Data and Usage Guideline Food Impacts Toolkit

Part of the IKI project
Climate Impacts of Food (CLIF)

Hamburg, 13. November 2024

corsus – corporate sustainability GmbH | Großneumarkt 50 | 20459 Hamburg | +49 40 398084-76 | www.corsus.de
Geschäftsführung: Dr. Ulrike Eberle, Dr. Anke Butscher | HRB 159654 | Gerichtsstand Hamburg

Contents

1 Introduction

Food production is a leading contributor of anthropogenic environmental impacts, particularly through agriculture's role in pushing several planetary boundaries. According to Campbell et al. (2017), agriculture is responsible for about 80% of changes to land systems and biodiversity loss, along with 84% of global freshwater use. The sector also increases nutrient pollution, accounting for around 85% of human-driven nitrogen flows and over 90% of phosphorus flows. Lastly, the authors estimate agriculture to contribute around 25 % to anthropogenic climate change.

While agriculture's impacts on the earth system are considerable, individuals' dietary choices can make a big difference on how large the resulting environmental impacts are (see e.g. Willett et al., 2019; Poore & Nemecek, 2018). Along with the reduction of food waste across all parts of the value chain, dietary change thus is one of the major levers for reducing food's environmental impacts.

To make more environmentally friendly choices, consumers and meal providers require information about the environmental impacts of food. However, this information is not commonly provided, and they consequently cannot easily make informed choices, even if they wish to improve their meals' environmental footprint. This is the focus of the IKI CLIF project: to equip consumers and meal providers with clear, accessible information they need to take steps towards an environmentally friendly way of eating.

One powerful tool for generating this information is Life Cycle Assessment (LCA), a well-established method for assessing the environmental impacts of products. LCA works by modelling emissions and resource use throughout a product's life cycle. This tool provides a comprehensive view of environmental impacts across categories such as climate change, resource depletion, ecosystem quality, and human health. This approach helps to avoid so-called burden shifting, where a product system is optimised towards reducing impacts in one environmental category (e.g. climate change) at the expense of unintentionally increasing impacts in other categories.

To effectively model the environmental impacts of food, extensive data is required, which is often not available. Because of this, databases of generic food products play an important role in modelling foods' impacts. Apart from providing impact data for different agricultural products, databases also apply consistent modelling assumptions across products, something that is very important for maintaining comparability between different products.

In the CLIF project, our goal was to provide information on food products in an easy, accessible and flexible way and relevant for users in across different regions including partners in Taiwan, South Africa, and Paraguay. To achieve this, we created a freely available, open-source tool called *Food Impacts Toolkit* (FIT). Currently a prototype, FIT calculates impact scores for individual products as well as recipes and delivers the data in a flexible way, so that it can then be displayed to different user groups. FIT is designed for global application by adapting regional datasets to estimate environmental impact values in different areas with minimal local data input.

Since data availability is limited, we used a publicly available French database (more on this in section 2.1). While this enabled us to provide many products for one geographic context, data for other regions was mostly missing. To be able to provide products from more regions, a proxy method was developed to better represent different production regions. Apart from this, we encourage the inclusion of further product data specific to additional geographies and production systems, always considering that consistent modelling principles need to be applied to maintain comparability.

To this end, this document will first outline the methodological choices made to arrive at the impact results that we provide. A brief LCA method overview is followed by specifics on what results we include and how we arrive at them. Apart from expanding on included impact categories and their aggregation into a single score, the section will also describe current limitations that are crucial for users to understand. The next section will then deal with FIT and describe what the aim of the tool is, which functionality it implements, how the proxy data the tool provides was created, and what the (current) limitations of the tool are. Crucially, the guideline also contains instructions on how to deploy the tool and how it can be interacted with.

2 Methodology

After giving a brief LCA method overview, this section details relevant methodological choices (normative decisions) in the context of FIT/CLIF. Lastly, methodological limitations are discussed.

2.1 Life cycle assessment method overview

Life Cycle Assessment (LCA) is an established and standardised approach for comparing products' impacts throughout their life cycle, i.e. including the sourcing of raw materials, production of precursor and the final product, involved transports, product use, maintenance, and end-of-life disposal. While LCA studies are commonly conducted to gauge and compare the environmental impacts of products, they can also be employed to study social impacts (social LCA).

In the context of environmental assessment, LCA studies cover a broad range of environmental issues, primarily to avoid overlooking environmental impacts by only focusing on a subset of them. LCA applies scientific principles and methods to quantify these impacts as precisely as is feasible and required within the context of any given study. The principal steps of each LCA study are outlined in Table 1.

Table 1: Overview of iterative phases of an LCA study.

LCA study phase	Description
Goal definition	The goal definition lines out why the study is performed, for whom it is performed and which specific questions it seeks to answer.
Scope definition	The definition of the scope follows from the formulated goal and includes choices such as ▪ The functional unit (FU) of the study, which quantifies the function or service for which impacts are being assessed ▪ Defining which processes and activities lie within the product system's system boundary and which parts of the life cycle need to be included to reach the study goal. ▪ Deciding which environmental impacts of the product system should be assessed and which impact assessment methods should be used. ▪ Deciding on the geographical and temporal boundaries of the product system.
Inventory analysis	During the inventory analysis, information on physical flows of the product system (inputs of resources, materials, precursor products, products, energy; outputs such as waste materials, substance emissions, valuable materials) is being collected and quantified for all processes and activities within the system boundary. Secondary data and generic data are commonly employed for background systems because of the size and complexity of the inventory. The finished analysis yields the life cycle inventory (LCI).
Impact assessment	Starting with the life cycle inventory, the impact assessment converts the product system's physical flows and interventions into environmental impacts by applying knowledge and models from environmental science. Commonly, these models are applied in sets. This is the case for the EF method, which combines 16 different impact assessment models and their indicators for environmental impact assessment.
Interpretation	The interpretation phase of an LCA aims to identify and evaluate the main environmental impacts of the product system, ensuring they align with the study's goals. This phase often involves iterating through previous phases, refining the scope, and updating the life cycle inventory (LCI) as needed. It may also require using additional or different impact assessment models or indicators.

General LCA principles are outlined by ISO standards.[1] However, these standards do not (and cannot) address every aspect of the method implementation for every kind of product or service – they inevitably leave room for interpretation. Because of this, many complementary standards and guidance have been developed by different bodies, seeking to further standardise and harmonise method implementation, thereby increasing comparability across different studies and particularly within product groups.

