METHODOLOGY

cm.chemicals database – methodology document

Version 1.01, September 2022

TÜVRheinland CERTIFIED
Certified Calculation Method
Regular Surveillance
www.tuv.com
ID 0000081021

carbonminds

Authors:

Laura Stellner, Dr. Arne Kätelhön, Oskar Vögler, Ronja Hermanns, Prof. Sangwon Suh, Prof. André Bardow, Raoul Meys

Citation:

Stellner, Laura; Kätelhön, Arne; Vögler, Oskar; Hermanns, Ronja; Suh, Sangwon; Bardow, André; Meys, Raoul. Methodology cm.chemicals. Version 1.01, September 2022. Carbon Minds GmbH, Cologne.

Acknowledgments:

Helpful comments to the draft version of this document have been received from Christian Zibunas, Marvin Bachmann, Leonard Müller, Birgit Himmelreich, and Johanna Kleinekorte. We thank them all for their kind input.

carbonminds

www.carbon-minds.com

https://www.linkedin.com/company/carbon-minds

For more information, you can contact us

info@carbon-minds.com

Version 1.01 2022, Carbon Minds

Content

1. Introduction

Life Cycle Assessment (LCA) is a powerful tool to determine the environmental impacts of products and services. LCA considers a product's interactions with the environment at all stages of the life cycle, from the extraction of raw materials through the production and transportation of products to the final disposal of wastes. LCA determines environmental impacts for a wide range of categories. By providing a holistic view across environmental impact categories and life cycle stages, LCA can pinpoint the sources of environmental impacts and avoid problem shifting between life cycle stages and environmental impact categories. By this means, LCA can provide robust evaluations of environmental impacts of products and enable environmental decision support for industry, politics, academia, and consumers.

To provide a reliable evidence base to decision-makers, LCA studies require a large amount of data on products and services across global production and utilization chains. This includes technical data on all production and utilization steps throughout the life cycle, as well as market and trade data, which enables flows of products and energy to be tracked throughout the global economy.

Gathering this data is especially challenging for industries with highly complex and globalized production chains. The chemical industry has long been recognized as one such case. Chemical production chains typically involve thousands of production sites worldwide. Each of these production sites has site-specific supplies of raw materials and energy, and for any given chemical, typically, more than one reaction pathway exists. These differences lead to sizeable variability in environmental performance. Basing LCA studies on data that accurately represents the chemical's actual production pathway is crucial for offering robust environmental decision support.

To support environmental decision-making on the production and use of chemicals, we have developed cm.chemicals, a comprehensive Life Cycle Inventory (LCI) database focusing exclusively on the production of chemicals. The database covers the production of a large number of chemicals in more than 190 production regions. Regional differences are represented in detail based on plant-level data.

The cm.chemicals database aims to provide LCA practitioners with data that enables them to conduct more representative and reliable LCA studies on the production and use of chemicals that are fully compliant with the ISO standards 14040 and 14044 for Life Cycle Assessment, and with ISO 14067 for Product Carbon Footprint calculations. To ensure the reliability of our data, we adhere to four strict principles:

- **Representative.** Chemical supply chains differ substantially in production regions and for different production technologies within a region. cm.chemicals aims to capture these differences as precisely as possible. Modeling the production of large-volume chemicals using plant-level data reveals differences in environmental impacts of suppliers and regions at an unprecedented level of detail. Furthermore, using an in-depth model of international trade flows in the chemical industry captures the interactions between regions at all production chain stages.
- **Consistent.** Environmental assessments commonly include comparisons aimed at informing a choice between products, technologies, or environmental impact reduction measures. For LCA practitioners to make fair comparisons, a consistent

methodology must be applied across the entire database that informs the assessment. To ensure comparable results, the cm.chemicals database is based on a single methodology, applied consistently to every single dataset in the database.

- **Quality assured.** Developing Life Cycle Inventory data for the chemical industry requires detailed technical data on all production technologies used within the chemical industry and market data and trade balances. Only with stringent data quality controls is it possible to ensure the integrity of the database. For cm.chemicals, this means that professionals have checked every single data and that the underlying methodology has been independently reviewed and certified by TÜV Rheinland. The certification also includes annual monitoring of potential changes.
- **Transparent.** The correct interpretation and LCA study's results require a detailed understanding of the underlying data. Transparency in modeling assumptions and data quality is, therefore, crucial for LCA-based decision-making. This document aims to illustrate the methodology and all modeling assumptions as transparently as possible.

The foundations of the cm.chemicals database lie in years of academic research at RWTH Aachen University. The current version of the database was built and is constantly updated by us, Carbon Minds, a data analytics company spun out of RWTH Aachen University. Delivering high-quality data on the production of chemicals is the principal mission of the company.

This document provides a detailed overview of the cm.chemicals database and illustrates the methodology used for the generation of the datasets. In the following Chapter 2, we will present the methodology used to construct the LCI model of the global chemical industry, which forms the basis for all LCI datasets. In Chapter 3, we illustrate the Goal and Scope of the LCI datasets, as well as the quality of the underlying data. Subsequently, Chapter 4 illustrates specific details of our modeling approach, including transportation, international trade, and waste incineration. Finally, Chapter 5 presents the documentation of the LCI datasets, as well as the data quality ratings specified for each LCI dataset.

This document and the methodology for the cm.chemicals database is reviewed by TÜV Rheinland Energy GmbH in an independent external review. In the annex, a copy of the Review Report by TÜV Rheinland on the ‘Critical Review of the Methodology for the LCI Database "cm.chemicals" by Carbon Minds’ is attached.

2. Overview of the cm.chemicals database

Chapter 2 provides an overview of the four major steps that are performed to provide representative, consistent, quality-assured, and transparent environmental data for chemical and plastic products to our customers. These four steps include (cf. Figure 1):

(1) The collection of state-of-the-art input data.
(2) The compilation of a consistent life cycle inventory model (LCI model) of the chemical industry.
(3) The generation of the cm.chemicals database that is provided to our customers.
(4) The continuous maintenance, updating, and reviewing of the input data, the LCI model, and the cm.chemicals database.

The following sections give a brief overview of each step, while Chapters 3 to 6 provide a more detailed discussion. Section 2.1 summarizes the state-of-the-art input data that is collected and required to build the cm.chemicals database. Section 2.2 gives an overview of how the input data is used to build a representative and consistent LCI model of the chemical industry. Afterward, Section 2.3 summarizes the scope of the datasets available in the cm.chemicals database and provided to our customers. Finally, Section 2.4 summarizes our approach to maintain, update and review input data, the LCI model, and the cm.chemicals database.

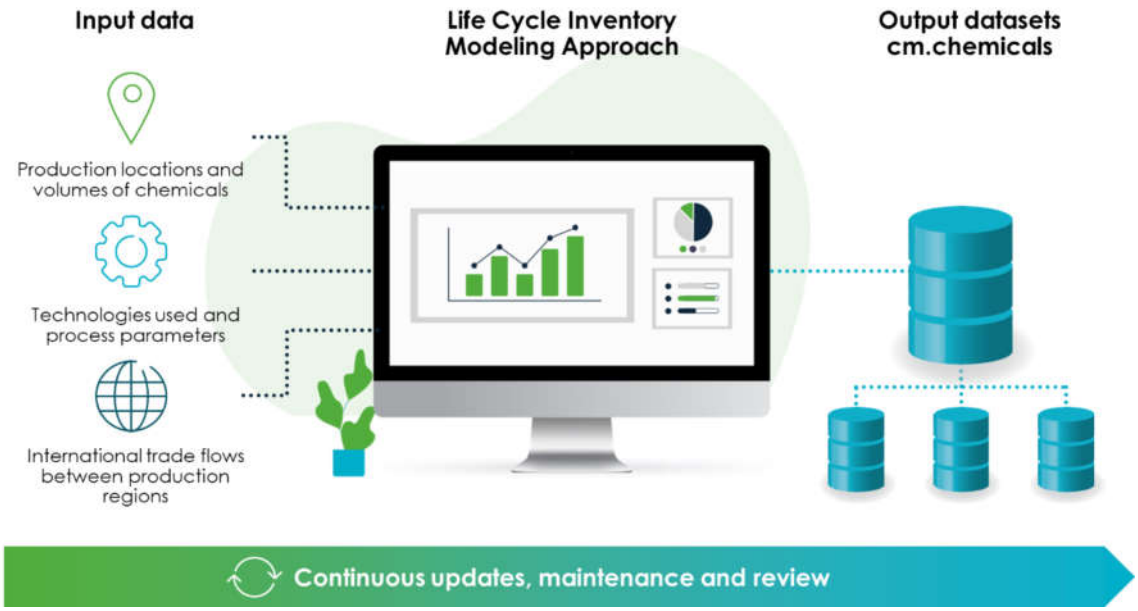

Figure 1: Major four steps to generate the cm.chemicals database and output datasets that are provided to our customers.

2.1 Collection of input data

In order to calculate a consistent LCI model of the chemical industry to evaluate the environmental impacts of chemical products, three types of data are collected:

(1) **Technology data.** This data depicts the full mass and energy balances for each production technology. For instance, this data includes information about the raw material consumption, utilities (e.g., energy use), resource extractions, emissions, co-products, and waste consumption of the steam cracking of naphtha. We use state-of-the-art data providers, extensive literature research, and our own modeling to obtain the data. All data is checked internally by our chemical engineering experts and extended to include data about direct emissions and waste streams. In particular, direct emissions and waste streams largely influence the environmental impacts of chemical production (cf. Section 4.4.3 for a detailed description of waste modeling).

(2) **Market information.** This data includes, for instance, how much ethylene is produced in Ludwigshafen via the steam cracking of naphtha. Furthermore, this data includes meta-information, like the company operating the plant (e.g., the BASF in Ludwigshafen) or the first year of operation. By including the data, we know which chemical is produced in which city, in which volume, by which company, and via which technology.
All market data is obtained from trusted providers and own literature research. Additionally, the input data is accompanied by our literature research to check and validate but also extend the respective data.

(3) **Trade data**. This data depicts, for instance, the imports of ethylene from the Netherlands to Germany. Including this data offers the possibility to understand which chemical is traded between countries.
The data about international trade flows is based on reported information by each country to the United Nations Statistical Division. The data are partly modified in a harmonization step to correct errors and increase data consistency.

More details about the data are provided throughout Chapters 3 and 4.

2.2 Life Cycle Inventory Model

To generate our LCI model, we automatically import and process the input data via Matlab and Python scripts. These scripts ultimately compile a consistent LCI model of the chemical industry (cf. Figure 1). The compilation is carried out in four main steps: In a first step, we explicitly model individual chemical plants. Subsequently, we model how individual plants interact in the context of integrated production sites. Afterward, we specify how individual plants and integrated sites contribute to national production mixes and consumption mixes. Finally, we add an extension layer to the model that depicts additional market dominant and industrially relevant technologies for a single chemical product. A detailed description of the modeling approach is provided in Chapter 4.

