Technical Helpdesk for National LCA Databases

Training on Data Acquisition and Dataset Development
Part 4 - Process Modelling for LCI Datasets

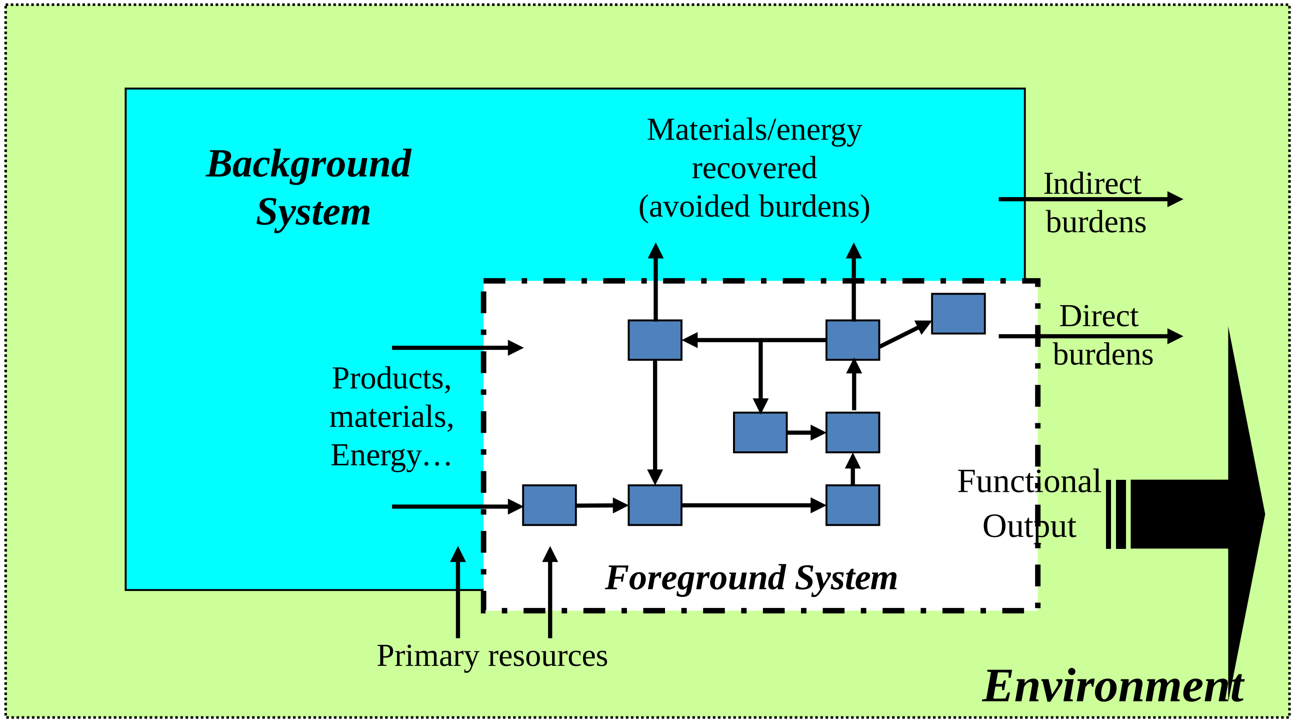
Content from Amir Safaei, ecoinvent

Managed by SETAC

August 2017 Version

Useful concepts
Foreground and Background system in LCA

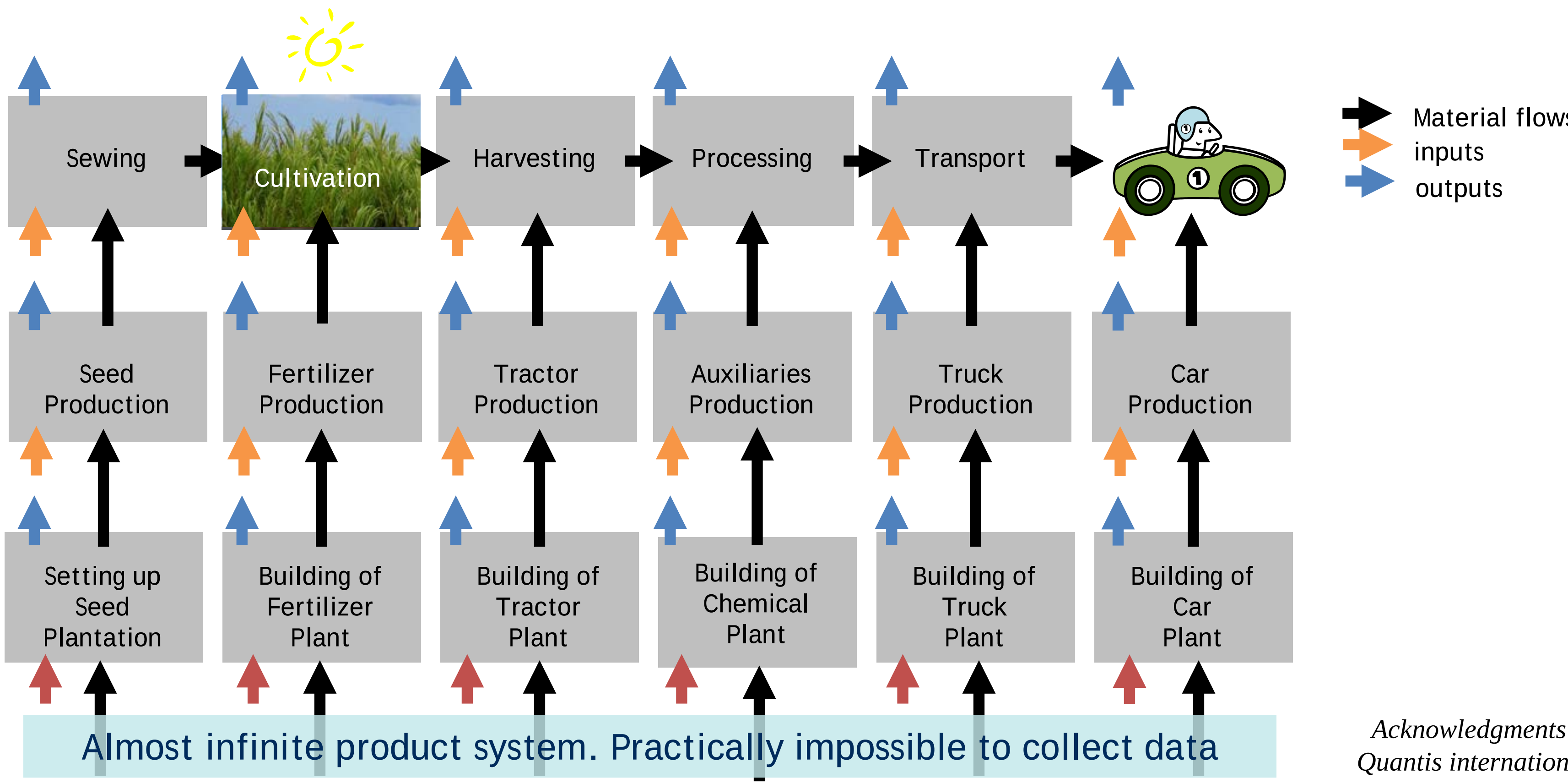
ecoinvent



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

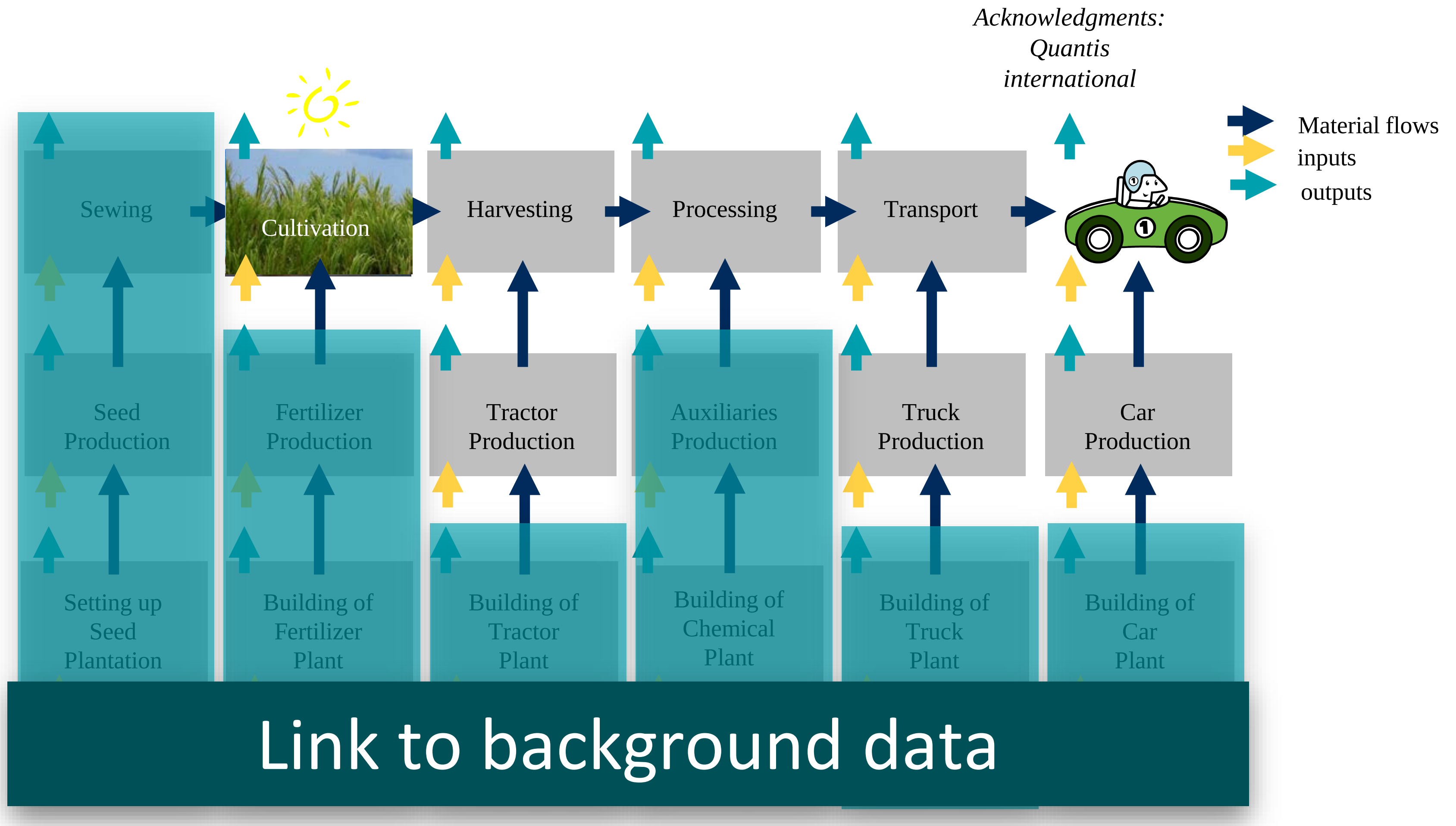
