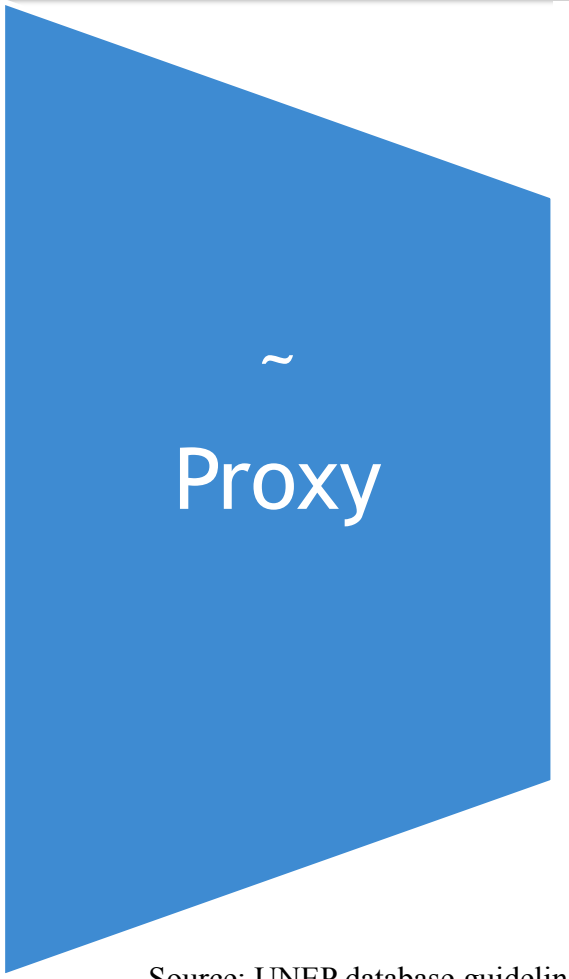
ecoinvent

Search:
Databases (link)
Literature

Ask:
LCA Discussion List

How to regionalize: Start with existing and adapt

ecoinvent



Source: UNEP database guidelines

Proxy Datasets

Existing similar (proxy) unit process (e.g. a related technology or the same technology for another region or another time period).

1) one can reasonably assume that the values indeed would be the same or very similar,
2) the flow is not contributing much to the impacts
3) no other data sources are available ("better than nothing" principle).

How to regionalize: Start with existing!

ecoinvent

~

Adaption of existing datasets

Example adaptation
only possible with unit processes!

1. What exists?

2. Identify what is important
(sensitivity/contribution analysis)

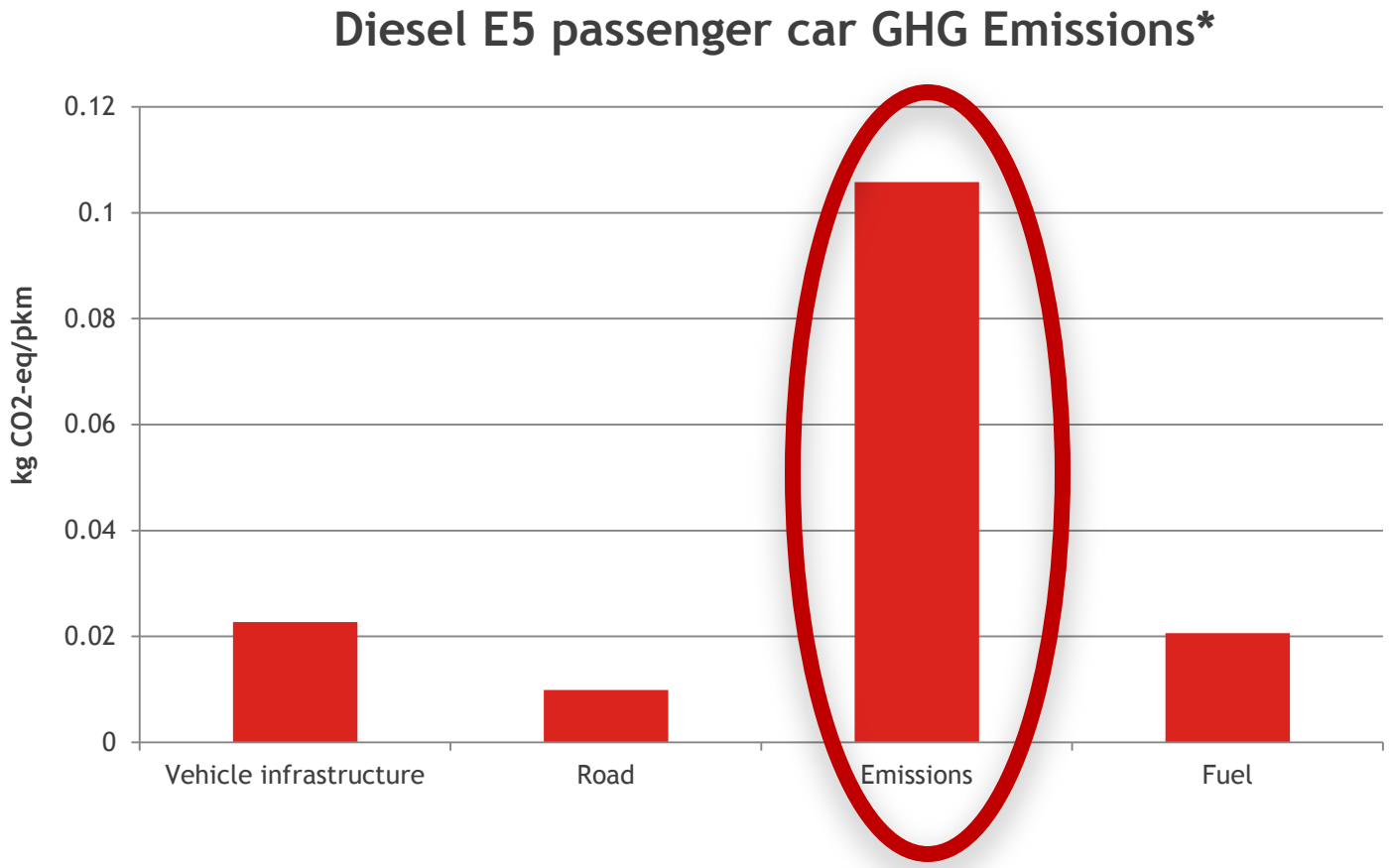
3. Adapt accordingly
(increase data quality)

Transport exchanges

ecoinvent

- Infrastructure
  - Vehicle (car, truck, train, boat, plane, etc.)
  - Road infrastructure
- Fuel input
- Use phase emissions
  - Fuel related (Co2, Sulphur, etc.)
  - Norm-regulated related (Euro 5, Euro 4, etc.)

Transport: relevant aspects (trends)



Adaptation of the use phase (emissions + fuel)

ecoinvent

- **Available data: Same technology / fuel**
  - Consumption adaptation?
    - Adaptation of fuel consumption dependent emissions (CO_2, Sulfur, etc.)
  - Same technology – different norm
    - Does an equivalence exist between the available Euro-Norms and other national standards?
    - Can differences in the norms be used to scale available data?
    - Where and how is the original energy feedstock produced (oil/natural gas/…)?

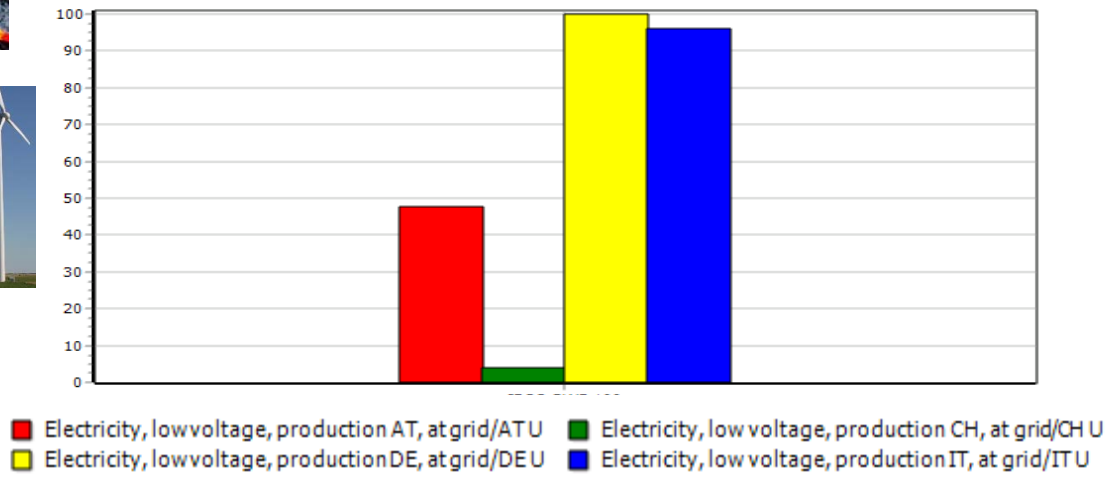
- **Available data: Different technology / fuel**
  - Simply changing the fuel type can lead to large deviations.
  - Natural gas/ biogas: ~97% of gas content is methane. Changing the CO_2 and CH_4 emissions to/ from biogenic can give a good first idea.

Electricity

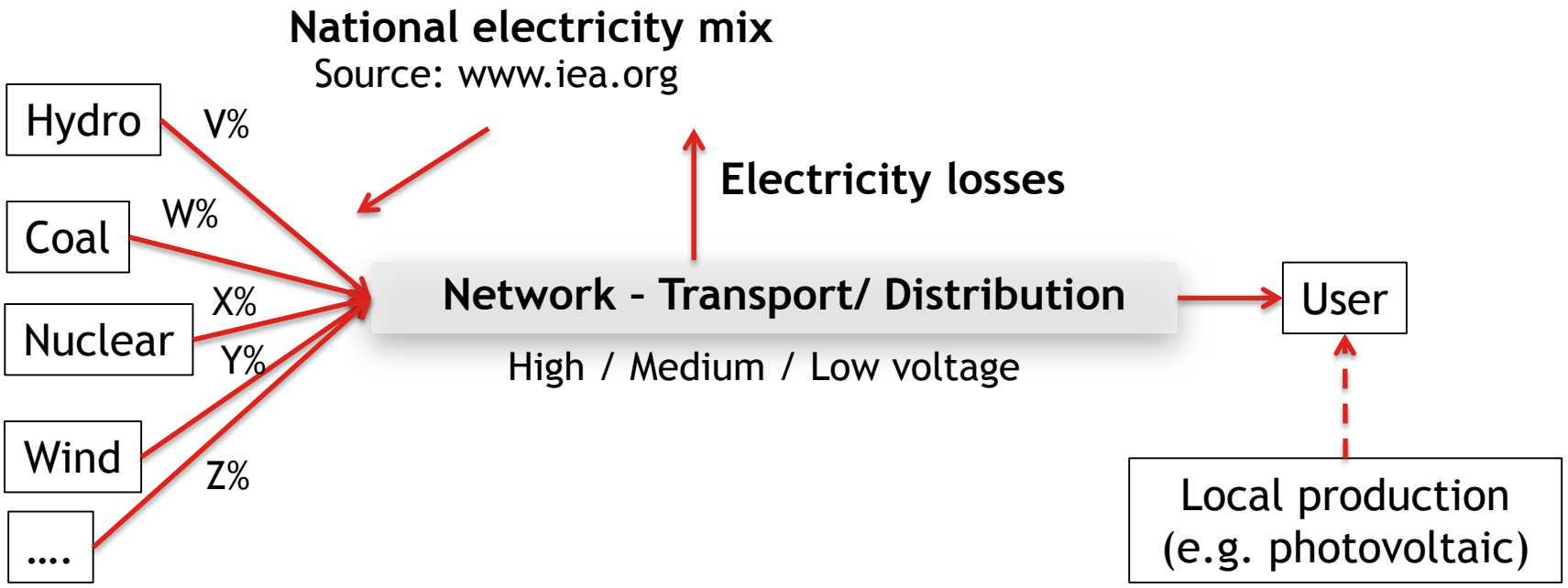
ecoinvent

- Energy related processes are present in almost any process
- Can be a key contributor to the impacts
- Strong regional variability

