Analysis of conformance of LCA databases with the Shonan Global Guidance Principles

Development of Shonan Guidance Conformance Criteria

DRAFT version September 2015, after-LCM conference

Version 2.1 for consultation

Copyright © UNEP DTIE (2015)

This publication may be reproduced, in whole or in part and in any form, for educational or non-profit purposes without special permission from the copyright holder, provided acknowledgement of the source is made. UNEP appreciates receiving a copy of any publication that uses this publication as a source.

No use of this publication may be made for re-sale or for any other commercial purpose whatsoever without prior permission in writing from the United Nations Environment Programme.

Disclaimer

The designations employed and the presentation of the material in this publication does not imply the expression of any opinion whatsoever on the part of UNEP concerning the legal status of any country, territory, city or area or of its authorities, or concerning delimitation of its frontiers or boundaries.

Moreover, the views expressed do not necessarily represent the decision or the stated policy of UNEP, nor does citing of trade names or commercial processes constitute endorsement.

Analysis of conformance of LCA databases with the Shonan Global Guidance Principles

Development of Shonan Guidance Conformance Criteria

Acknowledgements

Producer

This document has been produced by the UNEP DTIE

Supervision, technical editing and support

Llorenç Mila i Canals (UNEP DTIE)

Authors

Andreas Ciroth (GreenDelta), ciroth@greendelta.com

Franziska Möller (GreenDelta)

Contributors and reviewers (in alphabetical order)

Bruce Vigon (SETAC)

Guido Sonnemann (Bordeaux University)

Mary Ann Curran (BAMAC)

N.N. (tbc)

N.N. (tbc)

N.N. (tbc)

Design and lay-out

UNEP DTIE

Content

DRAFT

List of figures

No table of figures entries found.

DRAFT

List of tables

DRAFT

1 Executive summary

This document proposes a set of indicators which aim to reflect the Shonan Guidance Principles (SGP), [1], and are meant to be applied for an assessment of SGP conformance, on a data set level.

The indicators are provided on a data set and on a database level. Data sets are distinguished into unit and into aggregated data sets, since the SGP poses additional requirements for the aggregation procedure and also in parts for aggregated data sets.

The indicators are distinguished into “evaluative” indicators that evaluate SGP conformance. For these, in each case, a threshold is proposed that needs to be met by data sets to be fully SGP conformant. Other criteria are called informative, their result does not influence SGP conformance but they seem anyhow of relevance.

Indicators may have different measurement units; some are of a simple Boolean value (yes/no), many are evaluated in an ordinal scale, from 1 to 5, with 1 being best and 5 being the worst value.

Altogether the following criteria are proposed:

Table 1: Criteria for SGP conformance, for unit processes, aggregated processes, and entire databases

scope	criterion	Shonan conformance threshold	unit
data set, unit process	UP_1 data set goal and scope description available	yes	yes/no
	UP_2 suffiencient data set documentation	2	1…5
	UP_3 dataset quality assurance/validation performed	3	1…5
	UP_4 data set structure	yes	yes/no
	UP_5 uncertainties described		1…5
	UP_6 geographic context available		1…5
	UP_7 temporal context available		yes/no
	UP_8 product production volumes available		yes/no
	UP_9 technology level available		yes/no
aggregated data set	AP_1 data set goal and scope description available	yes	yes/no
	AP_2 sufficient data set documentation	2	1…5
	AP_3 dataset quality assurance/validation performed	3	1…5
	AP_4 data set structure	yes	yes/no
	AP_5 uncertainties described		1…5
	AP_6 geographic context available		1…5
	AP_7 temporal context available		yes/no
	AP_8 product production volumes available		yes/no
	AP_9 technology level available		yes/no
database	DB_1 harmonized lists	2	1..5
	DB_2 database protocol available	yes	yes/no
	DB_3 aggregated data sets clearly distinguished	yes	yes/no
	DB_4 supported ISO 14048 compliant data formats		yes/no

Some criteria need further discussion in the community. A first proposal is however detailed in this document, but they are clearly highlighted as “under discussion” and presented in a different color code.

- UP_5 Uncertainties described for the DS and AP_5 Uncertainties described for the DS
- UP_6 geographic context available and AP_6 geographic context available

- AP_7 temporal context available

As next steps, the proposed indicators need to be submitted to external feedback, and will then afterwards be applied on relevant databases, ideally in a guided and supervised self-assessment of database owners.

