Annex A: Life Cycle Inventory Dataset Review Criteria

Version 3.1

11 August 2015

Created by:

Andreas Ciroth, ciroth@greendelta.com

Jutta Hildenbrand

Alessandra Zamagni

Chris Foster

Reviewed and edited by:

Bruce Vigon

Llorenc Mila i Canals

Fayçal Boureima

Content

Preamble

This text is a version which is based on an initial version prepared by the author group, which received broad comments from a number of database providers in Life Cycle Assessment, and from experts in LCA, including:

- Bo Weidema
- Bruce Vigon
- Simone Fazio
- Rolf Frischknecht
- Thinkstep
- US Department of Agriculture (multiple experts)
- US Environmental Protection Agency (multiple experts)

Further feedback was provided in a phone conference, organised by UNEP and SETAC, with participants also from database operators in Thailand, Malaysia and Brazil, as well as Thinkstep and ecoinvent.

The current text reflects the feedback received. A separate text summarises the feedback and provides more detailed replies[1]. Overall, the feedback and responses were very positive. In the revised text, therefore, the principal structure, the specific criteria, and the evaluation of criteria in the ordinal 5 level scale ("pedigree") was confirmed.

A major change compared to the initial version of the text is that thresholds which were proposed for most of the criteria, to evaluate whether a dataset can pass the review process or not, have been removed, meaning that it is now the responsibility of the dataset provider to decide whether a dataset needs revision or is "fit for publication" (see reviewer guidance).

This document is a working version of a globally accepted set of criteria for reviewing LCI data sets, and which is practically applicable. The authors expect and hope that experiences, e.g. from the international review of 40 datasets in Malaysia, Thailand and Brazil will further refine and detail the criteria.

[1] Life Cycle Inventory Dataset Review Criteria Development, Comments and replies, version 2, 11 August 2015, Andreas Ciroth, Jutta Hildenbrand, Alessandra Zamagni, Chris Foster

1. Background and motivation

In an ongoing project related to the Technical Support of National and Regional Initiatives and Networks on LCA Databases and Related Capacity Development, Communication and Research Elements, Life Cycle Inventory (LCI) datasets have been created in various countries worldwide. The aim is for these data sets to become part of national Life Cycle Assessment (LCA) databases run in various countries, for example in Brazil, Malaysia and Thailand.

The different databases have the goal to be of high quality and to be interoperable. A Life Cycle Initiative workshop in Shonan Village, Kanagawa Japan has developed a Global Guidance Principles report[2] [1], and recently discussion has started on drafted Global Guidance Principles conformance criteria [2] to make the Global Guidance Principles report more operational.

The individual data sets intended to be published for international use should be complete and of known quality. Therefore the project supports a systematic review of these data sets against well-defined criteria, conducted by independent LCA experts, to ensure the data set quality. The international review of datasets is also a key demand from the participating countries, who see it as an essential element for them to make them internationally available.

However, a set of criteria for LCI dataset review that is commonly accepted does not exist yet, and proposed criteria schemes (some of whaich have been used internally by databased developers) are rather complex and therefore maybe not be suited for a project in the context of capacity development. Important contributions for LCI data set review are the ILCD data set review requirements [3], ecoinvent review criteria [4], GaBi DQIs and review scheme [10], the Global Guidance Principles criteria [2], and also review requirements set forth in ISO 14040 and 14044 [5-6], although the latter are addressing LCA studies and not LCI data sets.

The project team proposes criteria for review of LCI datasets with the aim to apply these for guiding reviewers of the data sets. The criteria are well-defined within a logical structure and rationale. They are developed by a group of international LCA experts, consisting of

- Dr. Andreas Ciroth, GreenDelta (chair)
- Mr. Chris Foster, Manchester University and EuGeos
- Dr. Jutta Hildenbrand, Chalmers University and Swerea, and
- Dr. Alessandra Zamagni, ENEA and Ecoinnovazione.

Developing the review criteria for LCI data sets is important, but not the only required aspect in a complete LCI dataset review procedure. Complementing the review criteria are:

- A review procedure is needed for performing the review of data sets. Different existing databases have each their own, slightly differing procedures, with typically a moderator and one or several independent reviewers performing the review, in an iterative loop with the data set creator.
- One or also several reviewers are needed for performing the review. Several parts of the review depend on the expertise of the reviewer(s), the ability to judge the data set in an unbiased, independent manner is important. In this project candidate reviewers were identified through a call for interest and their expertise in LCA methodology or specific economic sectors used to classify them for use in the review.