Sensitivity: mainly depending on generation technology (example GHG emissions)



How to adapt



Regionalizing the main emissions

Minimum requirements for electricity generation technologies

	hard coal	lignite	natural gas	oil	hydro	nuclear	wind	solar
exchange amount_carbon dioxide, fossil	x	x	x	x				
exchange amount_carbon monoxide fossil	x	x	x					
exchange amount_methane fossil	x	x	x					
exchange amount_nitrogen oxide	x	x	x					
exchange amount_NMVOC	x	x						
exchange amount_NOx retained	x	x	x					
exchange amount_PM<2.5	x	x	x					
exchange amount_PM>10	x	x	x					
exchange amount_PM>2.5<10	x	x	x					
exchange amount_SOx reatained	x	x	x					
exchange amount_sulfur dioxide	x	x	x					
land use_occupation					x			
land use_transformation					x			
exchange_nuclear waste						x		
parameter_amount_efficiency	x	x	x	x	x	x	x	x
parameter_amount_gross el prod	x	x	x	x	x	x	x	x
parameter_amount_losses (gross to net)	x	x	x	x	x	x	x	x
parameter_amount_Wind load hours							x	
parameter_amount_lifetime					x			
parameter_solar yield								x

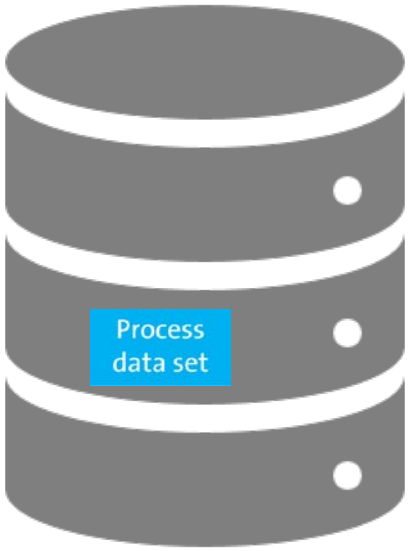
Part 7 - Storing datasets into databases

Content from Andreas Ciroth, GreenDelta

August 2017 Version

Managed by SETAC

Datasets in databases



Datasets in databases

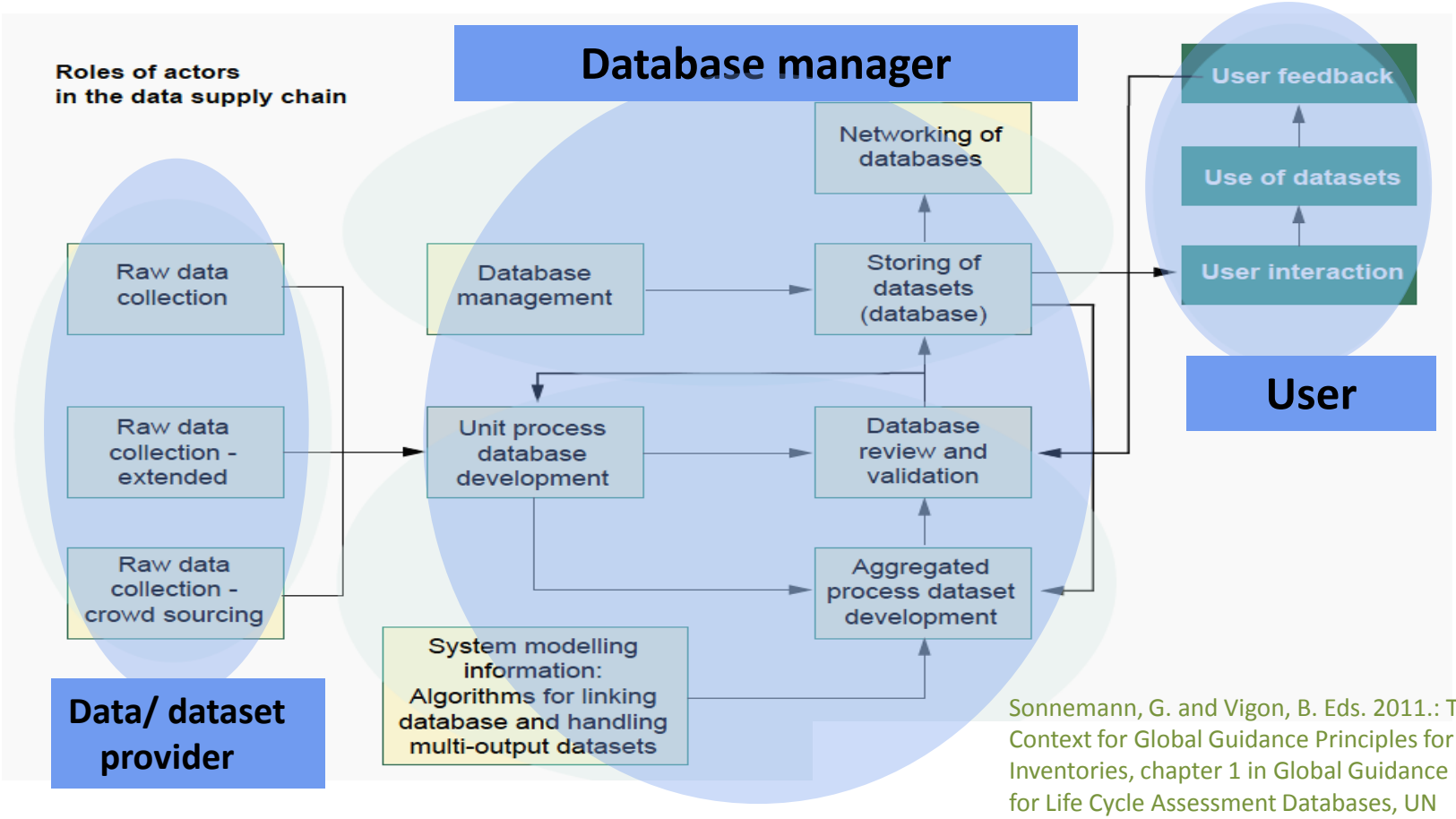
What is the issue?

LCA / LCI database characteristics

- Comprehensive input/output flows for datasets
- Consistent modeling approach and flow nomenclature
- Complete and consistent dataset documentation
- Often, capability beyond storing data, e.g. data collection and submittal support, review, aggregated process creation, …

Datasets in databases

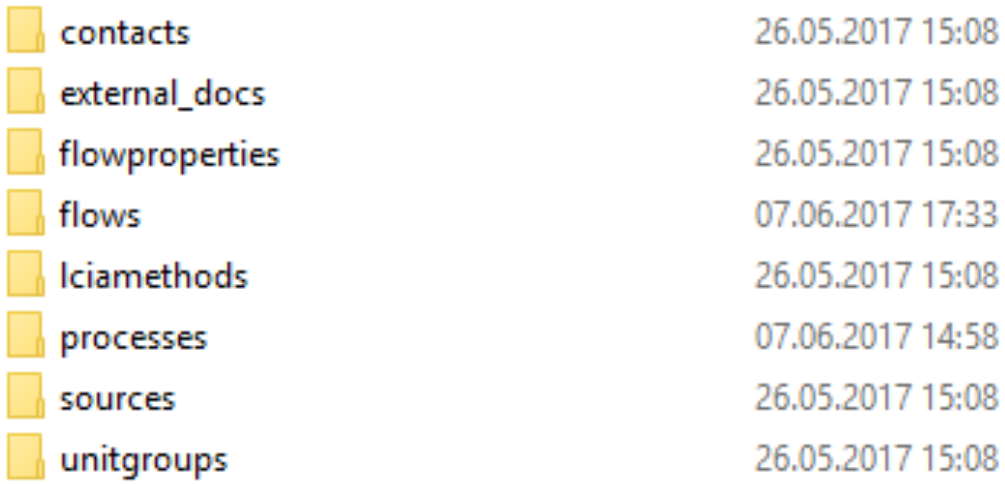
The whole picture:
data provider, DB manager, user



Sonnemann, G. and Vigon, B. Eds. 2011.: The Context for Global Guidance Principles for Life Cycle Inventories, chapter 1 in Global Guidance Principles for Life Cycle Assessment Databases, UN Environment Programme, ISBN: 978-92-807-3174-3, DTI/1410/PA.

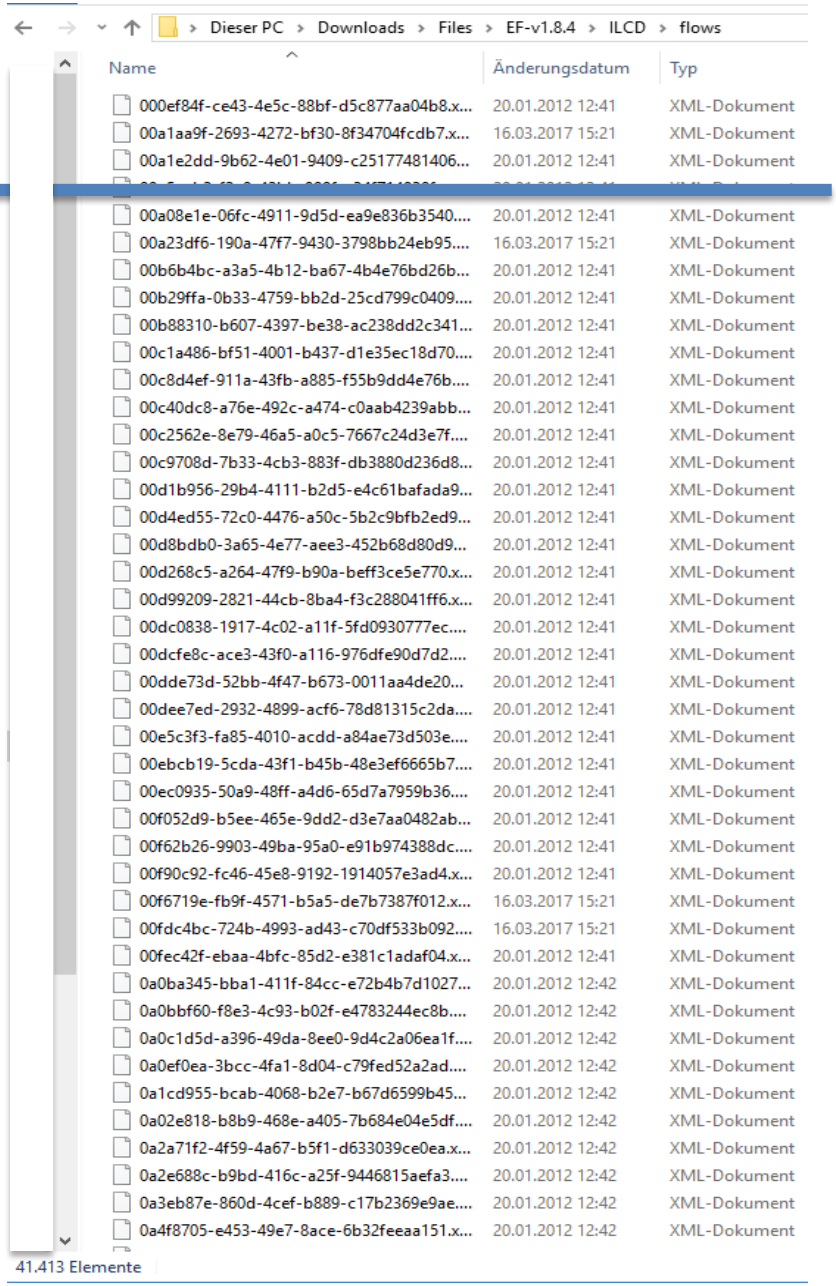
Datasets in databases

Common reference data: ILCD (LCDN, ELCD – JRC)