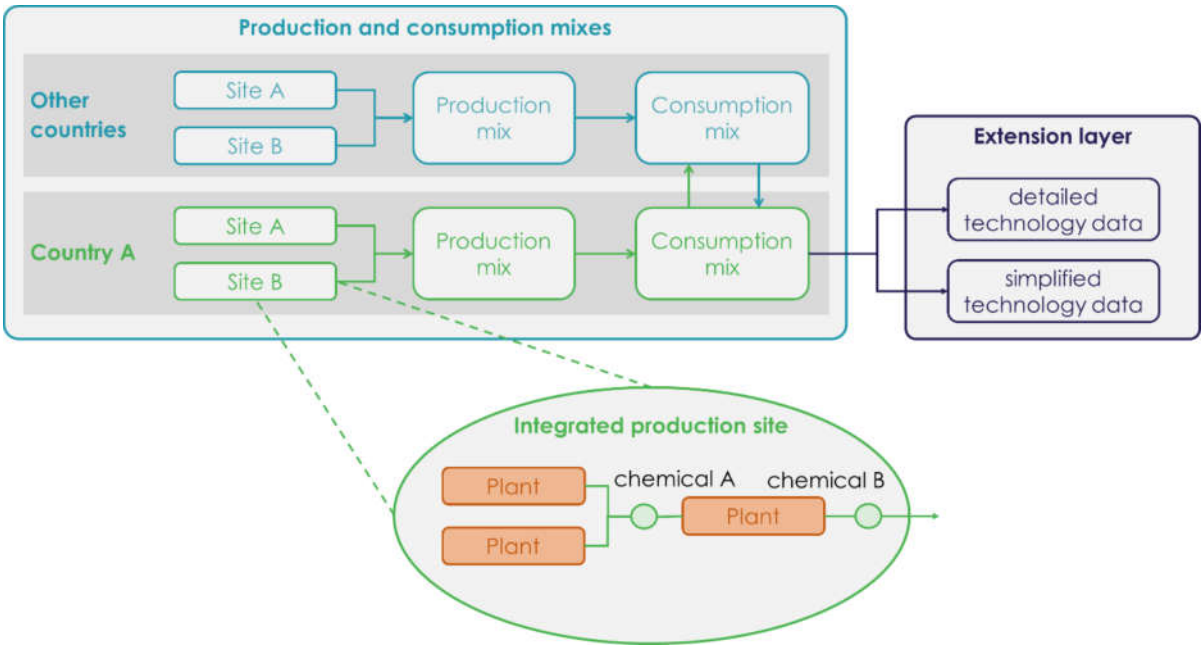

Figure 2: Simplified structure of the LCI model used to generate the cm.chemicals database.

Individual chemical production plants. To model individual chemical plants (cf. Figure 2, bottom, in orange), we use information on the production location for each chemical plant, the production volume, and the exact production technology used in the chemical plant. Afterward, detailed mass and energy balances are included for each chemical plant depending on the respective production technology. By this means, the raw material consumption, utilities (e.g., energy use), resource extractions, emissions, co-products, and wastes are included for each chemical plant.

Integrated production sites. After chemical plants are modeled individually, we model interactions between chemical plants within integrated production sites (cf. Figure. 2, bottom, green circles and lines). These interconnections are based on the mass- and energy balances of each individual chemical plant. Chemical plants located in the same city are assumed to be in the same integrated production site. The modeling of integrated production sites allows us to account for plant-specific supplies of raw materials within an integrated production site.

Production and consumption mixes. After modeling all production plants and integrated production sites in a country, we calculate the average national production and consumption mixes (cf. Figure. 2, top left). The production mix in a country is calculated from the output of all chemical plants, which produce a given chemical in that country. The production mix is calculated based on the country's proportional share of national production contributed by each chemical plant. However, the national production mix of a chemical does not necessarily reflect the consumption of that chemical in that country because parts of the amount consumed may be imported from other countries. Furthermore, parts of the national production may be exported to other countries. Consequently, the consumption mix is represented by the sum of a country's production mix, plus imported chemicals, minus exported chemicals.

Extension layer. The extension layer (cf. Figure. 2, top right) models the production of additional chemicals based on individual production technologies for which market

information is unavailable. Each production technology is represented by its technical flows (e.g., reactants, utilities, co-products, waste) and its elementary flows such as emissions and resource extractions. The technical and elementary flows are obtained either from detailed technology data or simplified technology data (cf. Section 4.4.4).

Further details about the modeling principles are provided in Chapter 4.

2.3 Datasets in the cm.chemicals database

Based on the LCI model of the chemical industry, the aggregated LCI datasets (i.e., system datasets) available in the cm.chemicals database are generated.

These datasets are:

- ***Plant-specific datasets*** represent the production of a chemical in a specific chemical plant in a given site by a given technology and producer. Examples:
  - The production of methanol in Ludwigshafen, Germany by BASF using steam methane reforming.
  - The production of acrylonitrile in Anquing, China by Sinopec using propylene ammoxidation.

- ***Supplier-specific datasets*** represent the production mass-weighted average of all plant-specific datasets for the chemical where the respective plants are owned by the specific supplier in the specific country. Examples:
  - The average production of methanol by BASF in Germany.
  - The average production of acrylonitrile by Sinopec in China.

- ***Technology-specific datasets (core layer)*** represent the production mass-weighted average of all plant-specific datasets for a chemical that utilize the same production technology in a specific country or a broader region. Examples:
  - The average production of methanol by steam methane reforming in Germany.
  - The average production of acrylonitrile by propylene ammoxidation in China.

- ***Production mix datasets*** represent the production mass-weighted average of all plant-specific datasets in a country or a broader region (e.g. Europe) that produce the same chemical. Examples:
  - The average production of methanol in Germany.
  - The average production of acrylonitrile in Europe.
  - The global average production of ethylene.

- ***Consumption mix datasets*** represent the production mass-weighted average of all plant-specific datasets in a country or a broader region (e.g., Europe) producing the chemical plus all mass-weighted imports to that country or broader region for the chemical. Thus, consumption mixes typically include both local production in the country or broader region, and imports from other countries or broader regions. Examples:
  - The average consumption of methanol in Germany.

- The average consumption of acrylonitrile in Europe
- The global average consumption of ethylene.

- ***Technology-specific datasets (extension layer)*** represent one production technology for a specific chemical in a country or broader region by assuming the consumption mixes or other technology datasets (extension layer) as inputs. Examples:
  - The production of chloroform in Germany using the thermal chlorination of methane.
  - The production of cyclohexane in China using the hydrogenation of benzene.

Besides the per default generated LCI datasets, the LCI model can be used to provide ***data-on-demand datasets*** for other geographical scopes, cross-company scopes, and many other options. For further information and discussion about the potential to generate your data-on-demand datasets, feel free to contact us at: info@carbon-minds.com datasets or pay a visit to www.carbon-minds.com.

2.4 Maintenance, updates, and review

New data and technologies, new scientific findings, new methods, or new user requirements lead to a constant opportunity and need to update and maintain the cm.chemicals database. Since maintenance requires constant work, we revise the input data and the model of the chemical industry continuously throughout the year. At least one expert for each input data type ensures that the respective input data is always up-to-date. Through our granular concept and modeling design, all parts of the input data can be maintained separately, and changes can be included in the model of the chemical industry throughout the year.

Based on the continuous maintenance of the input data and the model of the chemical industry, we provide a yearly update of the cm.chemicals database and all output datasets. Furthermore, we use versioning to enable the recalculation of older versions of the cm.chemicals database and respective output datasets.

The cm.chemicals database methodology is designed to provide data for ISO 14040/14044 compliant LCA studies and ISO 14067 compliant PCF studies. Furthermore, the compliance of the methodology to generate the cm.chemicals database with the ISO standards 14040 and14044 is reviewed by TÜV Rheinland Energy GmbH in an independent external review. The review covers the check of methodological approaches, a selected sample of primary and secondary input data, the documentation, the qualification of our employees, the calculation model, and the check of a selected number of output datasets. In the annex, a copy of the Review Report by TÜV Rheinland on the 'Critical Review of the Methodology for the LCI Database "cm.chemicals" by Carbon Minds' is attached.

3. Goal and scope of the cm.chemicals database

This chapter summarizes the goal and scope definition of the cm.chemicals database according to ISO 14040 and 14044. The following Section 3.1 summarizes the goal and the subsequent Section 3.2 the scope of the cm.chemicals database. In Section 3.3, we define data quality criteria, to specify the data quality of the datasets.

3.1 Goal of the cm.chemicals database

The goal of the cm.chemicals database is to provide a representative, consistent, quality-assured, and transparent source of LCI datasets representing the production of chemicals and plastics. By this means, the cm.chemicals database aims to enable LCA practitioners to conduct LCA studies on the production and use of chemicals in an ISO-compliant manner. Following the ISO standards, the following goal can be defined for the cm. chemicals database:

Intended application. The goal of all datasets is to reflect the environmental exchanges and the resulting environmental impacts associated with chemical production chains as precisely as possible. By this means, the cm.chemicals database can be used for any LCA, environmental assessment, carbon footprint assessment, or corporate carbon footprint calculation.

Reasons. Gathering LCI data is frequently seen as the major obstacle when performing LCAs for chemical products.[1] Providing representative, consistent, quality-assured, and transparent LCI data is, thus, the key to enabling LCA practitioners to conduct more representative and reliable LCA studies. Performing representative and reliable LCA studies is crucial for offering robust environmental decision support.

Intended audience. The cm.chemicals database was made for all LCA practitioners in, for instance, research and academia, consulting, politics, or industry.

Comparative assertions. The environmental assessments carried out based on our datasets can support various goals, including comparative assessments to be disclosed to the public. Nevertheless, the cm.chemicals database alone does not intend or support any comparative assertions to be disclosed to the public.

[1] Maranghi, S. and Brondi, C., 2020. Life Cycle Assessment in the Chemical Product Chain. Springer International Publishing.

3.2 Scope of the cm.chemicals database

This section covers the scope definition of the cm.chemicals database and the respective output datasets (cf. Section 2.3).

3.2.1 Function and functional unit

LCA quantifies the environmental impacts of a product system relative to its function, e.g., global warming impact per production of 1 kg of product. The so-called functional unit specifies and quantifies the function of a product system. The definition of a functional unit enables a fair comparison of different product systems serving the same function.

The functional units of all datasets are defined in relation to the datasets' reference products, i.e., the chemicals for which the datasets are compiled. The definition of the functional unit depends on the type of dataset, as shown in Table 1. In the cm.chemicals database, the amount of product, e.g., 1 kg, is used to depict the functional unit.

The functional unit is also highlighted in the documentation of each dataset generated from the cm.chemicals database (cf. Chapter 0)

Table 1. Functional units of different types of datasets.

Type of dataset	**Functional unit**
Plant-specific datasets	Production of 1 kg of the reference product
Supplier-specific datasets	Production of 1 kg of the reference product
Production mix datasets	Production of 1 kg of the reference product
Consumption mix datasets	Supply of 1 kg of the reference product
Technology-specific datasets	Production of 1 kg of the reference product

3.2.2 System boundaries

The system boundaries define which processes, material flows, and energy flows belong to the product system represented by a dataset. **All datasets in the cm.chemicals database have cradle-to-gate system boundaries.**

These cradle-to-gate system boundaries include all relevant processes needed for the production or supply of a respective chemical (cf. Table 1): from the extraction of raw materials ("cradle") through the production of all energy and material flows required to all final commissioning. The system boundaries also include transportation services related to international trade and waste disposal throughout the production chains, as illustrated in Sections 4.4.1 and 4.4.3.

