

HiQLCD: China's High-Quality LCI Database

1. The History of HiQLCD

The core dataset development of HiQLCD began in 2010, when the team initiated research on LCI data for the steel industry, accumulating initial industry experience and data resources. In 2015, the research scope was expanded to include raw materials in the upstream steel sector, further enriching the database content and coverage. By 2020, with the growing demand for carbon management in the energy sector, the database successfully extended to include electricity LCI data, broadening its industry applications and enhancing its comprehensiveness and versatility.

In 2021, in response to the increasing demand for carbon management, HiQLCD established a dedicated LCI data research and development team. Through a systematic and organized approach, the team has driven continuous optimization and improvement of the database, ensuring that both the quality and quantity of data consistently meet the growing market demands. At the end of 2023, the National Data Bureau and 17 other departments jointly released the Three-Year Action Plan for 'Data Factor X', marking a new phase in the development of China data element market and the increasing recognition of the value of data assets. In this context, HiQ SMART DATA was officially established, and the database development work transitioned into a more structured and systematic operational phase, laying a solid foundation for the professional development and market expansion of the database moving forward.

2. About HiQLCD

HiQLCD (High-Quality Life Cycle Database) is the largest and most advanced Life Cycle Inventory (LCI) database in China, designed to meet the unique demands of the Chinese market while serving global sustainability needs. Powered by cutting-edge AI, HiQLCD offers comprehensive and precise data across the entire lifecycle of hundreds of thousands of products, from cradle to grave. Its tailored insights provide accuracy for understanding supply chain carbon footprints, helping businesses refine processes, reduce emissions, and achieve their

environmental goals. For global companies, HiQLCD serves as an essential resource for evaluating China-based supply chains, building accurate regional databases, and driving eco-design initiatives. HiQLCD goes beyond data access with an AI-enhanced platform that simplifies analysis while continuously improving usability and precision. Supported by a team of sustainability experts, it offers actionable guidance for developing sustainable production strategies, optimizing resource utilization, and creating resilient supply chains. Whether addressing local challenges or global objectives, HiQLCD empowers businesses to make informed decisions, foster transparency, and take a leading role in building a sustainable future.

3. Database Overview

HiQLCD database is a Chinese LCI database developed based on the requirements of ILCD format. The database covers multiple product categories such as electricity, heat, coal, oil and gas, steel, non-ferrous metals, metal alloys, chemicals, building materials, photovoltaics, lithium batteries, electronic appliances, waste treatment, transportation, etc. It is a comprehensive LCA background database.

The HiQLCD database aims to provide users with background data that reflects the actual production process of Chinese industrial products, in order to support the implementation of data disclosure based on LCA results (LCA, carbon footprint report), certification (EPD, carbon labeling), product process improvement (ecological design, green supply chain management, etc.), policy research and other application scenarios. To ensure the reliability of data, four strict principles are followed:

Representativeness: Different products have significant differences in production regions and production technologies within a region. For example, there are differences in the way raw coal is mined in China. Of the 31 provinces and cities in China, only 23 have coal mines that can mine raw coal. 14 provinces and cities have both open-pit and underground mining, and 9 provinces and cities have only underground mining. HiQLCD aims to combine the actual situation of products and their supply chains in each industry in China as much as possible, reflecting the differences in regions (refine to the provincial level) and production technologies (specific technology and mixed technologies). By obtaining multidimensional data such as actual production data, industry statistics, and industry emission standards of various

enterprises in China in the past three years, the database reflects the local representativeness of China.

Consistency: HiQLCD database adopts unified rules for data information management, data processing flow, data information recording, etc., meeting ILCD format requirements to ensure format consistency; In addition, consistent methodologies are adopted between different datasets within the same industry to ensure consistency of data within the same industry. Furthermore, the data conforms to the principles of mass and energy balance, with inputs and outputs meticulously accounted for in every process. Relationships between different datasets are logical and coherent. The database employs rigorous validation rules to prevent the entry of illogical or erroneous data, further reinforcing its internal consistency.