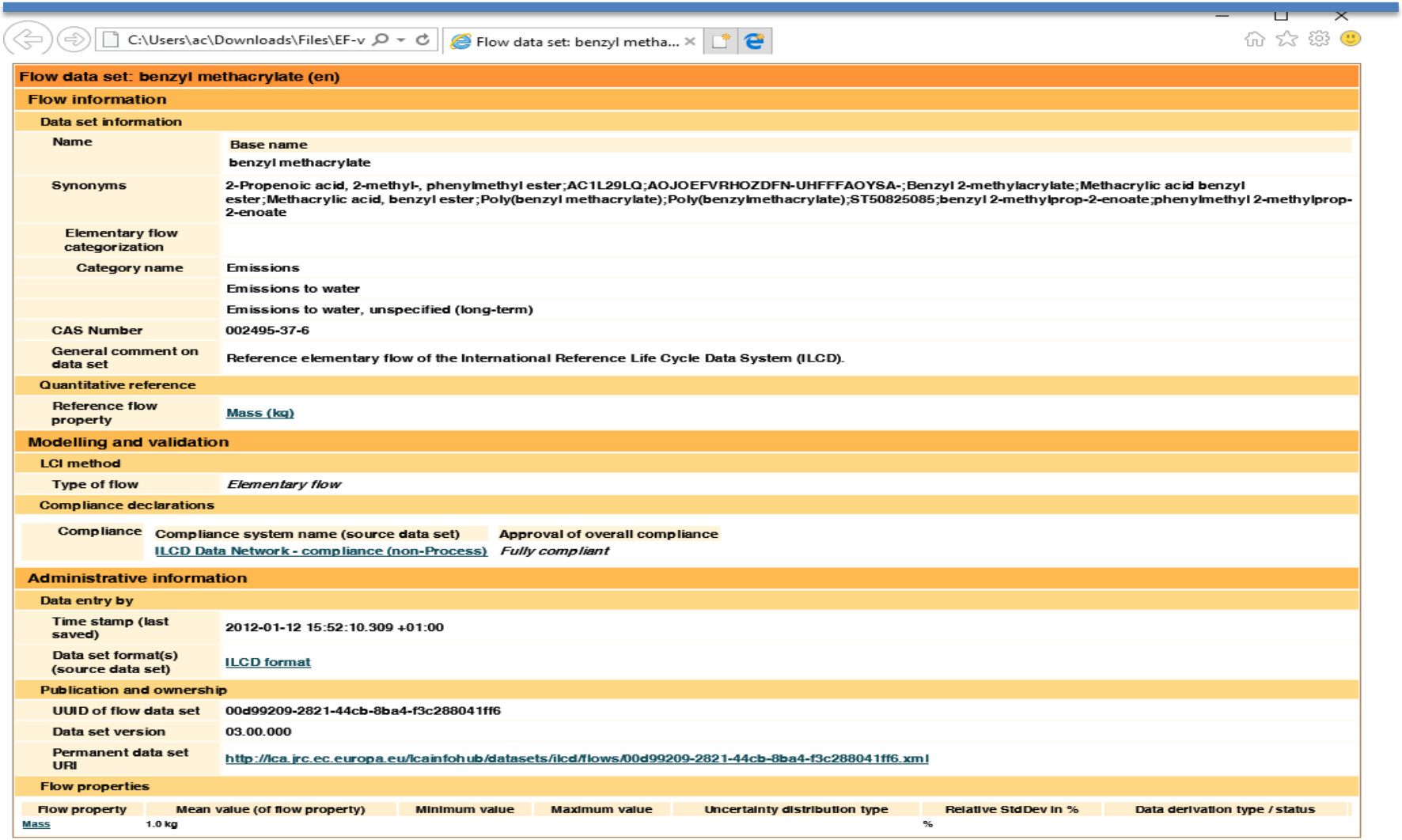


Datasets in databases

Common reference data: ILCD (LCDN, ELCD – JRC)



Datasets in databases



C:\Users\ac\Downloads\Files\EF-v | Flow data set: benzyl metha...

Flow data set: benzyl methacrylate (en)

Flow information

Data set information

Name	Base name benzyl methacrylate
Synonyms	2-Propenoic acid, 2-methyl-, phenylmethyl ester;AC1L29LQ;AOJOEFVRHOZDFN-UHFFFAOYSA-;Benzyl 2-methylacrylate;Methacrylic acid benzyl ester;Methacrylic acid, benzyl ester;Poly(benzyl methacrylate);Poly(benzylmethacrylate);ST50825085;benzyl 2-methylprop-2-enoate;phenylmethyl 2-methylprop-2-enoate
Elementary flow categorization	
Category name	Emissions
	Emissions to water
	Emissions to water, unspecified (long-term)
CAS Number	002495-37-6
General comment on data set	Reference elementary flow of the International Reference Life Cycle Data System (ILCD).

Quantitative reference

Reference flow property	Mass (kg)

Modelling and validation

LCI method

Type of flow	*Elementary flow*

Compliance declarations

Compliance	Compliance system name (source data set)	Approval of overall compliance
	ILCD Data Network - compliance (non-Process)	*Fully compliant*

Administrative information

Data entry by

Time stamp (last saved)	2012-01-12 15:52:10.309 +01:00
Data set format(s) (source data set)	ILCD format

Publication and ownership

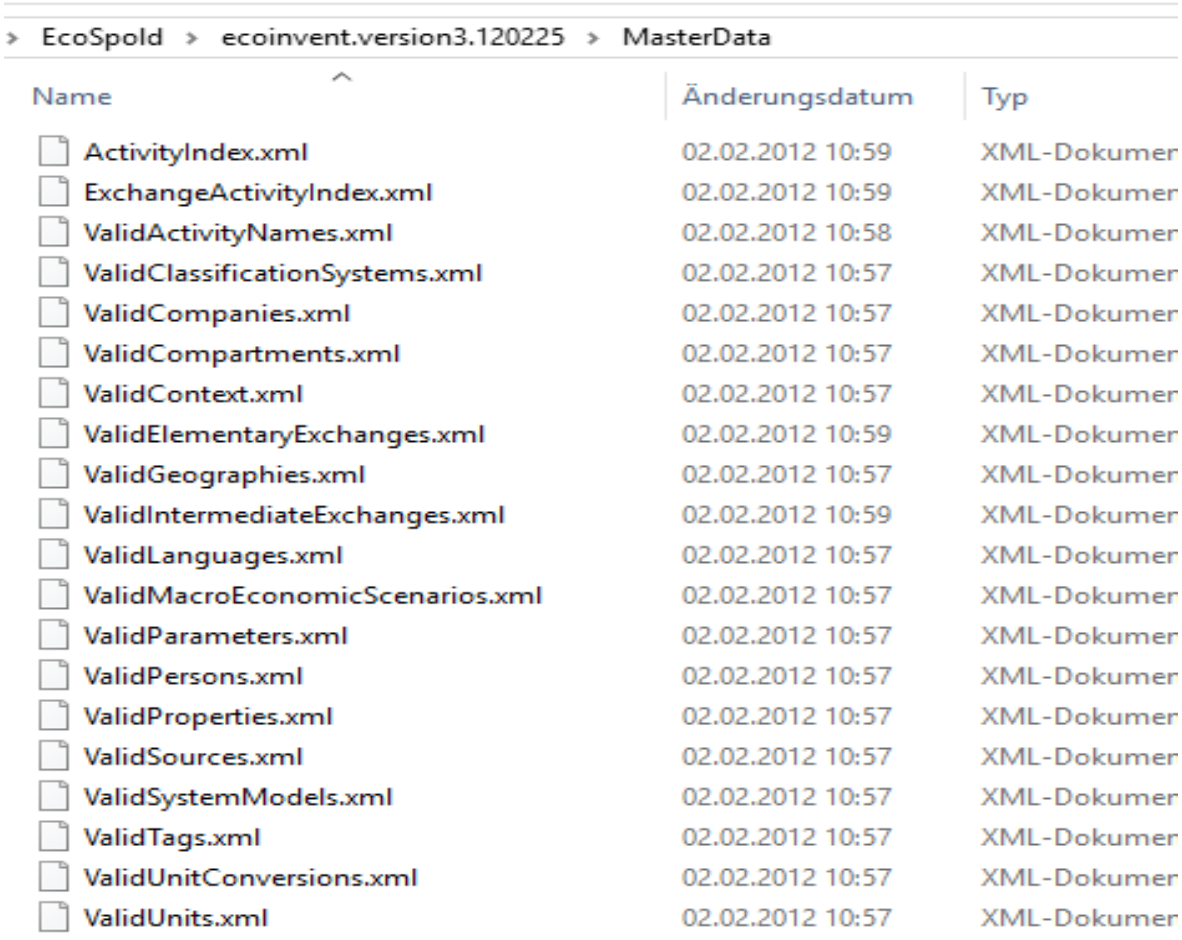
UUID of flow data set	00d99209-2821-44cb-8ba4-f3c288041ff6
Data set version	03.00.000
Permanent data set URI	http://lca.jrc.ec.europa.eu/lcainfohub/datasets/ilcd/flows/00d99209-2821-44cb-8ba4-f3c288041ff6.xml

Flow properties

Flow property	Mean value (of flow property)	Minimum value	Maximum value	Uncertainty distribution type	Relative StdDev in %	Data derivation type / status
Mass	1.0 kg				%	

Datasets in databases

Common reference data: ecoinvent



Datasets in databases

Format of the database

- In the past, long discussions in various emerging databases, e.g. Brazil, US, … about the “right” format
- Discussions are mainly about the exchange format (-> next slot; some few examples, from ecoinvent, JRC/EC, GreenDelta
- For the actual storage, the format is not so important but DB should be able to “deliver” in various broadly used formats
- Format in use for storage:
  - LCDN / EC: ILCD (only!)
  - USDA: openLCA
  - Ecoinvent: unknown

Datasets in databases

Review and QA

- Review “items” (what is reviewed, how)
- Reviewers
- Review workflow and organization

Datasets in databases

Review and QA

- Review “items” (what is reviewed, how)
- Reviewers
- Review workflow and organization
- On which type of processes should the review be performed (unit process, aggregated process?)