For consumption mix datasets, the "gate" represents the factory gate (entrance gate) of a potential production facility that consumes the respective chemical in the consumption mix region. Consumption mix datasets include transportation processes for both internationally traded intermediate products and the reference product's imports to the target market represented by the consumption mix. For all other datasets, the "gate" represents each "chemical plant's" factory gate (exit gate) producing the chemical represented by the dataset.

For example, for the consumption mix of methanol in Germany, the "gate" represents the factory gate of a chemical production plant in Germany which consumes the consumption mix of methanol. Thus, the consumption mix of methanol in Germany takes into account the regional production of methanol in Germany, imports of methanol from other countries to Germany, and the international transportation of these imports from the export countries to Germany.

In comparison, for the production mix of methanol in Germany, the "gate" represents the factory gates of all chemical production plants in Germany that produce methanol. Thus, no trade and international transportation of the reference product (here methanol) is included in the system boundaries for production mix datasets. However, international trade within the supply chain of the raw materials used for methanol production is considered.

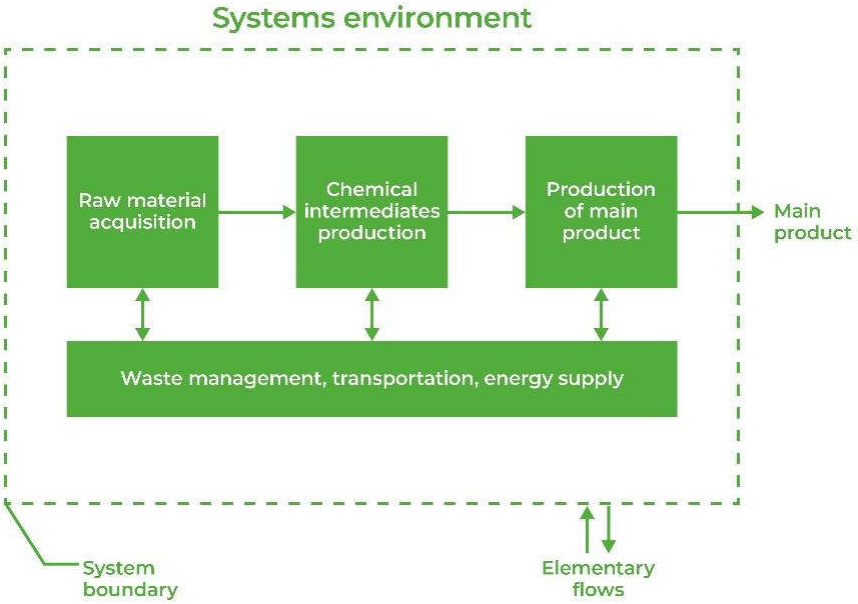

Figure 3. Overview of life cycle stages considered in the cradle-to-gate analysis of chemicals.

3.2.3 Cut-offs

Cut-off criteria are practical guidelines for identifying less relevant flows in a product system to be neglected in an individual assessment. We have neglected flows used in small quantities in a respective process and are not relevant for any other process in the LCI model. The sum of all cut-offs in a respective process is lower than 1% of the mass of all input flows, excluding cooling water.

- The flow is used in small quantities in the respective process (below 1 mass-% of all inputs, excluding cooling water)
- The flow is not relevant for any other process in the database after applying cut-off criteria
- The input data needed for modeling the production of the flow is not available to us

3.2.4 Supported LCIA methods

The calculation of environmental impacts is performed in the so-called Life Cycle Impact Assessment (LCIA). During the LCIA, all elementary flows are attributed with their specific influence on one particular environmental impact. By this means, the overall contribution of several elementary flows to one particular environmental impact can be calculated.

The datasets inside the cm.chemicals database include a list of elementary flows and thus, can be used for any relevant LCIA method. Thus, the LCA practitioners can use the LCIA methods relevant to the particular LCA case study.

3.2.5 Biogenic emissions

Biogenic emissions are greenhouse emissions that are based on a biological source. As an example, castor beans have a biogenic carbon content whereas fossil crude oil has a non-biogenic carbon content. This biogenic carbon content is based on the CO_2 emission (a biogenic emission) uptake during plant growth.

Due to the increased occurrence of biobased processes in the cm.chemicals database, a methodology has been developed to calculate the biogenic carbon content and biogenic carbon emissions of processes that are modeled in the core layer, extension layer, or simplified extension layer (cf. cm.chemicals database – methodology document). In our methodology, we subdivide the carbon dioxide, carbon monoxide, and methane emissions into biogenic carbon emissions and fossil carbon emissions, by tracking the carbon resources throughout the complete supply chain.

The methodology developed to calculate biogenic carbon contents and emissions includes five major steps:

1. For all chemicals, it is checked whether some of the raw materials are biobased. The biogenic carbon share is directly set to zero for all chemicals that are fully based on fossil raw materials and is set to one for all chemicals that are fully based on biogenic raw materials.

Example 1: The biogenic carbon share of carbon disulfide is directly set to 0 as it is produced from the raw materials natural gas and sulfur, which both have a completely fossil-based supply chain in our database.

Example 2: The biogenic carbon share of ethylamine is directly set to 1 as it is produced from the raw materials ethanol and ammonia. Ethanol, which has a completely bio-based supply chain in our database, supplies all carbon atoms for the ethylamine.

2. For all chemicals, which are partly based on biogenic raw materials, each chemical process is reviewed manually to quantify the percentage of carbon content originating from which raw material. This currently applies to around 5% of the chemicals in the cm.chemicals database, as the other 95% of the chemicals are either fully fossil-based or fully bio-based and thus do not need to be analyzed further.Example 1: Ethyl isothiocyanate is produced from ethylamine and carbon disulfide, thus it is partly based on fossil and partly based on biobased materials. Throughout a manual review, we find that 1/3 of the carbon is sourced from carbon disulfide, whereas 2/3 of the carbon is sourced from ethylamine (see table below).

Example 2: N,n-diethylthiourea is produced from ethyl isothiocyanate and ethylamine. Throughout a manual review, we find that 2/5 of the carbon is sourced from ethylamine, whereas 3/5 of the carbon is sourced from ethyl isothiocyanate (see table below).

				ethyl isothiocyanate process	n,n-diethylthiourea process
carbon disulfide	1	kg	CS2	-1/3	
ethylamine	1	kg	C2H7N	-2/3	-2/5
ethyl isothiocyanate	1	kg	C3H5NS	1	-3/5
n,n-diethylthiourea	1	kg	C5H12NS2		1

3. In the next step, a mathematical carbon supply chain model is built for the cm.chemicals database according to the mathematical calculation framework described in Section 4.5 of the cm.chemicals database methodology document.

Example: A small example of a matrix-based carbon supply chain model is shown in the table below. The carbon disulfide process and ethylamine process are aggregated because the biogenic carbon share was directly set in step 1.

	carbon disulfide process	ethylamine process	ethyl isothiocyanate process	n,n-diethylthiourea process
carbon disulfide	1		-1/3	
ethylamine		1	-2/3	-2/5
ethyl isothiocyanate			1	-3/5
n,n-diethylthiourea				1

4. After calculating the scaling vectors of the processes in the carbon supply chain model, as described in Section 4.5 of the cm.chemicals database methodology document, the biogenic and fossil carbon shares can be calculated by tracking the carbon atoms across the complete carbon supply chain. Moreover, the biogenic carbon content can be calculated by multiplying the carbon content with the previously calculated biogenic carbon share.

Example 1: For ethyl isothiocyanate, the scaling vectors of the carbon supply chain model are calculated (see table below). Afterwards the biogenic carbon share of ethyl isothiocyanate is calculated as follows: $\frac{1}{3} \cdot 0 + \frac{2}{3} \cdot 1 = 0.67$.

Example 2: For n,n-diethylthiourea, the scaling vectors of the carbon supply chain model are calculated (see table below). Afterwards the biogenic carbon share of n,n-diethylthiourea is calculated as follows: $\frac{1}{5} \cdot 0 + \frac{4}{5} \cdot 1 = 0.8$.

	ethyl isothiocyanate	n,n-diethylthiourea
carbon disulfide process	1/3	1/5
ethylamine process	2/3	4/5
ethyl isothiocyanate process	1	3/5
n,n-diethylthiourea process	0	1

5. Finally, based on the shares of biogenic and fossil carbon contents previously calculated in step 4, the carbon monoxide, carbon dioxide, and methane emissions to

air are recalculated. Accordingly, emissions that were previously attributed to fossil emissions are now partly or fully attributed to biogenic emissions.

Due to the differentiation between biogenic and fossil carbon emissions in the cm.chemicals database, biogenic and fossil-based environmental impacts can be calculated. However, not all LCIA methods have characterization factors defined for biogenic emissions. Therefore, depending on the LCIA method, biogenic emissions can be taken into account or can be neglected. Thus, an LCIA method should be selected carefully in order to properly cover the desired scope of an LCA study. However, the definition of LCIA methods depends on the detailed scope of an LCA study and is out of the scope of the cm.chemicals database (cf. cm.chemicals database – methodology document).

LCIA methods commonly used to calculate carbon footprints, more specifically the global warming potential for 100 years, are listed in Table 2. Table 2 shows that the LCIA methods EF v2.0 2018, EF v3.0, and EF v3.0 EN15804 provide a distinction between fossil-based and biogenic carbon footprints which is necessary to meet the requirements of ISO 14067 for carbon footprint calculations. Moreover, Table 2 lists other LCIA methods that can only be used for a total carbon footprint calculation and without any distinction between fossil-based and biogenic emissions.

Table 2. List of LCIA methods used to calculate carbon footprints (CF).

LCIA method	total CF	fossil-based CF	biogenic CF
EF v2.0 2018	yes	yes	yes
EF v3.0	yes	yes	yes
EF v3.0 EN15804	yes	yes	yes
IPCC 2013	yes	no	no
ReCiPe Midpoint (H)	yes	no	no
ReCiPe Midpoint (H) V1.13	yes	no	no
CML v4.8 2016	yes	no	no
EDIP2003	yes	no	no
TRACI	yes	no	no

3.3 Data quality requirements

We specify the data quality of our datasets based on data quality indicators. These data quality indicators represent six data quality criteria: Technological representativeness, Geographical representativeness, Time-related representativeness, Completeness, Reliability and Methodological Appropriateness, and Consistency. For each criterion, five data quality levels exist, where level 1 represents the highest data quality and 5 the lowest. The definitions of the data quality criteria and quality levels are based on the Product Environmental Footprint (PEF) Guide by the Joint Research Center of the European Commission[2] except for the criterion "Reliability". This criterion replaces the criterion "Parameter Uncertainty" specified in the PEF guide, which has not yet been assessed for cm.chemicals. Table shows the definition of each data quality criterion. Table gives an overview of the data quality assessment scheme for each data quality criterion and data quality level.