---

[1] The standards are ISO 14040:2006, including amendment ISO 14040:2006/Amd 1:2020, and ISO 14044:2006/Amd 2:2020, including amendments ISO 14044:2006/Amd 1:2017 and ISO 14044:2006/Amd 2:2020.

At the European level, the Environmental Footprint (EF) method was developed to this end. At the product level, the method is called Product Environmental Footprint (PEF), and it includes a growing set of rules for specific product categories, so-called category rules (PEFCR). The category rules specify how LCIs for specific types of products should be modelled. One salient example is the allocation of burdens: As soon as product systems have more than one function, as is commonly the case with animal product systems (e.g. the production of milk, meat and leather from bovine systems), common allocation principles and guidance are needed to maintain comparability across studies. Allocation also plays a role for plant-based product systems (e.g. between the main fruit and remaining plant parts that are sold as animal feed).

Apart from congruent rules, modelling the environmental impacts of food requires a lot of data. Because data availability is limited, LCA practitioners commonly resort to databases that contain generic version of the required processes. This is a perfectly valid approach for modelling the background system, components such as upstream goods and services (e.g. electricity, fuel, fertilisers, transportation, provision of raw materials) that are inputs to the foreground system, the main processes that make up the product system and for which data is more commonly available.

In an ideal world, we would have sufficient data and resources to carry out LCA studies for the food products we buy at the supermarket, considering the specific foreground system (production methods and sourcing modalities) that characterise the product in question. The current reality, however, is that we must resort to generic products because we do not have LCA data for most products on offer.

We decided to utilise the French public database Agribalyse[2] in the context of this project for several reasons. Firstly, Agribalyse offers a broad range of products as its aim is to cover the most important food products available in the French market. Today, Agribalyse contains around 2,700 products (including processed foods and ready-made meals). Secondly, Agribalyse is available free of charge, which greatly increases our work's accessibility. Furthermore, Agribalyse is largely aligned with the EF method, as well as with major databases it utilises (ecoinvent, World Food LCA Database).

The caveat of this approach is that to date, we mostly work with data that is specific to the French context and thus less applicable to other regions. To address this, we implemented a proxy calculation method as an interim solution in the context of the prototype (see section 3.3 for details). We would like to stress that further data is required to increase the validity of the impact assessment for an increasing number of geographies and production systems.[3]

---

[2] Agribalyse version 3.1 was utilised, it is available via the project homepage at: https://doc.agribalyse.fr/documentation-en (last visited 06.11.2024).

[3] In this context the European LIFE project ECO FOOD CHOICE is developing a method, how to extrapolate data from Agribalyse to other regions. In future, this method can be used to create national databases and thus, give more valid information on regional level.

2.2 Impact assessment and single score calculation

As outlined before, the impact assessment enables understanding the effects that the product system has on different aspects of the environment. Impacts are expressed in terms of impact category indicators (e.g. GWP_{100} for climate change) that utilise different units; an overview of the employed impact categories, their indicators and units is given in Table 2. Note that the table also contains a robustness assessment for each model (ranging from most robust (I) to least robust (III)). Robustness expresses the accuracy of the impact modelling.

Impact category results can be reported separately, but they can also be combined into a single score (or multiple sub-scores of related indicators) to increase the interpretability of results and to enable more straightforward comparisons of products' environmental performance. Various steps are required for the calculation of the combined score(s): first, the different indicators units need to be normalised (see section 2.2.3 for details). A second step involves assigning weights to the individual categories, i.e. determining their relative importance for the single score result. Optionally, method robustness can be considered, which is commonly done to increase the weight of more robust models relative to those which display larger uncertainty. These steps are highly normative as there is no objective way to carry them out: different stakeholders will judge the relative importance of environmental impacts differently. Therefore, in the context of the EF method, an extensive multi-stakeholder process involving experts and non-experts was carried out to derive the weighting factors (Sala et al., 2018).

Table 2: Impact assessment methods used by FIT; coincides with EF 3.1 set and most information is taken from the corresponding publication (Zampori & Pant, 2019); biodiversity impact category added; green shaded categories included following Delphi results.

Impact category (& abbreviation) (* = non-EF)	Impact category indicator	Unit	Characterisation model[4]	Robustness
Climate change (CC)	Radiative forcing as global warming potential (GWP100)	kg CO_2e	Baseline model of 100 years of the IPCC (based on IPCC 2013)	I
Ozone depletion (ODP)	Ozone Depletion Potential (ODP)	kg CFC-11 eq.	Steady-state ODPs as in (WMO 2014 + integrations)	I
Human toxicity, cancer (HTC)	Comparative Toxic Unit for humans (CTU_h)	CTU_h	USEtox model 2.1 (Fankte et al, 2017)	III
Human toxicity, non-cancer (HTNC)	Comparative Toxic Unit for humans (CTU_h)	CTU_h	USEtox model 2.1 (Fankte et al, 2017)	III

---

[4] In case this literature is not cited elsewhere in this document, please refer to Zampori & Pant (2019) for the references in question.

Particulate matter (PM)	Impact on human health	disease incidence	PM method recommended by UNEP (UNEP 2016)	I
Ionising radiation, human health (IRHH)	Human exposure efficiency relative to U^{235}	kBq U^{235} eq.	Human health effect model as developed by Dreicer et al. 1995 (Frischknecht et al, 2000)	II
Photochemical ozone formation, human health (OZF)	Tropospheric ozone concentration increase	kg NMVOC eq.	LOTOSEUROS model (Van Zelm et al, 2008) as implemented in ReCiPe 2008	II
Acidification (AP)	Accumulated Exceedance (AE)	mol H+ eq.	Accumulated Exceedance (Seppälä et al. 2006, Posch et al, 2008)	II
Eutrophication, terrestrial (EPT)	Accumulated Exceedance (AE)	mol N eq.	Accumulated Exceedance (Seppälä et al. 2006, Posch et al, 2008)	II
Eutrophication, freshwater (EPFW)	Fraction of nutrients reaching freshwater end compartment (P)	kg P eq.	EUTREND model (Struijs et al, 2009) as implemented in ReCiPe	II
Eutrophication, marine (EPM)	Fraction of nutrients reaching marine end compartment (N)	kg N eq.	EUTREND model (Struijs et al, 2009) as implemented in ReCiPe	II
Ecotoxicity, freshwater (ETFW)	Comparative Toxic Unit for ecosystems (CTU_e)	CTU_e	USEtox model 2.1 (Fankte et al, 2017)	III
Land use (LU)	▪ Soil quality index ▪ Biotic production ▪ Erosion resistance ▪ Mechanical filtration ▪ Groundwater replenishment	▪ Dimensionless (pt) ▪ kg biotic production ▪ kg soil ▪ m^3 water ▪ m^3 groundwater	Soil quality index based on LANCA (Beck et al. 2010 and Bos et al. 2016)	III
Water use (WU)	User deprivation potential (deprivation-weighted water consumption)	m^3 world eq.	Available WAter REmaining (AWARE) as recommended by UNEP, 2016	III
Resource use, minerals and metals (RUMM)	Abiotic resource depletion (ADP ultimate reserves)	kg Sb eq.	CML 2002 (Guinée et al., 2002) and van Oers et al. 2002.	III