Datasets in databases

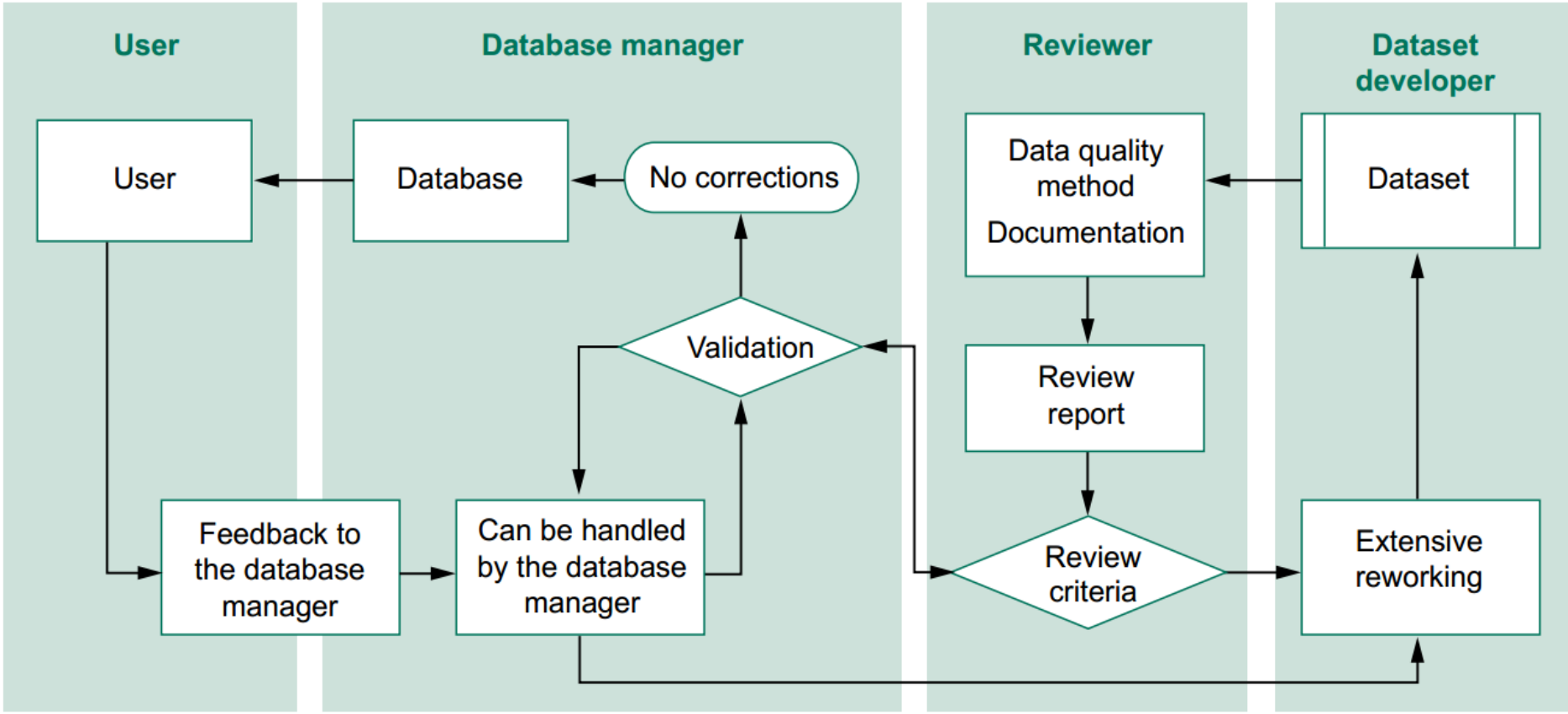
Review criteria

	Values	Scope		
		data set	flows/ exchanges	any other data set fields
Goal and scope completeness				
Reference time	yes/no	x		
Reference geography	yes/no	x		
Reference technology	yes/no	x		
Reference model completeness	yes/no	x		
Reference sample completeness	yes/no	x		
Sample approach (scientific, or expert-based)	scientic / expert	x		
Supported LCIA methods with version number	text	x		
Conformance				
Time related conformance	1..5		x	x
Geographical conformance	1..5		x	x
Technological conformance	1..5		x	x
Model completeness conformance, flows and documentation	1..5	x		
Sample conformance, correctness and reliability				
Sample conformance	1..5	x		
Accuracy of the provided information	1..5	x		
Precision of the provided information	1..4		x	
Reliability of the provided information	1..5		x	x
Consistency of the provided information	1..5		x	x
Materiality				
Mass- and energy balance in line with goal and scope	1..5	x		
LCIA results in line with goal and scope	1..5	x		
Order of 5 main drivers for main LCIA results in line with goal and scope	1..5	x		
Procedural and meta-information				
# of reviewers and their relation to data provider	1..5	x		
Data access	1, 3, 5	x		

Ciroth, A., , C. Foster,, J., Hildenbrand, and A., Zamagni, 2016.:Life Cycle Inventory Dataset Review Criteria Development, final report

Datasets in databases

Review procedure, example



Inaba, A., et al.: 2011. Data Documentation, review, and management, Chapter 4 in Sonnemann and Vigon (eds.) Global Guidance Principles for Life Cycle Assessment Databases, UN Environment Programme, ISBN: 978-92-807-3174-3, DTI/1410/PA.

Datasets in databases

Review, software support

- EcoEditor (fully integrated)
- ILCD validation tool (ILCD, for nomenclature, documentation completeness, … - not content)
- openLCA Collaboration Server (fully integrated)

(maybe more are available)

Datasets in databases

Version Control and Responsible Management Practices

- Typically, database users expect stability in their results
- ..but they also expect updates, bug fixes, extensions of the database

Datasets in databases

Version Control and Responsible Management Practices

- Typically, database users expect stability in their results
- ..but they also expect updates, bug fixes, extensions of the database
- Common practice:
  - Release versions
    - Ecoinvent: major, smaller release (3.3)
    - GaBi: Annual releases with mid-year update
  - Dataset versioning (e.g. ILCD hubs): each dataset has a version number and can be chosen accordingly

Datasets in databases

Common reference data
Format / storage
Review & QA

- Decisions necessary for creating, and especially for operating a database
- However, avoid “lock-in” – do not build on a solution that is bound to one specific format, provider, application
- Invent something new (new flow list, format, …) only where really needed

Datasets in databases

Using MS Excel for data collection and for building the database

+ Excel: commonly used and understood; fewer training necessary; flexible
- Excel: very flexible; but inconsistencies and mistakes possible
- Excel: does not follow the LCA object model (obviously); unit/flow/process

Part 8 - Linking databases and LCA software, and alternative dataset development methods

Content from Andreas Ciroth, GreenDelta and Amir Safaei, ecoinvent

August 2017 Version

Managed by SETAC

Linking databases and LCA software

LCA Software

- GaBi
- SimaPro
- openLCA
- Umberto

Linking databases and LCA software

LCA Software

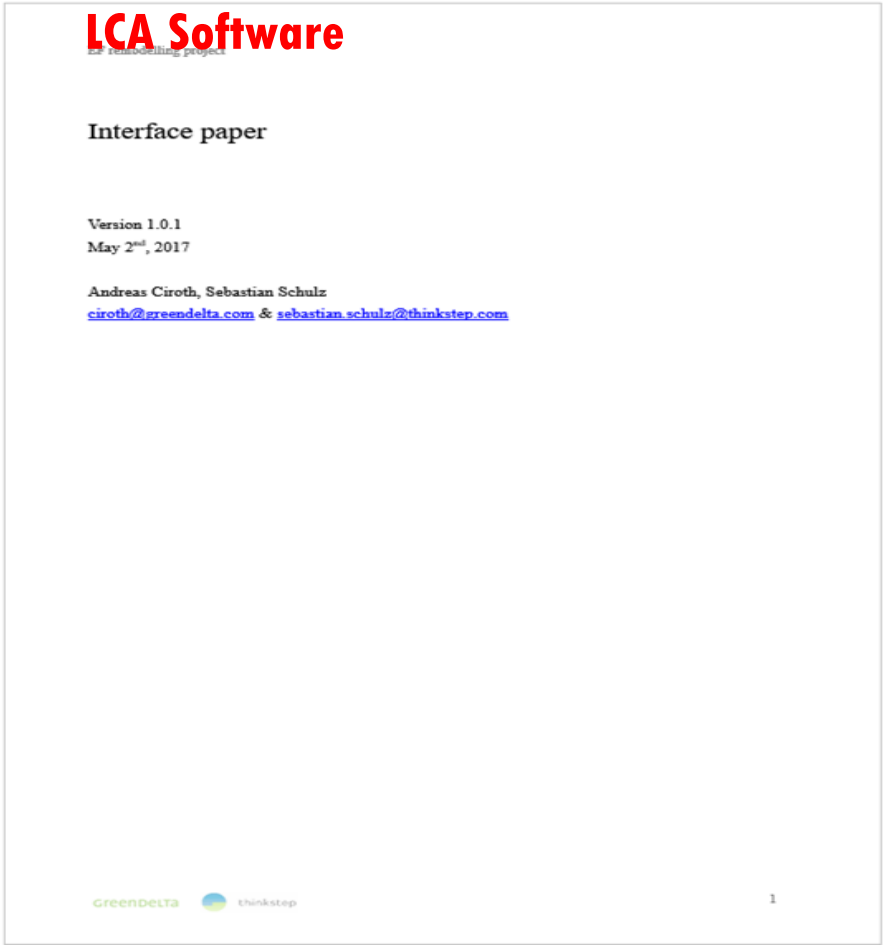
- GaBi
- SimaPro
- openLCA
- Umberto
- Freeware tools
- Specific tools for sectors

Linking databases and LCA software

LCA Software

- There is (of course) a common core of features and models that all main LCA software systems support, however:
- Differences:
  - Supported import and export formats
  - Modeling concepts and modeling features
    - Parameters
    - How processes are linked
    - …
  - Open source vs closed source tools

Linking databases and LCA software



LCA Software

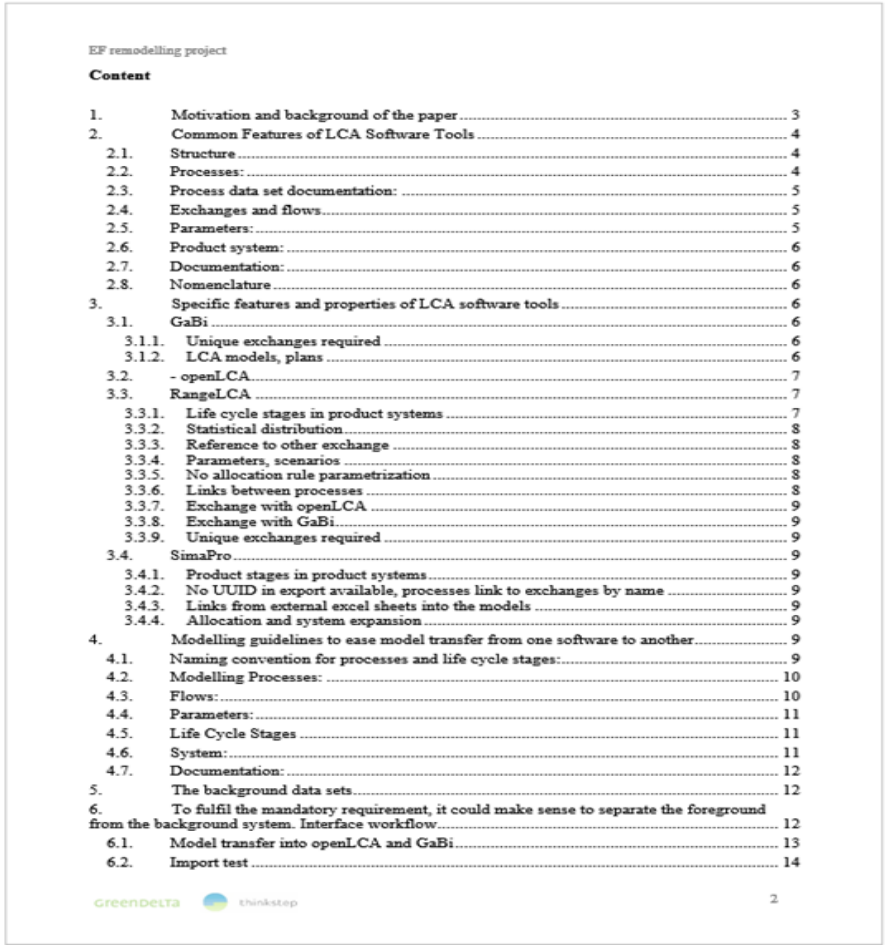
EF remodelling project

Interface paper

Version 1.0.1
May 2nd, 2017

Andreas Ciroth, Sebastian Schulz
ciroth@greendelta.com & sebastian.schulz@thinkstep.com