Background system: processes that actions can not be directly taken wrt the results of the LCA, often, external secondary data used

The need for background data
Example: Biofuels for Transportation



Acknowledgments: Quantis international

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



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

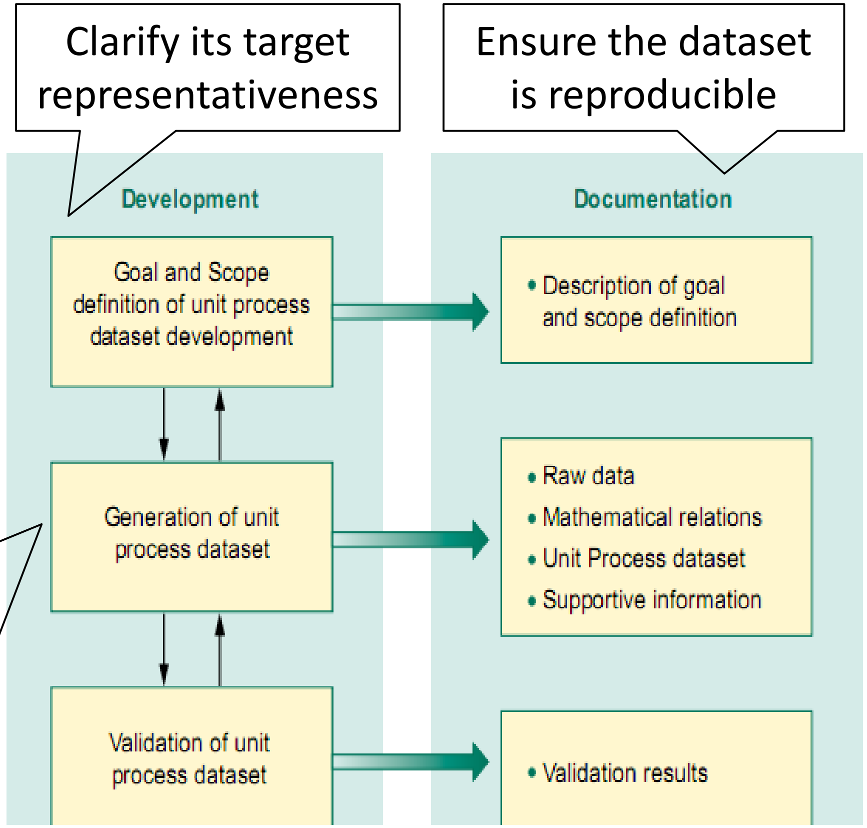
Step 2: Define the mathematical relationships