[2] https://ec.europa.eu/environment/eussd/pdf/footprint/PEF%20methodology%20final%20draft.pdf

Table 3. Definitions of data quality criteria.

Technological representativeness	Chemicals can often be produced by different production technologies using different reaction pathways and plant designs. Technological representativeness is an indicator for the degree to which the dataset reflects the true population of interest regarding production technologies applied throughout the supply chain.
Geographical representativeness	Chemical production chains differ among regions. Geographical representativeness describes the degree to which the dataset reflects the true population of interest regarding geography.
Time-related representativeness	Technical, market, and trade data change over time. Time-related representativeness refers to the degree to which the dataset reflects the specific conditions of the system being considered regarding the time/age of the data.
Completeness	Completeness indicates to which degree relevant flows are covered by a specific dataset. Completeness refers to both technical flows and elementary flows throughout the production chain.
Reliability	Input data can be obtained from different sources, including measurements, detailed modeling, simplified process calculations, and assumptions. This quality indicator rates the reliability of a dataset based on the underlying data sources.
Methodological Appropriateness and Consistency	Methodological consistency is crucial for comparable LCA results. Therefore, all datasets in the cm.chemicals database are compiled based on the same, consistent methodology described in this document if not stated otherwise. This quality indicator assesses the consistency of the methodology applied, as well as its appropriateness.

Table 4: Assessment scheme for the determination of data quality criteria and quality levels.

Quality level	1 - Very good	2 - Good	3 - Fair	4 - Poor	5 - Very poor
Technological representativeness	All relevant production technologies are considered for the main product under study, and for all major raw materials, e.g., complete production and consumption mixes are used where needed.	Production of one or more raw materials is not modeled based on all relevant production technologies and only the market dominant production technology is considered.	Production of up to 50% of the raw materials is modeled based on a production technology that is industrially relevant but not the dominant production technology in the market.	Production of more than 50% of the raw materials is modeled based on a production technology that is industrially relevant but not the dominant production technology in the market.	Production of the main product or one or more major raw materials is based on a technology that is known not to be representative.
Geographical representativeness	Data on the production of the main product and all major raw materials is fully representative of the respective region by including site-specific mixes, production mixes and consumption mixes. Fossil feedstock and energy supplies are rarely based on larger regional averages (e.g., European average for a specific country).	Data on the type of production technology and all major raw materials is (partly) based on the market dominant technology. Fossil feedstock and energy supplies are partly based on larger regional averages (e.g., European average for a specific country).	Data on the production technology for the main product is representative; generic process data is used for each production technology; supply of most raw materials (incl. chemical intermediates) is based on larger regional averages that include the region under study but are not fully representative.	Dataset is fully based on data for a different region; only the electricity mix has been adapted to represent the region under study. Fossil feedstock and energy supplies are partly based on larger regional averages (e.g., European average for a specific country).	Dataset is known to be not representative of the region under study.
Time-related representativeness	Representativeness has been checked and confirmed within the last 3 years.		Representativeness has been checked and confirmed within the last 3 years. Minor changes are known, but the dataset is still considered to be partly representative.		Data for substantial parts of the production chain is known to be outdated.
Completeness	All process data has been measured or modeled at a high level of detail, including all technical and elementary flows.	All technical flows and major elementary flows have been measured or modeled at a high level of detail. Potential data gaps have been closed based on additional modeling or calculations.	Only major technical and elementary flows are considered. It is possible that some relevant flows are missing.	Only some of the major technical and elementary flows are considered. Larger data gaps are likely.	Completeness has not been specified.

Reliability	The dataset is fully based on measurements at all relevant production sites (primary data). The results have been verified.[3]	The dataset is based on detailed process simulations. Potential data gaps are closed through thermodynamic calculations. The results have been verified.[3]	The dataset is based on simplified process calculations considering the underlying stoichiometric reaction. Default values are used for energy supplies and conversion efficiencies.	The dataset is based on qualified estimates or stoichiometric calculations, where energy supplies and conversion efficiencies are neglected.	The process data is based on non-qualified estimates.
Methodological appropriateness and consistency	3rd party verification of the compliance with a defined methodology or standard based on (at least) spot checks.	Dataset is compliant with the methodology specified in this document.	Requirements specified in ISO 14040 are mainly met.	Requirements specified in ISO 14040 are only partly met.	Methodological appropriateness and consistency are not specified.

[3] Verification can be carried out, e.g., by on-site checking, by additional modelling, through mass, energy, and elementary balances or by cross-checking with other sources.

4. The life cycle inventory (LCI) model

Our LCI model of the chemical industry includes thousands of individual production plants in multiple production regions. For this purpose, several data collection and validation steps are necessary (cf. Section 4.1). Based on the collected data, the construction of the model is carried out (cf. Section 4.2). After constructing the model, multifunctionality problems are solved according to a straightforward and consistent methodology. The procedure to solve multifunctionality is explained in Section 4.3. Additionally, some more specific modeling features are described in Section 4.4. Finally, all datasets are transferred to the mathematical matrix structure, to calculate LCI for each functional unit and output dataset. This procedure is outlined in Section 4.5.

4.1 Data collection and validation

All datasets are based on different types of input data, including technical data for all production processes, market information, and trade data (cf. Section 2.1 for a general description of input data). The quality of the resulting datasets depends on the quality of the input data and the quality of the methodology applied to compile the LCI model.

We continuously monitor available data sources to select the most appropriate data for the cm.chemicals database. This section illustrates our methods used to ensure the consistency of our input data during data collection.

Input data for the LCI model can be obtained from multiple sources. To ensure adequate data quality, however, it is essential to evaluate the consistency and quality of the respective data. Therefore, we perform several checks for all input data to be used for the cm.chemicals database. In particular, we conduct the following steps:

- *Market overview.* As a first step, we conduct market research to identify all relevant production technologies and regions. Subsequently, we collect data on every single production plant used to produce the respective chemical, including plant capacities, ownership, and production technology used.
- *Technological assessment.* As a next step, all production technologies are analyzed in detail by an expert of our chemical engineering team. This analysis includes the following steps:
  - Develop an understanding of the underlying process, including all reaction and separation steps based on relevant literature.
  - Identify potential data sources and select the data source with the highest expected data quality according to the data quality indicators illustrated below.
  - Assess the plausibility of the data based on general chemical engineering knowledge and benchmarking with similar processes (e.g., energy demands and conversion efficiencies)
  - In case of a positive plausibility check, transfer all relevant data from the selected data source to an internal data collection sheet developed by us. This data collection sheet is both human-readable to allow for further checks by one of our experts and machine-readable to avoid potential errors when integrating the data into the database.

- Calculate mass and elemental balances to identify potential data gaps in both technical and elementary flows.
- In case of missing elementary or waste flow, build a process model based suitable thermodynamic modeling approach to fill the data gaps, conduct suitable thermodynamic process calculations, or complete the data based on literature values (where available at sufficient quality). All related modeling exercises are performed by a chemical engineer with comprehensive process design and modeling expertise.
- Document all changes in the original data source in the data collection sheet.

- *Trade and transportation.* Finally, collect data on relevant bilateral trade flows for the respective chemical (cf. Section 4.4.2) and calculate all transportation distances needed (cf. Section 4.4.1).

4.2 Model structure

The LCI model consists of two layers: a core layer and an extension layer. In the **core layer**, thousands of individual production plants are explicitly modeled. International trade between all production regions is also modeled, based on detailed physical trade data. The resulting core layer of the LCI model represents the worldwide geographical distribution and technology mix of chemical production chains with the highest level of detail available in the cm.chemicals database. The chemicals included in the core layer account for about 80% of the greenhouse gas emissions of the global chemical industry.

The **extension layer** covers the production of additional chemicals by specific technologies on a country/region level. For these chemicals in the extension layer, detailed trade information and country-specific technology mixes are not available. Trade data is only included for some raw material supplies.

4.2.1 Chemical plant level

Chemical plants represent the production of a given chemical in a specific production site. We collect information on the production location (site) for each chemical plant, the production volume, and the exact production technology used in the plant.

We define the term *production technology* as the production techniques used in a specific chemical plant to produce a particular chemical, including the reaction pathway, reactor technology, and separation steps. We use detailed technical models for each production technology to determine the raw material consumption, utilities (e.g., energy use), resource extractions, emissions, co-products, and wastes.

4.2.2 Integrated production site level

After chemical plants are modeled individually, we model interactions between plants within integrated production sites. As integrated production sites offer a range of efficiency savings (such as reducing transportation distances, energy integration, and the use of co-products), they are standard practice in the chemical industry and should be considered whenever possible.

In our LCI model, individual plants are summarized to integrated production sites based on their location. Production plants located in the same city are assumed to be in the same integrated production site. The modeling of integrated production sites allows us to account for plant-specific supplies of raw materials within an integrated production site, as shown in Figure 4.

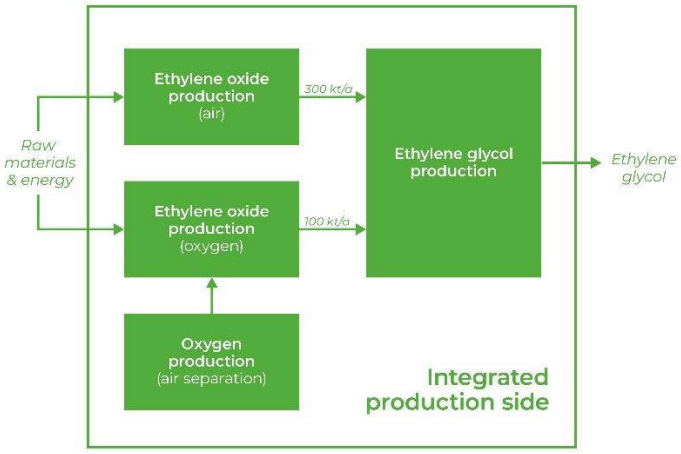

Figure 4: Modeling approach for plant-specific raw material supplies in integrated production sites in cm.chemicals.

Figure 4 illustrates the modeling of raw material supplies based on a simplified example. The figure shows an integrated production site with four individual plants. Two plants produce ethylene oxide via the oxidation of ethylene. One plant uses oxygen from the air for oxidation; the other uses pure oxygen. For on-site air separation, a third plant delivers pure oxygen to the ethylene oxide production plant that requires it. The fourth plant processes ethylene oxide to produce ethylene glycol.

In this example, the technology mix used to produce ethylene oxide within the integrated production site is calculated according to the weighted average production from both ethylene oxide plants. 75% of the ethylene oxide production mix inside the integrated production site is produced using oxygen from the air for oxidation. The other 25% is produced using pure oxygen from the air separation process.

Because one of the main intentions behind integrated production sites is the reduction of transportation distance and the use of co-products, we assume that site-specific production technology mixes are used to satisfy demands for raw materials within integrated production sites. If the production volume of a specific chemical intermediate within the production site is insufficient to satisfy the entire demand of that site, the remaining demand will be met by the national consumption mix (cf. Section 4.2.3). The national consumption mix also delivers all inputs which are needed by any plant within an integrated site but which are not produced inside of the integrated production site itself, e.g., the raw material and energy supplies of the ethylene oxide plants in Figure 4.