GreenDelta thinkstep 1



EF remodelling project

Content

GreenDelta thinkstep 2

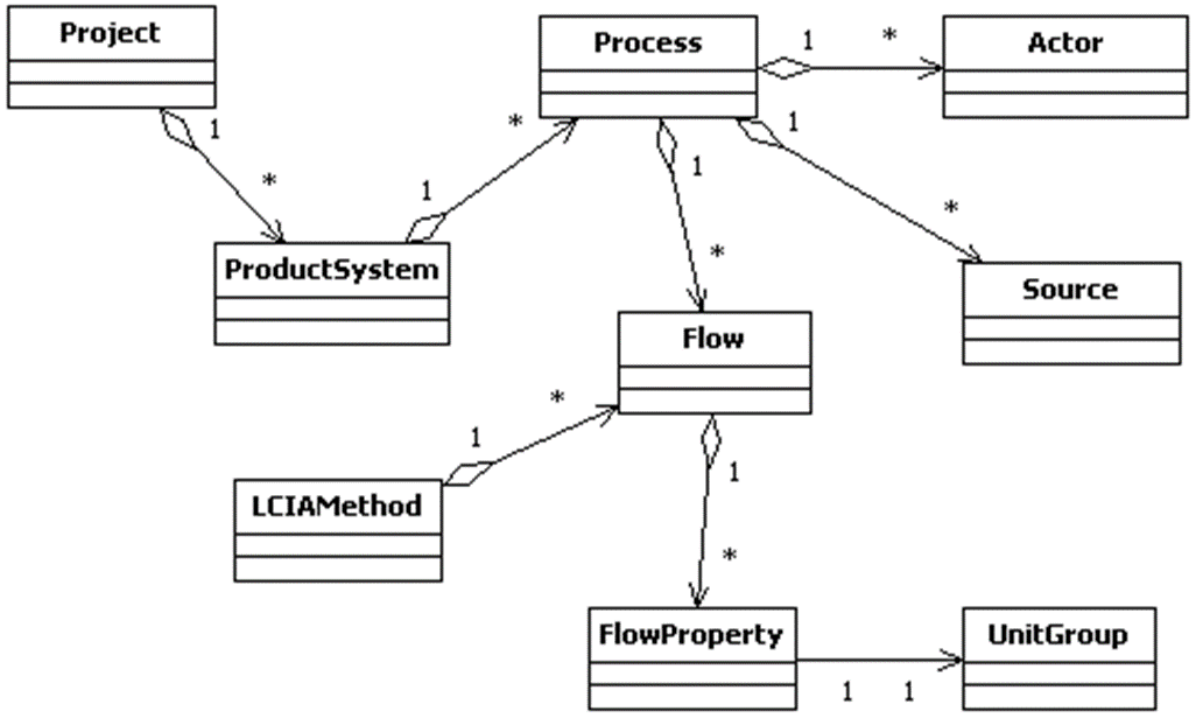
Linking databases and LCA software

Data formats?

- Define how data is stored and how data is exchanged between different applications
- Usually interesting for users (if at all): data ***exchange formats***, also in LCA

Data structures in databases

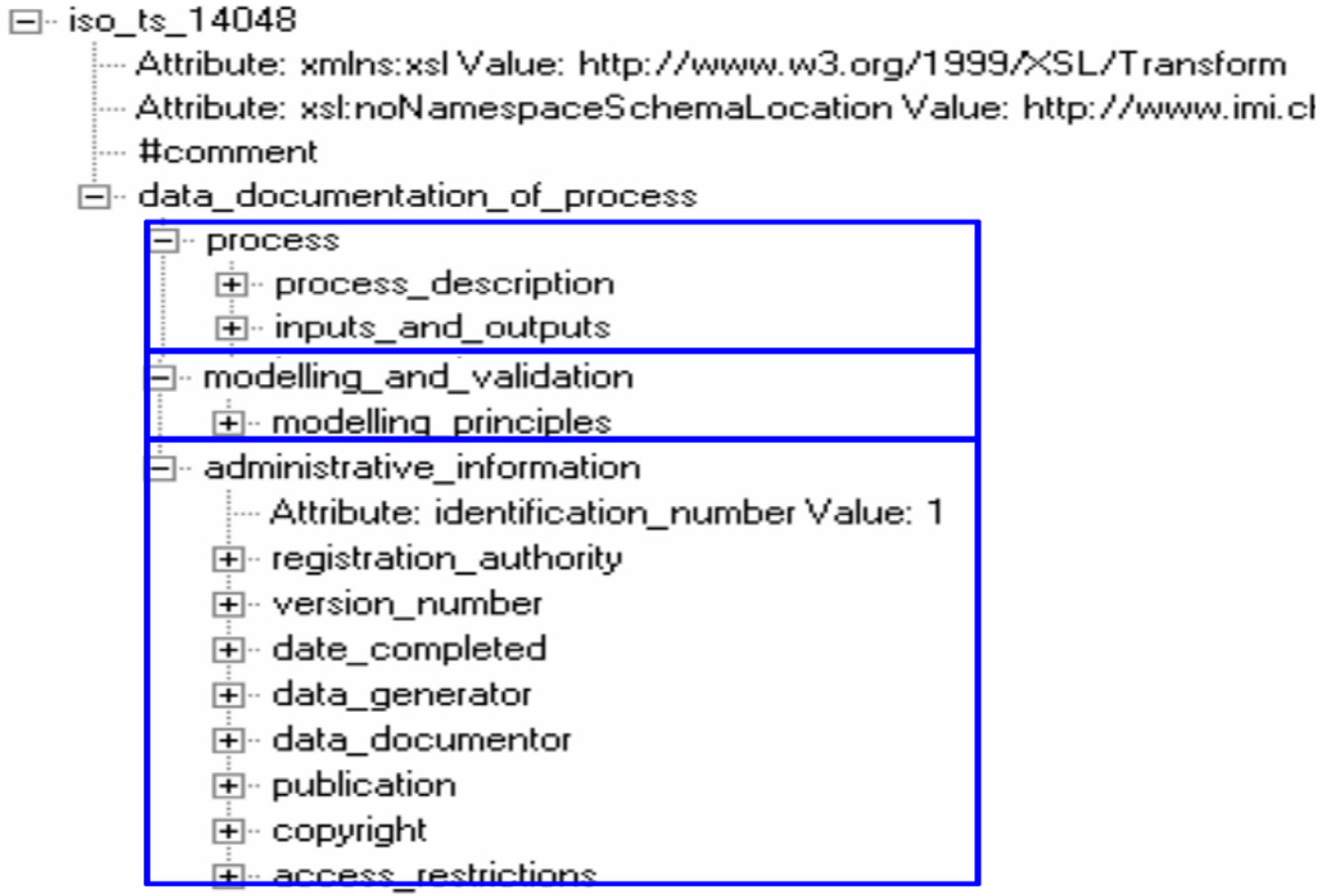
(Basic structure of early openLCA, as UML diagram → many of these elements are present in LCA data formats)



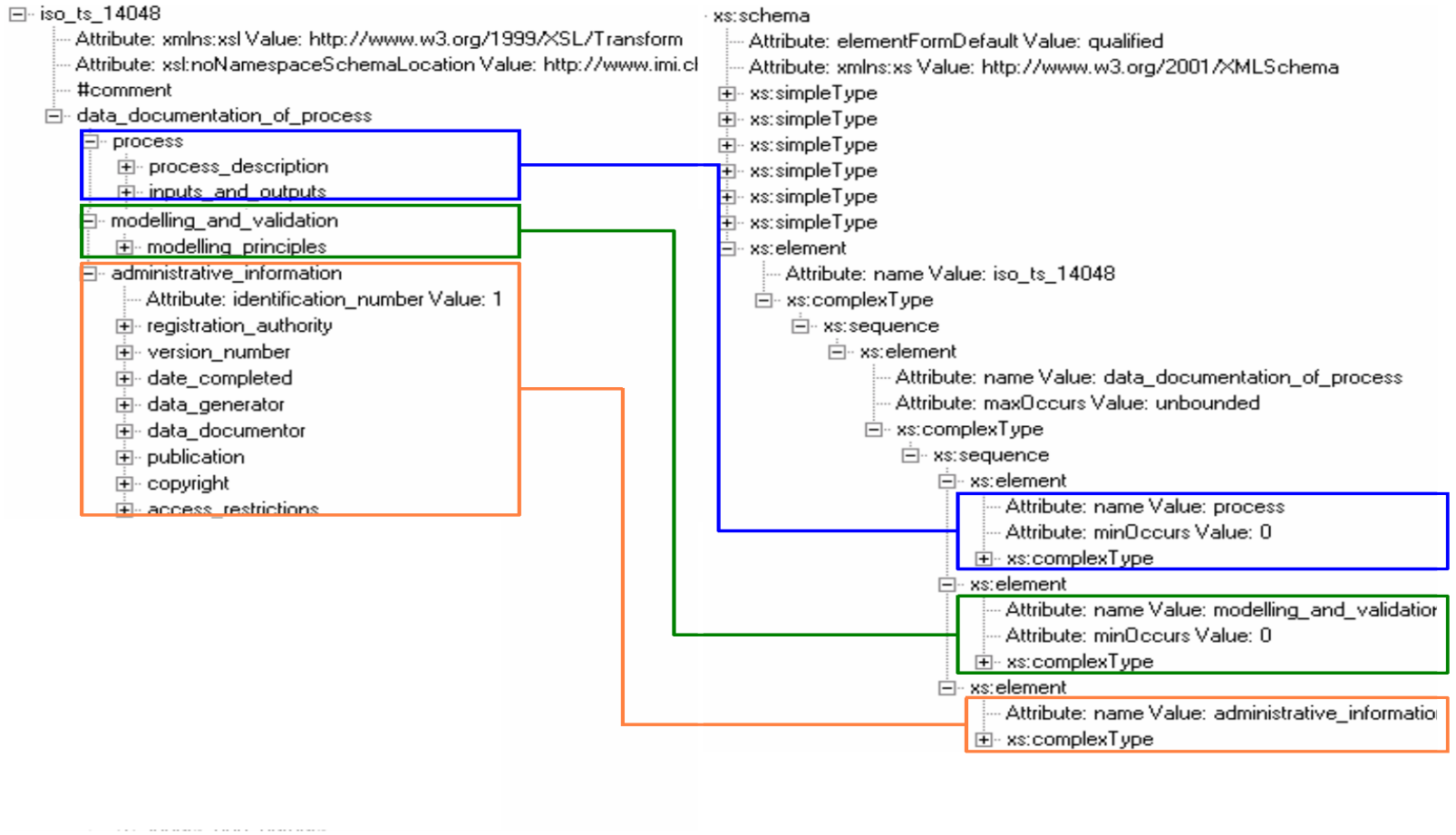
Data formats: ISO14048/TS

- Precursor of all modern LCA data formats
- Too generic to allow for smooth conversion between different, ISO14048-compliant LCA data formats