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

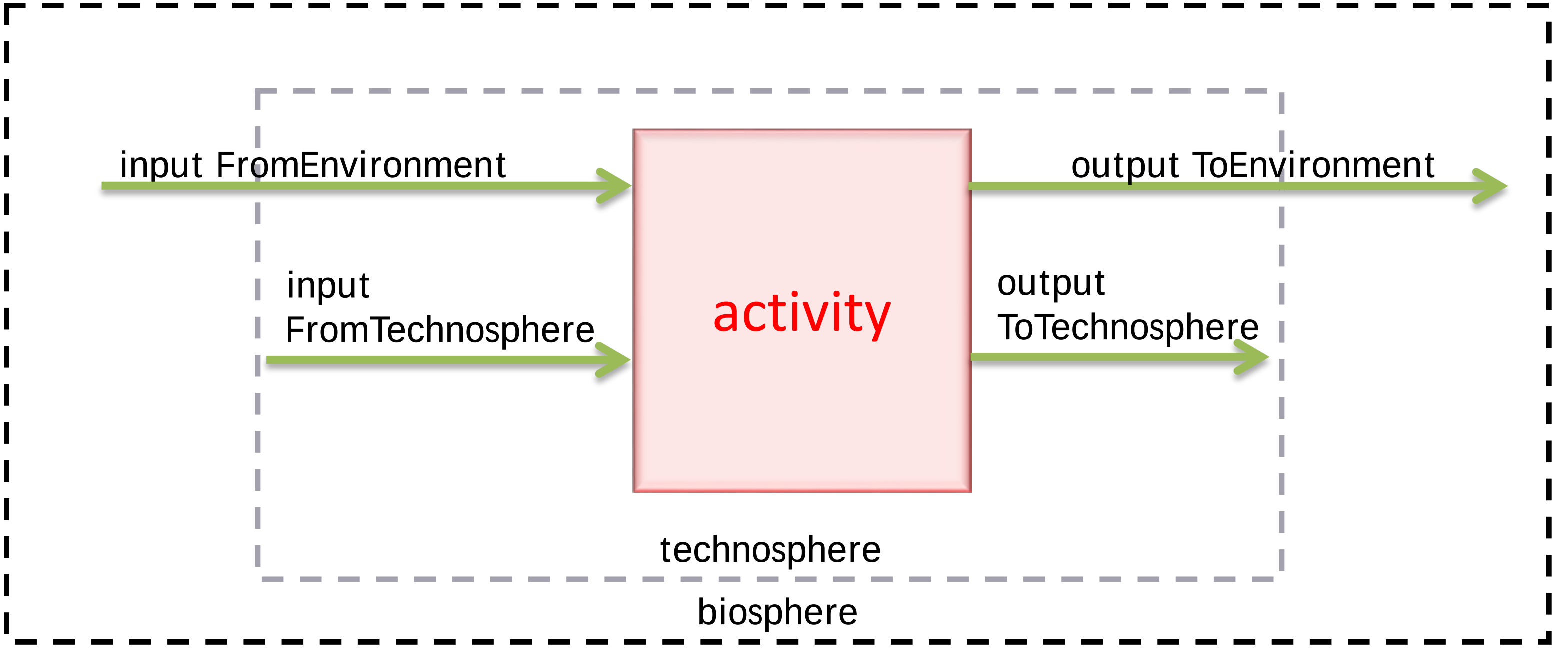
unlinked: the inputs and outputs are *not* yet linked to the corresponding supply chains.

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



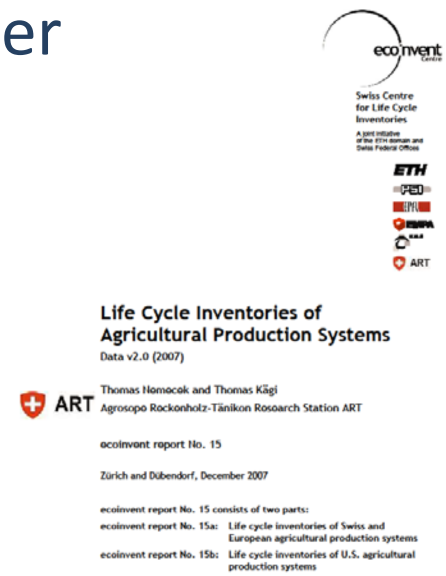
Explore similar datasets within the database

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

ETH

ART

Life Cycle Inventories of Agricultural Production Systems

Data v2.0 (2007)

Thomas Nemecek and Thomas Kägi

ART Agroscope Reckenholz-Tänikon Research Station ART

ecoinvent report No. 15

Zürich and Dübendorf, December 2007

ecoinvent report No. 15 consists of two parts:

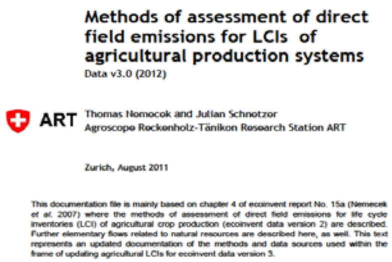
ecoinvent report No. 15a: Life cycle inventories of Swiss and European agricultural production systems

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...