By modeling integrated production sites, we can reveal to what extent the technology mix used to deliver intermediates within production sites differs from the country's average consumption mix where the site is located. By replacing national averages with explicit modeling, we obtain more representative data that can differ substantially from national averages.

4.2.3 National production and consumption mixes

After modeling all production plants and integrated production sites within a country, we calculate the average national production and consumption mixes.

The production mix in a given country is calculated from the output of all chemical plants, which produce a given chemical in that country (cf. Figure 5). The production mix is calculated based on the country's proportional share of national production contributed by each chemical plant.

The national production mix of a chemical, however, does not necessarily reflect the consumption of that chemical in that country because parts of the amount consumed may be imported from other countries. Furthermore, parts of the national production may be exported to other countries. Consequently, the consumption mix is represented by the sum of a country's production mix, plus imported chemicals, minus exported chemicals, as illustrated in Figure 5.

Production mixes are available for all countries where a specific chemical is produced. Consumption mixes are available for all countries that either produce a chemical and/or import it from other countries.

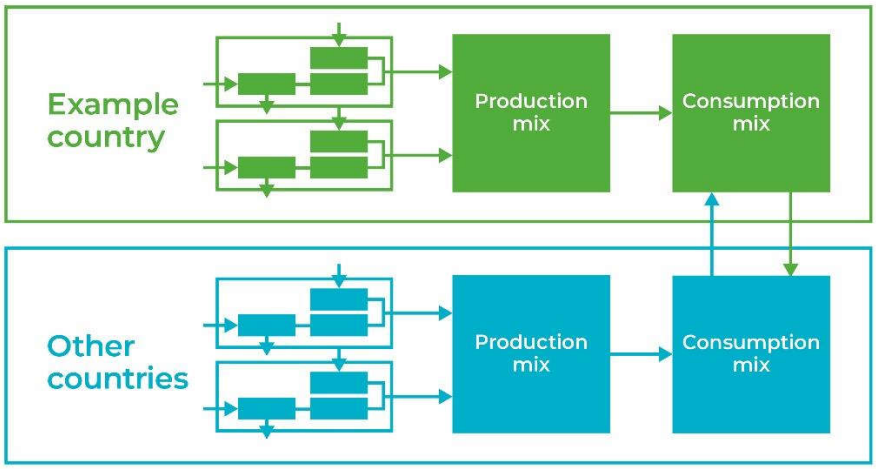

Figure 5: Determination of national production and consumption mixes based on plant-level data.

4.2.4 Extension layer

The extension layer models the production of additional chemicals based on individual production technologies for which plant-level data is not available. Each production technology is represented by its technical flows (reactants, utilities, co-products, waste) and its elementary flows such as emissions and resource extractions. The technical and elementary flows are either obtained from detailed technical models or the simplified modeling approach discussed in Section 4.4.4.

In the extension layer, input flows are provided either by consumption mixes from the core layer (where available) or by other production technologies from the extension layer representing the market's dominant technology.

Production technologies for chemicals in the extension layer are available in countries where all input flows are available in either the core or extension layer. By contrast, if

one or more input flow is missing in a country (e.g., because the country neither produces nor imports the respective chemical), no production technology is included for the specific chemical.

While the same level of technology detail is often used in the extension layer as in the core layer, in some cases technology data on the production technology is not available. In these cases, we use a simplified modeling approach to fill these data gaps (cf. Section 4.4.4).

4.3 Solving multifunctionality (allocation procedure)

Processes in the chemical industry often have more than one function and are therefore multifunctional. Functions include the production of a product and the treatment of waste. There are three types of multifunctional processes:

- Joint production of valuable outputs such as chemicals or fuels
- Joint treatment of multiple wastes
- Joint treatment of waste and production of one or more valuable products (e.g., recycling processes)

To calculate product-specific LCIs for products from multifunctional processes, the environmental exchanges of these processes over the life cycle need to be allocated between the processes' functions. The problem of how to allocate environmental exchanges between products is often called a multifunctionality problem.

Several methodological approaches exist in LCA methodology for solving the multifunctionality problem: sub-division, system expansion, and allocation using either an underlying physical relationship or an underlying other relationship. Not all approaches apply to every process, but typically more than one approach is technically feasible.

To guide the selection of methodological approaches, the ISO standard 14044 has defined a hierarchy among the methodological approaches. We apply this hierarchy to all multifunctionality problems. The following discussion is a brief summary of the methods for solving the multifunctionality problem. We recommend reading the more detailed description in ISO 14044 or related documents to those readers who are not familiar with the methods.

Step 1: Subdivision. Whenever possible, we solve the multifunctionality problem through subdivision. Subdivision is a methodological approach to address multifunctionality problems due to data aggregation. It can be applied when the data from different single-functional sub-processes are aggregated to one aggregated process. The aggregated (black-box) process then seems to be multifunctional only due to the level of aggregation. Subdivision solves this multifunctionality problem by collecting additional process data for all relevant underlying single functional processes and including only the relevant processes into the model.

Step 2: System expansion. If subdivision cannot solve the multifunctionality problem, we use system expansion via avoided burden in the next step. In this approach, credit is given for the joint provision of all functions not included in the functional unit. This credit represents the avoided environmental burden associated with the conventional way to provide these functions that would be used in the absence of the product system under study.

We use the method of system expansion via avoided burden for all fuels and steam outputs that are co-produced in chemical processes and not used internally in the process. We assume that all fuels are used for heat production and avoid the conventional production of heat based on natural gas. In the case of steam, we assume that conventional steam production based on natural gas is avoided. Consequently, the avoided burden represents the environmental burden associated with producing the same amount of heat from natural gas.

Step 3: Allocation. Finally, whenever system expansion via avoided burden cannot solve the multifunctionality problem, we apply allocation. Allocation divides the multifunctional process into processes with exactly one function. Then the environmental exchanges of the multifunctional process and its production chain are distributed to the functions reflecting either an underlying physical relationship or an underlying other relationship.

According to ISO 14044, an underlying physical relationship must be applied whenever possible by quantifying how inputs and outputs physically relate to the system's function. A way to determine physical relationships for processes producing more than one valuable product (functions) is to change the amount of one product produced while keeping the other products' production volume constant and observing how all other inputs and outputs change. Then the allocation of the inputs and outputs should reflect this quantitative change observed. The application of allocation based on physical relationships is documented in the respective dataset whenever it is used.

Suppose neither of the approaches can solve the multifunctionality problem. In that case, we allocate the environmental exchanges of the process and its supply chain in proportion to the energy content of the products. If one or more products have an energy content of zero, we apply allocation based on the mass content. Potential deviations from this approach are illustrated in the metadata of the respective dataset.

4.4 Specific-modeling features

The previous Sections 4.2 and 4.3 provide general information on the LCI model. This Section highlights more specific modeling details that are relevant for both the LCI model and the output datasets. These modeling details include the modeling of international trade, transportation, and waste incineration.

4.4.1 Transportation

Freight transportation has been considered for all internationally traded flows between two countries. The transport of chemicals is assumed to be weight-limited. Therefore, emissions and resource extractions are assumed to scale linearly with the mass transported over a given distance. Transportation is modeled from cradle to grave using background data from the latest Ecoinvent version (cf. Section 5.3).

We differentiate between two types of transportation:

1. **Transportation by land and sea.** The transportation by land and sea is a combination of sea transport between the countries' ports and transport via road from the ports to the inland.

2. **Transportation by land.** The transportation by land is assumed to be completely via road.

Transportation distances have been obtained from Sea Rates[4] and Openrouteservice[5].

Transportation by land and sea

For the calculation of sea distances between two countries, we have applied the following procedure:

- Identify the main ports in each country (minimum 1 port, maximum 2 ports).
- Determine sea distances between the two countries for all possible combinations of main ports
- Allocate each chemical production site to one of the main ports. Every site is allocated to the port with the shortest transportation distance.
- Example: Country 1 has two ports A and B and three sites I, II, and III with a respective production volume of 60%, 20%, and 20%. We calculate then the inland distances from all three sites to both ports. You can see the resulting inland distances in the table below. In a second step, select the closer port for each site: port A is closer to site I and II and port B is closer to site III. Thus, the share of trade going through port A is 80% and the share of trade going through port B is 20%.

	Site I (60%)	Site II (20%)	Site III (20%)
Port A_1	50 km	120 km	310 km
Port B_1	250 km	130 km	10 km

- Calculate the percentage of the total chemical production capacities of the country that is allocated to each port and assume that the same share of imports and exports will be shipped via this port.
- Calculate the percentage of bilateral trade between the two countries that is shipped via each of the combinations of main ports based on the shares of imports and exports shipped via the respective ports.
- Calculate the weighted average sea distance between the two countries using the percentage of trade flows shipped via each port combination as a weighting criterion.
- Example: Country 1 has port A1 and B1 and country 2 has port A2 and B2. First, the sea distances are calculated for each port combination, thus this results in 4 distances as shown in the table below. Furthermore, weighted values for the ports in country 1 (80% and 20%) and in country 2 (30% and 70%) were obtained in a previous step. In a second and third step, all distances are weighted and then summed up. This results in a final sea distance of 3000 km*0.8*0.3 + 3500 km*0.8*0.7 + 4000 km*0.2*0.3 + 3800 km*0.2*0.7 = 3452 km.

[4] www.searates.com
[5] www.openrouteservice.org

		Port A_1 (80%)	Port B_1 (20%)
		Country 1	Country 1
Port A_2 (30%)	Country 2	3000 km	4000 km
Port B_2 (70%)	Country 2	3500 km	3800 km

For the calculation of inland road distances between chemical production sites and ports, we have followed the following steps:

- Identify the nearest port for each chemical production site. The nearest port can either be one of the main ports identified for the calculation of the sea distances or another port located at the seaside or at an inland waterway.
- Determine the road transportation distance between each production site and its nearest port
- Calculate the weighted average transportation distance between production sites and ports using the production capacities of the production sites as weighting criterion.
- Example: Country 1 has two producing sites I and II with the respective production shares 30% and 70%. The API query from site I to the identified port results in an inland distance of 50 km in Country 1. The API query from site II to the respective port results in an inland distance of 300 km in Country 1. The total inland distance is thus calculated as follows: 0.3*50 km + 0.7*300 km = 225 km.

In case a country is not a producing county with no chemical production site, the inland road distance is calculated from the most economically important region (GeoDist[6] database) to a nearby port.

Transportation by land

Transportation by land is calculated only for country combinations that are on one continent (exception: Europe-Asia) and for combinations where neither of the two countries is an island.

Average land transportation distances between countries are calculated based on the following procedure:

- Determine the weighted average production location of all chemical production sites in each country. This weighted average production location is represented by the weighted average geo-coordinates of all production sites using the production capacities of the sites as the weighting criterion.

Determine the road transportation distance via openrouteservice between the weighted average production locations of the countries.

