An evaluation of sustainability assessments of food systems.

How can informed consumer food choices be facilitated?

Joost van Heiningen

An evaluation of sustainability assessments of food systems.

How can informed consumer food choices be facilitated?

by

Joost van Heiningen

Master Thesis: Environmental Sciences

Author: Joost van Heiningen

Registration number: 921209-319-030

Course number: ESA-80436

Supervisor: Dr. ir. Lenny van Bussel

Examiner: Prof. dr. Rik Leemans

Wageningen University and Research

Environmental Systems Analysis Group

Wageningen, May 2020

Summary

The global food system emits between one fifth and one third of the total anthropogenic greenhouse gases. In addition, it substantially contributes to the exceedance of several anthropogenic planetary boundaries. These impacts caused by the global food system can be mitigated through several macro strategies. One of these macro strategies is more sustainable food consumption by consumers. In many countries and especially The Netherlands, most consumer food is purchased in the supermarket. But how can consumers know which products are sustainable and which products are not? How are sustainable food products defined? In my thesis I aim to understand to what extent supermarket-food products can be assessed on sustainability criteria and if such assessment could be translated to the end consumer.

To fulfil this aim, a three-phase research plan was developed. First, an ideal food system sustainability assessment (FSSA) was conceptualized based on Sustainability Indicator and Life Cycle Assessment (LCA) literatures. Second, the edible oil FSSA literature was reviewed on methodological issues. The focus was on methodological issues because these limit the practical applicability of FSSAs. Third, three existing consumer FSSA platforms were analysed and compared with the ideal FSSA.

The results chapter covered three different sections. First, my results show that an ideal FSSA simply does not exist because of inconceivable trade-offs, although, an assessment close to ideal could conceptually be developed. Second, my results show that current edible oil FSSAs are far from this ideal. The edible oil FSSAs include a large variety of methodological applications and deal with several unresolved issues. The most important ones are the inability of LCAs to deal with field nitrous oxide (N_2O) emissions, biodiversity loss, changes in soil quality and land use change. Third, I show that consumer FSSA platforms all want to give consumers the possibility to make sustainable food purchases. However, the consumer FSSA platforms vary slightly in their aims and views on sustainability. These differences trickle down throughout each platform – large dissimilarities were found in the inclusion of indicators and impact categories, and in the aggregation and presentation of sustainability scores.

My thesis is the second publication that integratively reviews the bodies of LCAs of rape-, sunflower-, peanut- and olive-oil. Additionally, to my knowledge, my thesis is the first publication that compares multiple consumer FSSA platforms, and this provides a steppingstone for further research on these entities.

In terms of practical lessons learned, I advised consumers on more sustainable food consumption and I recommended several measures to my commissioner, a start-up that is developing a LCA-based FSSA platform for Dutch supermarkets. I advised the consumers to 1) align their view on sustainability with those offered by consumer FSSAPs; 2) to think about what type of sustainability score display they prefer as a consumer; and 3) to incorporate more non-airfreighted plant-based foods in their diets. In addition, I recommended the start-up 1) to better crystalize the organization's aim and views on sustainability; 2) to limit the drawbacks of data aggregation; 3) to consider excluding the social dimension in consumer FSSA platforms; 4) to develop strategies to deal with the main limitations of the LCA methodology; and 5) to create effective partnerships with some of the interviewees included in my research.

To Haddy, who showed us Mzungus how to joyfully dance.

To Jan, who never stopped introducing musical gems to his family and friends.

And to our dog Ko, who taught us how to love unconditionally.

Acknowledgements

Dear reader,

You just opened not my first, not my second but my third thesis. You would expect that this was the easiest one to write, but experience showed me the opposite – it was the most difficult one. This is mostly due to my choice to go completely out of my comfort zone, I am a social sciency person and in this thesis I had to be as to the point as possible… which feels like working with the less developed half of my brain. Anyway, I managed to put everything together in a product that feels out of my comfort zone but still good. Having pushed out most of my poststructuralist tendencies from the remainder of this thesis, I want to use this one-pager to elaborate on some more abstract thoughts.

The last seven months might have been the weirdest in my life. This because my perception of time bounced in all directions. I started off this thesis with the idea that I could finish it in a few months, pressuring myself to hurry up so I could start working after 7 years of studying. Ironically, this mindset froze me rather than making me productive. After two months of being half idle, we received the unfortunate message that our dear friend Haddy suddenly passed away. An event that really put me off track. Time was suddenly standing still – it was not there anymore; it was just us trying to deal with the unreal feeling reality. Nevertheless, good old-time pressure returned after some weeks. This was around Christmas. Normally, a time in which you can be idle, but for me, these were the weeks that I worked the most on this project. The feeling that you really *want* to work on something is way more productive than the feeling that you *have to* work on something. In 2020, I continued working with a somewhat normal pace while at the same time trying to setup our start-up, for which this thesis is a foundation. This went well until my parents shared the bad news that we had to put our beloved Kootje to sleep. At that time, I tried to spend as much time possible with my doggie, however, time was slipping through my fingers. Suddenly time was scarce. Thirteen days… one week… two days… 23 hours... 53 minutes… 3rd of March, 16:30 was there. The darkest hour of the clock. Was it though? It made me more appreciative to everyone – and ironically, also more productive. The last leg of this thesis I wrote from the 5th of March onwards. This productive period was even more so accelerated by the Covid-19 crisis, which suddenly introduced large insecurity in our lives. Rather than paralyzing me, it freed me from the burden of time pressure. The linear notion of time was broken. Which is logical – who came up with the linear notion of time anyway? When does time end? I realized that measuring time in hours is a bad reflection of how we actually perceive time individually. By measuring my output in my personal perceived feeling of time, I am proud of this product. By not feeling the pressure to achieve anymore, this project suddenly became a lot of fun rather than a thing I *had* to do for my degree. I hope this is also reflected in the end product. I guess the biggest personal lessons during the last 7 months are 1) screw time and 2) appreciate the things you already have instead of the things you still want, by doing so, you start living. #turnsoffpoststructuralism

Of course, I also want to thank a few people. First and foremost, I want to thank my girlfriend Neringa, who supported me continuously especially when I had difficulties saying goodbye to my dog. Second, my parents Ad and Hannie, my brother Wouter, his girl Lottje and their beautiful children Melody and Jercynio. Although I don't go to Beuningen that often anymore, it always feels like home thanks to you. Additionally, I want to thank my roommates Anna, Violet, Deli, and again Neringa, with whom I am in a voluntary lockdown now. It is very liberating that we support each other rather than being annoyed. Furthermore, a very big thank you to Frank, Martijn, Mike and Neringa for proofreading this thesis. Also, I want to thank my supervisor Lenny, who was always very understanding and putting me back on the positivistic science track when I was drifting off. Last, I want to thank the interviewees who participated in this research, I had very inspiring talks with you.

Contents

Overview of figures and tables

List of Acronyms

Acronym	Meaning
CH_4	Methane
CO_2	Carbon dioxide
FSSA	Food System Sustainability Assessment
FSSAP	Food System Sustainability Assessment Platform
FU	Functional Unit
GHG(s)	Greenhouse Gas(es)
GWP	Global Warming Potential
IPCC	Intergovernmental Panel on Climate Change
LCA	Life Cycle Assessment
N	Nitrogen
N_2	Dinitrogen
N_2O	Nitrous oxide
NH_3	Ammonia
NO_x	Nitrogen oxides
NO_3	Nitrate
OMWW	Olive Mill Wastewater
P	Phosphorus
PM	Particulate matter
RQ(s)	Research question(s)
SO_2	Sulphur dioxide

Introduction

1. Introduction

1.1. The global food system

The global food system causes many environmental issues that we currently are facing. Last summer, this was again underlined by the IPCC (2019). The global food system emits between one fifth and one third of the total anthropogenic greenhouse gases (GHGs), taking into account both pre- and post-production of food products (IPCC, 2019). Even more so, an increase of 50-90% of various environmental impacts including land use change, water use, nitrogen (N) and phosphorus (P) use, and GHG emissions can be expected as a consequence of the growing world population and income (Springmann et al., 2018). Consequently, these additional issues will lead to the exceedance of several anthropogenic planetary boundaries – threatening a safe living environment for humanity. Currently, four of these anthropogenic planetary boundaries are already exceeded as a consequence of human activities (Gladek et al., 2018). These exceeded planetary boundaries include climate change, alterations of nitrogen and phosphorous biogeochemical cycles, biodiversity loss and land use change. Due to these issues, feeding the world population sustainably is one of humanity's major challenges.