Resource use, fossils (RUF)	Abiotic resource depletion – fossil fuels (ADP-fossil)	MJ	CML 2002 (Guinée et al., 2002) and van Oers et al. 2002	III
* Biodiversity impact, terrestrial (BIOTER)	Biodiversity value increment (BVI)	BVI * m^2a	(Lindner et al., 2019)	III

2.2.1 Food specific impact assessment

The EF method aims at enabling comparisons between different product categories. Because of this, an inclusion of the full set of EF indicators makes sense. In the context of comparing food products, an effort was made to identify the most important environmental impacts of food and reduce the number of assessed impact categories accordingly.

As part of the CLIF project, a Delphi study[5] was carried out to identify these categories and thereby reduce the number of impact categories. The central question of the study was "Which are the most relevant environmental impacts of food?" Participants included food sector stakeholders from four countries (Germany, Paraguay, South Africa, and Thailand), as well as international LCA experts on food. The underlying hypothesis for a reduction of impact categories is that food production has a characteristic contribution to exceeding certain planetary boundaries (Campbell et al., 2017), and therefore not all environmental impact categories are equally significant for food. By reflecting this in our choice of impact categories, we give more relative weight to those categories that are central for food products, at the same time increasing the interpretability of results. Furthermore, the findings from the Delphi study were substantiated by statistically analysing the Agribalyse database using multiple linear regression analyses. These results underpinned that analysing only one impact category (e. g. climate change) is not sufficient to depict environmental impacts of food. At least, the impacts of water use need to be included in a single score, better also impacts on biodiversity and fourthly the impacts on eutrophication.

The indicators included following the results of the Delphi study are the ones that are shaded blue in Table 2. The weighting of the included categories was done relative to how frequently participants regarded them as relevant. The resulting weighting factors can be found in Table 4.

2.2.2 Including biodiversity impacts

One of the major shortcomings of the EF set of indicators is its lack of a biodiversity impact assessment. In order to address this gap, we turned to the terrestrial biodiversity impact assessment method developed by Lindner and colleagues (Lindner et al., 2019, 2020; Lindner & Knüpffer, 2020). The method uses management parameters to determine how far from a natural state a given area is due to an anthropogenic intervention (like farming or mining). This measure of area naturalness is combined with a region-specific ecoregion factor that expresses how valuable the area is for

---

[5] Publication of study results forthcoming.

biodiversity (e.g. species-rich, harbouring endemic or threatened species). Thereby, an activity's potential impact on biodiversity becomes quantifiable; the unit employed for this is the biodiversity value increment (BVI). Throughout this document, where reference to this method is made, it is referred to as BVI method.

The BVI method has been proposed as an addition to the Agribalyse database, and in 2022 a project was carried out to explore the feasibility of implementing it for the large number of products contained in Agribalyse (Lindner et al., 2022). We used the project's results, which are publicly available through the Agribalyse website[6], to complement the set of EF indicators contained in Agribalyse.

2.2.3 Normalization of impact category results

Since the impact category indicators have different units (cf.), normalization is necessary to set the results in relation to a common point of reference ("reference unit"). For this purpose, the annual impact of an average global citizen is determined for all included impact category indicators, i.e. "within the [Product Environmental Footprint] method the normalisation factors are expressed per capita based on a global value." (Zampori & Pant, 2019, Chapter 5.2.1) The corresponding normalization factors are part of the EF reference package 3.1 provided by the European Commission.[7]

---

[6] The results and the corresponding report can be found on the documentation pages at: https://doc.agribalyse.fr/documentation-en/agribalyse-data/documentation (last accessed 24.10.2024).

[7] Accessible via https://eplca.jrc.ec.europa.eu/LCDN/developerEF.html (last accessed 24.10.2024).

Table 3: Employed normalisation factors, as provided in the EF reference package 3.1; biodiversity normalisation factor provided by ADEME.

Impact category (* = non-EF)	Unit	Normalisation factor
Acidification	mol H^+ eq./person	5.56E+01
Climate change	kg CO_2 eq./person	7.55E+03
Ecotoxicity, freshwater	CTUe/person	5.67E+04
EF-particulate matter	disease incidences/person	5.95E-04
Eutrophication, freshwater	kg P eq./person	1.61E+00
Eutrophication, marine	kg N eq./person	1.95E+01
Eutrophication, terrestrial	mol N eq./person	1.77E+02
Human toxicity, cancer	CTU_h/person	1.73E-05
Human toxicity, non-cancer	CTU_h/person	1.29E-04
Ionising radiation	kBq U^{235} eq./person	4.22E+03
Land use	pt/person	8.19E+05
Ozone depletion	kg CFC-11 eq./person	5.23E-02
Photochemical ozone formation	kg NMVOC eq./person	4.09E+01
Resource depletion, fossils	MJ/person	6.50E+04
Resource depletion, minerals and metals	kg Sb eq./person	6.36E-02
Water use	m^3 water eq of deprived water/person	1.15E+04
* Biodiversity impact, terrestrial	BVI * m^2a	1.35E+14

2.2.4 Weighting of results

Weighting is carried out to assign relative importance to indicators before combining them and includes an optional robustness adjustment. The six weighting sets utilised in the context of this project are:

- Weighting in accordance with the **EF method**, following Sala et al. (2018) and utilising multiple variants:
    - Using **robustness factors ranging from 0.1 to 1.0**, which means that lower robustness reduces the weighting strongly
    - Using **robustness factors ranging from 0.5 to 1.0**, which means that lower robustness reduces the weighting less strongly

  - Using **no robustness factors**, which means that lower robustness has no impact on the weighting

- Weighting according to Delphi study findings and applying the same robustness factor variants as with the EF method (Table 4)

Biodiversity weighting

Since the EF set was complemented with the BVI method, a choice regarding the weighting of the indicator had to be made. It was decided to use the weighting proportion between climate change and biodiversity that was determined via the Delphi study, which means that the biodiversity impact is weighted virtually the same as the climate change impact in both the extended EF and the Delphi weighting schemes.