[6] http://www.cepii.fr/CEPII/en/publications/wp/abstract.asp?NoDoc=3877

Finally, for country combinations, for which both land and sea transportation is possible, we have chosen the option leading to the lower transportation costs. The transportation costs have been estimated based on cost factors provided by Maibach et al. (2006).[7]

Example: For the country combination between country 1 and country 2, we obtained the following results:

- For the land/sea distance, we obtained a transportation distance of 10 km road and 3000 km sea. This results in the following transportation costs:
  0.1 €/km *10 km + 0.01 €/km *3000 € = 31 €
- For the land distance, we obtained a transportation distance of 1200 km road and 0 km sea. This results in the following transportation costs:
  0.1 €/km *1200 km = 120 €

Thus, case 1 is the economically more advantageous distance and selected for the transport distances between country 1 and country 2.

4.4.2 Trade data

The LCI model is based on a detailed physical trade model. This trade model includes bilateral trade flows between all regions considered. The model is built from data directly reported by each country to the United Nations Statistical Division. The data has been harmonized by CEPII (Centre d'études prospectives et d'informations internationals) to eliminate data inconsistencies, e.g., contradictory trade data published by different countries.

Details on the methodology used for the harmonization are provided by CEPII.[8]

For most chemicals included in the LCI model, CEPII provides a specific 'HS code' which defines the bilateral trade flows of the respective chemical. However, in some cases, CEPII provides HS codes that cannot be assigned directly to a chemical included in the LCI model. Usually, these HS codes combine several chemicals in one HS code so they must first be separated from each other. In such case, the following hierarchy is used to determine which method is applied to obtain the correct trade flows for chemicals:

1. If for all chemicals or compounds listed under a specific HS code, market data on production volumes and locations is available in the LCI model, regionalized market shares are calculated based on these market data. These shares are then set to be the export shares of the compounds.
2. If there is a lack of market data in the LCI model for chemicals listed under a specific HS code, chemical reporting data (US EPA CDR) is used to calculate production shares for chemicals.
3. If none of the methods above can be applied, there are custom cases specified for the specific chemical. However, this case occurs rarely.

---

[7] Maibach, M.; Peter, M.; Sutter, D. (2006): Analysis of operating cost in the EU and the US. Annex 1 to Final Report of COMPETE Analysis of the contribution of transport policies to the competitiveness of the EU economy and comparison with the United States. Karlsruhe, Germany.
[8] http://www.cepii.fr/CEPII/en/bdd_modele/presentation.asp?id=37

4.4.3 Waste incineration

Waste incineration has been modeled based on a Life Cycle Inventory model developed by Doka (2003)[9]. The model determines LCIs for waste incineration depending on the composition of the waste. The consideration of the composition is crucial for the modeling of waste incineration in the chemical industry because waste compositions and resulting environmental impacts can differ substantially among chemical plants and production technologies.

The model considers both the incineration of the waste in an incineration plant and the separation and landfilling of the solid remains from incineration. A complete documentation of the model is provided in the original report.

We updated and adapted the model of Doka using primary data from hazardous waste incineration plants located in a chemical park in Germany.

4.4.4 Simplified modeling approach for the extension layer

The extension layer of the cm.chemicals database includes individual datasets for production technologies, for which no detailed process models are available. For these technologies, we have applied a simplified modeling approach. This simplified modeling approach enables the calculation of LCI data for the respective technologies. However, it leads to a lower data quality rating of the resulting LCI datasets compared to most other datasets in the core and the extension layer.

In our simplified modeling approach, unit process data is determined based on information on stoichiometric conversion. Determining LCI data based on stoichiometric conversion is a commonly applied method to fill data gaps in LCA studies.[10] We assume stoichiometric conversion according to the following Reaction 1:

$$v_{\mathrm{i}_1} M_{\mathrm{i}_1} + \cdots + v_{\mathrm{i}_\mathrm{n}} M_{\mathrm{i}_\mathrm{n}} \rightarrow v_{\mathrm{p}_1} M_{\mathrm{p}_1} + \cdots + v_{\mathrm{p}_\mathrm{m}} M_{\mathrm{p}_\mathrm{m}} \qquad (1)$$

[9] Doka G. (2003) Life Cycle Inventories of Waste Treatment Services. ecoinvent report No. 13. Swiss Centre for Life Cycle Inventories, St. Gallen, 2009.

[10] Georg Geisler, Thomas B. Hofstetter, and Konrad Hungerbuhler (2004). Production of fine and speciality chemicals: procedure for the estimation of LCIs. *The International Journal of Life Cycle Assessment*, 9(2):101–113.

Hirokazu Sugiyama, Ulrich Fischer, Konrad Hungerbuhler, and Masahiko Hirao (2008). Decision framework for chemical process design including different stages of environmental, health, and safety assessment. *AIChE Journal*, 54(4):1037–1053.

Akshay D. Patel, Koen Meesters, Herman den Uil, Ed de Jong, Kornelis Blok, and Martin K. Patel (2012). Sustainability assessment of novel chemical processes at early stage: application to biobased processes. *Energy & Environmental Science*, 5(9):8430.

In this reaction, v represents the stoichiometric coefficient, M the molar mass, indices $i_{1,\dots,n}$ reactants, and indices $p_{1,\dots,n}$ products. Based on Reaction 1, the mass flow m_i of reactant i that is needed per kilogram main product p can be estimated under the assumption of full conversion by:

$$m_{\mathrm{i}} = \frac{v_{\mathrm{i}} M_{\mathrm{i}}}{v_{\mathrm{p}} M_{\mathrm{p}}}. \tag{2}$$

Furthermore, to account for inefficiencies in process technology such as incomplete reactions or production losses, a product yield P of 95 wt. % is assumed following the recommendations by Hirschinger.[11] Thus, the mass flow m_i of each input can be determined:

$$m_{\mathrm{i_P}} = m_{\mathrm{i}} \cdot \frac{1}{P}. \tag{3}$$

In addition, chemical processes require energy for the conversion and operation, as well as for the subsequent separation or purification of the product. To include this energy demand, a simplified methodology by ecoinvent is applied. In this methodology, the missing energy inputs are approximated by the average energy demand of chemicals produced in the German chemistry park Gendorf, where more than 30 companies produce about 1500 chemicals.[12] These average demands for electricity and heat amount to 1.2 GJ and 2 GJ per ton product, respectively.

4.5 Mathematical calculation framework

The cm.chemicals database includes the aggregated LCI results per functional unit for all the output datasets listed in Section 2.3. To calculate the aggregated LCI results per functional unit, the LCI model of the chemical industry and the general matrix calculus of LCA is used[13].

In this matrix calculus, the exchange of intermediate flows between processes in the LCI model is described in the technology matrix A. In this matrix, rows represent intermediate flows, while columns represent processes. A process in the LCI model is, for instance, represented by a chemical plant or a production mix. Intermediate flows include, for instance, chemical raw materials, steam and electricity, or a solvent. A

[11] Roland Hischier, Stefanie Hellweg, Christian Capello, and Alex Primas (2005). Establishing life cycle inventories of chemicals based on differing data availability. *The International Journal of Life Cycle Assessment*, 10(1):59–67.

[12] H.-J. Althaus, M. Chudacoff, R. Hischier, N. Jungbluth, M. Osses, and A. Primas (2007). Life cycle inventories of chemicals. ecoinvent report no.8, v2.0. URL www.ecoinvent.org.

[13] Reinout Heijungs and Sangwon Suh (2002). The computational structure of life cycle assessment. Centre of Environmental Science Leiden University. Kluwer Academic Publishers Drodrecht. ISBN: 978-94-015-9900-9.

coefficient a_{ij} of the technology matrix A describes the intermediate flow i, which is produced (for $a_{ij} > 0$) or consumed (for $a_{ij} < 0$) by process j.

The net intermediate flows leaving the product system are specified in the functional unit vector f. For more information on the functional unit specified for the cm.chemicals database, please see Section 3.2.1. For an invertible technology matrix A and a given functional unit vector f, a scaling vector s can be calculated:

$$s = A^{-1} f. \tag{4}$$

The elementary flow matrix B describes the elementary flows of the processes. In the ISO standards on LCA (ISO 14040 and 14044), elementary flows are defined as "material or energy entering the system being studied that has been drawn from the environment without previous human transformation, or material or energy leaving the system being studied that is released into the environment without subsequent human transformation". In the elementary flow matrix, elementary flows are represented by rows, while the columns represent the same processes as in the technology matrix A. The matrix is defined such that a coefficient b_{ej} shows the elementary flow e of unit process j entering (for $b_{ej} < 0$) or leaving (for $b_{ej} > 0$) the process.

Multiplying the elementary flow matrix B with the scaling vector s yields the aggregated LCI result g representing the total elementary flows associated with the functional unit f:

$$g = B\, s = B\, A^{-1} f. \tag{5}$$

Thus, the cm.chemicals database is obtained by first creating the technology matrix A and the elementary flow matrix B for the complete global chemical industry model. This is done, by collecting and modeling data according to the principles explained in Chapter 4. In a final step, the aggregated LCI results are calculated for each output dataset based on formula (5).

5. Documentation of LCI datasets

5.1 Documentation principles and template

Aggregated LCI datasets are documented in the International Life Cycle Data System (ILCD) format. The ILCD format was developed by the European Commission and aims to facilitate the exchange of LCI/LCA datasets through an international standardized data format.

For each aggregated LCI dataset, we provide a broad list of process meta data that describe the process, modeling approaches and validation, administrative information, and the inputs and outputs for the respective aggregated process. Additionally to the process meta data, flow meta data, flow properties, unit group data, sources data, and contact data are provided for each aggregated LCI dataset.

In the following Table 5, the process meta data are described exemplarily for the consumption mix of methanol in Germany.

Table 5. Exemplarily description of process meta data that are provided for each documentation of an LCI dataset.