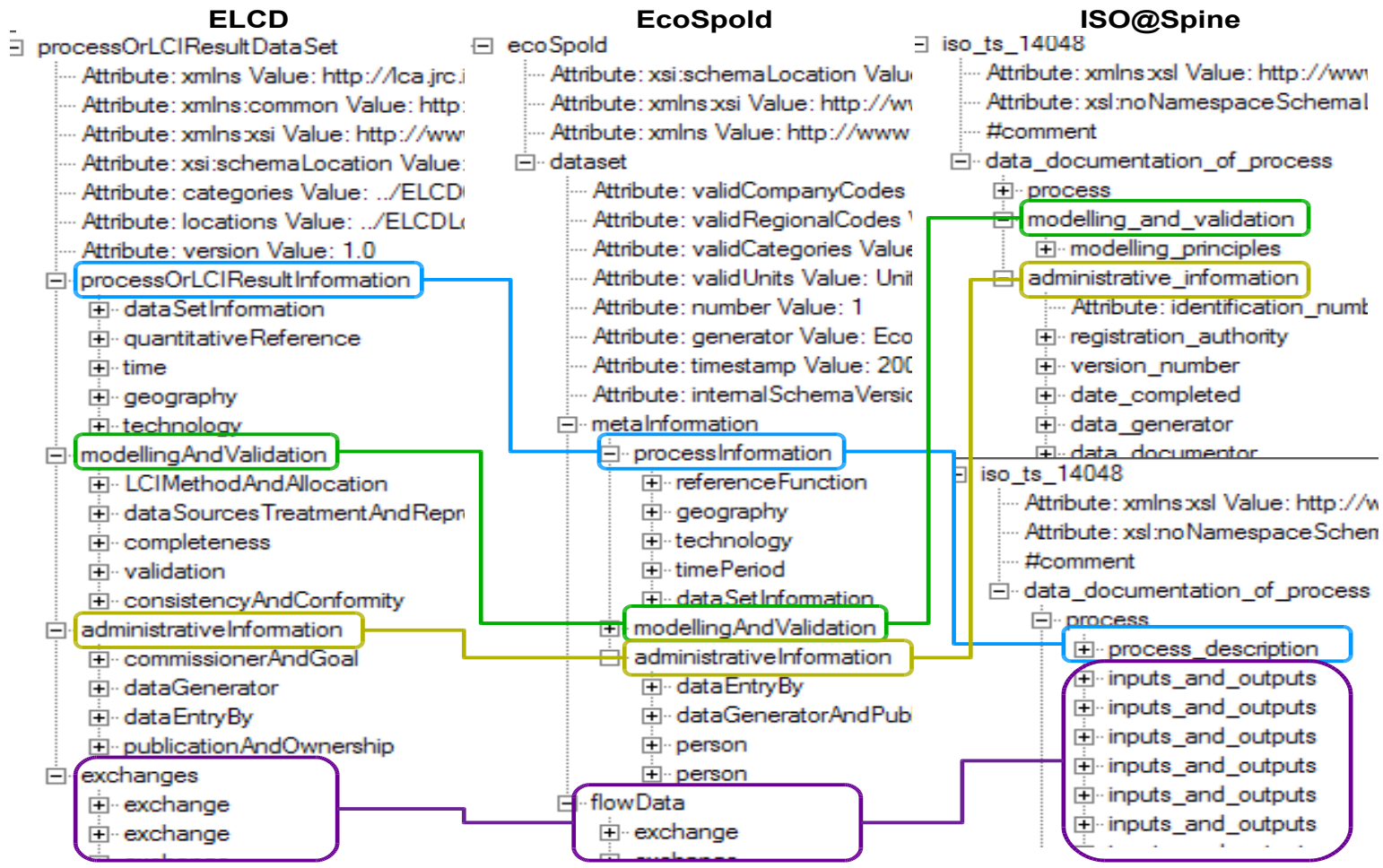
ISO14048 compliant data formats: ISO@Spine
Basic data set structure [Carlson, Palsson 2002]



ISO14048 compliant data formats: ISO@Spine
relation ISO14048 structure – ISO@Spine structure



ISO-14048 compliant data formats: relations



ISO-14048 compliant data formats: EcoSpold v.1

- **ecoinvent database,** broadly supported as exchange format

EcoSpold (v1) limitations

- No support for parameters in data sets
- Only two languages
- No (real) distinction betweeen process and flow/product

ISO-14048 compliant data formats: EcoSpold 2

ecoinvent database v3

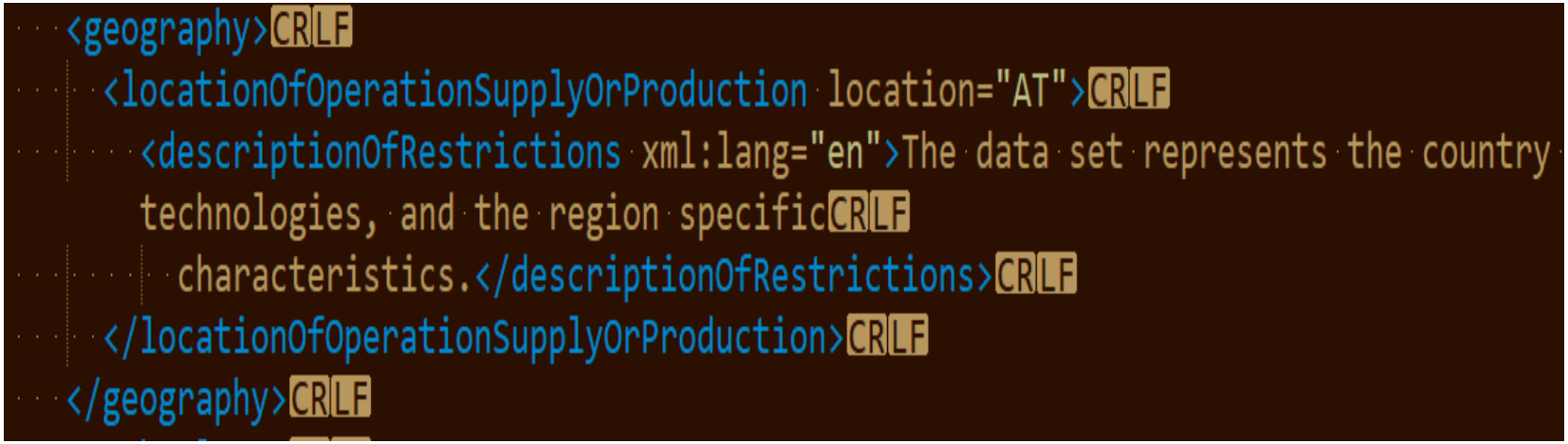
- Overcomes all limitations mentioned in previous slide
- In principle better alignment with ILCD data format (→ next slides)
- BUT adds new complexity (built „intelligence“ into the data format, parent-child processes)
- openLCA the only LCA software which has implemented EcoSpold2 so far
- Format used by ecoinvent 3 & ecoeditor is different from the public format and not documented (!)

ISO-14048 compliant data formats: ILCD formerly ELCD, by JRC, European commission

- Overcomes all limitations mentioned in ecoinvent 1 slide (parameters, multi-language, process/product)
- BUT is rather administrative (author for many single entries; many different languages possible in one data set, …)
- One data set consists of literally > 10,000 single xml files that are organised in a file structure, with one file missing the whole process data set is invalid (ELCD III example)
- GaBi provides good support for this format, openLCA too

All formats so far: XML formats

All formats so far: XML formats

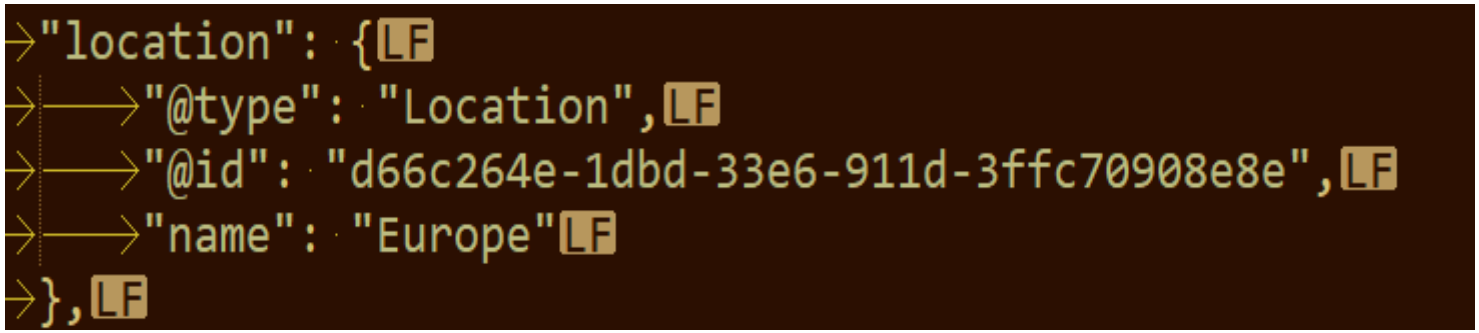


Recent Development: JSON-LD "2LD" data format

- ISO 14048/TS compliant but not an XML format
- Developed by GreenDelta for US EPA, in openLCA
- Motivation:
  - XML is inefficient, „too wordy", for larger amounts of data
  - Link to ontologies for data analyses and data creation
- JSON-LD: Developed /pushed by google, yahoo, IBM < 5 years ago -> optimised for not too well organised data, huge amounts of data

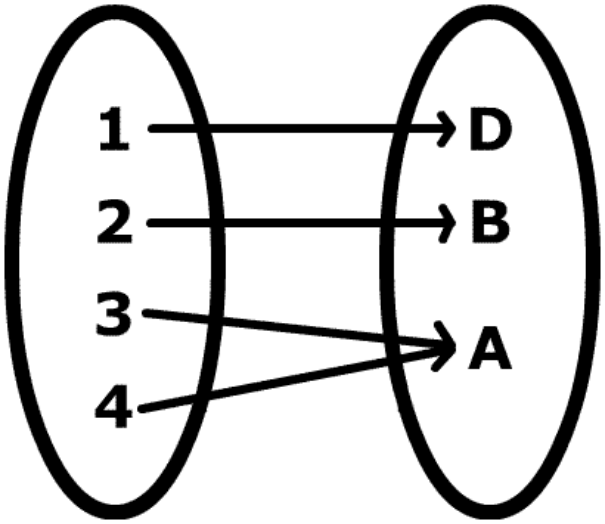
Recent Development: JSON-LD "2LD" data format

- http://greendelta.github.io/olca-schema/
- E.g., location:

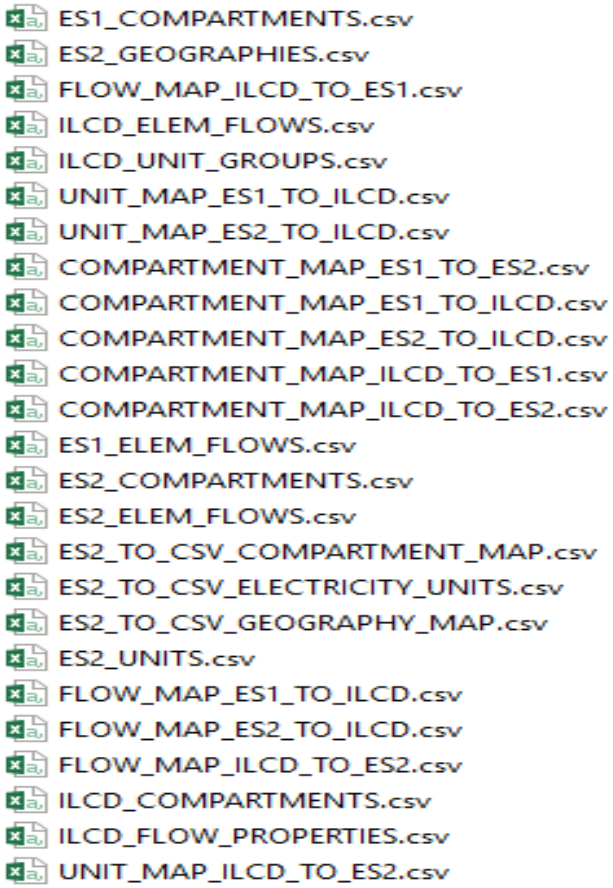


Format conversion

- Task: converting LCA data sets from one format to the other
  - Converting the data format fields
  - Converting the nomenclature systems: „mapping“ of reference data