Biodiversity robustness

Because the EF method does not provide a robustness factor for biodiversity and because the BVI indicator's robustness has been rated as III (low robustness) in the context of its Agribalyse trial, it was assumed the lowest robustness ratings in the respective ranges for the weighting variants that include robustness (0.17 and 0.57, respectively (Sala et al., 2018, Table 30)).

Table 4: Weighting and robustness factors employed for the modified EF method and the reduced indicator set derived from the Delphi study.

Impact category name	Delphi weighting	Modified EF weighting	Robustness factor (scale 0.5-1.0)	Robustness factor (scale 0.1-1.0)
Climate change	23.35	11.43	0.93	0.87
Ozone depletion	0.00	4.94	0.08	0.60
Human toxicity, cancer effects	0.00	6.03	0.57	0.17
Human toxicity, non-cancer effects	0.00	5.21	0.57	0.17
Particulate matter	0.00	4.86	0.93	0.87
Ionizing radiation, HH	0.00	5.05	0.73	0.47
Photochemical ozone formation, HH	0.00	4.22	0.77	0.53
Acidification	0.00	4.38	0.83	0.67
Eutrophication, terrestrial	5.14	2.61	0.83	0.67
Eutrophication, freshwater	5.56	2.83	0.73	0.47
Eutrophication, marine	5.12	2.61	0.77	0.53
Ecotoxicity freshwater	19.76	5.42	0.57	0.17

Land use	0.00	8.01	0.73	0.47
Water use	17.87	8.59	0.73	0.47
Resource use, mineral and metals	0.00	5.92	0.08	0.06
Resource use, fossils	0.00	6.53	0.08	0.06
Biodiversity, terrestrial	23.21	11.36	0.57	0.17

2.3 Methodological limitations and challenges

While LCA is a powerful tool for assessing environmental impacts, it has some notable limitations, especially in food systems:

- **Ecosystem services** (like soil health or biodiversity) are difficult to model and thus commonly not sufficiently quantified. As a result, an organic product's benefits may be undervalued, as organic farming often has advantages in these areas. We addressed this with the inclusion of the BVI method, but e.g. soil health is still not directly considered.
- As already mentioned, **limited data quality and availability**, especially regarding diverse farming practices and regional differences pose a challenge. Farming methods are variable depending on local practices and climates, making it hard to obtain representative data. This data variability means that LCA results can lack accuracy. If extensive or organic production systems are not modelled explicitly, their strengths are not captured.
- The choice of the **functional unit** in food LCAs also poses a challenge. LCAs often measure impact per kilogram of product, but this may not consider nutritional quality or benefits. Organic foods may have lower yields per hectare, which can look less efficient in LCA results. Proponents of using nutritional value as the functional unit argue that it would enable a more balanced comparison of organic and conventional options.

Overall, care should be taken while interpreting the results, always keeping in mind that the underlying data is generic and does not account for all kinds of production systems and regions, and that the impact assessment does not fully cover all environmental impacts of product systems (yet aiming at being as complete and relevant as possible). Furthermore, social impacts are not part of the assessment at all, as they are beyond the scope of the project.

3 Food Impacts Toolkit

The following section describes the Food Impacts Toolkit (FIT) prototype at a conceptual level, addressing its aim, what functionality it implements and how it provides proxy data for regions that are not part of the original dataset we employed. Lastly, it details current limitations and opportunities for further research and development.

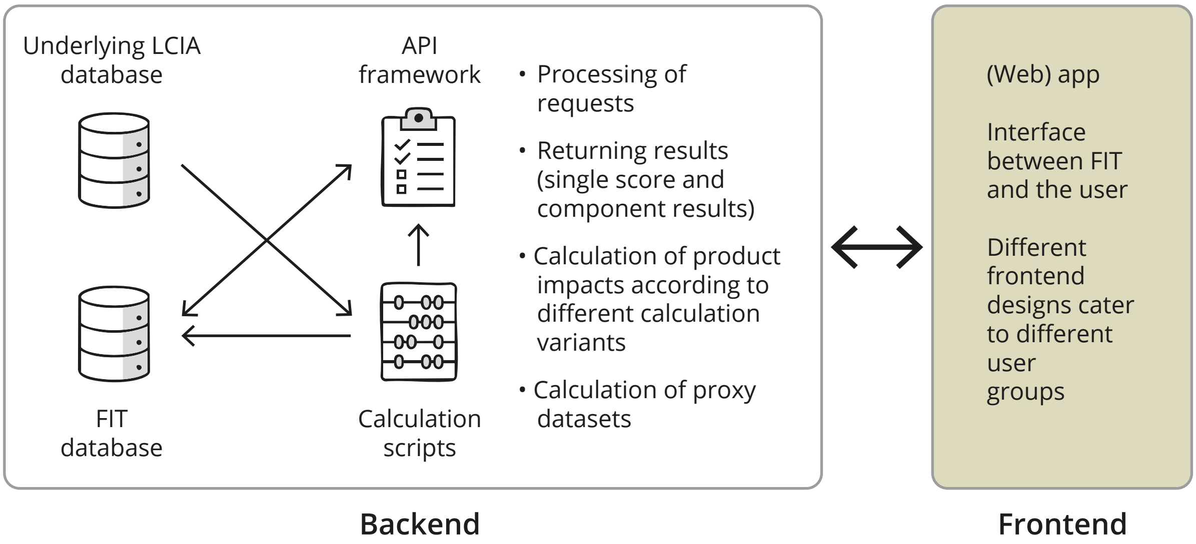

Figure 1: Graphical overview of the FIT backend and its relationship to the frontend.