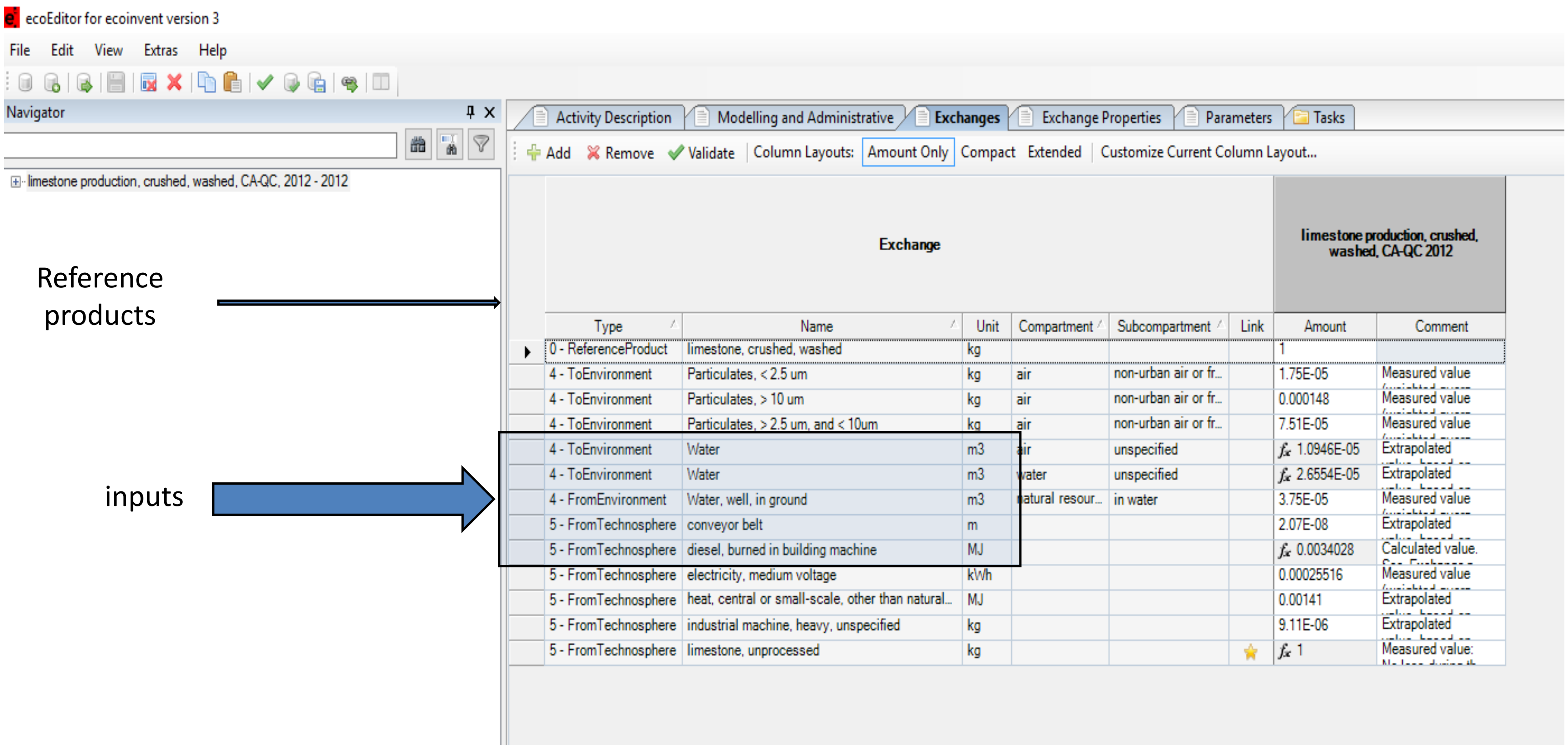
- When you have identified the similar processes in the DB, do a screening LCA or "process contribution" analysis
  - Choose a number of relevant LCIA categories
  - Do a "contribution analysis" to identify the most relevant flows and exchanges
  - You might be able get this info from the DB management

Example of process contribution

econvent

Reference products

inputs



Type	Name	Unit	Compartment	Subcompartment	Link	Amount	Comment
0 - ReferenceProduct	limestone, crushed, washed	kg				1	
4 - ToEnvironment	Particulates, < 2.5 um	kg	air	non-urban air or fr...		1.75E-05	Measured value
4 - ToEnvironment	Particulates, > 10 um	kg	air	non-urban air or fr...		0.000148	Measured value
4 - ToEnvironment	Particulates, > 2.5 um, and < 10um	kg	air	non-urban air or fr...		7.51E-05	Measured value
4 - ToEnvironment	Water	m3	air	unspecified		fx 1.0946E-05	Extrapolated
4 - ToEnvironment	Water	m3	water	unspecified		fx 2.6554E-05	Extrapolated
4 - FromEnvironment	Water, well, in ground	m3	natural resour...	in water		3.75E-05	Measured value
5 - FromTechnosphere	conveyor belt	m				2.07E-08	Extrapolated
5 - FromTechnosphere	diesel, burned in building machine	MJ				fx 0.0034028	Calculated value.
5 - FromTechnosphere	electricity, medium voltage	kWh				0.00025516	Measured value
5 - FromTechnosphere	heat, central or small-scale, other than natural...	MJ				0.00141	Extrapolated
5 - FromTechnosphere	industrial machine, heavy, unspecified	kg				9.11E-06	Extrapolated
5 - FromTechnosphere	limestone, unprocessed	kg			★	fx 1	Measured value:

Aggregation

ecoinvent

- 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

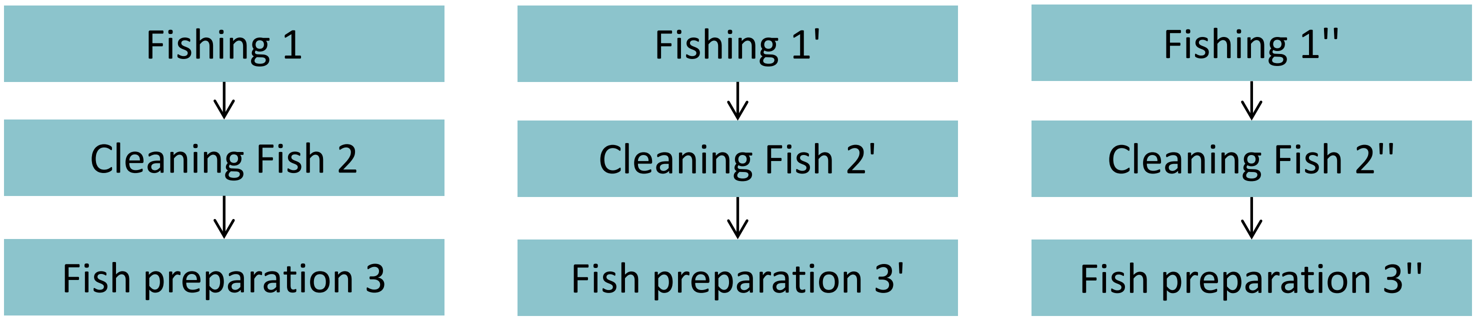
Aggregation

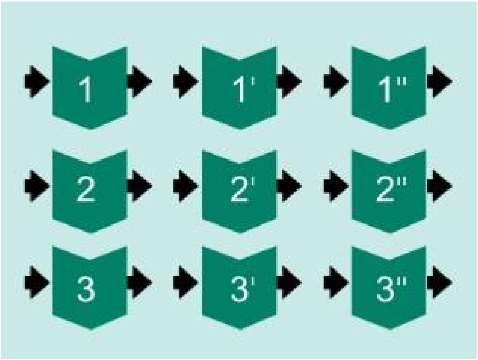
Typically LCI data vary based on ...

- 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.