During the last centuries, the growing food demand was partly met through the application of synthetic N and P fertilizers. This was made possible by two processes, namely 1) the invention of the Haber-Bosch process, in which atmospheric nitrogen (N_2) is transformed to reactive nitrogen; and 2) the mining of phosphorus (P), which is derived from phosphate-rich rocks (Westhoek et al., 2011; Zhang et al., 2015). The increase in N and P fertilizer application resulted in many benefits such as increased agricultural productivity and improved food safety (Lamichhane et al., 2016). However, this increased agricultural production also had negative consequences:

- Terrestrial biodiversity loss: multiple activities involved in agriculture put terrestrial biodiversity under pressure. Leip et al. (2015) estimated that an absolute loss of 34% of mean species abundancy can be accounted for by agriculture;
- Pesticide use: the application of pesticides has large toxic potential, not only for humans but also for terrestrial, fresh water and marine ecosystems (Notarnicola et al., 2004). Hence, pesticide use can be a major cause of biodiversity loss (Beketov et al., 2013);
- Acidification of soils and waters: the deposition of compounds like sulphur dioxide (SO_2), ammonia (NH_3) and nitrogen oxides (NO_x) can result in soil acidification. Soil acidification reduces plant growth and increases the mobility of toxic elements such as copper and zinc. Increased mobility of toxic elements can damage natural ecosystems and hamper agricultural productivity when not counteracted by liming (Leip et al., 2015);
- Eutrophication of land and water surfaces: the runoff and leaching of agricultural nutrient excesses can result in excessive algae growth. Nowadays, agriculture has become the largest contributor of excess nitrogen and phosphorous to waterways and coastal zones (Foley et al., 2005; Sutton et al., 2013). The excess nutrients cause large algae blooms that are harmful since these deplete water oxygen levels;
- Groundwater contamination due to nitrate (NO_3) leaching: high NO_3 levels in drinking water result in adverse impacts on human health (Erisman et al., 2013; Sutton et al., 2013);
- Overuse of fresh water sources: agriculture is the largest consumer of fresh water in Europe when compared to other economic activities (Shiklomanov & Rodda, 2004). More particularly, a large amount of water is used for irrigation purposes;
- Decrease in air quality: NH_3 and NO_x emissions can have adverse effects on human health (Leip et al., 2015). NH_3 emissions potentially form particulate matter (PM) and

photochemical smog, whereas elevated NO_x emissions could lead to higher levels of tropospheric ozone (Erisman et al., 2013; Sutton et al., 2013);

- Contribution to climate change: through the emissions of carbon dioxide (CO_2), methane (CH_4) and nitrous oxide (N_2O). Food systems contributed for 19-29% of the anthropogenic global greenhouse gas (GHG) emissions in 2008. Within food systems, agricultural production accounted for 80-86% of the total GHG emissions (Vermeulen et al., 2012); and
- Socio-economic inequality: because of power imbalances. For example, the booming quinoa demand in Europe resulted in a higher value of this crop in Bolivia. Wealthy Bolivian actors were able to mechanize their production and therewith increase their yield, whereas less resourceful farmers could not. This power imbalance contributed to growing inequality and socioeconomic upheaval in Bolivia (Carimentrand & Ballet, 2010).

The increased environmental pressures caused by the global food system and the growing demand for food creates a paradox by itself – food production must substantially increase and simultaneously also reduce its overall environmental impacts. Springmann et al. (2018) proposed four mitigation strategies to tackle this paradox:

1. Improvements in nutrient use efficiency along the whole food chain, through on-farm technological and management-related interventions, especially from soil to crop and from feed crop to animals;
2. Improved recycling of crop residues, and animal and human wastes;
3. Reductions in food loss/waste; and
4. Dietary changes towards more plant-based diets.

These strategies all have a substantial impact on one or several environmental impact categories, even more so when each strategy is applied simultaneously on a global scale. However, only one strategy has large enough magnitude that staying within anthropogenic planetary boundaries of climate change remains possible. This strategy is the dietary shift towards more plant-based diets (Springmann et al., 2018). Such a diet would be composed of: a large diversity of plant-based foods; low amounts of food produced from animals; foods containing unsaturated fats rather than saturated fats; and, only a low amount of highly processed foods and added sugars (Willett et al., 2019). So, a substantial shift to healthier and more sustainable diets by consumers helps to reduce environmental impacts – however, this does largely depend on the consumer's choice. Indeed, consumer demand itself ultimately shapes the global food system. Therefore, consumers could contribute to a more sustainable food system if they have the desire and ability to do so (Gladek et al., 2018). Hence, more sustainable consumer choices can be a promising pathway in addressing part of the food-environment puzzle.

1.2. People's diets: the consumer's choice

At the moment of writing, the scientific literature includes numerous studies that measure the sustainability of diets and food products (e.g. Chaudhary et al., 2018; Jones et al., 2016; Willett et al., 2019). Within food system assessment studies, Life Cycle Assessments (LCAs) are most frequently employed to assess the impacts of food production (Jones et al., 2016). However, these LCA studies often consider only one or just a few environmental indicators and the most popular indicator is GHG emissions (Khatri & Jain, 2017). In doing so, most studies ignore other environmental impacts like the forming of particulate matter and photochemical smog, and socio-economic consequences of dietary choices (Hayashi, 2000; Khatri & Jain, 2017).

The same focus on GHG emissions can be found outside the academic community. Many non-academic organizations developed initiatives to help consumers to make an informed sustainable choice. These initiatives include user platforms hosted by websites or phone applications. The user platforms facilitate sustainability ratings of consumer products, so that consumers can make more informed and sustainable choices. Some examples of these user platforms are Garaikoa, Ecolabel Guide, Buycot, Idemat, Goodguide, Carboncloud, Rank-a-brand and Evocco. Each platform shows varying concerns and impacts that might affect consumers. These include: sourcing of products, GHG emissions, metal contents, vegan produce, and water usage. If a consumer perceives these as equally important is questionable. Evocco (2019), who provides phone applications for supermarket food products, published a consumer survey and concluded the following:

> *"We surveyed a selection of our users to understand what they perceived to be the most important environmental factors when it came to food purchases. We discovered a strong preference towards waste reduction and the local sourcing of food. This could largely be due to these being the most tangible factors for consumers and the best understood.*
>
> *However, consumers don't know that most of the climate impact comes from the production of the food. This can range from 70-90% depending on the type of food. Food companies need to play a more active role in educating their consumers about the impact of food if they want to reduce their own impact."*

So, providing consumers with sustainable food choices probably is promising. One place to do so is in the supermarket. For example, in The Netherlands, almost four fifths of food expenditures is spent in supermarkets (Bie et al., 2012). As such, if the Dutch consumer would buy more sustainable products in the supermarket, it could contribute to an enormous positive environmental impact. However, will consumers consume more sustainably when they are bombed with sustainability information in the supermarket (*in situ*)? Kalnikaitė et al. (2013) argue that at the one hand consumers want more clarity about the consequences of their decisions, but at the other hand, they are also overwhelmed and confused by existing nutritional and environmental labelling in supermarkets. Von Reischach et al. (2010) stated a similar argument: people do not want to know everything – sometimes less *in situ* nudges are more effective. Supermarket consumers' decision-making processes are not exhaustive, but rather constrained by limited time and resources (Kalnikaitė et al., 2013). Supermarket purchases are based on 'low-cost' quick products that do not require a lot of contemplation like 'high value one-off' products such as clothes and electronics. Hence, supermarket purchase decisions are mainly based on only one or two reasons – most other information is often ignored. Two frequently stated reasons are price and quality (Kalnikaitė et al., 2013). As such, adding many more nudges in this decision making process can be ineffective: additional information has to be salient, simple, fast, frugal and personal (Kalnikaitė et al., 2013).

The aforementioned consumer platforms that show a wide variety of impact categories (e.g. sourcing, GHG emission, metal contents) may therefore be questioned. Is the presentation of several impact categories to consumers useful when these consumers only base their purchases on a limited amount of reasons? A solution to this overload of information can be the aggregation of environmental information in the simplest format possible (e.g. 1-5 star score). To create such a score, one must address many questions: can such aggregation contain all environmental and social impact information of a product? What are the most important environmental impacts? What are the most important social impacts? Who should decide upon what is important or not?

Currently, several applications and labels that show an impact category exist. Yet, to my knowledge, not many platforms exist that assess products on one fully integrated sustainability score, nor platforms that assess products on a complete range of sustainability indicators – although a few platforms manage to do this scoring to a considerable extent. This gap can also be identified within the scientific literature, in which products are frequently assessed on only a few impact indicators, with social economic consequences mostly being ignored (Jones et al., 2016). Additionally, to my knowledge, the scientific literature did not evaluate the dynamics behind consumer food system sustainability assessment platforms (FSSAPs) like the ones listed at the beginning of this section.

1.3. Problem statement

The global food system is a large contributor to many environmental problems. Several strategies can reduce the impacts of this system. A promising strategy is dietary changes of individuals. However, within The Netherlands, making 'right' food choices with regards to sustainability is extremely difficult for consumers. Sustainability information about production impacts is lacking or not communicated. This sustainability information constitutes of many facets and consumers simply do not have time nor the knowledge to consider these while doing groceries. Hence, more clarity on the wide range of negative production impacts is urgently needed. Transparency on how these impacts are aggregated in current consumer applications and, ultimately, a totally aggregated sustainability score that combines both this clarity and transparency in an easily understandable way, must be developed.