High Quality: The Compilation of datasets considers detailed technical data of all production technologies for specific products in China, as well as specific market and trade conditions on the consumer side. The database follows strict data quality control strategies, conducting data quality checks and verifications from multiple dimensions such as data integrity, consistency, and accuracy to ensure the integrity of the database. Each dataset in HiQLCD, including basic information and unit process dataset, has been reviewed by professionals to ensure the accuracy and reliability of the data. Integrating multidimensional data quality evaluation and various uncertainty analysis methods (such as pedigree matrix and Taylor series expansion) to ensure that the dataset has complete uncertainty results at both the flow and result levels. At the flow level, a pedigree matrix is used to score each flow across five dimensions, with default uncertainty assigned to each dimension's score. At the result level, the uncertainty of the final result is calculated using Taylor series expansion for additive propagation.

Transparency: The data provided by HiQLCD has clear source labels, and the origin of each data is clearly visible. Whether it is raw materials, production processes or carbon emission calculation methods, all information is clear at a glance. In this way, users can clearly understand the source of the data and its reliability, ensuring the transparency of the analysis process. HiQLCD supports full process traceability of each piece of data. From the initial entry, correction and merging of data, to model construction, calculation, verification, review, and final release and subsequent revisions, each step of the operation leaves a clear record. Through

these complete processes, users can understand in detail how the data is generated and changed, ensuring the credibility and scientific validity of the data.

4. Data Format and Categories

The HiQLCD database is provided in ILCD file package format for use in professional LCA software. The database contains three types of data: unit process dataset, LCI dataset, and LCIA results. The LCI dataset and LCIA results can be viewed and downloaded for application in HiQLCD data search platform. The most basic object for LCI calculation in HiQLCD database is the unit and system process dataset, which can output calculations containing various LCIA methods and their indicators such as IPCC2021, CML v4.8 2016, EF 3.0, Recipe, etc.

The dataset of HiQLCD database includes two major categories: product dataset and process dataset. Product dataset is a dataset that focuses on products and can be used independently; Process dataset is a dataset that focuses on the processing process. For example, a hot rolling process dataset. The hot rolling process is to roll a preheated steel billet into a flat plate with a certain thickness at a high temperature. Users can match the corresponding number of background datasets according to the weight of the steel billet processed by the hot rolling process being evaluated. This process should be used together with the production data of 1kg continuous casting billet, representing the hot rolling of 1kg continuous casting billet, corresponding to the generation of 0.949kg hot rolled coil. Compared with a product dataset, it does not consider the consumption of raw materials, but only considers the flow of categories such as energy, auxiliary materials, and emissions in the process. It is often used in conjunction with product class. In addition, the dataset distinguishes between production and market datasets based on different boundaries. Production datasets include the boundary range from product cradle to gate, while market datasets consider the downstream transportation process based on production datasets, integrating logistics and related transportation data within specific product areas.

5. Database Content Highlights

5.1 Electricity

Currently, the HiQLCD database's electric power sector contains over 1,000 datasets,

covering the entire process of the power system from production (including the construction of power plants and the manufacturing of major power generation equipment), transmission, conversion, and distribution to consumption. In terms of form of power generation, it involves five major forms: thermal power, hydropower, wind power, nuclear power, and photovoltaic power, representing 100% of statistically represented China's electricity generation. In terms of transmission and distribution (T&D) and power consumption options, it comprises voltage levels of 220 kV and above, 110 kV, 35 kV, 10 kV, and less than 1 kV, as well as providing options for high voltage, medium voltage, and unspecified voltage levels. In addition, it is fine-grained to the provincial level and reflects regional heterogeneity, while six regional grid levels (i.e., North China Grid, Central China Grid, East China Grid, Northeast China Grid, Northwest China Grid, and South China Grid) and national level data are provided to comply with various standards. According to the latest China Electricity Statistical Yearbook release, the reference year for power sector data is 2021.

5.2 Steel

China's steel industry is the largest steel production and consumption market in the world, and has a significant impact on global steel supply and demand. The industrial chain of China's steel sector covers the entire process from raw material acquisition, manufacturing, product sales, and recycling.