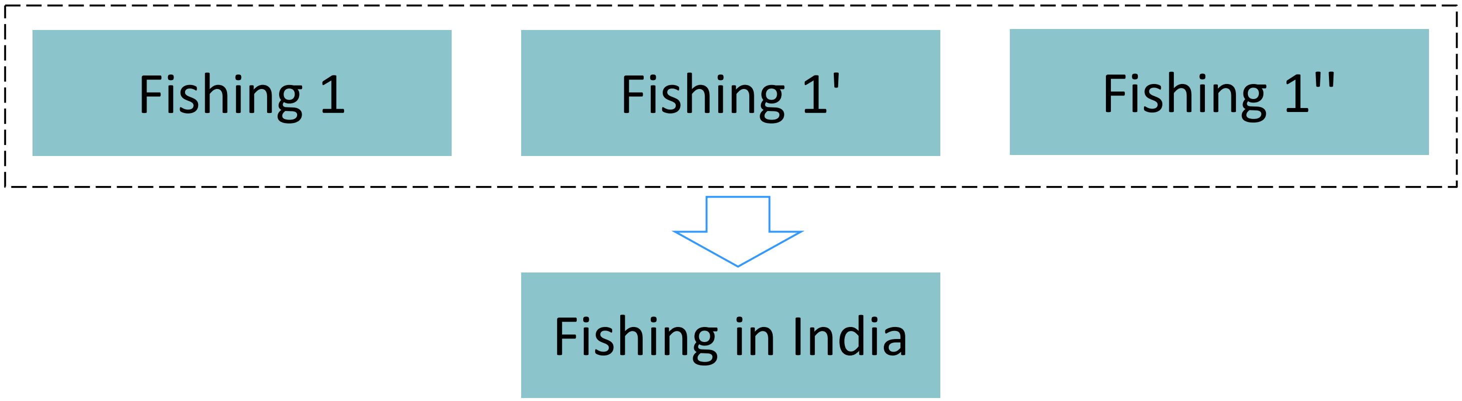


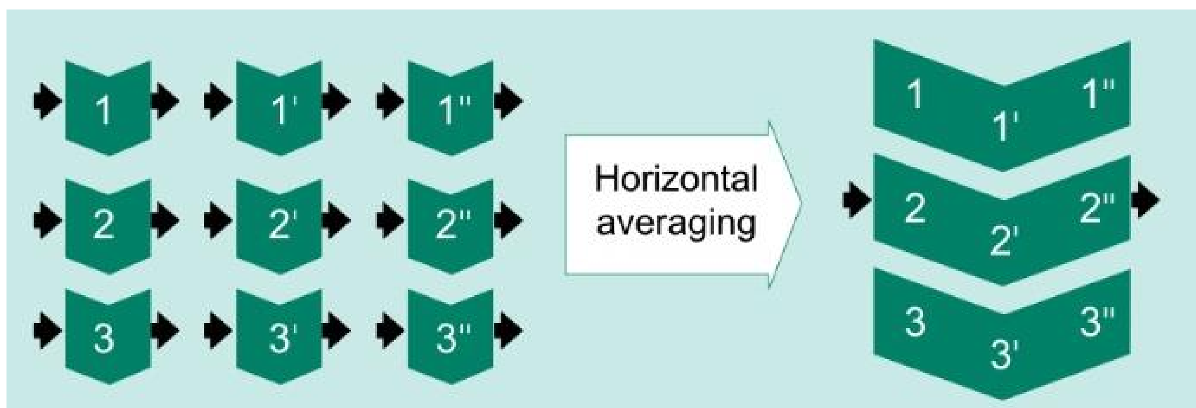
Source: Sonnemann, G., & Vigon, B. (2011).

Horizontal averaging

ecoinvent

- To represent, e.g. a regional average

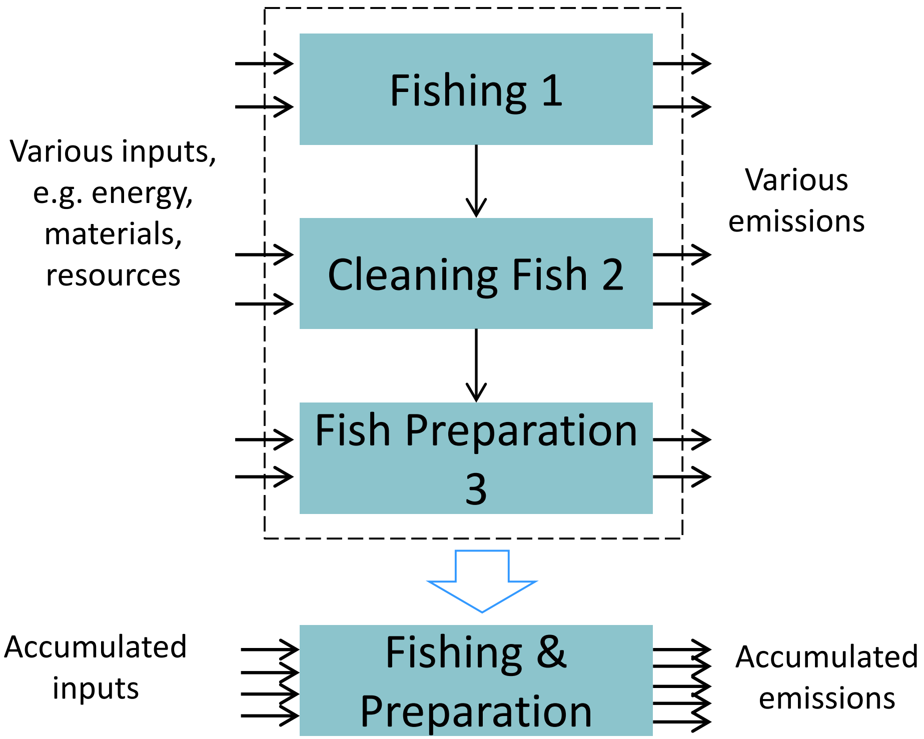


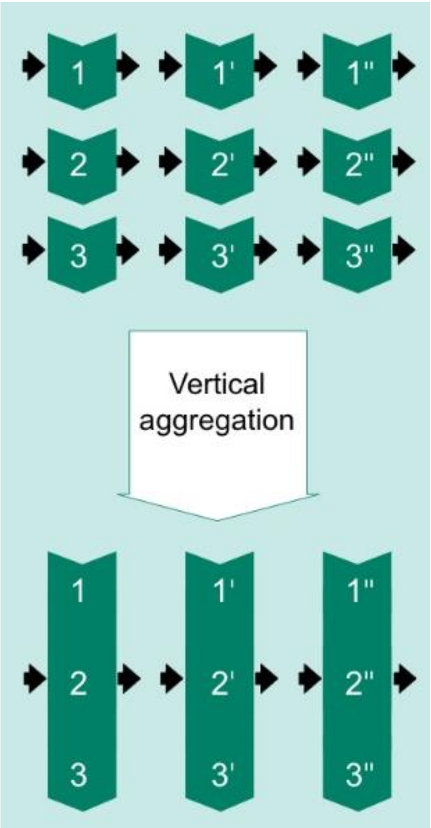


Source: Sonnemann, G., & Vigon, B. (2011)