1.4. Research aim

My research aims to determine to what extent supermarket-food products can be assessed on production impacts. These methodological insights are likely helpful to develop totally aggregated sustainability scores that assess the impacts of food products within Dutch supermarkets. Such aggregated sustainability scores are currently developed by my thesis commissioner. My commissioner, a start-up, wants to use the outcomes of my thesis to assist consumers in making more informed sustainable food choices. As such, my thesis has three audiences, an academic audience, a start-up that commissions this thesis and consumers. My thesis addresses the following research questions (RQs) to satisfy each audience and accomplish the research aim:

- RQ1: What would a complete and accurate (i.e. ideal) Food System Sustainability Assessment (FSSA) look like?
- RQ2: What methodological problems must be diagnosed when an ideal FSSA is compared to current scientific sustainability assessments on edible oil systems?
- RQ3: How do current providers of consumer FSSAPs assess sustainability?

1.5. Thesis structure

My thesis partially follows a standard research outline. The research topic is introduced, followed by a methodology. I chose not to include a separate theoretical framework chapter because this is mostly covered in the results section. The results section itself conceptualizes an ideal FSSA, then evaluates academic publications on edible oils and lastly analyses existing consumer FSSAPs. The subsequent discussion section nests my research within the existing literature. This existing literature is then critically evaluated by mirroring it with my findings. After this critical evaluation, the discussion section highlights the flaws of consumer FSSAPs and then provides both future research avenues and limitations of my research. The subsequent conclusion section briefly answers the RQs and the question posed in my thesis' subtitle. As a wrap-up, my thesis includes recommendations to consumers and my commissioner. I chose to include a recommendations section to ensure that every audience's expectations are satisfied.

Methodology

2. Methodology

The methodology chapter is divided in three sections. Each section covers the used methods, sampling techniques and procedures to answer a single research question.

2.1. Methods: Outlining an ideal FSSA (RQ 1)

To answer RQ 1, an ideal FSSA was outlined based on an exploratory literature review on food system sustainability assessments. I favoured this type of literature review over other methods because I was not very knowledgeable on the topic before starting my thesis. As such, I could get a better understanding myself before applying more elaborate methods. The publications included in this review were sampled in three ways:

1. The following two articles were recommended by my thesis supervisor and formed the backbone of the ideal FSSA:
    a. Van Asselt, E. D., Van Bussel, L. G. J., Van der Voet, H., Van der Heijden, G. W. A. M., Tromp, S. O., Rijgersberg, H., ... & Van der Fels-Klerx, H. J. (2014). A protocol for evaluating the sustainability of agri-food production systems—A case study on potato production in peri-urban agriculture in The Netherlands. Ecological Indicators, 43, 315-321.
    b. de Vries, M., & de Boer, I. J. (2010). Comparing environmental impacts for livestock products: A review of life cycle assessments. *Livestock science*, *128*(1-3), 1-11.
2. Additional articles were sampled from the reference lists of the two publications listed under 1. These additional articles were only selected if their title included 'sustainability assessment.'
3. Articles were sampled from Google Scholar's database by using the following search term: 'Sustainability assessment food systems.' The inclusion of this search term in publications' titles was the only criterion for selection. The sampling process itself continued until theoretical saturation.

By following this sampling procedure, 21 articles were included in the final sample. These articles were analysed through open and axial coding to filter out the main aspects an ideal FSSA must include. The ideal FSSA is presented in Section 3.1.

2.2. Methods: literature review on edible oil FSSAs (RQ 2)

To answer RQ 2, an extensive literature review on edible oils was carried out in two phases. I preferred this type of literature review over other methods since this would provide me a solid knowledge base before reaching out to the consumer FSSAPs. The two-phased literature review only included publications on edible oils. Both the rationale for selecting edible oils and additional methodological aspects concerning RQ 2 are outlined in this section.

2.2.1. Product selection

As part of the sampling procedure, the products of analysis were determined stepwise. As a first step, several product categories were selected based on the frequency of consumption in The Netherlands. On a weekly basis, the six most frequently consumed edible product categories in The Netherlands are 1*) non-alcoholic drinks, 2) cereals, 3) dairy products, 4) fats and oils, 5) meat products* and *6) vegetables* (Van Rossum et al., 2016).

Subsequently, these product categories had to be narrowed down so that it would remain feasible to review them during this thesis project. I narrowed these product categories by defining four additional criteria (Table 1). Based on these criteria, I excluded *non-alcoholic drinks* because it is not a food (Criteria 1). *Cereals* and *Dairy products* were also excluded since they both serve many different purposes in a Dutch daily diet (e.g. breakfast, snack, addition, etc.). *Meat products* were excluded for personal reasons: I am a vegetarian and therefore my interest in meat impacts is rather low (Criteria 3). Last, *vegetables* were also excluded for a personal reason: I already ran some project on vegetables and instead of repeating these projects, I wanted to challenge myself. As such, only *fats and oils* were selected for further analysis. For simplicity's sake, fats were also excluded since these have several differing qualities. So, *edible oils* became the product to be analysed.

Of these oils, only four were selected for further analysis so that it remained feasible to conduct this research in a reasonable timeframe. Initially, I attempted to select these oils based on frequency of consumption; however, data sources include a lot of ambiguity regarding the frequency of consumption in The Netherlands. This ambiguity is caused by oils being sold both as a product by itself and as a component of other products. As a result, the oils were not selected based on frequency of consumption. Instead, a list published by the Dutch Consumentenbond was used. This list suggested that the most important oils consumed in The Netherlands are olive oil, sunflower oil, rapeseed oil, peanut oil and salad oil (Rolvink, 2019). Of these oils, salad oil was also excluded since it is a mixture of different products. In the end, olive-, sunflower-, rapeseed- and peanut-oil were selected for further analysis in this thesis.

Table 1: Predefined criteria for exclusion of food product categories

Criteria #	Criteria
Criteria 1	The product category consists of food only
Criteria 2	The food products included in the product category serve a similar purpose in a Dutch daily diet
Criteria 3	The product category does not contain animal products
Criteria 4	The food products are sold in the largest Dutch supermarket chains, being the Albert Heijn

2.2.2. Literature review: Aims

The literature review consists of two phases. First, the review compares the ideal FSSA (RQ 1) to FSSAs published in the scientific literature in order to diagnose methodological issues. Second, the review evaluates these issues by presenting the frequency of occurrence. The results of this two-phased literature review are presented in Section 3.2.

2.2.3. Literature review: Sampling

The articles evaluated in the two-phase literature review were sampled by using specific search terms in Google Scholar. These search terms were as follows: ‘life cycle assessment’ OR ‘life cycle analysis’ OR ‘LCA’ OR ‘sustainability assessment’ OR ‘sustainability indicators’ AND the name of each selected edible oil. By using these search terms, many publications on biofuel were shown in Google Scholar; these were not included in the final sample since biofuels are not produced for human consumption.

The sampled publications were divided in two batches. A general batch that included all studies except of case studies and a second batch that only included LCA-based case studies. The general batch of articles was used to identify the main methodological issues in the literature. Subsequently, these methodological issued were evaluated on how frequently they occurred in the LCA-based case studies.

2.2.4. Literature review: Methodological issues and frequency of occurrence

The batch of general articles consisted of fourteen publications. These all focussed on olive oil sustainability assessments. The reason that these articles only focused on olive oil was simple: no literature regarding rapeseed oil, sunflower oil and peanut oil was found during the sampling procedure – except for the LCA-based case studies that were included in the second batch of articles. The results of the first reviewing phase were used to develop an evaluation sheet. The structure of this evaluation sheet was mainly based on two literature reviews carried out by Banias et al. (2017) and Khatri and Jain (2017). This structure was then expanded by additional publications included in the batch of general articles. A summarized version of the evaluation sheet can be found in Appendices 2 and 3.

During the second reviewing phase, the batch of LCA case studies was evaluated by means of the evaluation sheet. The evaluation mainly aimed on testing the frequency of occurrence of the methodological issues identified in the first reviewing phase. The batch of LCA cases included 33 edible oil LCAs, which can be broken down as follows: 16 LCAs on olive oil, 5 LCAs on rapeseed oil, 3 LCAs on sunflower oil, 2 on peanut oil and 7 mixed assessments that included multiple vegetable oil LCAs. An overview of the LCA-based cased studies is included in Appendix 3.

2.3. Methods: analysing consumer FSSAPs (RQ 3)

RQ 3 aimed to explore existing consumer FSSAPs by means of document analyses, interviews with representatives of consumer FSSAPs and subsequent analyses of the interview transcripts. The results are presented in Section 3.3.