Format conversion: mapping files



Format conversion: mapping files

FLOW_MAP_ILCD_TO_ES2.csv

```
"041f5cf4-6556-11dd-ad8b-0800200c9a66","0db53406-fb79-497d-a9e7-fc2c8d425b81",true,1
"041f5cf5-6556-11dd-ad8b-0800200c9a66","f228fcae-8e19-4a73-9d8c-60ec8adaea50",true,1
"041f5cf7-6556-11dd-ad8b-0800200c9a66","b066708e-5bab-45ee-9244-e1be2fa5b5f3",true,1
"041f5cf9-6556-11dd-ad8b-0800200c9a66","ef7bc7b8-f85b-4687-9cb1-eb60384650fa",true,1
"041f5cfa-6556-11dd-ad8b-0800200c9a66","651bd1c1-458e-4585-94e1-46409196de89",true,1
"041f5cfc-6556-11dd-ad8b-0800200c9a66","84cd61aa-92a0-40d1-aad5-e9e5c0d4c93e",true,1
"041f5d12-6556-11dd-ad8b-0800200c9a66","4aecbaff-32ae-4655-9587-db9337d0d350",false,1
"041f8403-6556-11dd-ad8b-0800200c9a66","a8a2f53b-c6d2-4302-a59f-502bd7e242fc",false,1
"041fab15-6556-11dd-ad8b-0800200c9a66","9cda96ce-e1cf-4192-8614-52be32d200f4",true,1
"041fd212-6556-11dd-ad8b-0800200c9a66","e4ff4151-febe-4c3f-bfaf-924d6f7bf101",false,1
"041fd214-6556-11dd-ad8b-0800200c9a66","bb656df7-0d8a-427c-b58c-02962e50df57",false,1
"041fd220-6556-11dd-ad8b-0800200c9a66","745503c2-768d-4f2c-b63c-bb75c1c5419c",false,1
"041fd248-6556-11dd-ad8b-0800200c9a66","1f5330ac-e81c-45f2-8c9c-335b531faab0",true,1
"041ff923-6556-11dd-ad8b-0800200c9a66","1f5330ac-e81c-45f2-8c9c-335b531faab0",true,1
"041ff94b-6556-11dd-ad8b-0800200c9a66","2fff3bda-1eec-4166-9969-bc9e757c8e8e",false,1
"041ff94d-6556-11dd-ad8b-0800200c9a66","909bc093-18b2-4a7e-8131-16f68eebc193",false,1
"04202031-6556-11dd-ad8b-0800200c9a66","f9727bf8-d8e3-432c-8feb-2dc43da4c37d",true,1
"04202046-6556-11dd-ad8b-0800200c9a66","8bc09c04-2190-4ee2-9ee2-ae988ccd4e0c",false,1
"04202047-6556-11dd-ad8b-0800200c9a66","c0dd7ccd-9e7a-42b3-b899-dfd18c2150ca",false,1
"04202048-6556-11dd-ad8b-0800200c9a66","a7ff17d4-d3fe-4a70-9f2e-392b34630772",false,1
"0420204a-6556-11dd-ad8b-0800200c9a66","57c71b25-4663-4fad-9167-7ce5be3e8268",false,1
"04202055-6556-11dd-ad8b-0800200c9a66","50f528ca-45e7-42d0-b399-998ea63ddabf",false,1
"0420474c-6556-11dd-ad8b-0800200c9a66","40022b1f-8906-4c54-a074-bf1a3fc9b13f",false,1
"04206e78-6556-11dd-ad8b-0800200c9a66","5eb2d548-1694-496d-b3bd-a439984fca7d",false,1
"08a91e70-3ddc-11dd-904b-0050c2490048","1aa8fd2e-9fbb-4701-8250-f4e31254b84a",false,1
"08a91e70-3ddc-11dd-9077-0050c2490048","b0cdbc78-e226-4e8b-97ff-3a557ab7ab55",false,1
"08a91e70-3ddc-11dd-9083-0050c2490048","7db70809-e30b-4633-992e-2cfe03891907",false,1
"08a91e70-3ddc-11dd-90ca-0050c2490048","863ebe6d-0f48-4c32-94df-a6369f963a2e",false,1
"08a91e70-3ddc-11dd-90d2-0050c2490048","d2885a24-b728-4832-8638-15697630eff9",false,1
```

Format converter and LCA software for converting data sets

- Format converter: Developed since 2007 by GreenDelta, with support of various stakeholders (PRé, thinkstep, ecoinvent, UN, JRC)
- Converts various LCA data formats
- Open source and free tool
- http://www.openlca.org/format-converter/

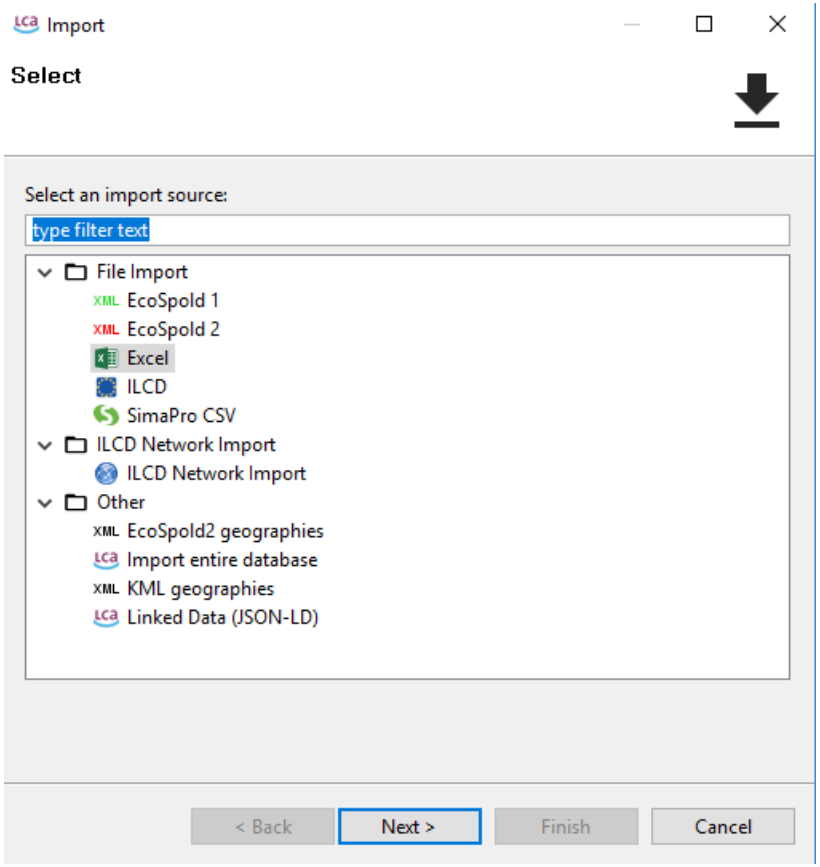
Format converter and LCA software for converting data sets

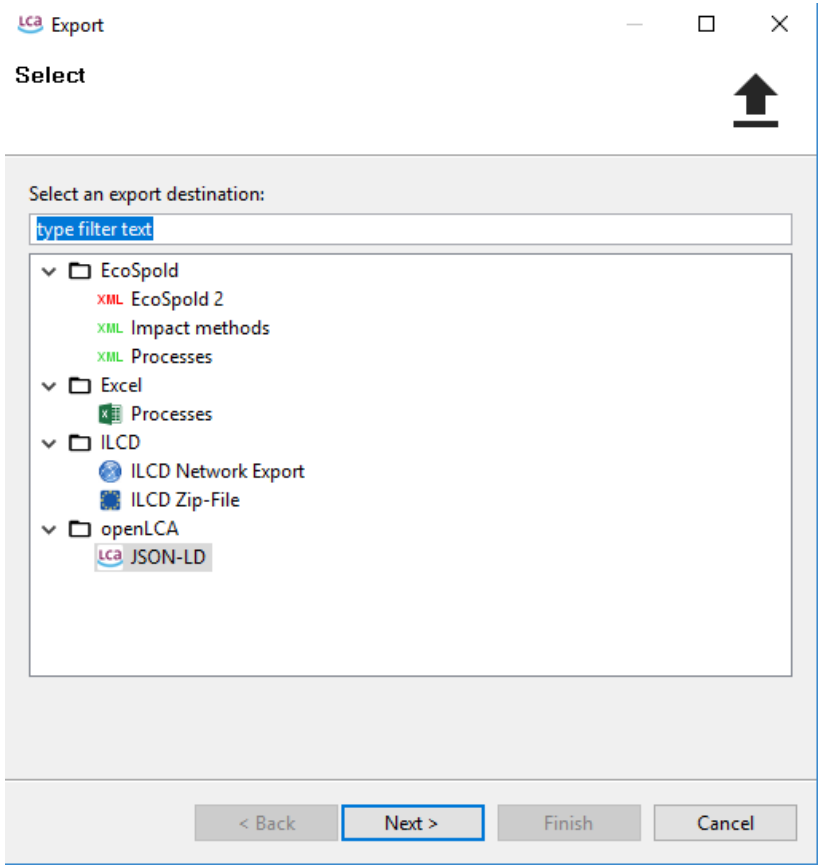
- Format converter: supported formats

from/to	EcoSpold 1	EcoSpold 1 (SimaPro)	EcoSpold 2	ILCD 1.1	CSV (SimaPro)
EcoSpold 1		x	x	x	-
EcoSpold 1 (SimaPro)	x		x	x	-
EcoSpold 2	x	x		x	x
ILCD 1.1	x	x	x		-
CSV (SimaPro)	-	-	x	-	

Format converter and LCA software for converting data sets

- openLCA: supported import and export formats



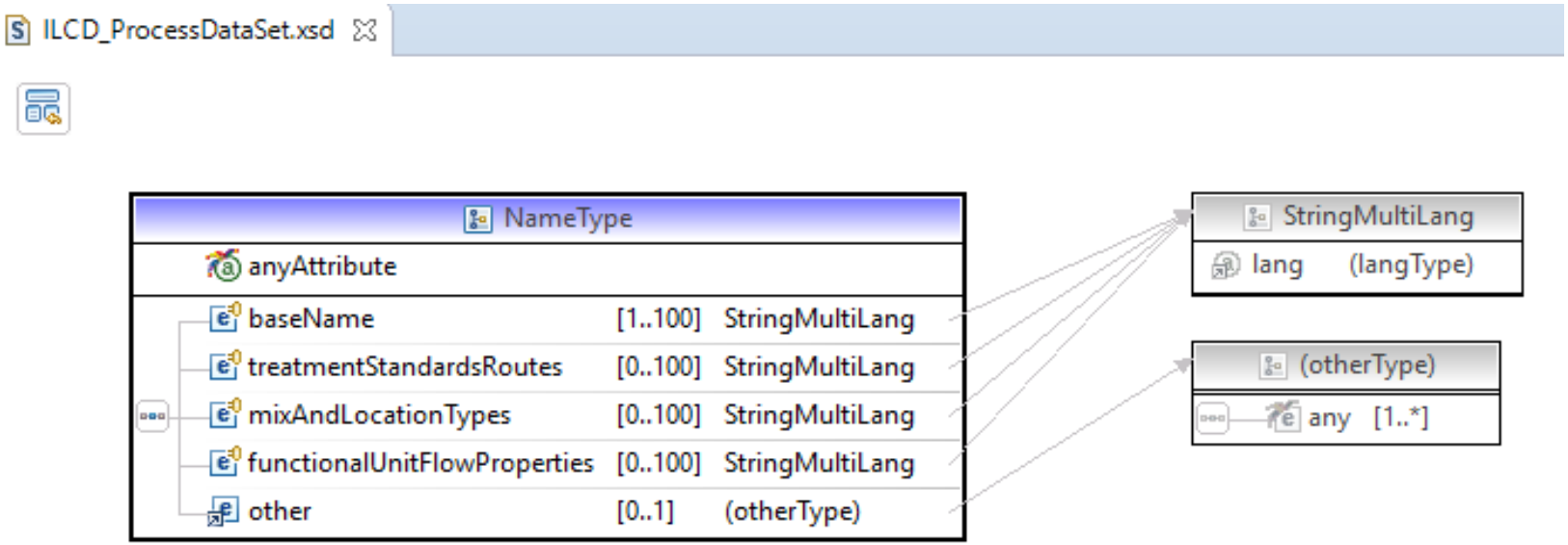


Format conversion: status

- Conversion is possible of course, tools like openLCA and the format converter can be used and can also be integrated into websites e.g.
- However, quite some fields and field requirements between formats are incompatible
  - Infrastructure:
    - required field in EcoSpold1 for a process data set
    - Not foreseen in ILCD

Format conversion: status

- Process name: 120 characters max in EcoSpold02, 100 characters in ILCD



Name	FieldID	SpoldID, version 1	Options	Type	Size	Multiple Occurence	Req
activityName	100	401-rename		TString120	120	No	Yes
	A name for the activity that is represented by this dataset.						