Process information	
Key Data Set Information	
Location	DE
Geographical representativeness description	The model is based on representative information on the production technology used in individual chemical plants along the entire supply chain. The term production technology refers to the production method in terms of reaction pathway, reactor technology, separation steps, etc. Each production technology is modeled based on detailed technical process data (e.g., mass and energy balances). International trade is modeled using trade data directly reported by each country to the United Nations Statistical Division and harmonized to align the exporter and importer declarations, which may differ in the original data. Country-specific fossil feedstock supplies are applied wherever possible. Otherwise, larger regional averages are used. Fossil feedstock, thermal energy, steam, electricity, and inorganic chemical (except chlorine and caustic soda) supplies are modeled based on data from the ecoinvent database, using the cut-off system model.
Reference year	2019
Name	methanol, consumption mix
Use advice for data set	The system boundary for this dataset is from cradle-to-gate. The dataset can be used to represent the environmental impacts of the respective commodity chemical. If the dataset is combined with other datasets, it can be used to produce user-specific LCAs.
Synonyms	CAS 000067-56-1
Classification	materials production / chemical
General comment on data set	The dataset represents a cradle-to-gate inventory with overall very good data quality (see cm.chemicals methodology document for details). All relevant production steps within the

	chemical industry are modeled based on representative data on the production technology used in individual plants along the supply chain, with data coverage ranging from 95% to 100% of global production capacities. This also includes the consideration of production plants that use old technologies but are still running today. The production of crude oil, naphtha and natural gas is covered on the basis of representative data at country level or at a larger regional level.
Quantitative reference	
Reference flow(s)	methanol; consumption mix – 1 kg (Mass)
Functional unit, Production period, or Other parameter	Production of 1 kg methanol
Time representativeness	
Data set valid until:	2023
Time representativeness description	Time representativeness is reviewed annually. Updates will be made for any data points identified as non-representative based on the quality ratings established for this data set. Details on the quality ratings are provided in the cm.chemicals methodology document.
Technological representativeness	
Technology description including background system	The consumption mix for methanol (Germany) is based on 50.75% regional production and 49.25% imports. Imports originate from: Netherlands (23.58%), Norway (6.51%), Belgium (4.66%), Equatorial Guinea (3.92%), Poland (3.57%), Trinidad and Tobago (3.33%), Saudi Arabia (1.93%), and other countries (1.75%). Background modeling: The data set represents a cradle to gate inventory, including all relevant process steps / technologies over the supply chain. The data set is based on different types of data: Process data is obtained from detailed process simulations. International trade volumes and regional production capacities are mainly based on primary data and complemented by secondary data where necessary. Electricity is modeled according to the individual country-specific situations, including national electricity grid mixes and imported electricity. Steam and thermal energy supplies take into account the country-specific situation, wherever possible. Otherwise, larger regional averages are used. The production of crude oil, naphtha, and natural gas is represented by either fully country-specific data or by partly representative data for a fully overlapping but not identical region (e.g., European average for a specific European country).
Modelling and validation	
LCI method and allocation	
Type of data set	LCI result
LCI method principle	Attributional
LCI method approaches	Allocation – net calorific value Allocation – mass
Deviation from LCI method principle / explanations	All datasets are based on the methodology defined in the cm.chemicals methodology document. Allocations are made along the entire supply chain based on the requirements of ISO standard 14040/14044 and the further provisions in the cm.chemicals methodology document in accordance with the ISO standard. The further

<table>
<tr><td></td><td>provisions include the use of the system expansion approach via avoided burden for the co-production of steam and thermal energy, and the use of mass as an allocation criterion if the use of calorific values is not appropriate.</td></tr>
<tr><td>Data sources, treatment and representativeness</td><td></td></tr>
<tr><td>Data cut-off and completeness principles</td><td>Cut-off criteria for unit process data are applied according to the following rules: 1. The flow is used in small quantities in the respective process (below 1 mass-% of all inputs, including all process utilities). 2. The flow is not relevant for any other process in the database after applying cut-off criteria. 3. The input data needed for modelling the production of the flow is not available to Carbon Minds. The sum of all flows that are neglected on a process level is lower than 1% of the mass of all inputs, including all process utilities.</td></tr>
<tr><td>Deviation from data cut-off and completeness principles / explanations</td><td>None.</td></tr>
<tr><td>Data selection and combination principles</td><td>The dataset is based on the consistent modeling approach for calculating life cycle inventories described in the cm.chemicals methodology document.</td></tr>
<tr><td>Deviation from data selection and combination principles / explanations</td><td>None.</td></tr>
<tr><td>Data treatment and extrapolations principles</td><td>A summary of the data treatment and extrapolations principles is provided in the cm.chemicals methodology document.</td></tr>
<tr><td>Completeness</td><td></td></tr>
<tr><td>Completeness product model</td><td>All relevant flows quantified</td></tr>
<tr><td>Validation</td><td></td></tr>
<tr><td>Review</td><td>Dependent internal review
<table>
<tr><th>Scope of review</th><th>Method(s) or review</th></tr>
<tr><td>Raw data</td><td>Validation of data sources
Sample tests on calculations</td></tr>
<tr><td>Unit process(es), single operation</td><td>Validation of data sources
Sample tests on calculations
Energy balance
Element balance
Cross-check with other source
Expert judgement
Mass balance
Compliance with ISO 14040 to 14044</td></tr>
<tr><td>LCI results or Partly terminated system</td><td>Cross-check with other source
Cross-check with other data set
Expert judgement
Compliance with ISO 14040 to 14044</td></tr>
<tr><td>LCIA results</td><td>Cross-check with other source
Cross-check with other data set
Expert judgement</td></tr>
</table></td></tr>
</table>

	<table><tr><td></td><td>Compliance with ISO 14040 to 14044</td></tr><tr><td>Documentation</td><td>Compliance with ISO 14040 to 14044</td></tr><tr><td>Life cycle inventory methods</td><td>Compliance with ISO 14040 to 14044 Expert judgement</td></tr><tr><td>LCIA results calculation</td><td>Compliance with ISO 14040 to 14044</td></tr><tr><td>Goal and scope definition</td><td>Compliance with ISO 14040 to 14044 Expert judgement</td></tr></table>
Data quality indicator	Technological representativeness: Very good Time representativeness: Very good Geographical representativeness: Very good Completeness: Very good Methodological appropriateness and consistency: Very good Overall quality: Very good
Review details	The LCI method applied is in compliance with ISO 14040,14044, and 14067. For details please see the cm.chemicals methodology document.
Reviewer name	Carbon Minds GmbH
Review	Independent external review <table><tr><td>Scope of review</td><td>Method(s)of review</td></tr><tr><td>Raw data</td><td>Validation of data sources Cross-check with other source</td></tr><tr><td>Unit process(es), single operation</td><td>Cross-check with other source Energy balance Element balance Mass balance Compliance with ISO 14040 to 14044</td></tr><tr><td>LCI results or Partly terminated system</td><td>Cross-check with other source Compliance with ISO 14040 to 14044</td></tr><tr><td>Documentation</td><td>Compliance with ISO 14040 to 14044</td></tr><tr><td>Life cycle inventory methods</td><td>Compliance with ISO 14040 to 14044 Expert judgement</td></tr></table>
Review details	The LCI method applied is in compliance with ISO 14040 and 4044, and has been verified by TÜV Rheinland Energy GmbH. As part of the review, TÜV Rheinland Energy GmbH assessed a selected number of datasets. The corresponding Review Report is attached in the annex of this document.
Reviewer name/ Institution	TÜV Rheinland Energy GmbH
Compliance declarations	
Compliance system name (source data set)	ILCD Data Network compliance
Approval of overall compliance	Not defined
Nomenclature compliance	Not defined
Methodological compliance	Fully compliant
Review compliance	Not defined
Documentation compliance	Fully compliant
Quality compliance	Not defined
Administrative information	

Data set generator / modeler	
Data set generator / modeler (contact data set)	Carbon Minds GmbH
Data entry by	
Time stamp (last saved)	2021-06-30 14:00:00
Data set format(s) (source data set)	ILCD format
Data entry by: (contact data set)	Carbon Minds GmbH
Publication and ownership	
UUID of Process data set	4be26daa-6c4a-4d51-a34c-7ec7c424407a
Date of last revision	2021-06-30T14:00:00
Data set version	01.01.001
Workflow and publication status	Data set finalised; entirely published
Owner of data set (contact data set)	Carbon Minds GmbH
License type	License fee
Access and use restrictions	Usage of datasets is subject to the General Terms and Conditions for the Provision of Data Sets of Carbon Minds GmbH, unless otherwise contractually agreed.
Inputs and Outputs	
Inputs	
Here, all inputs of the aggregated process data set are listed. This includes the definition of the Type of Flow, Classification of Flow, Flow name, Amount, and data source type.	
Outputs	
Here, all outputs of the aggregated process data set are listed. This includes the definition of the Type of Flow, Classification of Flow, Flow name, Amount, and data source type.	

5.2 Data quality indicators of datasets

Data quality rating. The data quality ratings are available for each dataset. Section 3.3 illustrates the data quality levels defined in more detail. The data quality rating for each dataset introduced in Section 2.3 are illustrated in Tables 5 to 7 below.

Table 6. Data quality ratings for plant-specific, supplier-specific, technology-specific (core layer), production mix and consumption mix datasets.

Quality level	Rating	Justification
Technological representativeness	1	All relevant production steps within the chemical industry are represented based on plant-level data covering between 95% and 100% of worldwide production capacities. The production of crude oil, naphtha, and natural gas is represented by data for production and consumption mixes from Ecoinvent (e.g., a specific European country or European average).
Geographical representativeness	1	Our model is based on representative information on which production technology is used in the individual chemical plants throughout the supply chain. Detailed technical process data is used for each production technology. Country-specific fossil feedstock supplies are applied whenever possible. Otherwise, larger regional averages are used. Fossil feedstock, energy, and electricity supplies are modeled based on data from the Ecoinvent database using the cut-off system model. Trade balances are based on data directly reported by each country to the United Nations Statistical Division and partly modified to correct errors or increase consistency (cf. Section 4.4.2).
Time-related representativeness	1	Representativeness is checked on an annual basis, and updates are carried out for all data points that have been identified not to be representative based on the quality ratings specified here.
Completeness	1	All technical flows and major elementary flows have been determined based on very detailed and sophisticated process modeling. Checks have been performed as discussed in the previous sections. Mass and elementary balances have been calculated for every chemical process to identify and subsequently close potential data gaps. Trade balances are based on data directly reported by each country to the United Nations Statistical Division and transformed into a harmonized physical trade model (cf. Section 4.4.2).

Reliability	2	Chemical process data has been obtained from detailed process simulations. Data gaps have been closed based on additional modeling. All process data has been verified through mass and elementary balances and – whenever possible - cross-checked with other sources.
Methodological appropriateness and consistency	1	All datasets are based on the methodology specified in this document. The application of the methodology leads to high levels of consistency.

Table 7. Data quality ratings for technology-specific (extension layer) with detailed technology data.

Quality level	Rating	Justification
Technological representativeness	2	The main product under study is modeled using the market's dominant production technology. Raw materials are partly based on consumption mixes (cf. Table 6) or the market's dominant production technology. The production of crude oil, naphtha, and natural gas is represented by data for production and consumption mixes from Ecoinvent (e.g., a specific European country or European average).
Geographical representativeness	2	Detailed technical process data (e.g., mass and energy balances) is used for each market's dominant technology. Part of the raw materials is based on national consumption mixes (cf. Table 6) or the market's dominant technology. Country-specific fossil feedstock supplies are applied whenever possible. Otherwise, larger regional averages are used. Fossil feedstock, energy, and electricity supplies are modeled based on data from the Ecoinvent database using the cut-off system model. Trade balances are only included for some raw materials.
Time-related representativeness	1	Representativeness is checked on an annual basis, and updates are carried out for all data points that have been identified not to be representative based on the quality ratings specified here.
Completeness	2	All technical flows and major elementary flows have been determined based on very detailed and sophisticated process modeling. However, the main product is represented by only the market's dominant technology and no trade data has been included for the main product and parts of the raw materials. Checks have been performed as discussed in the previous sections. Mass and elementary balances have been calculated for every chemical process to identify and subsequently close potential data gaps.
Reliability	2	Chemical process data has been obtained from detailed process simulations. Data gaps have been closed based on additional modeling. All process data has been verified through mass and atom balances and – whenever possible - cross-checked with other sources.
Methodological appropriateness and consistency	1	All datasets are based on the methodology specified in this document. The application of the methodology leads to high levels of consistency.

Table 8. Data quality ratings for technology-specific (extension layer) with simplified technology data.