[2] That said, the Shonan Guidance Principles report is so far only a report in paper format and cannot be directly used for assessing interoperability of LCI databases. In a recent project sponsored by SETAC, Shonan Guidance conformance criteria have been developed and proposed to make the Shonan Guidance Principles report operational. For details, see e.g. Ciroth, A.: Conformance of LCA databases to the Global Guidance Principles, presentation, SETAC Barcelona, May 2015.

- Finally, for a larger amount of data sets, it is probably possible to design a smarter, more efficient review procedure, which could possibly cluster data sets and treat them in a similar manner, avoiding e.g. the need to look into each of them "from scratch". Also these more advanced procedures are out of scope of the current project and report.

2. Review criteria set development principles

For developing the set of criteria, the following aspects were considered.

- Basically, every single process data set is a model of reality; as with every model, it is a simplification of reality, but it should reflect important aspects of reality in a correct way.
- As a process data set, the model will be used as part of a Life Cycle Assessment, and consists of input and output flows and meta-information, following principles of ISO 14048 [7]
- What is more specifically important for the process dataset depends on the intended application and on the goal and scope of the dataset and, if relevant, also of the database where the dataset is integrated.
- The overarching set of criteria use the definitions and review aspects of ISO 14040 and 14044 as a starting point. Existing proposals by ILCD and ecoinvent and GaBi and others were considered as well.
- The criteria set should be workable, as simple as possible, flexible, and yet not pointless or open to abuse. This means also that the criteria application should be, ideally, reproducible, and that their application leaves little room for interpretation.
- For the set of criteria, the following "items" are provided:
    - a list of criteria with definition and motivation,
    - rules and procedures how to assess the criteria, for a given data set, with possible assessment results, and
- The review criteria distinguish between unit processes and aggregated processes, simply because aggregated processes require an additional evaluation of the aggregation procedure, and its documentation.

A review of datasets has the aim to assess the data quality of a dataset; a dataset that performs well during the review should be of known and expected high data quality; as already stated in ISO 14040 [5], data quality is fully application dependent (data quality: ability to satisfy stated requirements, ISO 14040); this aspect is considered in the specification of the review criteria.

Mandatory and optional criteria have not been pre-determined. Reviewers are being asked to indicate which of the criteria they consider more or less important to the assessment.

The threshold where a dataset is acceptable depends on the application and is in part database and user specified. For one application a, an "average" performance in a given criterion might suffice, while for another application b, only "good" or better performance is acceptable.[3]

[3] This implies that it is more demanding to create „general purpose data sets", as they need to satisfy probably different application contexts.

3. The set of review criteria in detail

3.1. Two generic, basic requirements to be used as a starting point

The application-dependency of data quality, as defined in ISO 14040 (and also broadly accepted elsewhere, as a similar definition can be found e.g. in ISO 9000) leads to the problem that "anything goes" for the data quality assessment criteria. This is not useful in a practical application, and it also contradicts common practice of LCA dataset use.

Therefore, we are proposing two generic, basic requirements that will be used for a more detailed specification of review criteria. These requirements are derived from the first two of the above-mentioned "review criteria set development principles":

1) every single process data set is a model of reality; as every model, it is a simplification of reality, but it should reflect important aspects of reality in a correct way.
2) As a process data set, the model will be used as part of a Life Cycle Assessment, and consists of input and output flows and meta information, following principles of ISO 14048

This means, the data set should be:

- Correct
- Understandable and clear
- Reflective of basic LCI data set structure as described in ISO 14048, with input and output flows and meta-information.