3.1 Aim of the project

The project aim for FIT is developing a functional prototype of freely available API that provides data on foods' environmental impacts in a flexible way. Users of the API can then implement different ways of presenting the information to cater to different target groups (e. g. end consumers, meal providers).

Within the context of the project, it was decided to implement FIT at the level of impact assessment results, which has two main advantages: firstly, it reduces the cost of implementation as dealing with inventory data would have been significantly more complicated. This results in that FIT was limited in the adjustments of input datasets.[8] The second advantage is that of licensing: using Agribalyse data at an inventory level would require users outside of France to purchase an ecoinvent licence, which is not the case with the impact assessment results which were utilised, which are made available under the French Etalab licence (comparable to the Creative Commons Attribution licence).

The choice of Agribalyse and employing the EF method also aims at maintaining a high degree of alignment with existing databases and standards. Instead of re-inventing the wheel, FIT was embedded into the context of existing efforts, albeit without being prescriptive and while complementing existing assessment approaches (by including terrestrial biodiversity impacts and offering alternative weighting schemes, see section 2.2).

On a technical level, it was sought to implement the API using state of the art technologies to facilitate ease-of-use (such as building a containerised application and using commonplace data exchange formats).

---

[8] However, more complex adjustments to the data can of course be made before making the resulting impact assessment results available for use in the API.

3.2 Implemented functionality

The API provides an in-depth assessment of environmental impacts for individual products or recipes, based on LCIA data. Designed to support flexible input and a detailed display of results, the API enables users to submit recipes with multiple items, specifying quantities and optional weighting schemes, and returns a comprehensive impact assessment across environmental categories and life cycle stages.

Key features include:

- **Flexible data input**: Users can submit recipes by listing items with specific quantities and optionally choose a weighting scheme. Each item is identified with relevant geographic data to ensure tailored impact calculations.
- **Environmental impact assessment**: The API calculates the environmental impact of each item and the entire recipe, including breakdowns across life cycle stages (such as agricultural stage and processing) and impact categories (like climate change).
- **Aggregation, scoring, and grading**: Through methods of normalization, weighting, and scaling, the API delivers a summarised view of each item's impact, as well as the recipe's overall environmental impact. This includes a single score for high-level assessment, detailed scores per life cycle stage and impact category, and graded performance indicators (e.g., A, B, C) for each category, ensuring accessible and clear interpretation of results.
- **Proxy data**: The API provides datasets for more regions than is originally the case with Agribalyse. This increases the results' relevance in different geographical contexts. At the same time, these modifications are made transparent: If any data is estimated or inferred, the API indicates it, allowing users to see which results are based on actual data and which involve approximations.
- **Organized output**: The API's response includes a structured dataset covering general recipe information, single and aggregated scores, and details by life cycle stage and impact category. This dataset is designed for easy integration into other analysis tools or reporting frameworks.
- **Deployment and versioning**: Managed through version control and containerization, the API is packaged for consistent deployment and straightforward updates. Docker creates isolated environments to ensure reliable deployment across different systems, while git version control enables seamless collaboration and tracking of updates over time.

3.3 Creation of proxy datasets

High-quality data on the environmental impacts of food products is scarce, making it challenging to assess these impacts accurately across different regions. To address this gap, new LCIA values were estimated for products in the Agribalyse database by creating proxy datasets. These proxies involve recalculating LCIA data, adjusting it with known impact factors from specific regions to improve regional relevance and accuracy.

3.3.1 Regional factors

A **Regional Factor** is essential in proxy calculations because it adjusts environmental impact data to reflect the specific conditions of a particular location. The proxy calculations in FIT used the general framework of the Agribalyse data but modified the regional factors to more accurately represent the environmental conditions of the target region, as illustrated by the following equation:

$$Impact_{iB} = \frac{Impact_{iA}}{RF_{iA}} \times RF_{iB} \tag{1.0}$$

Where:

- $Impact_{iB}$ is the estimated impact i of the product in the new region B
- $Impact_{iA}$ is the known impact i of the product in the original region A
- RF_{iA} is the Regional Factor for the impact assessment indicator i in the region A
- RF_{iB} is the Regional Factor for the impact assessment indicator i in the new region B.

Within the FIT framework, new values for two impact categories were estimated: the water footprint and biodiversity footprint. For the Water Footprint estimation, the AWARE factor (Boulay et al., 2018) was applied and adjusted by replacing the original Regional Factor with the factor of the target region. Similarly, for the biodiversity footprint, the Biodiversity Value Increment (BVI, Lindner et al., 2019) was used and adjusted by replacing the Ecoregion Factor of the original region with that of the desired region.

3.3.2 Suitable datasets for proxy creation

To create proxy datasets that accurately reflect the environmental impacts of agricultural products in new regions, datasets were selected that could be linked to existing crops available in the FAOSTAT database. The datasets were chosen based on the following criteria:

- **Lightly processed mono-products**: The selected datasets should primarily represent lightly processed mono products, such as raw or minimally processed agricultural items. For example, shelled and roasted peanuts are a suitable dataset because they involve only basic processing steps that do not significantly alter the upstream agricultural practices involved.
- **Exclude animal products**: Animal husbandry introduces complexities that make animal products unsuitable for this type of proxy creation. The main environmental impacts for animal products are not necessarily linked to the cultivation of the animals themselves but to the feed they consume, making them incompatible with agricultural product proxies.

- **Exclude most beverages**: Products like plain water, alcoholic drinks, and some soft drinks were excluded from proxy creation. A notable exception are plan-based drinks, which are commonly compared with milk. Compared to food products, beverages often have low impacts because of their high-water content. Because of this, comparing them to solid foods using the same grading scheme yields few meaningful insights. A beverage-specific rating was not implemented in the context of this project.
- **Threshold for areatime contribution**: Only used products were used whose largest contributing process accounted for 50 % or more of their overall land occupation (according to the *BVI to Agribalyse* data (Lindner et al., 2022)). It was assumed that the remaining percentage of land occupation caused by other processes happened in the same region as that of the largest process.

After selecting the appropriate datasets and regional factors, including the water scarcity index and ecoregion factors, FAOSTAT yield data were used to determine where larger quantities of the respective crops were produced, employing a global production threshold of 80 %. This means that the largest producing regions have been included until a cumulative production of 80 % of the global annual total was reached. The environmental impacts were then recalculated (only the agricultural stage's water and biodiversity impacts) for these regions, thereby creating the proxy datasets.