Format conversion: status

- Format conversion is possible, but works only to some extent.
- Tool support is available, but typically, for a full conversion, special „treatment“ of datasets is needed in addition to better capture the content
  - Store information in other fields
  - Decide about default values
  - Adjust mappings
  - …

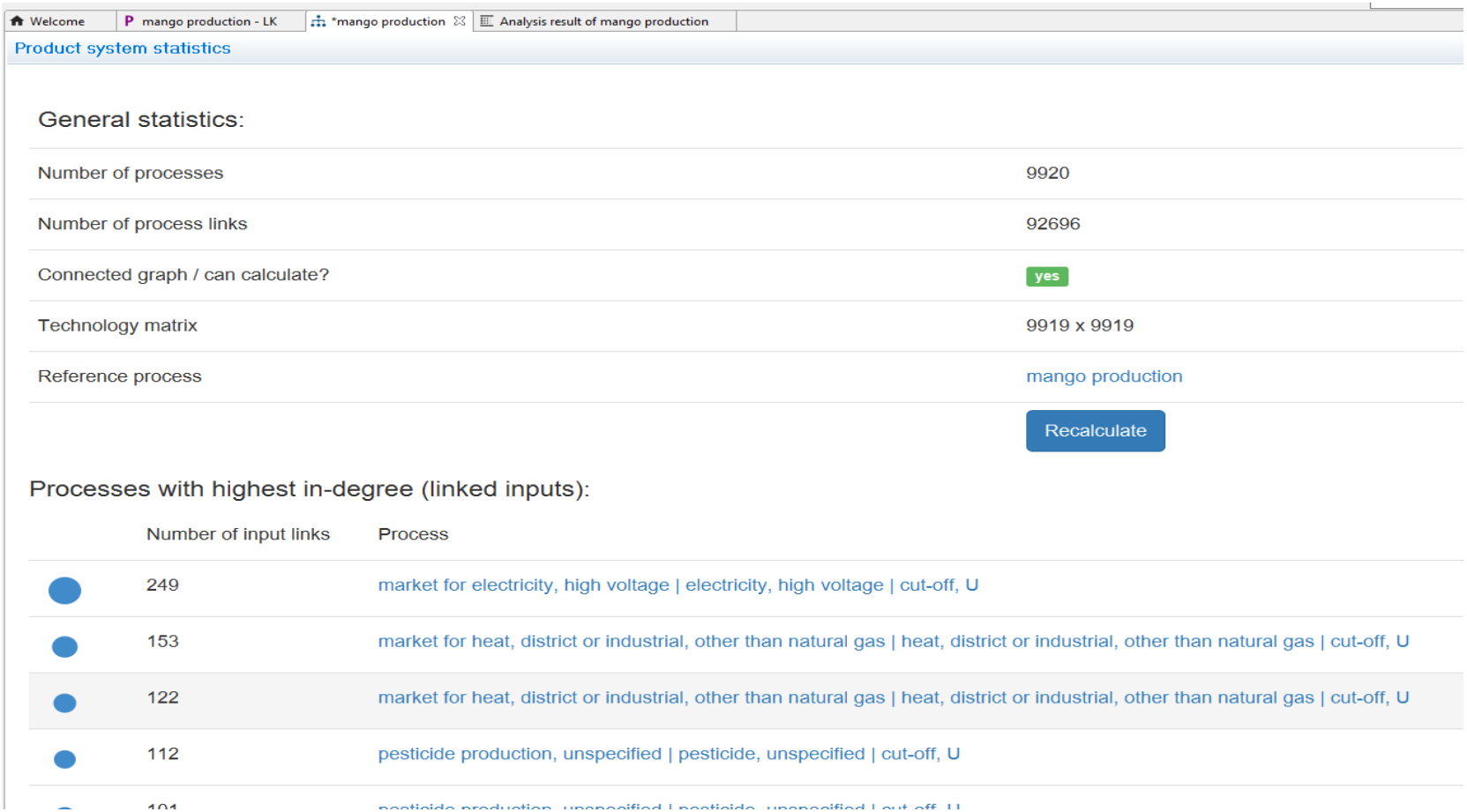
Walkthrough using the mango dataset example: EcoSpold02 data set created in the ecoEditor

Name

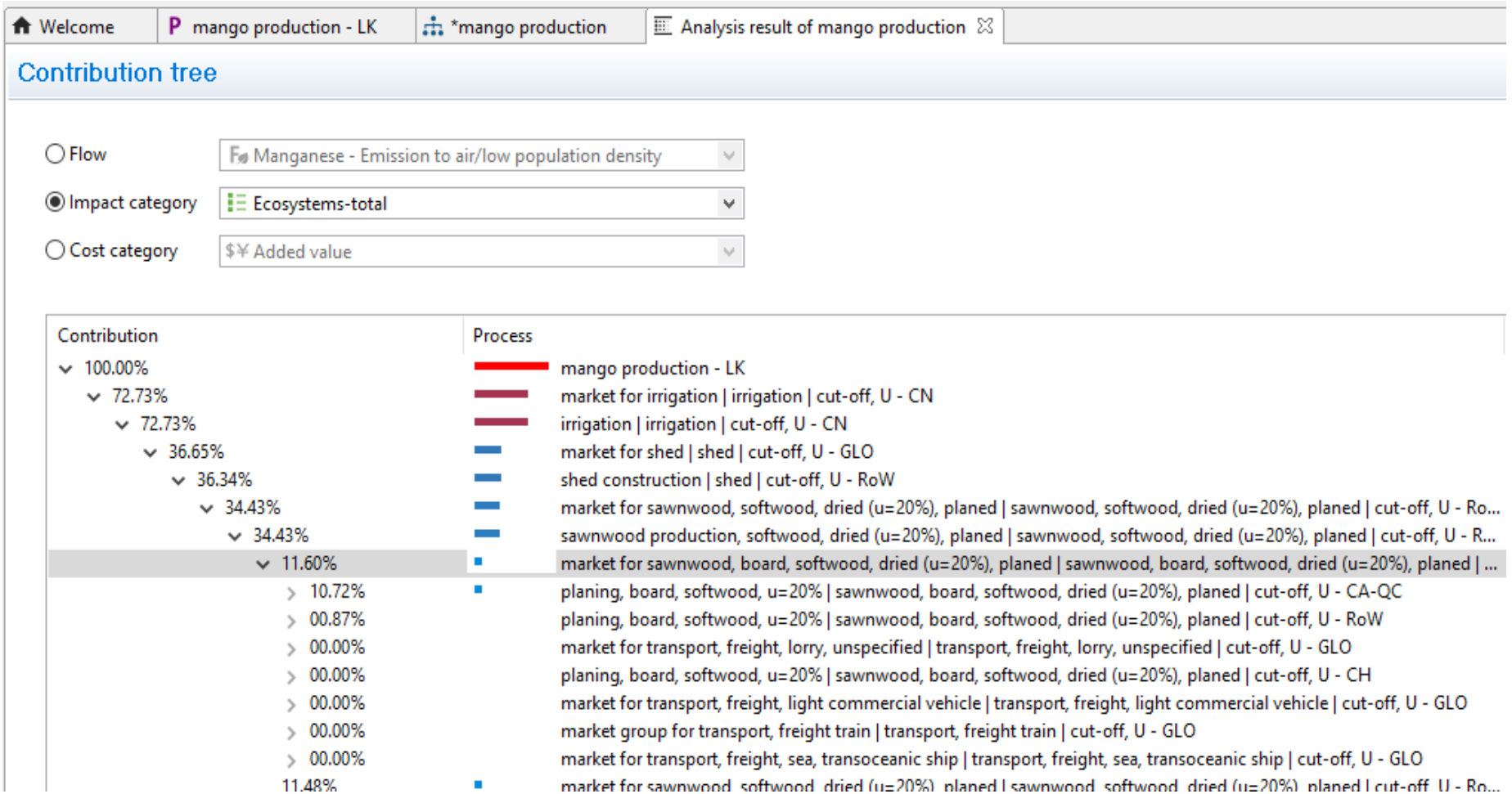
mango production_LK, 2016 - 2016_created during the course.spold

```
<ecoSpold xmlns="http://www.EcoInvent.org/EcoSpold02">CRLF
  <activityDataset>CRLF
    <activityDescription>CRLF
      <activity id="1a091cc1-8544-4509-92f7-0c8b59d04d82" activityNameId="b606483a-3de2-4599-85e8-eb065d97c756" inheritance
      type="1" specialActivityType="0">CRLF
        <activityName xml:lang="en">mango production</activityName>CRLF
        <includedActivitiesStart xml:lang="en">from prepraring the field </includedActivitiesStart>CRLF
        <includedActivitiesEnd xml:lang="en">with harvesting the mangos</includedActivitiesEnd>CRLF
        <generalComment>CRLF
          <text xml:lang="en" index="1">This datseset refers to the production of Mango in Sri Lanka.</text>CRLF
        </generalComment>CRLF
      </activity>CRLF
      <geography geographyId="0c13823c-7d7e-11de-9ae2-0019e336be3a">CRLF
        <shortname xml:lang="en">LK</shortname>CRLF
      </geography>CRLF
      <technology technologyLevel="3" />CRLF
      <timePeriod startDate="2016-01-01" endDate="2016-12-31" isDataValidForEntirePeriod="true" />CRLF
      <macroEconomicScenario macroEconomicScenarioId="d9f57f0a-a01f-42eb-a57b-8f18d6635801">CRLF
        <name xml:lang="en">Business-as-Usual</name>CRLF
      </macroEconomicScenario>CRLF
    </activityDescription>CRLF
    <flowData>CRLF
      <intermediateExchange id="bc72e57d-5045-482f-85d6-066eaccfdcb8" unitId="487df68b-4994-4027-8fdc-a4dc298257b7" amount=
      intermediateExchangeId="1c819f15-1202-4fa9-8ef9-72196c2dfb85">CRLF
        <name xml:lang="en">apple</name>CRLF
        <unitName xml:lang="en">kg</unitName>CRLF
        <property propertyId="6d9e1462-80e3-4f10-b3f4-71febd6f1168" amount="0.83" unitId="487df68b-4994-4027-8fdc-a4dc29825
          <name xml:lang="en">water in wet mass</name>CRLF
```

Walkthrough using the mango example: Import EcoSpold02 data set in the ecoEditor to openLCA



Walkthrough using the mango example: Imported EcoSpold02 dataset in openLCA



Capstone User Exercise

A user exercise can be done if you have or acquire certain tools[(a)]:

- the ecoEditor from ecoinvent (http://www.ecoinvent.org/data-provider/data-provider-toolkit/ecoeditor/ecoeditor.html)
- and openLCA from GreenDelta (http://www.openlca.org/)

Exercise:

1. Load the mango dataset into ecoEditor (or even better use ecoEditor to build one of your own datasets)
2. Import the test dataset from ecoEditor into openLCA
3. Connect and build a complete life cycle, and calculate results in openLCA.

[(a)] The LCA Databases Helpdesk does not endorse these tools, but simply identifies them as development capabilities and learning resources.

Thank you

Training on Data Acquisition and Dataset Development

August 2017 Version

Managed by SETAC