I chose to use interviews over other methods because interviews fitted the exploratory nature of RQ 3. The interview type used was semi-structured interviews. Semi-structured interviews were preferred because it gave me the possibility gain insights into the representatives' views on their consumer FSSAPs, while maintaining some freedom to deviate from the interview guide. Additionally, the use of fully structured interviews was omitted because it could jeopardize the depth of the collected data, whereas non-structured interviews were omitted since these might lack comparability with each other (Boeije, 2009).

The selected case studies were sampled from a list of consumer FSSAPs made by me. This list included organizations that all had the same purpose, namely, providing consumers with sustainability information on food products. The list included nine organizations, which were all approached by e-mail. Three organizations replied and were willing to be interviewed – these three organizations are included in the mini-case studies presented in Section 3.3. Two interviewees were in leading positions whereas the third interviewee was an organization's spokesperson. Prior to the interviews, I asked the interviewees to fill out a consent form (Appendix 5). Two interviewees returned the form and one provided oral consent during the interview. I prepared each interview by 1) developing an interview guide; and 2) conducting a document analysis. I based the interview guide on the results of RQ 1 and RQ 2. Before each interview, I slightly tailored the interview guide to each organization based on the document analyses. This document analyses consisted of a thorough scan of items published by the organizations. These items included websites, and documents and videos published on these websites. So, besides a few tailored questions, the interview guides were identical, except for the language in which they were written. The base interview guides can be found in Appendix 4. The main questions I aimed to answer with each interview guide were:

- What indicators are consumer FSSAPs employing? Why these?
- How do they obtain the values for these indicators?
- What is the uncertainty of these values?
- What are overall gaps in their methods? Why?
- Do these food assessment platforms weigh their impact categories? Why so/not?

The interviews were all conducted online. Each interview was audio recorded. Two interviews were fully transcribed and one partially (Appendices 7, 8 and 9). The partial transcription is a direct consequence of technical issues with both my voice recorders during one interview. To solve this, I jotted notes during the interview on which I immediately elaborated afterwards. Statements based on my notes are triangulated through the partial transcription and documents published by this organization (Tatanka). Nonetheless, I still had to leave out some statements made by the organization's representative, because I was not able to triangulate them. In addition, the quotes derived from the other two interviews were also triangulated as much as possible through the organizations' published documents. In doing so, I aimed to draw a picture of each organization rather than solely describing the interviewees' opinions.

The documents and transcripts were analysed through two levels of qualitative coding, namely open and axial coding. The goal of these coding exercises was to filter out the main points from the interview transcripts in a comparable way (Boeije, 2009). These coding exercises were used to develop three summaries, which are presented in Section 3.3.

The organizations included in my thesis are not mentioned by name in order to ensure anonymity of the interviewees. Instead, the following pseudonyms are used in my thesis: Tatanka, Gaia and Endymion. To maintain this anonymity, I separated the consent forms, transcripts and coding schemes from this file. These documents can only be found in Appendices 5-9, which I submitted in a separate file to my thesis assessors.

The quotes derived from the transcriptions were formatted using APA guidelines as much as possible; however, these differed somewhat per source. Some interpunction used in the quotes is not completely self-explanatory, the guidelines used for these is as follows: ellipses '…' were used when text was left out; square brackets and a full stop '[.]' was used to indicate a pause; and square brackets '[ ]' were used when I added additional meaning to the quote, this was only done in cases where quotes needed extra clarification.

Results

3. Results

This chapter is divided in three parts. First, an ideal FSSA is build stepwise by outlining findings from existing publications (Section 3.1). Second, the ideal FSSA is re-evaluated by outlining methodological issues and choices that occur in academic edible oil FSSAs (Section 3.2). Third, Section 3.3 gives insights in how consumer FSSAPs deal with the methodological issues identified in the preceding Sections. Last, Sections 3.1, 3.2 and 3.3 will be aligned and compared in a final synthesis (Section 3.4).

3.1. The ideal FSSA

An ideal FSSA is outlined in this chapter, I mean with ideal that it is complete and balanced. The context of this ideal FSSA is framed first by discussing some overarching general aspects related to FSSAs.

3.1.1. Overarching aspects

Exploring sustainable agriculture

If the sustainability of agricultural food systems has to be assessed, having a basic understanding of what is meant by this is necessary. Many definitions of sustainable agriculture are based on the concept of sustainable development that became widespread due to a report called 'Our common future', which is a publication by the World Commission on Environment and Development (Brundtland, 1987). In this report, Brundtland (1987) pioneered the triple perspective of social justice, economic viability and environmental stability as sustainability's core. Considering this triple perspective as sustainability's core, sustainable agriculture would then consist of an activity that would not negatively impact these three dimensions for an indefinite period of time (Gómez-Limón & Sanchez-Fernandez, 2010). This rather broad conceptualization makes the assessment of sustainable agriculture ambitious, yet ambiguous. For example, the importance of each of the sustainability dimensions differs in varied area-based contexts (Roy & Chan, 2012). These contexts themselves are also affected by time (Bélanger et al., 2012). As such, using the three sustainable development pillars as described in the Brundtland report as a baseline brings new challenges with it. How to operationalize the perspectives of different stakeholders? Do they vary in importance? Whom decides this? The answers to these questions are depending upon how a FSSA is framed.

Horizontal, vertical and time boundaries

Several publications on agri-food system sustainability assessments (e.g. Bélanger et al., 2012; Van Asselt et al., 2014; Van Cauwenbergh et al., 2007) only focus on on-farm activities of the production cycle. This means that upstream activities (e.g. fertilizer production and feed import) and downstream activities (e.g. transport, processing and packaging) are excluded. I question the exclusion of these activities, since some of the larger production impacts might take place upstream or downstream. For example, milk production in The Netherlands causes multiple environmental impacts because of off-farm activities, such as the deforestation caused by the production of soy meal that is used as animal feed (Thomassen et al., 2008). This example illustrates that excluding parts of the production chain can result in overly optimistic FSSAs. Hence, I argue that an ideal FSSA must include a complete production chain. At the other hand, the choice to only focus on parts of chains is often substantiated. For example, Gómez-Limón and Sanchez-Fernandez (2010) elaborate upon their choice to only assess on-farm activities by referring to existing policies that mostly only focus on on-farm activities. On the contrary, authors sometimes exclude on-farm activities due to the inapplicability of larger area-based indicators to asses local farm contexts (Dantsis et al., 2010). As such, an ideal FSSA would assess both complete production chains and more local units of analysis, which requires a combination of assessment methods.

With regards to the unit of analysis of FSSAs, a horizontal and vertical component can be considered (Van Cauwenbergh et al., 2007). With regards to the horizontal component, the literature seems to be somewhat ambiguous. For example, Van Cauwenbergh et al. (2007) include parcel, farm and landscape as units of analysis, whereas Dantsis et al. (2010) use farm data and therefrom derived regional units of analysis. Gómez-Limón & Sanchez-Fernandez (2010) list plots (production unit), farms (management units), agricultural systems (landscape or ecological units) and regions or nation states (political units) as a possible hierarchy. In terms of the vertical component, Van Cauwenbergh et al. (2007) refer to the rhizosphere (e.g. soil acidification) and the four earth spheres, including the biosphere (e.g. biodiversity loss), geosphere (e.g. nutrient leaching), hydrosphere (e.g. runoff) and atmosphere (e.g. GHG emissions). Figure 1 depicts the life cycle boundaries and spatial units of analysis that one could consider within FSSAs according to Van Cauwenbergh et al. (2007). An ideal FSSA would consider all these units of analysis.

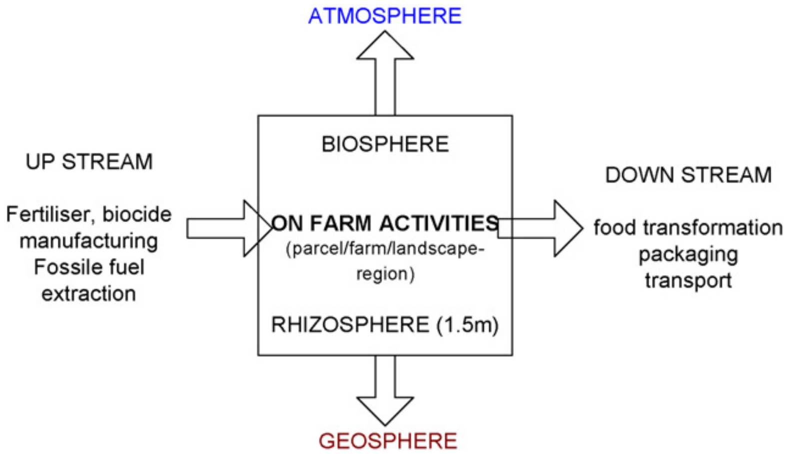


Figure 1: Possible product life cycle and earth spheres one could consider in FSSAs. Derived from Van Cauwenbergh et al. (2007)

An additional system boundary that one must consider in a FSSA is time. Time is essential since the meaning of sustainability morphs not only in space but also in time. Hence, extrapolation of FSSAs results should always be done with utmost care (Gómez-Limón & Sanchez-Fernandez, 2010). Indeed, FSSAs should avoid being a single snapshot in time (Meul et al., 2008). A simple example is the application of manure: even if manure management is appropriate, phosphorous values in soil might still be way too high due to accumulation in the past (Bélanger et al., 2012). Therefore, an ideal FSSA should be able to track changes in social, environmental and economic trends by means of time-integration (Van Cauwenbergh et al., 2007).