DRAFT

2 Background and motivation

The UNEP Shonan Guidance Principles [1] are one important result of an international Pellston workshop in 2011 in Shonan Village, Kanagawa Prefecture, Japan, which was organised by the UNEP/ SETAC Life Cycle Initiative. In this workshop, around 50 well-known, experienced LCA scientists and practitioners convened and discussed Life Cycle Inventory (LCI) databases and agreed on interoperability and management principles.

Aim of this text is to "operationalise" these Principles (SGP, in the following) by proposing a set of criteria which can be used for measuring and assessing conformance of datasets with these SGPs. This will in the end allow assessing whether and why databases and data sets in databases are conform with the SGPs, which is not possible with the existing, broadly accepted SGP report text [1].

The criteria presented in the text fully rely on the SGP report. Aspects not covered in this report, for example the share of unit process data sets in a database, or the supported Life Cycle Impact Assessment methods, are not reflected in the criteria although they might be of practical relevance.

This current, final version reflects input from several stakeholders and also input from a workshop held directly after the LCM conference in Bordeaux, France. It is therefore termed "final", although a further discussion of criteria once they are tested more in detail could make sense.

DRAFT

3 Developing the Shonan Guidance Principles criteria

The Shonan Guidance Principles document distinguishes unit processes and aggregated process data set development, and data management and review, in different chapters. These three elements can be seen as fundamental for an LCI database which is based on processes[1]. The criteria reflect this distinction.

The criteria are developed following three main principles, being:

1. Comprehensive (capture spirit and content of SGP as much as possible)
2. Clear (easy to understand, as short as possible, no double-counting and overlaps)
3. Concrete (easy to measure and process)

The criteria will have different values and scopes: Some criteria have Boolean values (yes/no), some criteria have ordinal values with a score ranging from e.g. 1 to 5, 1 being the best value, 5 the worst.

Finally, the criteria are of different priority. Some criteria are considered as being essential for reflecting the SGP, and if the investigated database or data set fails to score above a certain threshold regarding one of these criteria, the database or data set is considered as being non-conforming to the SGP. These criteria are called evaluative. Other criteria are for information only, their score does not influence SGP performance. These criteria are called informative in the following.

Apart from the SGP report, ISO standards for LCA, specifically ISO 14040 [3], 14044 [4], 14048 [2] are relevant for developing the criteria.

DRAFT

[1] In contrast to Input/Output databases which are based on sectors.

4 The Shonan Guidance Principles criteria in detail

The criteria are provided separately for unit processes, aggregated processes, and the entire database.

4.1 Unit processes

Unit processes are defined in the SGP as “smallest element considered in the life cycle inventory analysis”, referring to ISO 14040 [3], [1, p. 54].

The following criteria are proposed.

4.1.1 Evaluative criteria

UP_1. Data set goal and scope description available

Description and motivation	A goal and scope description is essential for assessing various aspects of data set data quality.
SGP reference	• Several locations, • e.g. [1, p 62]: • “a completeness check at the unit process dataset level is the process of verifying whether information from that dataset is sufficient to reach conclusions that are in accordance with the goal and scope definition of the unit process dataset” • [1, p. 77]: • “one should clearly define and document the goal and scope as follows: • product (good or service) […]. In particular, information should be provided on - properties and functions of the product; - geographical, temporal, and technological validity of the dataset (e.g., steel produced in Europe in 2010 via electric arc furnace); and • modelling approach and the types of uses for which the […] dataset is suitable. • intended level of verification or review of the […] dataset. • LCIA requirements to be met and the elementary flows to be included. • data quality requirements. • intended audience.”
Unit	yes/no
Unit values	yes: understandable and sufficiently complete goal and scope description exists; no: otherwise
Shonan conformance threshold	Yes