3.4 Limitations of FIT

While the FIT tool provides a straightforward and flexible approach to assessing food-related environmental impacts, it is important to acknowledge its limitations. These limitations stem from data availability, the scope of the project, and ongoing research needs. Below a few of them are described.

Data availability and coverage

One of the primary limitations of the FIT tool is the reliance on generic datasets for agricultural impact analysis, since product-specific LCIA datasets are not commonly available. The tool therefore primarily uses data from sources such as Agribalyse (including data of both ecoinvent and the World Food Life Cycle Assessment Database), and FAOSTAT. Depending on the question the user seeks to answer and their geographical location, its scope and accuracy can be limited: While these databases provide useful insights, they are not always comprehensive or fully up to date. Additionally, many agricultural datasets do not cover all the regions or impact categories required for detailed assessments. For instance, data on specific, non-conventional production systems (extensive, organic), is often scarce or unavailable. This means that FIT is constrained by the quality and completeness of the data in the underlying databases. Product-specific LCIA data can of course be integrated where available, provided that it has been generated following the Agribalyse methodology.

Scope of the project

FIT is designed to handle LCIA data, rather than inventory (LCI) data. This decision was made to streamline the project in accordance with its scope. However, it means that datasets could not be modified in-depth, so that result modification decreased the accuracy of results more than would have been the case with manual modifications of the LCI data. It also means that FIT is not built for modifying LCI data, but modified inventory's LCIA results could of course still be fed to the tool. The modifications and LCIA calculations would simply need to happen upstream, employing tools such as any graphical (GUI) LCA software or the Brightway LCA framework.[9]

Ongoing research and method development

To improve the FIT tool's accuracy and utility, continuous research and method development are essential. For example, there is a need for more region-specific data, especially regarding biodiversity, water use, and soil health. More robust methods for assessing biological soil quality and other environmental factors will enhance the overall assessment and provide a more complete picture of agricultural impacts across different regions. Further work is required to integrate these factors into the FIT tool's existing framework.

Data quality and uncertainty assessment

As of now, a comprehensive data quality and uncertainty assessment is not part of the FIT prototype. However, proxy datasets are indicated as such and advise users that their accuracy might be reduced. This notwithstanding, working towards a more comprehensive and quantitative uncertainty assessment in the future is recommended.

4 How to use FIT

The following section is somewhat technical and aims at providing users with a high-level understanding of how to run FIT. The required code and data can be found in two repositories:

(1) **FIT_scripts**, which is a collection of (interactive) scripts that transform and combine the required data into a SQL database (apart from the proxy datasets, no new products are created, however)[10]

(2) **FIT_API_public**, which contains the code required to run the API, employing the database previously created[11]

It is assumed that readers have a basic understanding of the employed technologies (Python programming language, containerisation, API endpoints).

---

[9] See the project site for more information: https://docs.brightway.dev/en/latest/.

[10] Access via https://github.com/corsus-GmbH/FIT_scripts.

[11] Access via https://github.com/corsus-GmbH/FIT_API_public.

4.1 Deployment

Launching and hosting the API can be done either locally for development or on a server for production, with each setup designed to ensure consistent behaviour and easy transition from development to deployment. The API is containerised using Docker, creating an isolated and reproducible environment that simplifies both local testing and production deployment. This approach enhances reliability across different systems and makes scaling straightforward for more demanding applications.

4.1.1 Local deployment

For local deployment, Docker is used to create a controlled, isolated environment where the API and its dependencies run consistently. This setup ensures that any configurations or changes made during development can seamlessly be adapted to production.

- **Install Docker**: Ensure Docker is installed and running on your machine.
- **Build and launch**: Using the provided Dockerfile, you can build the container with a single command. Docker will handle running all required setup scripts automatically, enabling you to run the API locally with the same configuration expected in production.
- **Access the API**: Once the container is running, you can access the API locally to test and develop, simulating its performance in a server environment.

Server-Side Deployment

For production deployment, the containerized setup allows the API to run on any server or cloud environment that supports Docker, providing flexibility, consistency, and scalability to meet production demands.

- **Server Setup:** Install Docker on your server or use a cloud provider's managed container service for simplified deployment.
- **Build and deploy:** Transfer the Docker configuration files to the server and build the container using Docker. The Dockerfile will handle all setup processes automatically, launching the API and configuring necessary dependencies.
- **Scaling and load balancing:** For environments that need high availability or must handle large volumes of traffic, Kubernetes (or other orchestration tools) can be used to manage and scale multiple instances of the container. Kubernetes handles load balancing, health checks, and auto-scaling, making it a robust option for maintaining consistent API performance in production.

4.1.2 Configuration files

The API relies on configuration files located in the config/ directory, which allow for adjustable settings across deployment environments such as development, testing, and production. Key configurations include:

- **Database engine configurations:** Specify the database type, host, user credentials, and connection settings tailored to each deployment.
- **Debugging options:** Enable or disable debugging output based on the environment, making it easy to switch between development and production modes.

When deploying locally, ensure that these configurations align with your local development needs. In server environments, Docker reads these configurations directly, enabling smooth transitions from development to production. For Kubernetes-based setups, configurations can be loaded as environment variables or ConfigMaps, ensuring consistency across clusters.

This setup, combining Docker and optional Kubernetes orchestration, provides a reliable and scalable way to develop, deploy, and run the API across environments, from local development to full-scale production.

4.1.3 Collecting necessary data

Data collection within the FIT project centres around supporting accurate LCA calculation. Each data source contributes specific types of information needed to calculate environmental impacts for agricultural products across different regions.

- **Agribalyse database** – Selected as impact data, Agribalyse offers comprehensive LCA data, ensuring methodologically consistent information on many products. However, as it focuses on France, adjustments were made using regional factors to represent other locations more accurately by way of calculating proxy datasets.
- **FAOSTAT** – Maintained by the Food and Agriculture Organization (FAO), FAOSTAT supplies global agricultural statistics, including crop production, yields, and trade data across different countries. FAOSTAT was used to inform the creation of proxy datasets by answering the question which production regions are the most relevant for each crop.
- **AWARE factors** – Available Water Remaining (AWARE) factors[12] quantify the relative availability and scarcity of water resources within specific regions, providing a way to assess water-related impacts. Used to refine the Water Footprint calculations in different regions. By replacing French data in Agribalyse with the AWARE factor values specific to target regions, FIT captures more accurate water scarcity-related impacts.
- **Crop and country- specific ecoregion factors** – The Biodiversity Value Increment (BVI) is a metric designed to assess potential impacts on biodiversity across different ecoregions, using their respective ecoregion factors, which quantify the regions' biodiversity value. Crop- and country

---

[12] Available at https://wulca-waterlca.org/aware/.

specific factors[13] have been used to adapt biodiversity impact calculations to different regional contexts, allowing FIT to better reflect local biodiversity risks and ecosystem pressures.