Local factors: Social context

FSSAs are not solely an objective matter, but also rooted in a subjective environment. Different factors matter within different social settings. A practical example of this social rooting is the current nitrogen debate within The Netherlands, in which excessive nitrogen pollution was circumnavigated by policy makers for years. However, the nitrogen pollution was suddenly brought very high on the political agenda by a ruling of the Dutch Raad van State (i.e. the Council of State) (Winterman, 2019). This example illustrates that the social point of view is influenced by time. What is considered important now might have been of low importance two years ago or in the future. Hence, an ideal FSSA is considers local circumstances and is easily adaptable due to the ever-changing nature of local agendas.

Local factors: farm environment conditions

Local farm conditions can also vary within food systems. A good example is local soil conditions in which specific conditions such as phosphorous contents and soil organic matter, could greatly affect the interpretation of assessments (Bélanger et al., 2012). An example of these differences in local farm conditions is illustrated in several LCAs on pork and chicken production system carried out by de Vries and de Boer (2010). The authors found that both farm management and climatic conditions caused large variation in acidification potential (de Vries & de Boer, 2010). As such, an ideal FSSA must consider local farm characteristics such as soil conditions, farm management and field emissions.

Food system sustainability assessment: indicators

The FSSA literature includes a large variety of procedures to carry out assessments of which the development of Sustainability Indicators is frequently advocated as favourable (e.g. Bélanger et al., 2012; Van Cauwenbergh et al., 2007).

In contrast to other procedures like the Cost Benefit Analysis and Environmental Impact Assessment, Sustainability Indicators can deal flexibly with the three sustainability pillars as described in the Brundtland report (Jeswani et al., 2010). This is shown by several authors who applied Sustainability Indicators for food system sustainability assessments in various ways. For example, Van Cauwenbergh et al. (2007) proposed the consistent use of Principles, Criteria and Indicators (PC&I), a framework that is designed to assess a hierarchy of multiple spatial units of analysis (parcel, farm and higher spatial units of analysis). Another application is proposed by Gómez-Limón & Sanchez-Fernandez (2010), who use 'composite indicators' to quantify and operationalize sustainable agriculture. A composite indicator is an indicator that is composed of multiple underlying multi-dimensional indicators. Applying composite indicators brings multiple advantages such as comparability, visible data reduction and applicability to various spatial units of analysis. However, the intentional reduction of data also comes with disadvantages such as possible misleading policy messages, loss of data transparancy, oversimplifications and the exclusion of difficult to measure concepts within assessments (Gómez-Limón & Sanchez-Fernandez, 2010). Besides these examples of sustainability indicators, there are many more applications that all have one thing in common – the applications are all highly pragmatic (Bélanger et al., 2012). It is this applicability that is also described in the research purpose of this thesis. As such, the ideal FSSA is mostly constructed on the body of 'Sustainability Indicators' literature.

3.1.2. Scoping, principles and criteria

A general context of the system under scrutiny is required for Sustainability Indicator assessments. Defining this context is done in varied ways by authors. For example, Van Asselt et al. (2014) started by defining the system boundaries and possible scenarios for their case study in collaboration with a policy maker. On the contrary, Gómez-Limón and Sanchez-Fernandez (2010) contextualized their study by first conducting a theoretical scan of principles and attributes within FSSAs. Likewise, Van Cauwenbergh et al. (2007) proposed to first determine some principles and concrete criteria to achieve sustainability. Each of these examples set the context for the remainder of assessments, oftentimes in collaboration with local stakeholders. It is this context that is an essential basis for indicator generation and interpretation; hence, it should be part of an ideal FSSA.

3.1.3. Indicators

Indicators: considerations

The selection and development of indicators comes with many considerations as indicators should represent complex systems accurately (Roy & Chan, 2012). First, the indicators included in an ideal

FSSA should cover each sustainability dimension. However, most sustainability assessments only focus on environmental sustainability, omitting all socio-economic aspects (Hayashi, 2000). In doing so, food systems' complexity is not well addressed, since that would require more holistic assessments (Smith et al., 2000).

Second, the number of indicators has to be considered. To use as many indicators possible to cover sustainable agriculture's complexity might sound logical. However, using as many indicators possible results in problems with data collection, validation and interpretation (Von Wirén-Lehr, 2001). At the other hand, by only considering a small number of indicators, assessments probably overlook critical aspects and not take certain system trade-offs into account (Roy & Chan, 2012). Therefore, Bélanger et al. (2012) propose to keep indicator sets as small as possible whilst capturing the maximum amount of indicators possible. Although this limited number of indicators could result in less completeness, the authors reason that *"it is better to be roughly right than precisely wrong"* (Goodland, 1995, p. 19, in Bélanger et al., 2012). The same reasoning is applied by authors who argue for using so-called 'essential indicators' (e.g. Bossel, 2002; Roy & Chan, 2012). Although there is no clear outline on how these essential indicators are exactly developed, Roy and Chan (2012) proposed some standard criteria: 1) indicators should accurately and objectively measure a system's capacity to meet the sustainable development goals, 2) local users should be able to apply them, 3) they should be scientifically valid, 4) they should be measurable, 5) data should be available, and 6) they should be cost-efficient.

Last, the notions of measurability and data availability also have an inherent flaw – they are fundamentally inconsistent with the inclusion of stakeholder perceptions; since stakeholder perceptions are difficult to quantify. A few authors support some form of qualitative indicators (e.g. Van Asselt et al., 2014), although qualitative indicators are seldom mentioned in the literature reviewed.

In line with Roy and Chan (2012) and Van Asselt et al. (2014), I agree that an ideal FSSA would include a small set of essential indicators. These essential indicators should consist of 1) generic concepts for comparability's sake and 2) tailored indicators to assess specific aspects of the food system under scrutiny. These tailored indicators should allow for a wide spectrum of (possibly qualitative) concepts in order to ensure inclusivity of stakeholder perceptions.

Indicators: development

Numerous methods for indicator development are proposed in the academic literature. These development methods differ in their degree of quality with regards to clarity, complexity and stakeholder inclusiveness (Roy & Chan, 2012). An often-used method for development is the so-called 'authors appraisal', in which experts deem themselves suited to develop the indicators (e.g. Dantsis et al., 2010). However, authors appraisal is criticized on its relevance since 1) experts might bring with them their own biases, and 2) experts cannot always ensure that indicators are useful for local situations (Roy & Chan, 2012). Another method put forward frequently is 'participative indicator selection', in which local stakeholders are empowered through consultation (e.g. Reed et al., 2006). Although more inclusive, participative indicators selection's drawbacks include 1) a too large variation in participants' perception of sustainable agriculture, and 2) the overrepresentation of specific individuals or institutions their agendas (Roy & Chan, 2012). As such, some authors plea for a middle ground (e.g. Roy & Chan, 2012; Van Asselt et al., 2014) – the use of a combination of bottom-up and top-down indicator development. By using this combination, academics can ensure scientific validity, and external experts and stakeholders can ensure local relevance. Following Roy and Chan (2012) and Van Asselt et al. (2014), I define a combination of bottom-up and top-down indicator development as ideal.

3.1.4. Validation of indicators

Not validating indicators affects the credibility and quality of any FSSA. However, the literature on how to validate FSSAs indicators only includes a few publications (Roy & Chan, 2012). One validation method was proposed by Meul et al. (2008), who elaborated on three forms of possible validity. First, 'design validity' covers the level of scientific grounding of the indicators. Second, 'output validation' assesses indicators on if they measure what they were meant to measure. Last, 'end-use validation' is used to screen the usefulness of indicators for potential decision making. Similarly, Roy and Chan (2012) also plea for indicator validation through combined expert appraisal and participation of stakeholders. Combining expert appraisal and stakeholder participation was even more expanded in the three stage validation method proposed by Cloquell-Ballester et al. (2006). The three stage validation method consists of: 1) the self-validation stage in which the 'work team' reflects on the correct performance of the indicators; 2) the scientific validation stage in which independent experts evaluate the indicators; and, 3) the social validation stage in which relevant stakeholders' judgements are integrated. Following Cloquell-Ballester et al. (2006), I agree that the indicators included in an ideal FSSA's are validated by a research team, independent experts and relevant stakeholders.

3.1.5. Evaluation

Multiple evaluation methods could be used to define a set of indicators its 'level of sustainability'. These evaluation methods include the comparison of reference values, expert appraisal, applying weights, and sensitivity analyses (Petit et al., 2018; Roy & Chan, 2012; Van Asselt et al., 2014). Each of these methods comes with certain pros and cons, which are outlined below.