Vertical aggregation

- To cumulate and represent multiple LC stages





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

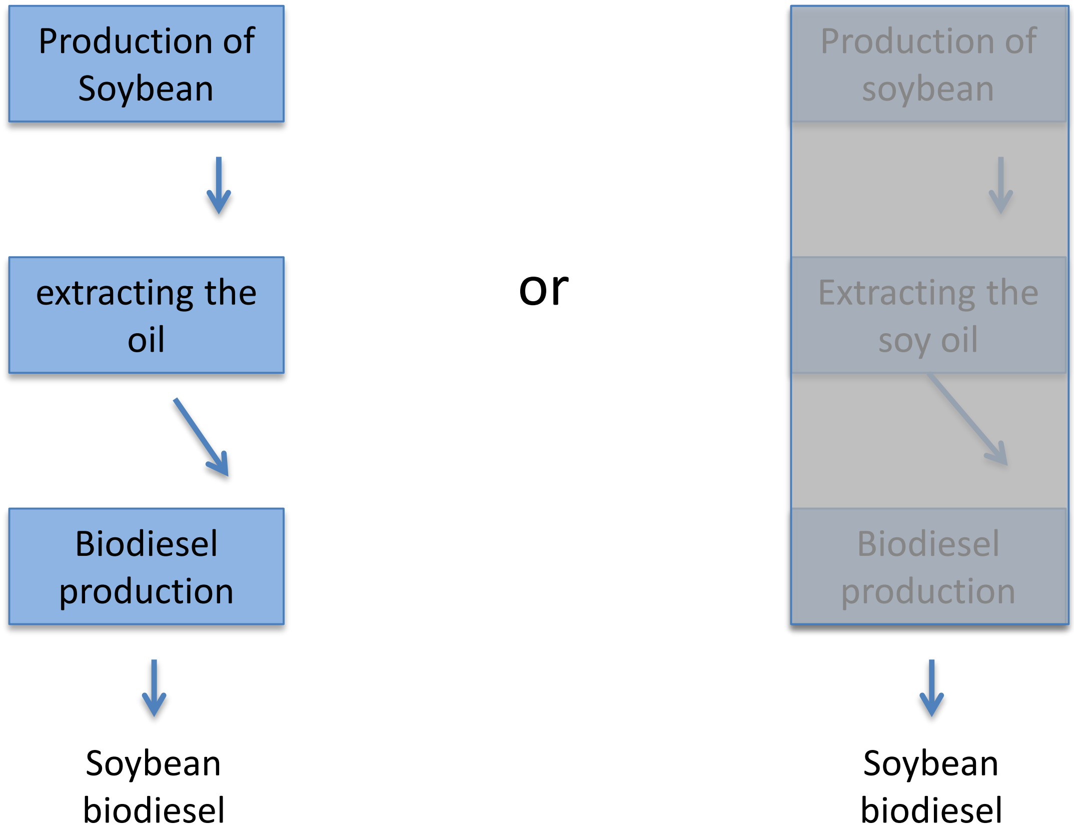
Vertical aggregation

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, **<u>better be adjusted</u>** to the unit process level and the aggregation level of the existing datasets

Vertical aggregation example

ecoinvent



Which one is suitable for hot spot analysis?

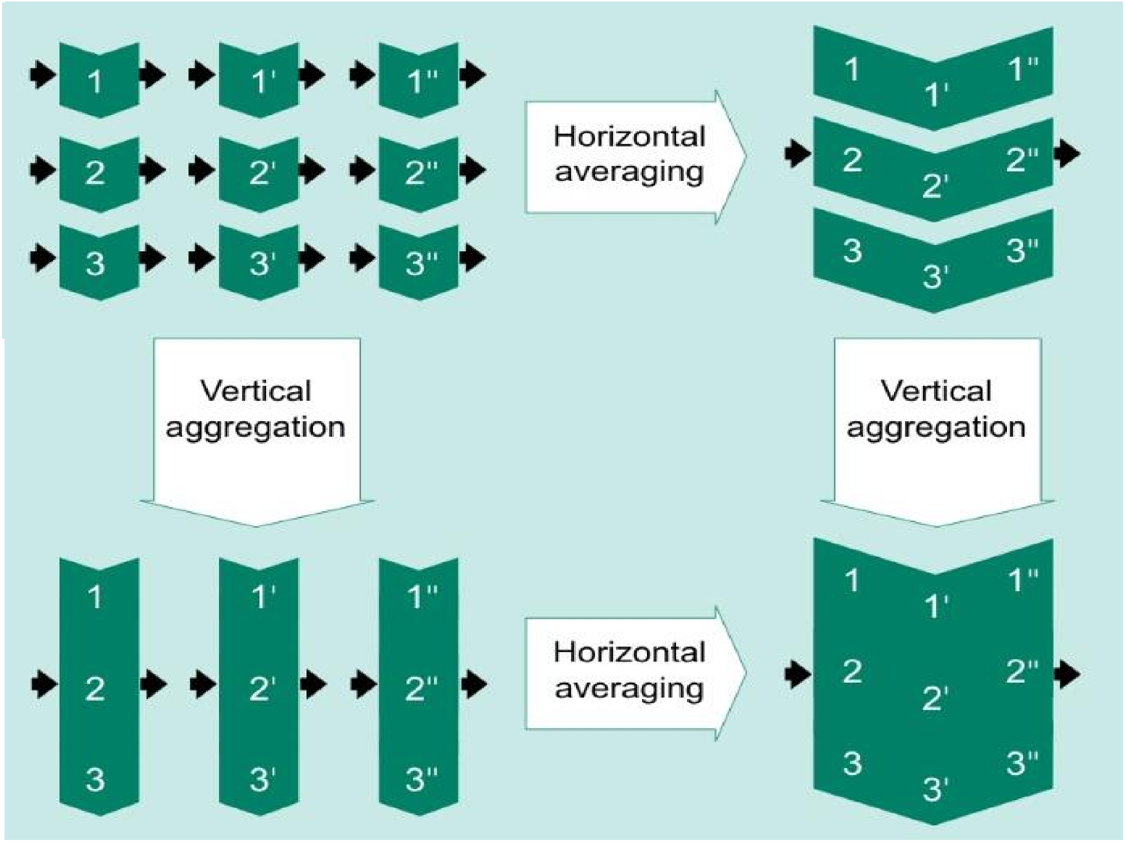
Modelling your data according to the DB (vertical aggregation)