UP_2. Sufficient data set documentation

Description and motivation	Only with sufficient documentation of modelling procedure, sources, and limitations, the data set can be used correctly, and understood by an LCA practitioner and modeller. For multi-functional processes, information about the allocation or system expansion procedure applied is also important.
SGP reference	• Several locations, e.g. [1, p 62]: • "The dataset documentation must appropriately describe the process and the achieved accuracy, precision, and completeness, as well as any limitations" • [1, p 86]: • "In a perfect dataset, all of the information or metadata needed to describe the quality and usability of a dataset for a given purpose would be included […] We strongly recommend that each dataset be given a unique name and a unique ID that includes a version number as well as a product or process description, […] a classification, such as by NACE code, […] a system description (e.g., included processes, intended downstream use of the system outputs, specific single-plant or market average representation, suitability for consequential modelling ." • [1, p 87]: • "We strongly recommend that, for each flow within a dataset, the origin of the data be documented and references be provided. For primary data (for instance, collected by means of interviews, surveys, questionnaires, bookkeeping, tools, or measurements), the origin shall be denoted as measured, calculated, or estimated. For secondary data (for instance, assembled by means of interviews, statistics, or literature review), the references shall be appropriately cited." • [1, p 87]: • "We strongly recommend that the documentation included with the dataset specify whether the process is unallocated or allocated. If allocated, the description is to include which allocation methods have been used: system expansion, that is, avoiding allocation, with details on replaced processes (and source of data); physical causality, such as mass, energy, or stoichiometric allocation; economic allocation, with cost information and source of data; or end of life or recycling, for example, 1:1 assignment to current and subsequent life cycle. Further, we strongly recommend that the documentation included with the dataset specify whether there are any unallocated flows remaining. Finally, to facilitate review and allow sensitivity analysis, we recommend that when datasets are allocated the associated unallocated datasets also be provided." • [1, p. 60]: • "It is good practice to distinguish missing values from zero. When the data is not clear, it should be entered as '?' which should be distinguished from the entry '0' when clearly not used or emitted How missing values are dealt with should be documented.
Unit	ordinal, 1-5

Unit values	1 sufficient documentation easily available for the data set (either within the dataset or in a separate document) for modelling procedure, including allocation, sources, system boundary setting, limitations, treatment of missing values; for sources also on the flow level; for allocated data sets, also the unallocated data sets are available 2 one of the aspects modelling procedure, allocation, sources, system boundary setting, limitations, treatment of missing values insufficiently described or not easily available for the data set, for sources also on the flow level 3 two of the aspects modelling procedure, allocation, sources, limitations, treatment of missing values insufficiently described or not easily available for the data set 4 insufficient documentation (three or more of the aspects modelling procedure, allocation, sources, limitations, treatment of missing values lacking; or all not easily available for the data set) 5: no documentation available
Shonan conformance threshold	2

UP_3. Data set quality assurance / validation performed?

Description and motivation	[1, p 61]: "Validation is understood as the procedure of ascertaining that the developed unit process dataset represents the "real" process dataset well, by comparing the behaviour of the developed process to that of the real one". Without validation it is not clear whether a process data set reflects what it should. The SGP list several techniques relevant for a validation: Completeness check, plausibility check, investigation of sensitivity and uncertainty, and consistency check [1, pp 61] The validation techniques can be applied in different levels of detail. Using statistical tools and balance checks, considering mathematical relations, are seen as detailed plausibility and validation checks.
SGP reference	see above, description
Unit	ordinal, 1-5
Unit values	1 completeness check, plausibility check, investigation of sensitivity, consistency check performed and documentation available, or review according to the UNEP/SETAC dataset review criteria and procedure performed, 2 completeness check, plausibility check, investigation of

	sensitivity, consistency check performed 3 plausibility check and completeness check performed 4 plausibility or completeness check performed or the data set is only partially checked 5 no information available or no validation performed
Shonan conformance threshold:	3

UP_4. Data set structure

Description and motivation	It is good practice and e.g. suggested by ISO 14048 that a data set contains the following structure: Administrative information; modelling and validation; process description; inputs and outputs. Further, a functional unit or one or several (for multi-product processes) quantitative reference flows. This structure is also supported by all modern LCA data exchange formats (e.g. ILCD; EcoSpold01, EcoSpold02). A data set that does not follow this generic structure can only be used with lots of additional modification effort in any modern LCA software, independently of the specific data format.
SGP reference	e.g. [1, p. 87]: "Each dataset must have one of the following: • functional unit in case of a (partial) product system with a defied use (which need not be quantified as a flow, but can be any quantified use, e.g., m_2),• reference flow in the case of a single-output or allocated process, or • several reference flows in case of an unallocated multi-output process."
Unit	Boolean, yes/no
Unit values	yes: data set structure contains, and consists of, the following elements: administrative information; modelling and validation; process description; inputs and outputs. Further, a functional unit or one or several (for multi-product processes) quantitative reference flows. no: data set structure does not contain, and does not consist of, the following elements: administrative information; modelling and validation; process description; inputs and outputs. Further, a functional unit or one or several (for multi-product processes) quantitative reference flows.
Shonan conformance threshold:	Yes