Reference values

Two types of reference values are absolute and relative reference values. Absolute reference values constitute of values derived from the published scientific literature, legal norms and policy targets (Van Asselt et al., 2014; Van Cauwenbergh et al., 2007). Indicators and reference values have to be assessed on the same spatial units of analysis when compared. Since, not every absolute reference value makes sense on a smaller area of extent (e.g. economic criteria). Therefore, a comparison between reference values and indicators might only be possible when larger regional or national units of analysis are applied (Van Cauwenbergh et al., 2007). Besides absolute reference values, there are also relative reference values. Relative reference values are based on best practices, regional averages and trends over time (Van Cauwenbergh et al., 2007). Problems that can occur with relative reference values are the lack of available data and the difficulty to define quantifiable values for some economic and social indicators (Roy & Chan, 2012).

Author and expert appraisal

Another form of evaluation is author appraisal, in which the research team evaluates a developed set of indicators themselves (e.g. in Bélanger et al., 2012). Some authors state they rather avoid this method – they would only use it when no other method suffices (e.g. Meul et al., 2008). A drawback of author appraisal is the researcher's subjectivity. For example, an environmental researcher alone might be ill fitted to define evaluation criteria for non-environmental factors, such as local social factors. Therefore, Roy and Chan (2012) advocate for a more inclusive appraisal, which is an interplay between policy-makers, scientists, local stakeholders and communities.

Weighing

Weighing of different indicators is seen as the most important step within FSSAs by some academics. For example, Gómez-Limón and Sanchez-Fernandez (2010) stated that weighing of indicators reinforces subjective views on sustainability, since sustainability ultimately is a 'social construction' that is always affected by relative importance of individual indicators.

On the positive side, weighing exercises can incorporate society's preferences and therewith make assessments more politically relevant, if desired. However, this socio-political relevance is heavily reliant on the weighing method used by the researcher. An example of a weighing method that incorporates society's preferences is transdisciplinary weighing (Van Asselt et al., 2014). Transdisciplinary weighing incorporates not only top-down actors (e.g. experts) but also bottom-up stakeholders (e.g. farmers and local politicians). In doing so, transdisciplinary weighing is both scientifically and socio-politically relevant. However, the incorporation of multiple stakeholder groups in weighing exercises also comes with a large drawback – different stakeholder groups will assign different weights to indicators, possibly resulting in different assessment outcomes. As such, an assessment's outcome can be highly influenced by something simple like the people that were–or were not—invited to a weighting workshop (Castellini et al., 2012). When a decision is made on the final weights, the weights can be tested on sensitivity through a sensitivity analysis. In doing so, the consequences of including and excluding certain weights can be easier understood, and adjusted if necessary (Petit et al., 2018).

Seemingly, every evaluation method comes with some drawbacks; hence, to state an ideal evaluation method is difficult. Therefore, I deduce that an ideal FSSA would at least meet the following requirements: 1) the FSSA include an evaluation method in the first place, 2) the drawbacks of this evaluation method are reflected upon by the researcher, and 3) the evaluation method are screened by a sensitivity analysis if possible.

3.1.6. Data collection

LCAs can be used to collect data for Sustainability Indicators. The LCA is a systematic and robust tool that quantifies the environmental impacts of a product, process or an activity (Jeswani et al., 2010). The LCA is time and location independent and standardized according to ISO's requirements. Although the LCA is widely applicable, most LCA applications focus only on environmental impacts, neglecting the importance of the social and economic domains (Jeswani et al., 2010). In addition, the published body of LCAs shows some flaws: only a small number of agricultural products is covered, (allocation) methods vary widely, agricultural chains are often only partially covered, and the focus is mostly Western European (Poore & Nemecek, 2018). Nonetheless, the LCA is potentially an appropriate tool to collect generic environmental data in an ideal FSSA – but only if the LCA's defects are covered by supplementary means of assessment.

3.1.7. The ideal FSSA: conceptual framework

The discussed literature on FSSAs shows a recurring pattern that causes trade-offs – choices must be made in which one good option has to be sacrificed over another good option. Many methodological considerations can be placed in this dualistic line of thought. Can an assessment be fully objective or should a certain level of subjectivity be incorporated? Are less data intense less refined assessments preferred over locally important granular assessments? Should qualitative indicators be excluded from assessments for comparability's sake? Even though these same qualitative indicators might be important for understanding the local context? Unfortunately, no clear-cut answers exist to these dualistic questions.

Due to these methodological discrepancies, depicting what an ideal FSSA would look like is impossible – the ideal FSSA simply does not exist. During an assessment, a researcher has to make choices. Due to these choices, assessments cannot be fully complementary. Indeed, the ideal food system assessment simply does not exist due to the paradoxical choices that researchers have to make. This paradoxical space in which FSSAs are utilized is depicted in Figure 2. It shows that reaching the ideal FSSA is impossible since choice made by a researcher automatically sacrifices another desired option. Consequently, a FSSA can just be pulled around in the space of trade-offs that lurks under the ideal possibility – which is an impossibility.

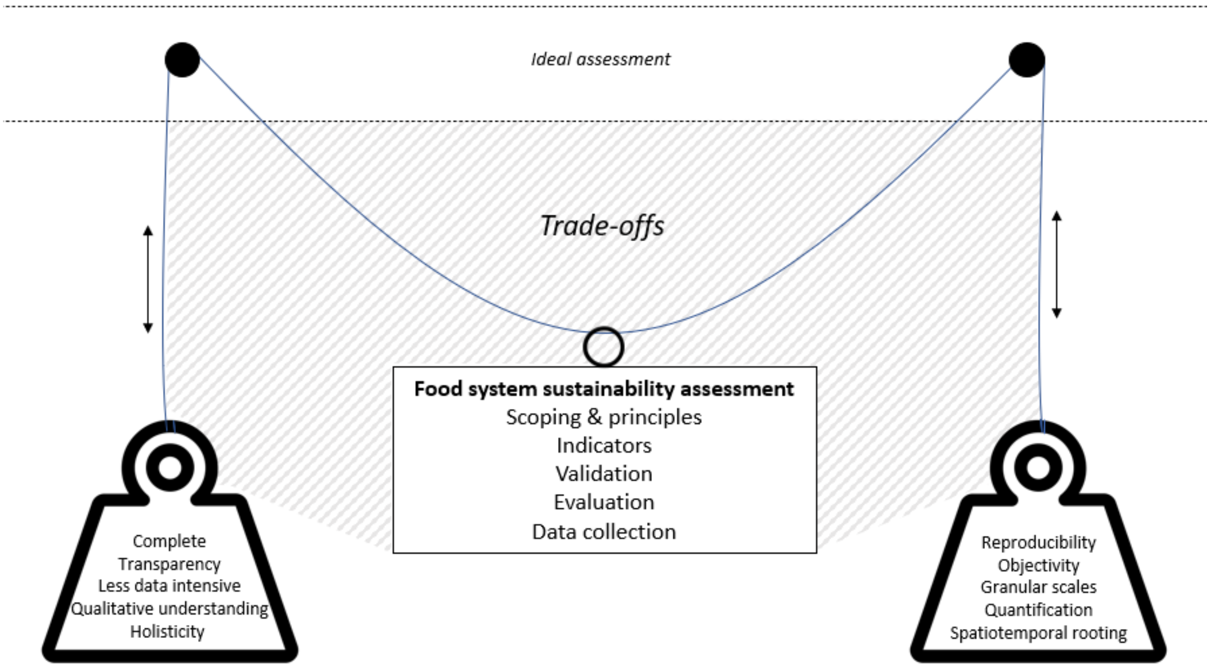


Figure 2: Conceptual framework: The ideal FSSA that is by default hampered by trade-offs

However, when the limiting barriers of research projects such as time, access and money are ignored, the outline of an ideal FSSA can be sketched. In doing so, the choice trade-offs can be reduced to some extent. Instead of choices between this *or* that, an ideal FSSA could include this *and* that. For example, 'comparability *and* locally adapted' instead of 'comparability *or* locally adapted.'

The ideal FSSA would then consist of a combination of methods that assesses complete production chains, considering both specific and generic factors. As such, the ideal FSSA would be a holistic assessment that is tailored to the food system under scrutiny. The factors considered in this assessment should include the social, economic, and environmental dimensions of sustainability. With regards to the environmental dimensions, the most concerning effects on the biosphere, geosphere, hydrosphere and atmosphere should all be included in the ideal FSSA for completeness sake. Also, the ideal FSSA needs to be tailored to local needs and factors. These local factors include socioeconomic needs determined by local agendas, as well as physical factors such as field emissions, farm management and soil conditions. Additionally, the ideal FSSA also needs to incorporate some form of time-integration in order to avoid the assessment becoming a single snapshot.