3.2. The set of review criteria

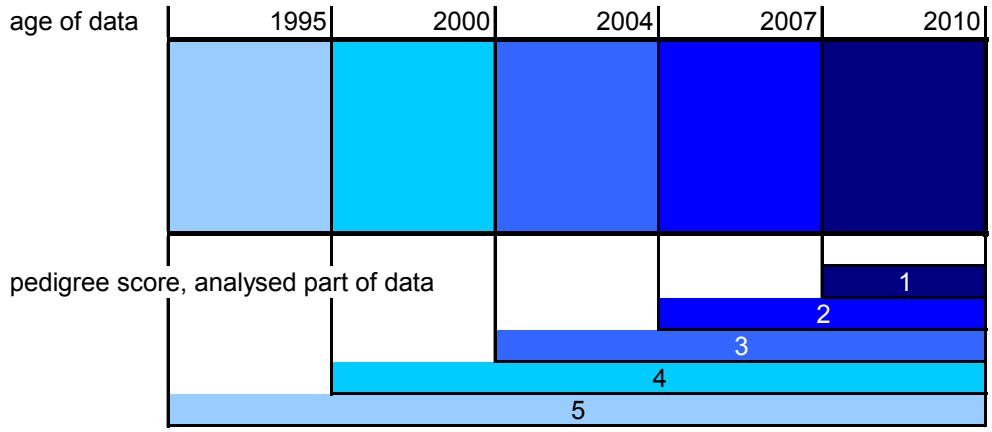
3.2.1. 3+1 Review criteria cluster

Starting from the basic requirements (3.1), the review criteria in detail are all derived from, and can be grouped into, three review criteria clusters:

Cluster I: Conformance: conformance criteria address how well the data set reflects aspects foreseen in its goal and scope description or in goal and scope description of the database(s) where it is meant to be integrated, if relevant. These "aspects" are stated reference time, location, technology-related, and completeness requirements. This principle is explained in fig. 1, below, for the time reference [9].

Cluster II: Correctness & reliability: Correctness and reliability criteria assess, as the name suggests, how correct and reliable the information provided in the data set is. This is important to know since otherwise, information provided in the dataset can claim that it is perfectly conformant, while in reality it is not.

Cluster III: Procedural & meta-information: These criteria serve to document and assess aspects of the review procedure itself, for example the number of reviewers involved, the level of access to the data sets, the level of independence of the reviewers. Ideally, a reviewer should have full access to the models underlying the data set; due to practical limitations, this is sometimes not possible.



pedigree score definitions

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

Figure 1: Assessing time conformance for data sets, principle example for a data set with a reference time in 2010, according to goal and scope [5, p. 18]

Additionally, goal and scope documentation completeness of the data set is important. All the information evaluated in the criteria needs to be documented, especially, the goal and scope of the dataset needs to be provided, since otherwise, all the conformance aspects cannot be evaluated. Documentation completeness is therefore an essential additional "cluster" with one single criterion.

3.2.2. The review criteria, their definition and evaluation

This section explains and defines the different review criteria that are proposed. For each criterion, the assessment rules are also provided. These rules are simple on purpose, and in general consist of a qualitative, ordinal assessment which links to an assessment score. We thereby follow the principle of a pedigree assessment which is already broadly used for LCA dataset quality assurance, for example by ecoinvent [4], GaBi [10], and in the ILCD system [3].

For each criterion, also the "assessment scope" is noted: Some criteria are meant to be assessed on the level of the entire data set, while others should rather be assessed for, e.g., each exchange / flow separately. It is the responsibility of the reviewer to apply the appropriate level of detail here. It is common practice in LCI data set creation to combine information (raw data) from several sources and possible different reference times, geographic locations and so forth into one data set, but this combination is – again according to common practice as perceived by these authors – typically relevant only for some exchanges. These different exchanges will need to be addressed by the "flow specific" assessment.

For example, when modelling an engine in a truck, it may be that the fuel consumption (for a certain payload, and street type) per distance is taken from one source and from one year, while the emissions of the engine are available from another source from a slightly different year. It is useful if these aspects (differences in data sources and reference year) are documented. In this example the assessment of the dataset should be done on the flow level.

The criteria are grouped into the 3+1 clusters identified in the previous section.

Goal and scope documentation completeness

Definition and motivation: Goal and scope for the dataset, i.e. its intended application, needs to be clear, either by documentation that is specific for the dataset or by references to documentation that is available on the database level. The goal and scope documentation needs to contain:

- Reference time
- Reference geography
- Reference technology
- Reference model completeness
- Reference sample completeness

Scope: dataset

Assessment:

1 All reference elements are specified for the dataset

5 At least one of the reference elements is missing

Conformance criteria

Time related conformance

Definition: Time difference between the reference time for the dataset, according to documentation, and the time period for which data were surveyed (i.e. the period of the initial data collection).