4.1.2 Informative criteria

UP_5. Uncertainties described for the data set

This criteria requires further discussion

Description and motivation	All flows in LCA are uncertain; without any information of the uncertainty related to the flow and/or to the entire process data set, a user is not able to understand the quantitative reliability of the data set. Uncertainty of different flows will be different, thus uncertainty information should be provided for the different flows ideally, not only for the data set as a whole. Uncertainty is thereby defined in the SGP somewhat loosely as "uncertainty: Uncertainty of the information (e.g., data, models, and assumptions)" [1, table 4.1]. According to the SGP, this is the ISO description.
SGP reference	[1, p 63]: "[…] uncertainty analyses can also be used to assess the reliability of each flow of the unit process dataset derived from raw data and mathematical relationships."
Unit	ordinal, 1-5
Unit values	1 quantitative uncertainty assessment on the level of flows, based on empirical information 2 uncertainty assessment partly based on expert judgement, on the level of flows 3 expert-judgement based uncertainty on the flow level. 4 expert-judgement based uncertainty on the data set level. 5 no information available or no uncertainty assessment performed
Shonan conformance threshold:	-

UP_6. Geographic context available

This criteria requires further discussion

Description and motivation	The geographic context is relevant for evaluating the data quality when the data set is used in any LCA study, and it is also useful for consistency checks within the data set
SGP reference	[1, p. 55]: "In particular, the [process data set] developer should consider the following: […] geographical area covered;" [1, p 87]: "The following items should be included [in the data set documentation]: […] geographic context, including spatial, reference region, or site."
Unit	ordinal, 1-5

Unit values	1 KML or GML Geocode or similar 2 longitude latitude available for the process 3 country code following ISO 3166-1 alpha-2 or ISO 3166-1 alpha-3 4 free text 5 not available
Shonan conformance threshold:	-

UP_7. Temporal context available

Description and motivation	The process time reference (or temporal context) is relevant for evaluating the data quality when the data set is used in any LCA study, and it is also useful for consistency checks within the data set
SGP reference	[1, p. 55]: "In particular, the [process data set] developer should consider the following: […] time period covered" [1, p. 87]: "In particular, the [process data set] developer should consider the following: […] temporal context, including - reference year of data collection; - year of calculation; - daily, seasonal, or annual variations, as necessary; - other temporal information, such as the temporal profile of emissions (e.g., carbon provenance); - in case of combined references, the year best represented; and - a temporal validity statement may be included (e.g., an expiry date »valid until …«)."
Unit	Boolean, yes/no
Unit values	Yes: Process time reference available No: Process time reference not available
Shonan conformance threshold:	-

UP_8. Product production volumes available

Description and motivation	Production volume information is needed for creation of market mixes and in consequential modelling also for determining the overall market size and thus whether increased demand can be satisfied by a given market. This information can also be obtained from other sources, but it is helpful if it is provided already with the LCA data set, in sufficient quality.
SGP reference	[1, p 60]: "the specification of numerical annual production volumes is essential for linking of datasets into production or consumption mixes"
Unit	Boolean, yes/no
Unit values	Yes: Annual production volume information for the product available, for a product, time, and region that fits to the product of the data set No: Annual production volume information for the product, for a product, time, and region that fits to the product of the data set, not available
Shonan conformance threshold:	-

UP_9. Product technology level available

Description and motivation	The product technology level is relevant for consequential modelling and also for evaluating the data quality when the data set is used in any LCA study. Product technology level is understood as quantitative or qualitative description of the relative competitiveness, level of sophistication, or assessment of the state of development of the production technology used.
SGP reference	[1, p 60]: "the specification of numerical annual production volumes is essential for linking of datasets into production or consumption mixes" [1, p 90]: "we recommend that the process [data set] data include a quantitative or qualitative description of the relative competitiveness, level of sophistication, or assessment of the state of development of the production technology used"
Unit	Boolean, yes/no
Unit values	Yes: Product technology level description available No: Product technology level description not available
Shonan conformance threshold:	-

4.2 Aggregated processes

The SGP report contains a chapter on unit processes and a separate chapter on process aggregation, but, somehow surprisingly, not on aggregated processes, although these are mentioned, see e.g. [1, Fig. 2.1, p. 54]. A footnote in the report [1, p.54, footnote 1] even claims that aggregated and unit processes in the understanding of ISO are both considered as unit processes[2], while of course the ISO standard clearly distinguishes unit processes from aggregated processes. A unit process is "smallest element considered in the life cycle inventory analysis for which input and output data are quantified" (ISO 14040, clause 3.34). Unit processes are further characterised as follows (ISO 14040, clause 4.4): "Unit processes are linked to one another by flows of intermediate products and/or waste for treatment, to other product systems by product flows, and to the environment by elementary flows." And the SGP defines aggregated dataset as "An activity dataset showing the aggregated environmental exchanges and impacts of the product system related to one specific product from the activity" [1, glossary].