Based on this industry background, the HiQLCD database has established a comprehensive dataset covering the entire life cycle of steel products, from raw material acquisition and processing in the upstream to the production and sale of finished steel products in the downstream. The classification is more refined, and the variety of products is more extensive. Compared to other databases, the HiQLCD database has detailed key parameters for different products and steel grades, including material input/output and energy consumption. A total of over 6,000 datasets have been created for various parts of the steel industry, meeting the specific needs of different sectors. This includes, but is not limited to, core processes such as sintering ore, pelletizing ore, blast furnace ironmaking, direct reduced iron, converter steelmaking, electric arc furnace steelmaking, hot rolling, cold rolling, and other production processes and corresponding products. It also highlights the differences in specific production technologies and mixed techniques. The dataset categorizes different steel types, such as carbon

steel, low-alloy steel, alloy steel, stainless steel, silicon steel, bearing steel, and special steels, down to individual steel grades.

To ensure more comprehensive and accurate data sources, the HiQLCD database collects data from the actual production data of more than 50% of the large steel enterprises across China. These data are cross-verified from bottom to top. Additionally, the geographic coverage is detailed down to the provincial level, reflecting the impact of regional differences in raw material grades and other factors on China's steel industry structure. This provides users with precise regional information and regional disparity analysis.

In addition to the raw materials and product data directly used in steel production, the HiQLCD database also includes a comprehensive dataset of auxiliary materials required by the steel industry. For example, HiQLCD database for the steel industry is exceptionally comprehensive, consisting of over 10,000 datasets that cover all the materials needed by the steel industry. This extensive collection of data helps to form a complete steel industry data package, enabling deeper insights and better decision-making across the entire steel production chain, from raw material procurement to final product manufacturing and recycling.

5.3 Coal

The HiQLCD coal industry dataset contains over 600 datasets, divided into production and market datasets, and comprehensively covers the major stages of the coal lifecycle, including mining, processing, transportation, and sales.

The production dataset provides detailed data on mining processes, reflecting the specific mining technologies used in China, including underground mining (shaft, inclined, and horizontal tunnels) and open-pit mining methods. It also includes energy consumption data for different types of mines, as well as methane emissions from high-gas and gas-outburst mines. The categories including anthracite, bituminous coal, and lignite, and further divided into thermal coal and coking coal based on their uses. The processing section focuses on China's diverse coal washing techniques, including wet dense medium separation, jigging, flotation, and dry processing, with detailed records of energy consumption, material flows, and emissions for each process, meeting the specific needs of various industries.

The market dataset examines coal production and consumption patterns across 31 provinces in China, as well as the import and export of coal. The transportation dataset focuses

on China's unique coal transport network, primarily rail-based with supplementary water and road transport. It highlights the typical "West-to-East" and "North-to-South" transportation routes, along with the energy consumption and emissions characteristics associated with these transport modes.

The HiQLCD coal industry dataset is carefully tailored to reflect the realities of China's coal industry, providing high-quality localized data that supports precise and comprehensive LCA (Life Cycle Assessment) analysis, helping users make informed decisions in carbon footprint calculation, green design, and supply chain optimization.

5.4 Iron Ore

The HiQLCD iron ore industry dataset contains over 500 datasets, mainly divided into production and market datasets, covering key lifecycle stages of iron ore, including mining, beneficiation, and transportation.

The production dataset includes data on iron ore extraction, which covers three main mining ways: open-pit mining, underground mining, and combined open-pit and underground mining. The dataset is further subdivided into various mining techniques such as caving, filling, and shallow-hole leave-mining. The input-output inventory system reflects key metrics such as recovery rate, dilution rate, and resource utilization efficiency for each mining method. The beneficiation section focuses on different processing techniques for various ore types (e.g., magnetite, hematite). For magnetite ore, the processing methods include stage grinding, weak magnetic separation followed by reverse flotation, full magnetic separation, ultra-fine grinding with wet magnetic separation tailings treatment, and composite selection for low-grade magnetite. Hematite ore typically requires more complex beneficiation processes, often using combined methods to improve concentrate grade and recovery rate, such as stage grinding, coarse-fine separation, gravity separation-magnetic separation-anionic reverse flotation, and continuous grinding with magnetic separation and anionic reverse flotation. The input-output inventories for different ore types and beneficiation methods vary significantly, leading to differing carbon emission characteristics.