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

ecoinvent



To retain

- 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

- 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

- 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

- 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

ecoinvent

Rice production, IN

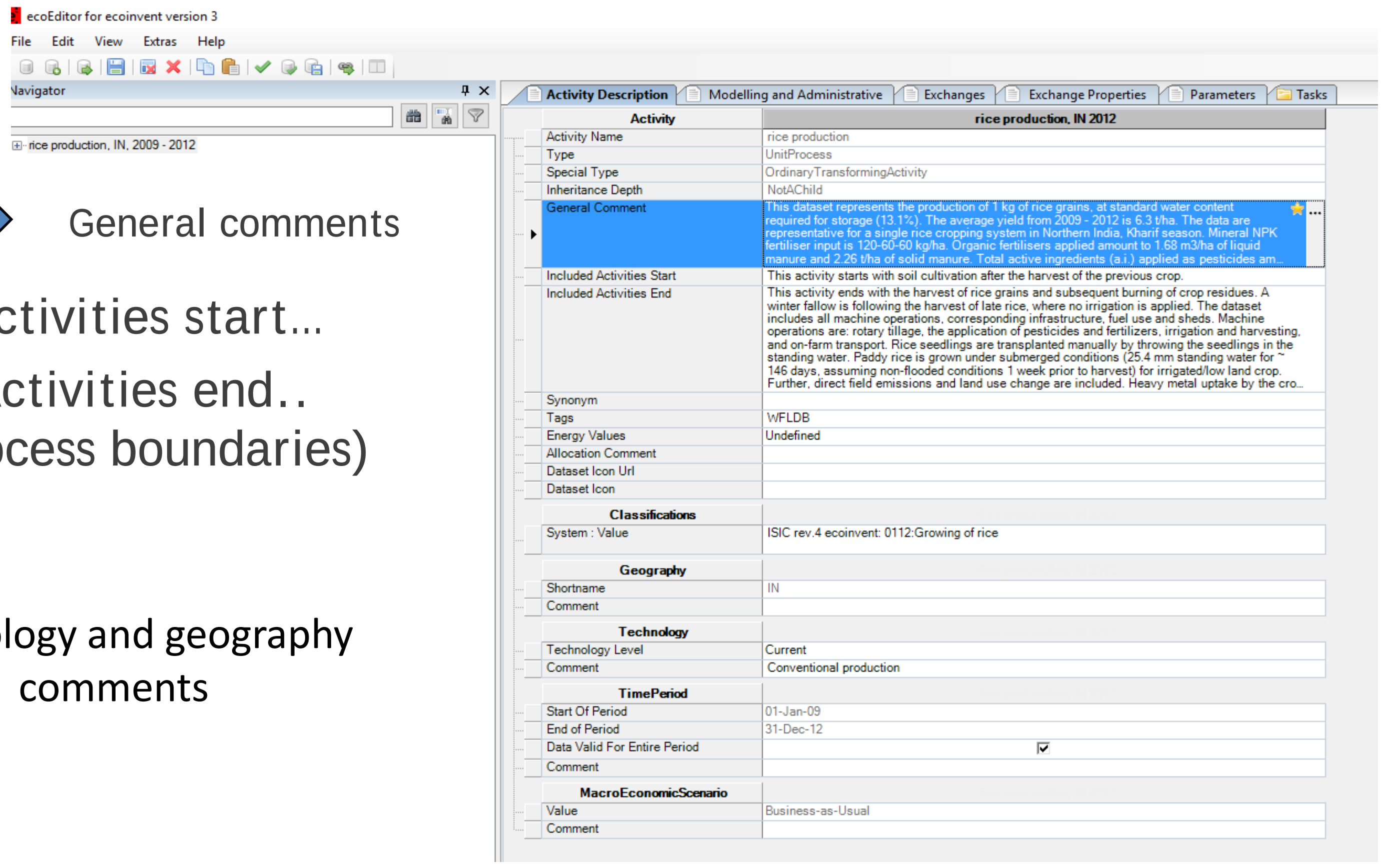
General comments

Activities start…

Activities end..

(process boundaries)