Aggregated processes are result of an aggregation procedure and therefore some aspects of the aggregation procedure reflect in these processes. Since the procedure itself cannot be checked when analysing databases, focus for the following criteria are the process data sets, and here especially the aggregated process data sets in the sense of ISO, i.e. those data sets where an aggregation has been performed.

Many of the criteria described above for unit processes apply also for aggregated processes. For these, however, some of the evaluation values are different. These differences are highlighted in **bold blue**.

The following criteria are proposed, again distinguished into evaluative and informative criteria.

4.2.1 Evaluative criteria

AP_1. Data set goal and scope description available

Description and motivation	A goal and scope description is essential for assessing various aspects of data set data quality.
SGP reference	• several locations, e.g. [1, p 62]: • "a completeness check at the unit process dataset level is the process of verifying whether information from that dataset is sufficient to reach conclusions that are in accordance with the goal and scope definition of the unit process dataset" • [1, p. 77]: • "[…] one should clearly define and document the goal and scope as follows: • product (good or service) that the aggregated process dataset will represent (reference flow). In particular, information should be provided on

[2] [1, p. 54, footnote 1]: "Unit process" is defined as "smallest element considered in the life cycle inventory analysis" in ISO 14040. Therefore, when so-called "unit process datasets" and "aggregated process datasets" in this database guidance are applied in an LCI analysis, both of them will be unit processes."

	- properties and functions of the product; - geographical, temporal, and technological validity of the dataset (e.g., steel produced in Europe in 2010 via electric arc furnace); and - **in the case of horizontally averaged datasets, a clear statement as to whether the average represents a production or a consumption mix for the region.** • **motivation for aggregation** • **type (horizontal, vertical, engineering-based) and level (e.g., gate-to-gate, partially- or fully-terminated) of aggregation.** • modelling approach and guidelines to be followed in the aggregation (e.g., attributional, consequential, in line with database-specific guidelines) and the types of uses for which the aggregated dataset is suitable. • intended level of verification or review of the aggregated dataset. • LCIA requirements to be met and the elementary flows to be included. • data quality requirements. • intended audience."
Unit	yes/no
Unit values	Yes: understandable and sufficiently complete goal and scope description exists; No: otherwise
Shonan conformance threshold	Yes

AP_2. Sufficient data set documentation

Description and motivation	Only with sufficient documentation of modelling procedure, sources, **aggregation procedure**, system boundary setting, and limitations, the data set can be used correctly, and understood by an LCA practitioner and modeller. For multi-functional processes, information about the allocation or system expansion procedure applied is also important. **Sufficient documentation covers also sufficient documentation regarding the unit processes used in the aggregation.**
SGP reference	• several locations, e.g. [1, p 62]: • "The dataset documentation must appropriately describe the process and the achieved accuracy, precision, and completeness, as well as any limitations" • [1, p 86]: • "In a perfect dataset, all of the information or metadata needed to describe the quality and usability of a dataset for a given purpose would be included […]. We strongly recommend that each dataset be given a unique name and a unique ID that includes a version number as well as a product or process description, […] a classification, such as by NACE code, […] a system description (e.g., included processes, intended downstream use of the system outputs, specific single-plant or market average representation,

	suitability for consequential modelling." • [1, p 87]: • "We strongly recommend that, for each flow within a dataset, the origin of the data be documented and references be provided. For primary data (for instance, collected by means of interviews, surveys, questionnaires, bookkeeping, tools, or measurements), the origin shall be denoted as measured, calculated, or estimated. For secondary data (for instance, assembled by means of interviews, statistics, or literature review), the references shall be appropriately cited." • [1, p 87]: • "We strongly recommend that the documentation included with the dataset specify whether the process is unallocated or allocated. If allocated, the description is to include which allocation methods have been used: system expansion, that is, avoiding allocation, with details on replaced processes (and source of data); physical causality, such as mass, energy, or stoichiometric allocation; economic allocation, with cost information and source of data; or end of life or recycling, for example, 1:1 assignment to current and subsequent life cycle. Further, we strongly recommend that the documentation included with the dataset specify whether there are any unallocated flows remaining. Finally, to facilitate review and allow sensitivity analysis, we recommend that when datasets are allocated the associated unallocated datasets also be provided." • [1, p. 60]: • "It is good practice to distinguish missing values from zero. When the data is not clear, it should be entered as '?' which should be distinguished from the entry '0' when clearly not used or emitted How missing values are dealt with should be documented.
Unit	ordinal, 1-5
Unit values	1 sufficient documentation easily available for the data set for modelling procedure, including allocation, sources, **aggregation procedure, aggregated processes,** system boundary setting, limitations, treatment of missing values; for sources also on the flow level; for allocated data sets, also the unallocated data sets are available 2 one of the aspects modelling procedure, allocation, sources, **aggregation procedure, aggregated processes,** system boundary setting, limitations, treatment of missing values insufficiently described or not easily available for the data set, for sources also on the flow level 3 two of the aspects modelling procedure, allocation, sources, **aggregation procedure, aggregated processes,** system boundary setting, limitations, treatment of missing values insufficiently described or not easily available for the data set 4 insufficient documentation (at least three of the aspects modelling procedure, allocation, sources, **aggregation**

	procedure, aggregated processes, system boundary setting, limitations, treatment of missing values lacking; or all not easily available for the data set) 5: no documentation available
Shonan conformance threshold	2