4.1.4 Creating databases

All the scripts and files necessary for the FIT project are available in the GitHub repository: https://github.com/corsus-GmbH/FIT_scripts. This section provides a brief overview of the repository structure, explains where to find the files, and outlines best practices for using the FIT scripts.

Repository organization

The repository is structured into four Jupyter notebooks and an auxiliary Python script, each serving a specific function in the process of building FIT database. The repository includes the following components:

- *01_fao_data_conversion.ipynb*: Uses FAO annual crop production statistics to derive 3-year-average global production values and to determine major production regions for each crop.
- *02_fit_derive_bvi_aware_rfs.ipynb*: Derives the regional factors (crop- and country-specific ecoregion factors, regional water scarcity factors) necessary for proxy calculations.
- 03_fit_proxy_calculations.ipynb: Performs the core proxy calculations using the derived factors.
- 04_fit_create_db_tables.ipynb: Combines Agribalyse, BVI and the generated proxy data to compile the tables that populate the FIT database.
- helper.py: Contains auxiliary functions used across the notebooks.

Additionally, the repository includes specific folders for input and output files necessary for the script processes:

- *input_data/*: Contains the primary input data required for crop production averages and other necessary regional factors. This folder includes datasets needed for initial processing.
- *proxies_input_data/*: Includes additional input data required specifically for proxy calculations.
- *intermediate_outputs/*: Stores the data generated during the intermediate stages of data processing (i.e. outputs of notebooks 1-3).
- *results/*: This folder stores the final output files (output of notebook 4), including the CSV tables representing the final FIT database and supplementary files detailing included and excluded products.

---

[13] These factors were derived as part of a research on the German environmental impacts of foods, see the publication Eberle and Mumm (2024).

The data stored in the input folders are part of the repository and are required to run the scripts. The repository also contains the file *input_data_documentation.xlsx*, which lists the data sources for each input file.

Tables created and their purpose

The scripts mentioned above create several key tables that are essential for the FIT database:

- **Regional factor tables**: Includes Table_AWARE_RF.csv and Table_BVI_RF.csv created by *02_fit_derive_bvi_aware_rfs.ipynb*. These tables contain the regional factors for the modification of biodiversity and water use impacts. They are critical for adjusting the proxy calculations based on geographical differences.
- **Proxy calculation tables**: Including impact_proxy_aware_df.csv and impact_proxy_bvi_df.csv. These tables are created by *03_fit_proxy_calculations.ipynb* and represent the food production and consumption impacts across different regions, utilizing the regional factors.
- **Database tables (final output)**: Generated by *04_fit_create_db_tables.ipynb*, these final tables in CSV format represent the structure and content of the FIT database and are needed to run the FIT API.

Best practices for using FIT scripts

When using the scripts, a set of best practices are advisable to ensure an efficient and effective workflow for using the FIT repository to create the necessary databases for the FIT project:

If you plan to simply use the repository:

- **Install dependencies** - Use the *requirements.txt* file to install the required Python packages (e.g. using Python's package manager *pip*).
- **Run scripts sequentially** - The scripts are dependent on one another. It is necessary to run them in order (script 01 through to script 04).

Furthermore, if you want to make changes to the repository:

- **Use feature branches and create pull requests**: When modifying or testing new features in the scripts, always create a new branch. Create a pull request afterwards if you want your changes to be integrated into the repo's main branch.
- **Verify data integrity**: Before using the tables generated from a script you modified, ensure that all input data is correctly formatted and accurate. This will help prevent errors in the final outputs.
- **Document changes**: If you make any changes to the scripts, document them clearly in the code itself and update any relevant documentation (such as *input_data_documentation.xlsx*), so others can understand the modifications.

4.2 Use by front end

The API is designed for easy access to environmental impact information for individual products and recipes. These can then be displayed in a front end tailored to the respective target group. Users provide recipe details with items and quantities, and the API calculates and returns detailed impact scores, breakdowns by life cycle stages, and environmental categories.

To use the tool:

1. **Submit a recipe**: Prepare a recipe submission by listing items and their quantities. Each item is uniquely identified to allow accurate data retrieval.
2. **Choose weighting (optional)**: You may specify a weighting scheme to influence how environmental categories are prioritized in the assessment (see section 2.2.4). If not specified, a default scheme is applied.
3. **View results**: After submission, the API provides a structured assessment with an overall score, detailed breakdowns for individual items and life cycle stages, and information on proxy data usage if applicable (see below).

For further guidance on request formats, advanced settings, and specific response fields, please refer to the README in the repository, which provides more detailed technical instructions and examples.

4.2.1 How to obtain information from FIT

To obtain information from the API, users can interact with the two main endpoints to access item details and calculate environmental impacts. To get started:

1. **Fetch Item Details:** Use the /items/ endpoint to retrieve a comprehensive list of all items in the database. Each item is detailed with its name, country of origin, and category information, alongside a flag indicating if the data is actual or a proxy. This is helpful for exploring available data and preparing recipes.
2. **Calculate Environmental Impact of Recipes:** Use the /calculate-recipe/ endpoint to assess the environmental impact of a recipe. Simply submit a list of items with their quantities, and optionally include a weighting scheme to customize the calculation. The API will return an organized report with impact scores across life cycle stages and environmental categories, with additional scores and grades to indicate environmental performance.

For specific request formats and examples, please refer to the README file and comprehensive documentation.

4.2.2 What does the result say?

The results returned by the API are structured into two sections. In the following, a quick rundown is provided. The first section (“Recipe Info”) contains general information about the request, as well as results for the entire recipe. The second section (“Item Results”) shows the component results for individual items.

Recipe Info has a unique section called “General Info”, which contains request-specific metadata. It specifies the utilised weighting scheme, whether the results include proxy datasets and what the overall mass of the recipe ingredients is.