Both the scoping of an assessment and the generation and validation of indicators should be carried out from a bottom-up and top-down perspective in the ideal FSSA. So, the ideal FSSA includes voices from local stakeholders, external experts, and researchers. Inclusion of these voices ensures local socio-political relevance and scientific validity. Furthermore, the number of indicators should be relatively small. Favourably, the final set of indicators should cover 1) essential generic concepts for comparability's sake and 2) specific indicators tailored to the local food system under scrutiny. The final set of indicators should allow for a wide spectrum of possible indicators, including qualitative indicators, to ensure local stakeholders' perceptions. Ideally, the final set of indicators is evaluated after validation and tested on sensitivity by a sensitivity analysis. The researcher is also transparent on the flaws of these evaluation method(s). Last, generic environmental data can be collected through LCAs – if supplementary measures are taken to assess additional local and specific considerations.

3.2. Edible oil system sustainability assessments: methodological choices and issues

This section outlines several methodological problems that were identified in the literature regarding edible oil system sustainability. First, the scopes of the studies reviewed are discussed. Second, more specific methodological issues and choices are highlighted. Every paragraph follows the same structure: 1) an issue raised in a publication is outlined and 2) the issue is evaluated by analysing the frequency of occurrence in edible oil LCAs (peanut, sunflower, canola, and olive). A summary of this evaluation can be found in Appendices 2 and 3.

Some paragraphs are specifically related to olive oil systems. Paragraphs that focus on olive oil systems are highlighted by mentioning olive oil in italics at the start of the paragraph. The focus on olive oil was an unintended outcome of the thin body of assessment literature on peanut-, sunflower- and canola-oils.

3.2.1. Scope of the studies reviewed

The environmental, social, and economic dimension of sustainability

The majority of the *olive oil* FSSAs only focus on the environmental dimension of sustainability (Menozzi, 2014). As a result, knowledge on socio-economic sustainability of the olive oil sector is limited. However, socio-economic functions of olive oil regions are regarded as most important by local stakeholders (Egea & y Pérez, 2016). Olives are sometimes even referred to as a 'social crop' (Gomez-Limon & Riesgo, 2010). This local socio-economic importance of the olive crop and the environmental interests of the scientific community depict a gap between local interests and the focus of scientific communities.

A partial explanation for this gap is the difficulty to assess the economic and especially the social domain (Guillaume Busset et al., 2013; Fleskens et al., 2009). For the economic domain, high quality data might be available, but confidentiality could hamper organizations' willingness to share these (Petit et al., 2018). For the social domain, multiple aspects make assessment difficult. First, social impacts are difficult to relate to a specific unit of produce, for example a litre of olive oil (Petit et al., 2018). Additionally, social indicators that are difficult to quantify are oftentimes made qualitative – resulting in lower accuracy of assessments (Petit et al., 2018).

In terms of the *edible oil* LCA case studies reviewed, all 33 papers included the environmental dimension of sustainability (Appendix 3). Additionally, 5 papers included the economic dimension and 2 the social dimension.

Scope of analysis

The ultimate aim of a LCA is to identify the overall environmental impacts of a production system. However, LCA publications regarding *edible oils* frequently have a more limited scope (Banias et al., 2017). In the literature review on *edible oils* carried out by Khatri and Jain (2017), the goal and scope were provided in all LCAs evaluated. However, the goals stated in the sample of Khatri and Jain (2017) varied: 69% of the studies attempted to assess a multi-fold of impacts, whereas individual goals included GHG emissions (20%), energy analysis (14%), land use (3%), water demand (6%) and economic analysis (14%).

The outcome of my evaluation of LCA studies (Appendix 3) is somewhat similar with the results found by Khatri and Jain (2017): in my evaluation, 26 papers aimed to assess environmental impacts, 3 papers constituted of an economic analysis, 4 were an energy analysis and 15 papers had more specified scopes. The 15 papers with more specified scopes focussed on: evaluations of different methods of weeding and fertilizing (respectively De Luca et al., 2018; Noorhosseini & Damalas, 2018), exploring residue and waste management techniques (see Cossu et al., 2013; Guermazi et al., 2017), identifying the consumption of raw materials and natural resources (see Avraamides & Fatta, 2008; Iraldo et al., 2014), integration and comparison of multiple assessment methods (see Cavallaro & Salomone, 2010; Notarnicola et al., 2004; Schmidt, 2004), assessing LCA parameter uncertainty (see Figueiredo et al., 2017) and assessing the inclusion of land use change in LCAs (see Figueiredo et al., 2012).

3.2.2. Methodological issues and choices

In their literature review, Khatri and Jain (2017) showed that methodological variation is a key challenge within the body of *olive oil* system assessments. Khatri and Jain (2017) explained that the methodological variation they found in several publications expressed itself by the different assumptions made by academics regarding olive oil systems, differences in methodological choices, data used, impact categories analysed, and allocation criteria applied. Indeed, the body of olive oil system assessments includes wide-ranging results for the same product (Khatri & Jain, 2017; Petit et al., 2018). Consequently, the comparability of olive oil FSSAs is heavily reduced. The coming paragraphs will highlight the varied choices made in both *olive oil*- and other *edible oil*-system sustainability assessments.

System boundaries

LCAs on *olive oil* systems are criticised for being too limited in their system boundaries – a majority of the publications only focus on the farming and manufacturing phases (Banias et al., 2017; Khatri & Jain, 2017). These narrow system boundaries are frequently justified by the fact that the agricultural and manufacturing phases have the most substantial impacts on the environment (Banias et al., 2017; Khanali et al., 2018). However, other phases in the production chain also result in environmental pressure but are not always included in FSSAs. Some of the main phases that are often neglected are packaging, warehousing, transportation, reusing and handling of waste streams (reverse logistics); and temporal boundaries. For example, warehousing was only mentioned twice in the 98 papers reviewed by Banias et al. (2017). In the two papers that mentioned warehousing, the phase was merely associated with waste – meaning that no managerial insights could be drawn by Banias et al. (2017).

With regards to the *olive oil* production chain, a phase that needs extra mentioning is the treatment and reusing of olive oil by-products, also referred to as reversed logistics. The wrong handling of these by-products is one of the main environmental issues in the olive oil production chain (Banias et al., 2017). The most concerning of these by-products is olive mill wastewater (OMWW) – OMWW can be very damaging for local ecosystems when improperly treated. Hence, the reversed logistics phase is essential in olive oil system sustainability assessments.

Regarding the *edible oil* LCAs I reviewed, the system boundaries mainly included the agricultural and manufacturing phases (Figure 3). Notably, most of the publications included the agricultural phase (31 papers out of 33). The two papers that excluded the agricultural phase explicitly mentioned that the focus of the paper was on other phases in the product chain namely: reversed logistics of wastes (see Cossu et al., 2013) and packaging of vegetable oils (see Accorsi et al., 2013).

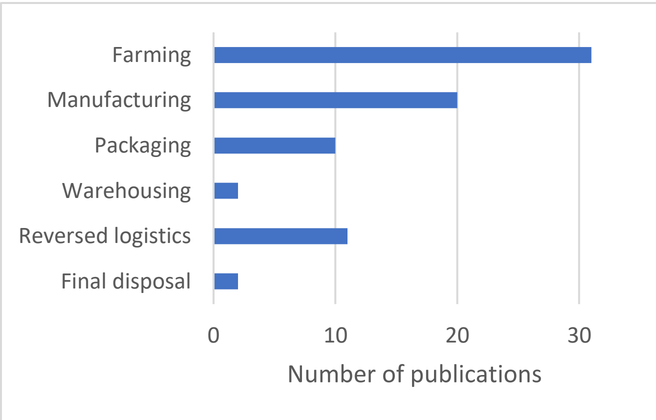


Figure 3: Production chain phases included in reviewed LCA studies

Temporal boundaries

Time-based boundaries were reported to be scant in *edible oil* LCAs. Only 14 out of 36 studies reviewed by Khatri and Jain (2017) reported any temporal system boundaries. I found similar results in my evaluation: only 11 out of 33 articles clearly stated temporal boundaries (Appendix 3). An additional eight articles did mention time boundaries indirectly, although not explicitly, through seasons and the period in which a survey was distributed.

Data allocation

Allocation of input-output flows between products and their co-products is one of the major methodological issues within the agri-food literature (Poore & Nemecek, 2018). The issues with allocation methods are important since allocation criteria have substantial impact on the results of assessments. In their review on *edible oils*, Khatri and Jain (2017) found that most heterogeneity of researchers' choices is indeed found in allocation criteria. Even more so, authors oftentimes completely ignore allocation criteria, or simply reject the importance of these criteria (Khatri & Jain, 2017). For example, Guillaume Busset et al. (2013) state: "*the olive oil production process is simple so no allocation rule is needed"* (p. 4). However, I argue that olive oil can be associated with several by-products such as olive oil pomace; or, co-produced secondary products such as almond, vine, cereals, firewood and grazing area for animals in mixed farming systems (Fleskens et al., 2009).