AP_3. Data set quality assurance / validation performed?

Description and motivation	[1, p 61]: "Validation is understood as the procedure of ascertaining that the developed […] dataset represents the "real" process dataset well, by comparing the behaviour of the developed process to that of the real one". Without validation it is not clear whether a process data sets reflects what it should. The SGP list several techniques relevant for a validation: Completeness check, plausibility check, investigation of sensitivity and uncertainty, and consistency check [1, pp 61]
SGP reference	See above, description, and also: [1, p. 82]: "**The generation of the aggregated datasets should be validated. This means that all unit processes and their interlinkages should be checked with regard to data plausibility and completeness, uncertainty, and methodological consistency**."
Unit	ordinal, 1-5
Unit values	1 completeness check, plausibility check, investigation of sensitivity, consistency check performed and documentation available, or review according to the UNEP/SETAC dataset review criteria and procedure performed, **for the aggregated unit processes and for the aggregated data set** 2 completeness check, plausibility check, investigation of sensitivity, consistency check performed, **for the aggregated unit processes and for the aggregated data set** 3 plausibility check and completeness check performed, **for the aggregated unit processes** 4 plausibility or completeness check performed, **for the aggregated unit processes** 5 no information available or no validation performed
Shonan conformance threshold:	2

AP_4. Data set structure

Description and motivation	It is good practice and e.g. suggested by ISO 14048 that a data set contains the following structure: Administrative information; modelling and validation; process description; inputs and outputs. Further, a functional unit or one or several (for multi-product processes) quantitative reference flows. This structure is also supported by all modern LCA data exchange formats (e.g. ILCD; EcoSpold01, EcoSpold02). A data set that does not follow this generic structure can only be used with lots of additional modification effort in any modern LCA software, independently of the specific data format.
SGP reference	e.g. [1, p. 87]: "Each dataset must have one of the following: • functional unit in case of a (partial) product system with a defied use (which need not be quantified as a flow, but can be any quantified use, e.g., m_2),• reference flow in the case of a single-output or allocated process, or • several reference flows in case of an unallocated multi-output process."
Unit	Boolean, yes/no
Unit values	yes: data set structure contains, and consists of, the following elements: administrative information; modelling and validation; process description; inputs and outputs. Further, a functional unit or one or several (for multi-product processes) quantitative reference flows. no: data set structure does not contain, and does not consist of, the following elements: administrative information; modelling and validation; process description; inputs and outputs. Further, a functional unit or one or several (for multi-product processes) quantitative reference flows.
Shonan conformance threshold:	Yes

4.2.2 Informative criteria

AP_5. Uncertainties described for the data set

This criteria requires further discussion

Description and motivation	All flows in LCA are uncertain; without any information of the uncertainty related to the flow and/or to the entire process data set, a user is not able to understand the quantitative reliability of the data set. Uncertainty of different flows will be different, thus uncertainty information should be provided for the different flows ideally, not only for the data set as a whole.

	Uncertainty is thereby defined in the SGP somewhat loosely as "uncertainty: Uncertainty of the information (e.g., data, models, and assumptions)" [1, table 4.1]. According to the SGP, this is the ISO description. **For aggregated processes, the uncertainty comes from the aggregated unit processes and additionally from the aggregation procedure.**
SGP reference	[1, p 63]: "[…] uncertainty analyses can also be used to assess the reliability of each flow of the unit process dataset derived from raw data and mathematical relationships." [1, p. 89]: **"We strongly recommend that the following additional information be provided for aggregated process datasets: […] an assessment of the uncertainty […]";**
Unit	ordinal, 1-5
Unit values	1 **uncertainties described for the aggregated data set, starting from quantitative uncertainty assessment on the level of flows** for the unit processes, based on empirical information, 2 uncertainty assessment partly based on expert judgement, on the level of flows, **starting from quantitative uncertainty assessment of unit processes** 3 expert-judgement based uncertainty on the flow level, **starting from uncertainty assessment of the aggregated unit processes**. 4 expert-judgement based uncertainty on the data set level. 5 no information available or no uncertainty assessment performed
Shonan conformance threshold:	-