Scope: Flow / exchange level and individual data set field level

Assessment:

1 Less than 3 years of difference to the time period of the dataset (reference time)

2 Between 3 and 6 years of difference to the time period of the dataset (reference time)

3 Between 6 and 10 years of difference to the time period of the dataset (reference time)

4 Between 10 and 15 years of difference to the time period of the dataset (reference time)

5 Age of data unknown or more than 15 years of difference to the time period of the dataset (reference time)

Geographical conformance

Definition: Geographical difference between the reference area for the data set, according to documentation, and the area for which data were surveyed

Scope: Flow / exchange level and individual data set field level

Assessment:

1 Data from an area under study

2 Average data from larger area in which the area under study is included

3 Data from area with similar production conditions

4 Data from area with slightly similar production conditions

5 Data from unknown or distinctly different area (North America instead of Middle East,

OECD-Europe instead of Russia)

Technological conformance

Definition: Technological difference between the reference technology foreseen for the data set, according to documentation, and the technology for which data were surveyed. The criteria list distinguishes organisation, process (of creating or producing the product), and material. Material is not only quality but the sheer material, e.g. “steel” vs. “aluminium”.

Scope: Flow / exchange level and individual data set field level

Assessment:

1 Data from organisations, processes and materials under study

2 Data from processes and materials under study (i.e. identical technology) but from different organisations

3 Data from processes and materials under study but from different technology or application (e.g. data on the recycling of a Li-ion battery used in electronic devices used for the recycling of Li-ion batteries used in electric vehicles)

4 Data on related processes or materials

5 Data on related processes on pilot or laboratory scale or from different technology

Remark: The “and” in the scores is understood as a logical ”and”.

Model completeness conformance, flows and documentation

Definition: Availability of information required for practitioners to select a dataset for application, beyond the more narrow time, geography and technology criteria. This criterion is to combine all model-related aspects, which includes the reference flows used, and other dataset modelling aspects as well.

Scope: Data set level

Assessment:

1 sufficient documentation easily available for the data set for modelling procedure, sources, aggregation procedure, system boundary setting, limitations; for sources also on the flow level

2 one of the aspects modelling procedure, sources, aggregation procedure, system boundary setting, limitations insufficiently described or not easily available for the data set, for sources also on the flow level

3 two of the aspects modelling procedure, sources, limitations insufficiently described or not easily available for the data set

4 insufficient documentation (two of the aspects modelling procedure, sources, limitations lacking; or all not easily available for the data set)

5 no or very minimal documentation available

Sample conformance

Definition: Representativeness of sample data used in the enquiry for the foreseen reference market.

Scope: Flow / exchange level and individual data set field level

Assessment:

1 Representative data from all sites relevant for the market considered, over an adequate period to even out normal fluctuations; usually 12 months cover seasonal fluctuations and can be considered adequate

2 Representative data from >50% of the sites relevant for the market considered, over an adequate period to even out normal fluctuations

3 Representative data from only some sites (<50%) relevant for the market considered, over and adequate period of time to even out normal fluctuations, or >50% of sites but from shorter periods

4 Representative data from only one site relevant for the market considered or some sites but from shorter periods

5 Representativeness unknown or data from a small number of sites and from shorter periods

Remark: 'Representative data from all sites' means that for the sample, all sites are considered. The actual data can come from few sites, which are selected in a way that the selection is demonstrated to be representative. One example for LCA is described in [11]. This is scientific best practice in other fields. Access to all sites is of course not always possible, therefore a lower percentage is addressed in the scores 2 and 3. With 'market' we mean the product or several comparable products which are provided (to a next process data set in a database for example). For specific data, the required "sample" is only the one specific data set, and the market is only one provider / producer.