Quality level	Rating	Justification
Technological representativeness	3	The main product under study is modeled using an industrially relevant production technology that is not necessarily the market's dominant technology. Raw materials are partly based on consumption mixes (cf. Table 6), the market's dominant production technology, or an industrially relevant production technology. The production of crude oil, naphtha, and natural gas is represented by data for production and consumption mixes from Ecoinvent (e.g., a specific European country or European average).
Geographical representativeness	2	Simplified technical process data (e.g., mass and energy balances) is used for each industrially relevant production technology. Part of the raw materials is based on national consumption mixes (cf. Table 6), the market's dominant production technology, or an industrially relevant production technology.

		Country-specific fossil feedstock supplies are applied whenever possible. Otherwise, larger regional averages are used. Fossil feedstock, energy, and electricity supplies are modeled based on data from the Ecoinvent database using the cut-off system model. Trade balances are only included for some raw materials.
Time-related representativeness	1	Representativeness is checked on an annual basis, and updates are carried out for all data points that have been identified not to be representative based on the quality ratings specified here.
Completeness	3	All technical flows and major elementary flows have been determined based on simplified process modeling. The main product is represented by only an industrially relevant production technology and no trade data has been included for the main product and parts of the raw materials. Checks have been performed as discussed in the previous sections. Mass and elementary balances have been calculated for every chemical process to identify and subsequently close potential data gaps.
Reliability	3	Chemical process data has been obtained from simplified process simulations. Data gaps have been closed based on additional modeling. All process data has been verified through mass and atom balances and – whenever possible - cross-checked with other sources.
Methodological appropriateness and consistency	1	All datasets are based on the methodology specified in this document. The application of the methodology leads to high levels of consistency.

5.3 Meta information about background data used

This section summarizes the currently used background data versions.
Ecoinvent: Version 3.8
Trade data: 2019
Market data: 2019
Technology data: 2019 to 2020
Transportation distances data: 2021

Annexes

Annex A. Review report by TÜV Rheinland

In the following, a copy of the Review Report by TÜV Rheinland on the 'Critical Review of the Methodology for the LCI Database "cm.chemicals" by Carbon Minds' is attached.

For more information on the review, you can also check the certificate database Certipedia of TÜV Rheinland by using the review ID 0000081021 (https://www.certipedia.com/quality_marks/0000081021?locale=en).

TÜVRheinland®
Precisely Right.

Review Report

15 November 2021

Report Number: CF-2021-09-21253190

Critical Review of the Methodology for the LCI Database “"cm.chemicals" by Carbon Minds

Carbon Minds GmbH
Eupener Str. 165
50933 Cologne

Documents for Review (sent between 22 June 2021 and 1 September 2021):

Methodology
- *cm_chemicals_methodology_V1_2021_01.docx*

Internal technical documentation
- *INTERNAL_METHANOL_docu.docx*

Input Data
- *final_distances.xlsx*
- *plant_production_volumes.xlsx*
- *SystemExpansion_coreLayer_V2.xlsx*
- *SystemExpansion_extensionLayer_V2.xlsx*
- *trade_volumes.xlsx*

Output Data
- *Folder ILCD_sample_TUEV20210722163632 containing sample data for 7 consumption mixes, 7 production mixes, 25 technology mixes in the core layer, 15 technology mixes in the extension layer and ILCD data set template*

Publications
- *Relevant documents on conferences, marketing and scientific publications*

Qualifications
- *Relevant documents on internal meetings, participation in conferences, CVs and certificates*

TÜV Rheinland Energy GmbH
Am Grauen Stein
D-51105 Köln

Tel +49 221 806-5200
Fax +49 221 806-1349
Mail tre-service@de.tuv.com
Web www.umwelt-tuv.de
www.eco-tuv.com

Geschäftsführung und Sitz der Gesellschaft:

Geschäftsführer: Dirk Fenske

Sitz der Gesellschaft: Köln
Amtsgericht Köln HRB 56171
Ust.-Id-Nr.: DE 814653989

Project:	Critical Review of the Methodology for the LCI Database “cm.chemicals”
Client:	Carbon Minds GmbH, Eupener Str. 165, 50933 Cologne
Contact person:	Raoul Meys (Carbon Minds GmbH)
Date of creation:	01.09.2021 (final documents)
Evaluated data:	Life Cycle Inventory (LCI) data for chemical industry
Inspector:	TÜV Rheinland Energy GmbH
Editor:	Susanne Jorre (Sustainability Expert) Susanne Dunschen (Sustainability Expert)
Test criteria:	Reliability, transparency, relevance and representativeness of the methods and data used; Conformity with ISO Norms 14040 and 14044
Scope:	Carbon Minds’ methodology covers Life Cycle Inventory (LCI) data for approx. 1,000 products from the chemical sector in about 190 regions
Review checklist	Checkliste Methodenzertifizierung_Carbon Minds_TÜVRh_12.08.2021_CM_revTR.xlsx
Certificate	C01-2021-09-21253190, valid until 31.12.2022
Certipedia ID	81021

TÜV Rheinland Energy GmbH
Am Grauen Stein, 51105 Cologne

Susanne Jorre
susanne.jorre@de.tuv.com
0221 806 4501

Susanne Dunschen
Susanne.dunschen@de.tuv.com
0221 806 4568

Content

1 General Information and Background of the Study

Carbon Minds GmbH (herafter “Carbon Minds”) has developed a comprehensive Life Cycle Inventory (LCI) database for the chemical sector. TÜV Rheinland Energy GmbH (hereafter “TÜV Rheinland”) has carried out an independent review of Carbon Minds’ LCI methodology.

Scope of the review is to check if Carbon Minds’ methodology for compiling the LCI database are in accordance with the requirements of ISO 14040:2006/Amd.1:2020 and 14044:2006 + Amd 1:2017 + Amd 2:2020. In addition to the actual database, the motivation, method description(s), the technical suitability of the developers with regard to LCI/LCA topics as well as any existing templates and assumptions used, data sources and quality are also critically reviewed. TÜV Rheinland's services include an intensive review of the LCI methodology based on data sampling and comparison with relevant references. Carbon Minds is responsible for the content, methodology and its application. TÜV Rheinland’s conclusions are based on the assumption that the information and data provided by Carbon Minds is complete and accurate.

Carbon Minds’ methodology document and the completed audit as well as review checklist are considered to be framework documents defining the LCI methodology. These documents reflect the goal and scope of the cm.chemicals database, the LCI model, the documentation of LCI datasets and provide the motivation and the current status of implementation by Carbon Minds. The methodology document is publicly available and the review checklists are intended for internal use only.

The aim of this methodology review is to prove the consistency, transparency, relevance and representativeness of the underlying methods and data.

Relevant documents such as the methodology document, the underlying data and samples of LCI datasets were submitted between 22 June 2021 and 1 September 2021. Recommendations concerning the content of the documents, the LCI methodology, data used and open questions were discussed during an online conference call (30 August 2021).

Responsible for the LCI methodology review at Carbon Minds are Dr. Arne Kätelhön, Raoul Meys and Laura Stellner

2 Standards and Criteria

The review is carried out according to the international standards ISO 14040:2006 + Amd.1:2020 and 14044:2006 + Amd 1:2017 + Amd 2:2020.

The review shall ensure that:

- the methods used to model LCI data are consistent with this International Standard,
- the methods used to model LCI data are scientifically and technically valid,
- the data used are appropriate and reasonable in relation to the goal of the study,
- the interpretations reflect the limitations identified and the goal of the study, and
- the LCI dataset documentation is transparent and consistent.

TÜV Rheinland's scope comprised furthermore the following aspects:

- Carbon Minds' intended purpose of the LCI methodology,
- limited desktop review of the used background data,
- limited desktop review of Carbon Minds' data quality management,
- sample based review of technology data, market information, trade date and final LCI datasets.

3 Results of the Critical Review

The LCI data is modeled as required by ISO 14040 and ISO 14044. Carbon Minds has implemented a process to review data on relevance and accuracy, which takes place once a year for the cm.chemicals database and all output datasets. Updates are carried out for all data points that have been identified not to be representative based on Carbon Minds' quality ratings.

The quality assessment of the data used (technology data, market information and trade data) as well as the LCI dataset template for documenting LCI data were discussed in the review meeting. These were assessed as appropriate and suitable by the reviewers. The reviewers' recommendations for optimization of the template were adopted by Carbon Minds.

The reviewers recommend to update the methodology regularly to reflect future developments with regards to technology and data availability within the chemical sector.

The data base for modelling an LCI dataset consists of three types of data:

- Technology data depicting the full mass and energy balances for each production technology;
- Market information including production volumes of chemicals at specific sites for specific processes as well as meta-information on the company operating the plant or the first year of operation;
- Trade data reflecting import and export activities for the chemicals between countries. The data about international trade flows is based on reported information by each country to the United Nations Statistical Division

All data is derived from secondary data sources such as state-of-the-art data providers and literature research. The major sources are renowned databases, peer-reviewed studies and research papers. The different types of data, the data acquisition process as well as the process to derive LCI datasets are described in the methodology report. Regarding the available datasets, Carbon Minds differentiates between plant-specific, supplier-specific, and technology-specific datasets, as well as production mix and consumption mix datasets, depending on the level of detail and geographic scope needed by database users.

Altogether the data quality of the used sources and the quality of the resulting LCI datasets are estimated to be very high. It can be assumed that used data is appropriate. Nonetheless, a verification or quality check of all adduced studies and databases for sourcing data was not performed by the reviewer. To ensure the traceability of data, the modelling and documentation methods were explained within the provided documents. Altogether the data quality seems to be high in relation to the objective of the methodology.

It was noted by the reviewers, that the description of goal and scope, data acquisition, data modelling and documentation is presented in a detailed, consistent and transparent way.

A certification, which displays that the LCI methodology at Carbon Minds' meets the requirements of ISO 14040:2006 + Amd.1:2020 and 14044:2006 + Amd 1:2017 + Amd 2:2020, is approved by the reviewers.

4 Summary

In relation to the standards ISO 14040:2006 + Amd.1:2020 and 14044:2006 + Amd 1:2017 + Amd 2:2020 the reviewers conclude that the LCI database methodology developed by Carbon Minds GmbH is scientifically based and reflects the state of the art. The approach and principles behind the methodology are generally appropriate for the development of LCI datasets of the target industry. Furthermore the data used are appropriate for the goal and scope of the method. Necessary recommendations for the documentation and the datasets were discussed during the video conferences and implemented by Carbon Minds GmbH. For the future, TÜV Rheinland recommends to update the methodology in line with the developments in science and technology and to adapt the methodology document accordingly.

Literature

ISO 14040: 2006 + Amd.1: 2020: International Standard (ISO)
European Committee for Standardization (CEN),
Environmental management – Life cycle assessment – Principles and framework

ISO 14044:2006 + Amd 1:2017 + Amd 2:2020: International Standard (ISO)
European Committee for Standardization (CEN),
Environmental management – Life cycle assessment – Requirements and guidelines

Cologne, 15. November 2021

Susanne Jorre

susanne.Jorre@de.tuv.com
0221 806 4501

Susanne Dunschen

susanne.dunschen@de.tuv.com
0221 806 4568