AP_6. Geographic context available

Description and motivation	The geographic context is relevant for evaluating the data quality when the data set is used in any LCA study, and it is also useful for consistency checks within the data set. **For aggregated data sets, the geolocation of the final process producing the reference product and also the location of the processes that were aggregated over the supply chain is relevant.**
SGP reference	[1, p. 55]:

	"In particular, the [process data set] developer should consider the following: […] geographical area covered;"
Unit	ordinal, 1-5
Unit values	1 KML or GML Geocode for the reference process **and the aggregated processes** 2 longitude latitude available for the reference process **and the aggregated processes** 3 country code following ISO 3166-1 alpha-2 or ISO 3166-1 alpha-3 for the reference process **and the aggregated processes** 4 free text or only information for the reference process 5 not available
Shonan conformance threshold:	-

AP_7. Temporal context available

This criteria requires further discussion

Description and motivation	The process time reference (or temporal context) is relevant for evaluating the data quality when the data set is used in any LCA study, and it is also useful for consistency checks within the data set. **For aggregated processes, the time reference should be available for the reference process and for the aggregated processes in the supply chain, for example in the form of a time span.**
SGP reference	[1, p. 55]: "In particular, the [process data set] developer should consider the following: […] time period covered" 1, p. 87]: "In particular, the [process data set] developer should consider the following: […] temporal context, including - reference year of data collection; - year of calculation; - daily, seasonal, or annual variations, as necessary; - other temporal information, such as the temporal profile of emissions (e.g., carbon provenance); - in case of combined references, the year best represented; and - a temporal validity statement may be included (e.g., an expiry date »valid until …«)."
Unit	Boolean, yes/no
Unit values	Yes: Process time reference available for the reference process and **the processes in the aggregated supply chain** No: Process time reference not available **or only available for**

	the reference process
Shonan conformance threshold:	-

AP_8. Product production volumes available

Description and motivation	Production volume information is needed for creation of market mixes and in consequential modelling also for determining the overall market size and thus whether increased demand can be satisfied by a given market. This information can also be obtained from other sources, but it is helpful if it is provided already with the LCA data set, in sufficient quality.
SGP reference	[1, p 60]: "the specification of numerical annual production volumes is essential for linking of datasets into production or consumption mixes"
Unit	Boolean, yes/no
Unit values	Yes: Annual production volume information for the product available, for a product, time, and region that fits to the product of the data set No: Annual production volume information for the product, for a product, time, and region that fits to the product of the data set, not available
Shonan conformance threshold:	-

AP_9. Product technology level available

Description and motivation	The product technology level is relevant for consequential modelling and also for evaluating the data quality when the data set is used in any LCA study. Product technology level is understood as quantitative or qualitative description of the relative competitiveness, level of sophistication, or assessment of the state of development of the production technology used.
SGP reference	[1, p 60]: "the specification of numerical annual production volumes is essential for linking of datasets into production or consumption mixes" [1, p 90]: "we recommend that the process-level data include a quantitative or qualitative description of the relative competitiveness, level of sophistication, or assessment of the state of development of the production technology used"
Unit	Boolean, yes/no

Unit values	Yes: Product technology level description available No: Product technology level description not available
Shonan conformance threshold:	-

4.3 Database

4.3.1 Evaluative criteria

DB_1. Harmonized lists

Description and motivation	Using different data sets from one database that do not contain the same set of elementary flows will rarely lead to a consistent life cycle model. Ideally, different databases use the same reference list. Also other database elements should be consistent (units, contact names, asf.). For the calculation result, harmonization of flows is most relevant. Finally, also the category system a database uses should be harmonized and consistent.
SGP reference	[1, p. 89]: "We recommend the use of a globally harmonized reference list of elementary flow names as the primary condition for interoperability of datasets and databases"
Unit	ordinal, 1-5
Unit values	1 Harmonized lists for flows, units, contact names, process names, using either globally harmonized lists or lists where international accepted mapping files exist 2 Harmonized lists for flows, units, contact names, process names, database-wide 3 Harmonized lists for flows, database-wide 4 Harmonized lists for impact-assessment-relevant elementary flows, database-wide 5 No harmonized lists
Shonan conformance threshold:	2

DB_2. Database protocol available

Description and motivation	[1, p. 85]: "To facilitate understanding by the user community and to encourage preparation of datasets for inclusion, each database manager should prepare a protocol document describing the contents, format of datasets, method for feedback on datasets, and requisites for inclusion of datasets within the database"