It then details the **single score results**, as well as the **results per stage** and per **impact category**. The same is done per item. Notice, however, that the single score result section of the individual items also contains the information on whether the item is a proxy. Recipe Info give this information as part of its general information section instead.

The **scaled values** express the result in relation to the logarithmic minimum and maximum of the results distribution. If desired, this enables freely calculating ‘sub-grades’ for arbitrary combinations of impact categories, e. g. by calculating the arithmetic mean the categories scaled values as a basis for assigning grades. The given LCIA_values are always normalised values (see section 2.2.3).

The following list shows the nested structure of the response and its constituent pieces of information.

- “Recipe Info”
  - “General Info”
    - “Weighting Scheme”: weighting scheme name
    - “contains_proxy”: true or false
    - “Overall Mass”: x kg
  - “Single Score”
    - “Single Score”: x mPt
    - “Grade”: e.g. A – E
    - “Scaled Value”: value between 0.0 and 1.0
  - “Stages”
    - “Agriculture”
      - “lcia_value”: single score of life cycle stage in mPt
      - “Grade”: e.g. A – E
      - “Scaled Value”: value between 0.0 and 1.0
    - “Transformation”
      - [as above …]
    - “Transport”
      - [as above …]
    - “Supermarket and distribution”
      - [as above …]
  - “Impact Categories”
    - “Climate change”
      - “lcia_value”: normalised impact score of impact category across stages
      - “Grade”: e.g. A – E
      - “Scaled Value”: value between 0.0 and 1.0
  - [< Impact category > …]
    - [as above …]
- “Item Results”
  - < Item ID >
    - “Single Score”
      - “Single Score” : x mPt
      - “Grade”: e.g. A – E
      - “Scaled Value”: value between 0.0 and 1.0
      - “contains_proxy”: true or false
    - “Stages”
      - [as above …]
    - “Impact Categories”
      - [as above …]
  - [< Item ID > …]

5 References

Boulay, A.-M., Bare, J., Benini, L., Berger, M., Lathuillière, M. J., Manzardo, A., Margni, M., Motoshita, M., Núñez, M., Pastor, A. V., Ridoutt, B., Oki, T., Worbe, S., & Pfister, S. (2018). The WULCA consensus characterization model for water scarcity footprints: Assessing impacts of water consumption based on available water remaining (AWARE). *The International Journal of Life Cycle Assessment*, *23*(2), 368–378. https://doi.org/10.1007/s11367-017-1333-8

Campbell, B., Beare, D., Bennett, E., Hall-Spencer, J., Ingram, J., Jaramillo, F., Ortiz, R., Ramankutty, N., Sayer, J., & Shindell, D. (2017). *Agriculture* production as a major driver of the Earth system exceeding planetary boundaries. *Ecology and Society*, *22*(4). https://doi.org/10.5751/ES-09595-220408

Eberle, U., & Mumm, N. (2024). Reduction potential of German environmental food impacts due to a planetary health diet. *The International Journal of Life Cycle Assessment*. https://doi.org/10.1007/s11367-024-02352-4

Lindner, J. P., Fehrenbach, H., Winter, L., Bischoff, M., Blömer, J., & Knüpffer, E. (2020). *Biodiversität in Ökobilanzen.* Weiterentwicklung *und vergleichende Studien* (No. 575; BfN Skripten). Bundesamt für Naturschutz. https://doi.org/10.19217/skr575

Lindner, J. P., Fehrenbach, H., Winter, L., Bloemer, J., & Knuepffer, E. (2019). Valuing Biodiversity in Life Cycle Impact Assessment. *Sustainability*, *11*(20), 5628. https://doi.org/10.3390/su11205628

Lindner, J. P., & Knüpffer, E. (2020). *LC.biodiv.IA Guideline*. https://www.ibp.fraunhofer.de/content/dam/ibp/ibp-neu/en/documents/publications/life-cycle-engineering/guideline-lcidivia.pdf&ved=2ahUKEwjw-um6uMCHAxV42AIHHelkD84QFnoECBEQAQ&usg=AOvVaw1SeyK411fWT8uq2zfQhlRv

Lindner, J. P., Koch, P., Fehrenbach, H., & Buerck, S. (2022). BVI to Agribalyse—Bringing the Biodiversity Value Increment method to Agribalyse. Hochschule Bochum, ecolysis GmbH, IFEU Institut, ADEME.

Poore, J., & Nemecek, T. (2018). Reducing food's environmental impacts through producers and consumers. Science, 360(6392), 987–992. https://doi.org/10.1126/science.aaq0216

Sala, S., Cerutti, A. K., & Pant, R. (2018). Development of a weighting approach for the environmental footprint (JRC Technical Reports). EC JRC. https://data.europa.eu/doi/10.2760/945290

Willett, W., Rockström, J., Loken, B., Springmann, M., Lang, T., Vermeulen, S., Garnett, T., Tilman, D., DeClerck, F., Wood, A., Jonell, M., Clark, M., Gordon, L. J., Fanzo, J., Hawkes, C., Zurayk, R., Rivera, J. A., Vries, W. D., Sibanda, L. M., … Murray, C. J. L. (2019). Food in the Anthropocene: The EAT–Lancet Commission on healthy diets from sustainable food systems. The Lancet, 393(10170), 447–492. https://doi.org/10.1016/S0140-6736(18)31788-4

Zampori, L., & Pant, R. (2019). *Suggestions for updating the Product Environmental Footprint (PEF) method* (Nos. JRC115959, EUR 29682 EN; JRC Technical Reports). EC JRC.

6 Imprint

Published by:

corsus-corporate sustainability GmbH
Großneumarkt 50
20459 Hamburg

Authors:inside:

Marius Rödder, Talita F. Amado (PhD), Paul Appel

Contact:

Marius Rödder, m.roedder@corsus.de

Responsible:

Dr. Ulrike Eberle

Design / Layout:

www.schierrieger.de

Status:

November 2024

Implementing partners

WWF

corsus
CORPORATE SUSTAINABILITY

TMG
ThinkTankforSustainability
Töpfer Müller Gaßner

Supported by:

Federal Ministry
for the Environment, Nature Conservation,
Nuclear Safety and Consumer Protection

IKI
INTERNATIONAL
CLIMATE
INITIATIVE

based on a decision of
the German Bundestag