The market dataset includes comprehensive research on iron ore and iron concentrate imports and exports across China's 31 provinces. It also provides detailed insights into the supply characteristics of major exporting countries such as Australia, Brazil, India, and South

Africa, and explores the application scenarios of imported ores in the Chinese market.

While China's total iron ore resources are relatively abundant, the amount of exploitable high-grade ore is limited. The country's iron ore industry is characterized by low-grade ores, a high proportion of poor-quality ores, many small to medium-sized mines, and complex ore types. The HiQLCD iron ore industry dataset is centered on the specific characteristics of China's iron ore industry, providing accurate and professional data to support life cycle assessments (LCA) and supply chain optimization. This enables users to conduct environmental impact assessments, carbon footprint calculations, and make informed industry decisions.

5.5 Waste Treatment

The HiQLCD database includes over 3,000 datasets related to waste management, covering Solid Waste Management (SWM) and Wastewater Treatment (WWT). Solid Waste Management encompasses recycling processes and end-of-life disposal (such as incineration and landfilling), and covers four major stages of waste: generation, collection and transportation, treatment, recycling, and final disposal. For municipal solid waste, HiQLCD has developed a regionalized underlying calculation model, focusing on three core features: waste composition, treatment region, and treatment method. By integrating regional pollutant emission survey data from China, a provincial and municipal-level waste treatment dataset is generated. In collaboration with universities, HiQLCD has also developed materials recycling datasets, supporting users in flexibly selecting recycling process chains in compliance with LCA modeling requirements.

For the wastewater treatment datasets, HiQLCD has surveyed the process usage of wastewater treatment plants across China in 2022. The process data sets are combined to facilitate region-specific modeling of wastewater treatment, with the ability for parametric updates to reflect actual regional conditions. Pollutant discharge information is sourced from China's regional pollutant survey data. The wastewater treatment dataset includes both market datasets that consider pipeline transportation and treatment datasets that do not consider pipeline transportation, for users to choose according to their needs. The treatment models cover three typical scales of treatment plants: urban medium-sized, county-level small, and township-level small wastewater treatment plants.

5.6 Oil and Gas

The HiQLCD database focuses on the field of oil and gas, with its datasets covering all aspects from oil and gas extraction to refining as well as information about related products. It provides over 500 regional datasets in China, including datasets for crude oil, natural gas, and downstream products, as well as market datasets that fully consider factors such as regional consumption, transportation, and material loss emissions. The datasets of "Crude Oil Market, China" and "Natural Gas, high pressure Market, China" combine international trade reality to accurately present the situation of China's oil and gas consumption.

This oil and gas dataset is based on the theories of petroleum and systems engineering to construct a bottom-layer model accounting system for the whole process, attaching great importance to the heterogeneity of petroleum resources. The data collection comprehensively considers multiple core indicators. It carefully selects 242 oilfield blocks that were in production in 2021, with their 2P reserves exceeding 50 million tons. The oil production of these blocks in that year accounted for more than 70% of the total domestic production, and the average annual oil production during the historical production period accounted for more than 85% of the total domestic production. These blocks are distributed in major oil and gas basins such as the Songliao Basin, Ordos Basin, and Tarim Basin, and are operated by 24 oilfield companies under major oil and gas enterprises such as PetroChina, Sinopec, CNOOC, Yanchang Petroleum, and private companies. The basic data related to oil production is sourced from the collation of academic literature and public reports, which, through rigorous organization and collection, effectively ensure the authority and reliability of the data.

6 Data Provision and Services

As the data provider, HiQLCD is committed to offering customers high-quality, high-standard LCA datasets through LCA Tools like openLCA, jimuLCA, etc.

6.1 Other Technical Support

1) If users need to deploy data on other tool platforms, they can seek additional technical support from HiQ Smart Data.

2) Users can choose whether to require data update and iteration services when purchasing the data package. Haike Data will provide regular updates and maintenance services for the data package. For more information, please contact info@HiQLCD.com.