	[1, p.93]: "Some suggested elements of the protocol include • formats in which a dataset is available to the users. • format in which the datasets should be provided to the database (post-review with full documentation). • documentation required with a dataset (see also above), including - required metadata, for example, boundaries, processes, representativeness of technology, allocation, aggregation, geography, completeness, time; - comprehensive list of types of processes generally found in the database, regions typically covered, date of last full review; and - quality assurance process for datasets and database, specifically validation and review procedures for dataset inclusion."
SGP reference	See description
Unit	Boolean, yes/no
Unit values	Yes: Database protocol available No: Protocol not available
Shonan conformance threshold:	Yes

DB_3. Aggregated data sets clearly distinguished

Description and motivation	This criterion exists for technical reasons only; databases that do not clearly distinguish aggregated from other data sets do not allow separate application of criteria for aggregated data sets
SGP reference	-
Unit	Boolean
Unit values	Yes/no
Shonan conformance threshold:	Yes

4.3.2 Informative criteria

DB_4. Supported ISO 14048 compliant data formats

Description and motivation	Data formats were explicitly excluded from the SGP. However, it seems relevant to include information about the different, ISO 14048-compliant data formats a database is able to deliver data sets in. This is independent from how the data sets are stored in the database since these formats are exchange formats only.

SGP reference	[1, p. 85]: "[…] each database manager should prepare a protocol document describing the […] format of datasets […]"
Unit	Boolean
Unit values	Yes/no
Shonan conformance threshold:	-

DRAFT

5 Next steps

As next steps, the suggested criteria will need to be evaluated and reviewed by interested stakeholders and afterwards, the criteria will be applied on relevant LCI databases internationally, ideally following a supervised self-assessment of different databases.

DRAFT

6 References

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

[2] ISO/TS 14048:2002, Environmental management -- Life cycle assessment -- Data documentation format

[3] ISO 14040:2006, Environmental management -- Life cycle assessment -- Principles and framework.

[4] ISO 14044:2006, Environmental management -- Life cycle assessment -- Requirements and guidelines.

[5] www.openlca.org

[6] www.ecoinvent.org

[7] www.ecoinvent.org/news-events/news/newsdetails/view/new-cut-off-system-model-for-ecoinvent-version-3/

[8] http://unstats.un.org/unsd/cr/registry/isic-4.asp

[9] www.ecoinvent.org/fileadmin/documents/en/Data_Quality_Guidelines/01_DataQualityGuideline_v3_Final.pdf

[10] www.ecoinvent.org/data-providers/how-to-submit-data/

[11] http://www.gabi-software.com/international/databases/

[12] https://nexus.openlca.org/database/GaBi

[13] http://www.gabi-software.com/support/gabi/gabi-modelling-principles/

[14] http://www.gabi-software.com/fileadmin/gabi/Modelling_Principles/GaBi_Modelling_Principles_2014.pdf

[15] http://eplca.jrc.ec.europa.eu/ELCD3/downloads/datasets/ELCDIII_1.zip

[16] http://cpmdatabase.cpm.chalmers.se/Document/CPM_Report_1999_1_LCI-DocumentationIntroGuide.pdf

[17] http://cpmdatabase.cpm.chalmers.se/AboutDatabase.htm

[18] http://cpmdatabase.cpm.chalmers.se/Document/ManualCPMDB.htm

[19] http://cpmdatabase.cpm.chalmers.se/Scripts/AdvGeneral.asp

[20] http://cpmdatabase.cpm.chalmers.se/ILCD/ILCDinfo.htm

[21] http://eplca.jrc.ec.europa.eu/uploads/MANPROJ-PR-ILCD-Handbook-Nomenclature-and-other-conventions-first-edition-ISBN-fin-v1.0-E.pdf

[22] http://eplca.jrc.ec.europa.eu/uploads/ILCD-Data-Network-Compliance-Entry-level-Version1.1-Jan2012.pdf

[23] http://eplca.jrc.ec.europa.eu/uploads/ILCD-Handbook-Specific-guide-for-LCI-12March2010-ISBN-fin-v1.0-EN.pdf

[24] http://eplca.jrc.ec.europa.eu/?page_id=126

DRAFT

[25] Cochran, W.G.: Sampling Techniques, 3rd edition, John Wiley, New York 1977.

[26] Schwarz, H. Stichprobenverfahren, Oldenbourg, München 1975.

[27] 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).

[28] http://cpmdatabase.cpm.chalmers.se/Document/ESA-DBP_Report_ESA_CPM_2011-1.pdf

DRAFT