Technology and geography comments



Documentation, example

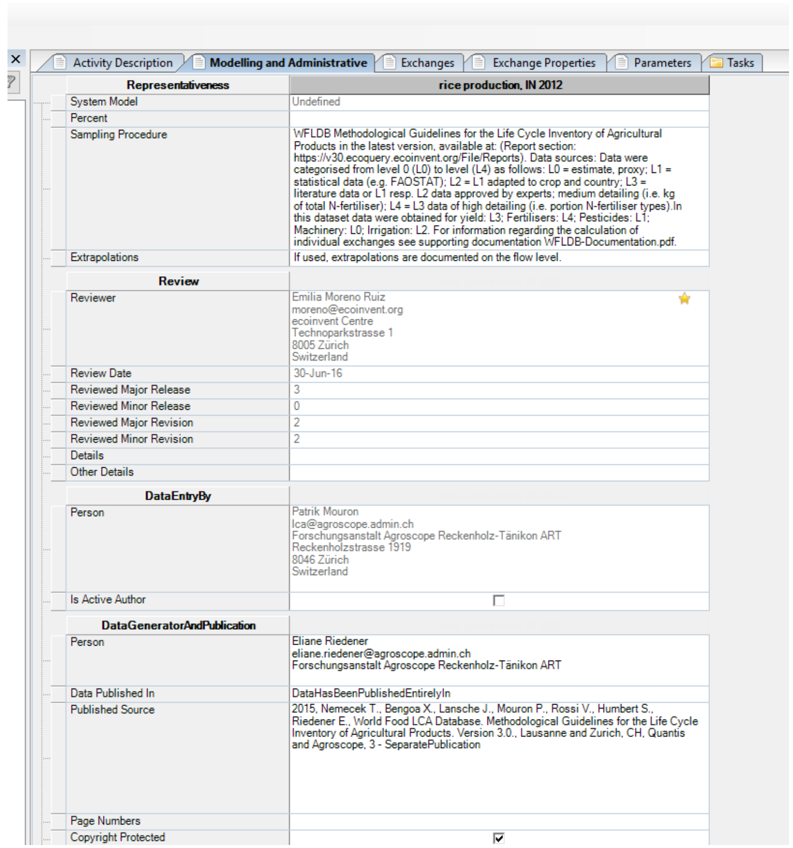
ecoinvent

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



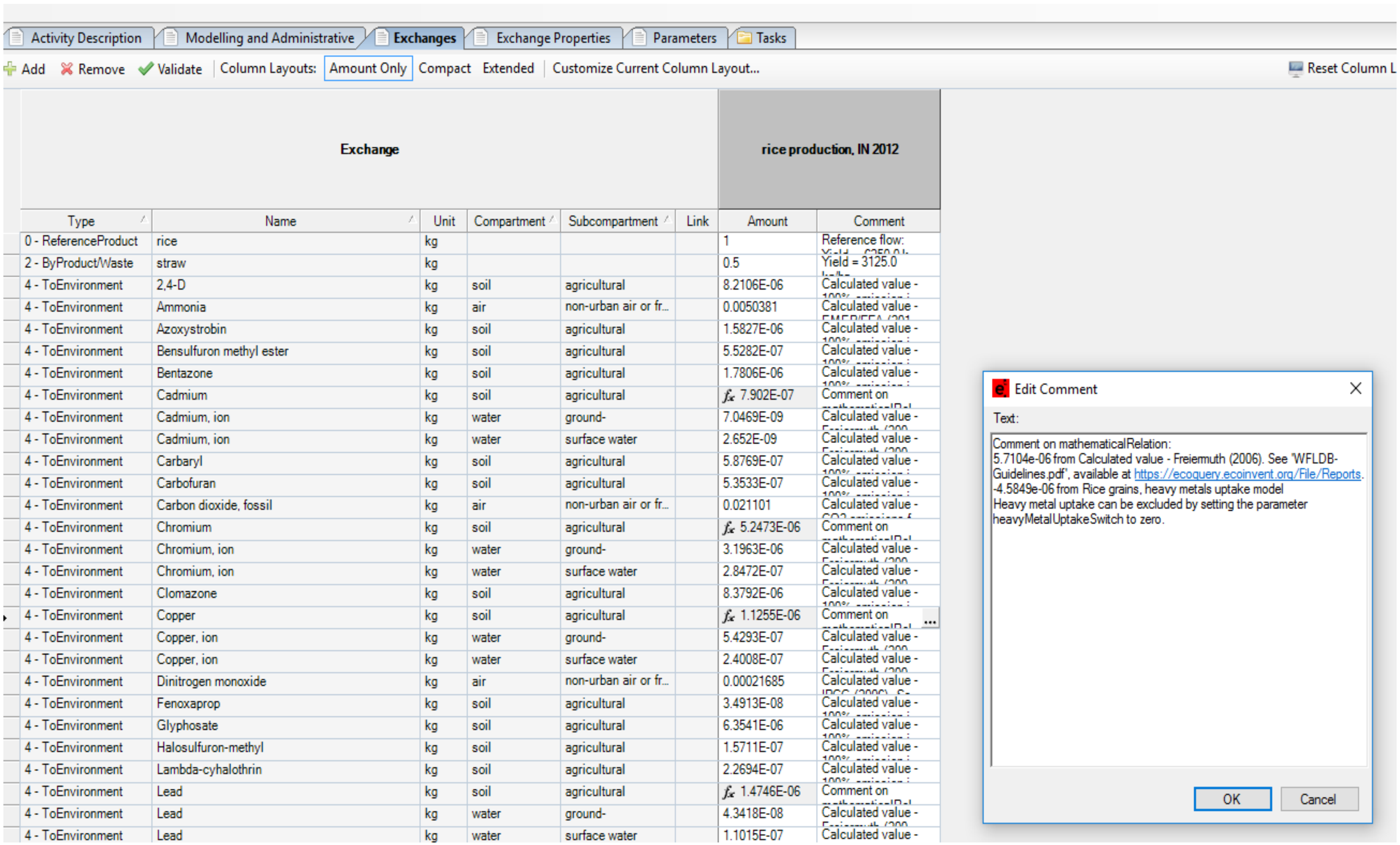
Documentation; system boundary

ecoinvent

- *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

ecoinvent



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
  - CO_2 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

ecoinvent

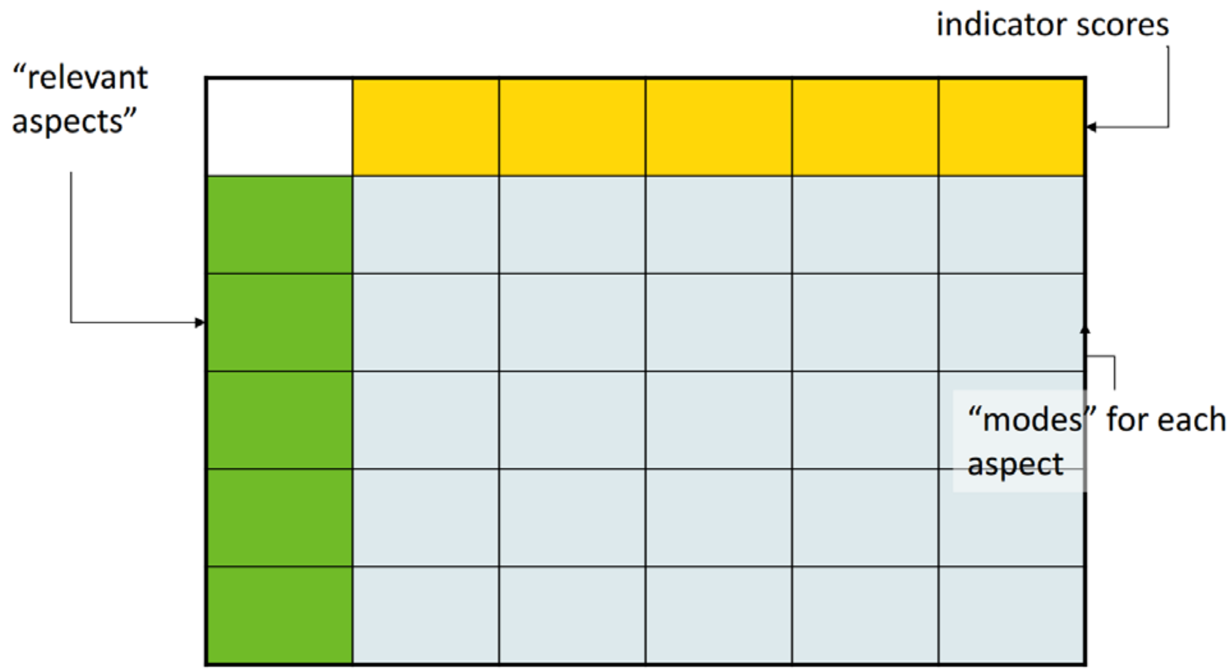
- Uncertainty is present in all phases of LCA
  - Goal and scope definition
  - LCI
  - LCIA
- Consequently, all the numerical values contain an associated uncertainty
- The approaches to calculate the uncertainty can be quantitative (Monte Carlo simulation e.g.) and qualitative
- Once again, each DB 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

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

Uncertainty field

All numerical fields have an uncertainty field.

Activity Description | Modelling and Administrative | **Exchanges** | Exchange Properties | Parameters | Tasks

Add | Remove | Validate | Column Layouts: Amount Only | Compact | Extended | Customize Current Column Layout... | Reset

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

- 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

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

econvent

- 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.

For helpdesk assistance –

- Become a Helpdesk member:
  - To access the Helpdesk exchange space (or any other Clearinghouse area), you will need to create an account in the Clearinghouse (www.scpclearinghouse.org):
  - Toward the bottom of the homepage you will see a button labeled 'Join the Community now'. Click on this link and open a form to allow you to create a login and profile.
  - Once logged in, you can modify or update your profile or explore the various SCP topic areas.
  - Go to 'About' and then to 'Exchange Spaces' where you will see Lifecycle Approaches in the drop down menu and one menu level below that is the Technical Helpdesk.
  - The Technical Helpdesk space will be available to any visitor, logged in or not. Without being logged in and joining the helpdesk space, any visitor can look at the various sections of the helpdesk space, but cannot contribute any content.
  - In order to become a member of the helpdesk space, on the homepage under the summary, is "Request space membership". Click here, you will automatically be given rights of a members to contribute content, since it is a public group.
  - For your next login, you go directly to http://spaces.scpclearinghouse.org/ and then choose the Technical Helpdesk space in the dropdown list.
- Helpdesk Manager - Bruce Vigon, Consultant to SETAC,
- Helpdesk Coordinator – Kristina Bowers, UN Environment, Economy Division