Correctness & reliability criteria

Accuracy of the provided information

Definition: This criterion serves to evaluate whether the data set deviates from an unknown true value. Accuracy is challenging to be evaluated per se unless the sample size compared with the population is very high and the uncertainties are provided. Even here the true value is a probability estimate. Inconsistent data are possibly not accurate. Similar to consistency, comparison to other reliable data sets and to technical domain expertise and natural laws for example can reveal whether a data set is inaccurate. Accuracy is difficult to assess for a data set and typically cannot be "proven" with one single aspect of a data set; accuracy directly relates to the fundamental requirement formulated in the introduction that the data set should reflect important aspects of reality in a correct way.

Scope: Entire dataset

Assessment:

1 Accuracy of the data set is supported by the supplied statistical information and defined comparisons to reliable datasets

2 Likely minor inaccuracies based on the provided information that do not affect the overall accuracy of the dataset

3 Likely minor inaccuracies based on the provided information but where the overall impact

on the data set is not clear

4 Inaccuracies found which are likely to have a moderate effect on the dataset accuracy (dataset may be usable where its contribution to the results are of low sensitivity)

5 The data set is highly inaccurate or information to evaluate accuracy is not provided

Precision of the provided information

Definition: The criterion serves to evaluate the availability of information regarding the variance of a value given in the dataset. Note that this criterion does not include an assessment whether the data given are inaccurate (not providing a true value).

Scope: Flow / exchange level; focus is here on the quantitative amounts of exchanges

Assessment:

1 Information is listed as datapoint and distribution with parameters

2 Information is listed as a range between value x and y

3 Information is listed as a data point

4 Information is listed as greater than x

Reliability of the provided information

Definition: Origin of the information used to create the dataset, according to documentation, flows and meta-information.

Scope: Flow / exchange level and individual data set field level

Assessment:

1 Verified measurement (verified either requires that the measurement was conducted using accepted protocols and all relevant error sources have been evaluated and documented or that the measurement was performed multiple times and the precision is very tight (e.g. +/- % of mean)

2 Unverified measurement or verified calculation (verified calculation means that the result was independently found to be correct and the calculation procedure – model or equation – is appropriate)

3 Unverified calculation

4 Documented estimate

5 Undocumented estimate

Remark: Overall, the data set must meet its goal and scope in order to provide overall reliable and consistent information. Obviously for a data set, many different "information pieces" are provided, which have different relevance for the overall data set and for its reliability. For example, the description of the product and quantitative reference is usually of higher relevance than information about publication references. However, the authors think it is not possible to assess in a generic way whether the complete data set is sufficiently reliable, since e.g. also an estimated unit or unit conversion factor can potentially affect the reliability of the entire data set. Therefore, this criterion does not have a predefined threshold but it rather needs to be evaluated by the reviewer.

Consistency of the provided information

Definition: Consistency addresses the requirement that the information provided in a data set is not contradictory, internally but also in comparison to other reliable data sets and to technical domain expertise and natural laws for example. Other reliable data sets are for example data sets that have passed the review already.

Scope: Data set

Assessment:

1 No inconsistencies found

2 Minor inconsistencies found that do not affect the overall reliability of the data set

3 Some inconsistencies found where the overall impact on the data set is not clear

4 Inconsistencies found which are likely to have an effect on the data set

5 The data set is inconsistent

Procedural & meta-information

Number of reviewers and their relation to the data provider

Definition: This criterion serves to document how many reviewers have participated in a review and whether they were independent or internal. For this purpose independent means the reviewer has no organizational affiliation with the dataset developer. Internal means there is a common employer or the individuals who developed the dataset and those reviewing it work for the same organization, even if in different locations. The employment arrangement also does not matter as full time employees or temporary project-specific workers are considered internal. For sake of simplicity, the expertise of the reviewer is not assessed.

Scope: Dataset

Assessment:

1 Two or more independent reviewers

2 One independent *and* one or more internal reviewers-

3 One independent or two or more internal reviewers

4 One internal reviewer

5 No reviewer

Data access

Definition: This criterion is especially important for a review of aggregated data sets; often, it is important for the reviewer to have access to the processes that were aggregated to the data set to understand the aggregation procedure, reference time and geography for the aggregated processes, and so forth. Depending on the supply chain model, sometimes also access to a selected number of processes provides good insight, for example in cases where parts of the supply chain where no access is possible are of minor importance (e.g., small mass flows).

For review of a unit process (unaggregated) data set, by definition always full access to the supply chain is granted.

Scope: Dataset

Assessment:

1 Full access to the supply chain model

2 Limited access to the supply chain model, access to more than 80% of all flows and more than 80% of mass

3 Limited access to the supply chain model, access to more than 60% of all flows and more than 60% of mass

4 Limited access to the supply chain model, access to more than 40% of all flows and more than 40% of mass

5 No access to the supply chain model

3.2.3. Aggregating review criteria results

After the individual review criteria have been assessed, they will need to be aggregated when the review criteria are meant to be applied for different flows of the data set

For aggregating the criteria results, several approaches can be envisaged, for example calculation of an arithmetic mean [3], creation of a "virtual uncertainty contribution" of different data quality criteria [4], calculating an arithmetical average with specific consideration of outliers to take into account overcompensation [8], and a mere check whether the data set fails in a specific criterion, i.e. if for at least one criterion, a threshold is violated.

For this review please simply calculate an arithmetic average per criterion evaluated at the flow level. Do not hide or remove the individual scores in the results form.

4. Outlook and next steps

This report is a first attempt at defining an overall, broadly accepted approach for reviewing life cycle inventory data sets. It is in its current state a proposal of the author group with feedback from a set of stakeholders. It needs to be more broadly discussed, also at an international level. It will certainly be useful to test the approach in practical cases as is planned in the UNEP technical assistance project to the Thai, Malaysian and Brazilian databases.

Once consolidated, it would further make sense to provide more practical guidance in terms of electronic checklists and maybe automated, web-based tools to take over some tasks required for the review of a data set, so that the review experts can focus on those parts where their expertise and judgement is needed.

5. References

[1] http://www.unep.org/pdf/Global-Guidance-Principles-for-LCA.pdf

[2] Ciroth, A.: UNEP Shonan Guidance Principles put into practice, presentation, SETAC conference Barcelona, May 2015

[3] European Commission- Joint Research Centre - Institute for Environment and Sustainability: International Reference Life Cycle Data System (ILCD) Data Network - Compliance rules and entry-level requirements. Version 1.1, 2012. EUR 24380 EN. Luxembourg. Publications Office of the European Union; 2012; http://eplca.jrc.ec.europa.eu/uploads/ILCD-Data-Network-Compliance-Entry-level-Version1.1-Jan2012.pdf

[4] Weidema B P, Bauer C, Hischier R, Mutel C, Nemecek T, Reinhard J, Vadenbo C O, Wernet G. (2013). Overview and methodology. Data quality guideline for the ecoinvent database version 3. Ecoinvent Report 1(v3). St. Gallen: The ecoinvent Centre

[5] ISO, ISO14040. "Environmental management–life cycle assessment–principles and framework." *International Organization for Standardization* (2006).

[6] ISO, ISO14044. "Environmental management—life cycle assessment—requirements and guidelines." *International Organization for Standardization* (2006).

[7] ISO/TS 14048:2002 "Environmental management -- Life cycle assessment -- Data documentation format" ." *International Organization for Standardization* (2002).

[8] Ciroth, A., Franze, J.: LCA of An Ecolabeled Notebook - Consideration of Social and Environmental Impacts Along the Entire Life Cycle, ISBN 978-1-4466-0087-0, Berlin, 2011.

[9] Ciroth, A., Muller, St., Weidema, B.: Refining the pedigree matrix approach in ecoinvent, project report, project commissioned by the ecoinvent centre, 2012; http://www.greendelta.com/fileadmin/Pedigree_report_final_May2012.pdf

[10] Baitz M, Makishi C, Kupfer T, Florin J, Schuller O, Kokborg M, Köhler A, Thylmann D, Stoffregen A, Schöll S, Görke J, Rudolf M, Liedke W (2014). GaBi Database & Modelling Principles 2014, Chapter 3.5.2., Version 1.0, November 2014, thinkstep AG, http://www.gabi-software.com/fileadmin/gabi/Modelling_Principles/GaBi_Modelling_Principles_2014.pdf

[11] Ciroth, A., Srocka, M.: How to Obtain a Precise and Representative Estimate for Parameters in LCA: A case study for the functional unit, in: Int J of LCA 13 LCA (3) pp 265 - 277 (2008).