The same heterogeneity of allocation methods was found in my evaluation of *edible oil* LCAs (Figure 4). In 14 out of 33 papers reviewed no allocation method was specified at all. Additionally, 3 papers did acknowledge allocation methods as a common procedure but did not apply them in their final work. The remainder (16 papers) used a variety of allocation methods, or multiple at the same time, including: system expansion (4), mass allocation (4), economic allocation (7), energetic allocation (2), substitute calculations in consequential modelling (2) and the 'next cycle approach' (1).

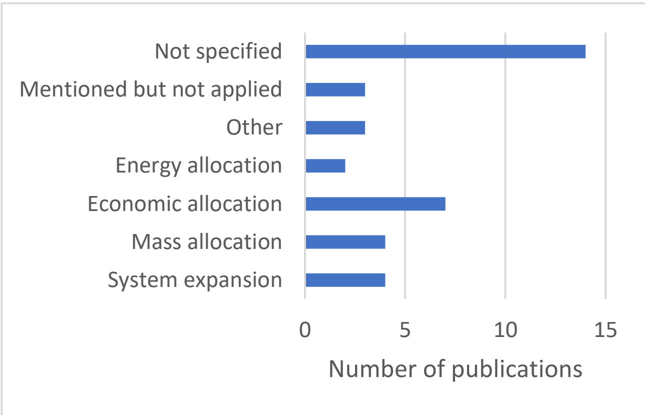


Figure 4: Allocation methods used in reviewed LCA studies

Inventory

I identified several issues with LCA inventories in the literature. First, Banias et al. (2017) found that fuel is consumed in practically every phase of the *olive oil* life cycle. However, the impacts of fuel use are rarely assessed despite the prominent presence of fuel in the production chain (Banias et al., 2017). Second, the olive tree is a species that does not require irrigation per se. However, when irrigated, the olive tree substantially improves its yield, emphasizing the need for inclusion of this inventory item (Salomone & Ioppolo, 2012). Last, multiple publications on *edible oil* LCAs highlighted a limited availability of data on fertilizer inputs, pesticides inputs, the use of capital goods and human labour (Khatri & Jain, 2017).

In my literature review, I also evaluated the presence of inventory items in *edible oils* LCAs (Figure 5). Most noteworthy, my evaluation shows that water use was the least mentioned inventory item (20 out of 33). In addition, I also evaluated OMWW because of its large relative environmental impacts in olive oil systems, even though OMWW is an inventory item that only applies to olive oil production chains. OMWW was included most of the times (11 out of 12 papers) in the *olive oil* LCAs that considered the manufacturing phase.

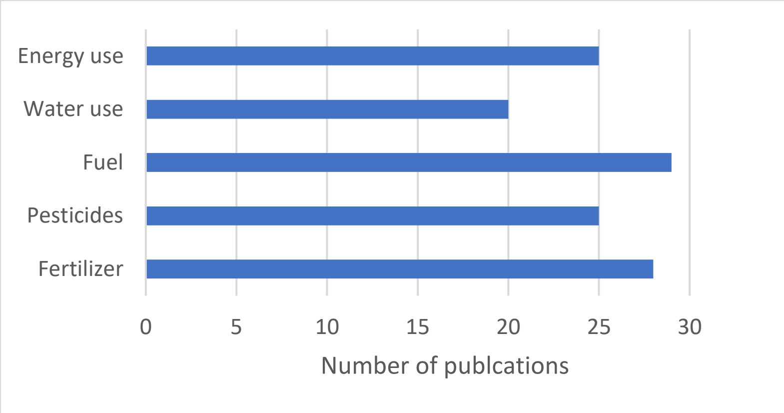


Figure 5: Life Cycle Inventory items included in reviewed LCA studies

Farming systems

Farming systems themselves are also varied – a fact that is often ignored in assessment studies despite its importance in food systems (Basset-Mens et al., 2010). When included, the farming system in which a product is grown can determine the outcome of a LCA.

I evaluated 33 *edible oil* LCAs for this thesis, of which 31 LCAs included the agricultural phase. Of these 31 papers, 16 did not specify the type of farming system examined (Figure 6). The main farming systems specified by the other studies were organic-, integrated-, conventional- and traditional-farming. Additionally, some authors examined specific traits of farming systems such as rainfed versus non rainfed systems (see Figueiredo et al., 2017; Figueiredo et al., 2012; Kazemi et al., 2016), high density versus low density systems (see De Gennaro et al., 2012), extensive systems (see Guermazi et al., 2017), different residue management applications (see Mousavi-Avval et al., 2017; Queiros et al., 2015), and different frequencies of tillage (see Queiros et al., 2015).

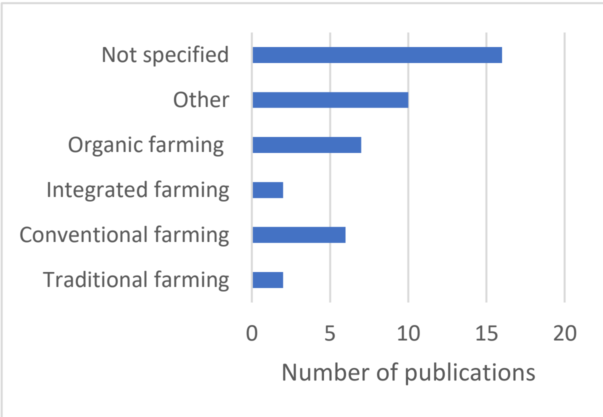


Figure 6: Farming systems considered in reviewed LCA studies

Soil characteristics and management

The outcomes of a LCA can also be affected by area-based characteristics of a production system. For example, in Italy, some olive varieties perform very well on one farm while performing poorly on another farm in the nearby vicinity – which depicts the influence of different soils, climates and exposure in different areas despite the relative small distance (Menozzi, 2014).

Of these area-based characteristics, I evaluated the occurrence of soil characteristics in my literature review. Soil characteristics were rarely mentioned in the 31 *edible oil* LCAs that included the agricultural phase (8 out of 31 papers). Of the 8 papers that included soil characteristics, 7 indicated a (range of) of soil type(s). The remaining paper included a complete soil analysis (see Noorhosseini & Damalas, 2018). In addition, 13 papers included the effects from soil management as an inventory item – the soil management mainly constituted of ploughing.

Field N_2O emissions

Field N_2O emission caused by fertilizer use contributes largely to a farm's GHG emissions (Schmidt, 2004). However, methods and data to estimate field N_2O emissions accurately do not exist (Basset-

Mens et al., 2010). Oftentimes, field N_2O emission are estimated through generic emissions factors or general inventories. This type of estimation results in substantial uncertainties of FSSAs results (Basset-Mens et al., 2010).

The lack of accuracy of the two main estimation methods (the IPCC method and Brentrup's method) was indeed often reported in the *edible oil* LCAs I evaluated in my review. However, alternatives to these estimation methods were not put forward. From the publications that included the agricultural phase (31), 11 did not mention field N_2O emissions. Of the remaining 20 papers, 10 used the IPCC's method to calculate field N_2O emissions (see Eggleston et al., 2006), 9 authors used the 'Brentrup' method (see Brentrup et al., 2000) and 1 used an average value presented in an earlier publication (Appendix 3).

Impact categories

Translating LCA input-output flows into impact categories is common practice. Although a wide variety of impact categories exists, the LCA publications regarding *edible oils* frequently only focus on a limited number (Banias et al., 2017; Khatri & Jain, 2017). According to Khatri and Jain (2017), the most frequently applied impact categories include global warming potential (GWP), acidification potential, eutrophication potential, ozone layer depletion, various ecotoxicity categories, photochemical oxidant formation, human toxicity, energy use and land use related impacts.

However, multiple other impact categories are seldom included in LCAs even though these neglected impacted categories can cause substantial damage to local landscapes. Examples of these neglected impact categories include non-point water pollution (Banias et al., 2017; Guillaume Busset et al., 2013), soil erosion (Fleskens et al., 2009), land use change (Schmidt, 2015) and biodiversity loss (Gomez-Limon & Riesgo, 2010). Several authors pointed out that these impact categories are essential to draw a complete picture of environmental sustainability. One reason for exclusion of these impact categories is the inability of LCA methods to deal with them (Khatri & Jain, 2017). A second reason for this exclusion is researchers' tendency to base relevancy of impact categories on reliability. As a result, impacts that are difficult to measure in a reliable way are oftentimes excluded from assessments (e.g. in Badey et al., 2013).

My evaluation of *edible oil* LCAs confirmed the neglection of several impact categories and a strong focus on others (Figure 7). In my evaluation, climate change potential was included in all publications except one, followed by acidification potential (24 out of 33) and eutrophication potential (22). Additionally, some impact categories were considered regularly by authors, namely: ozone layer depletion (15), ecotoxicity (14), human toxicity (14), photochemical oxidant formation (13), land use related impacts (16) and abiotic depletion (14). Noteworthy is the low amount of papers that considered biodiversity loss (2), water pollution (2); and, the complete absence of soil erosion.

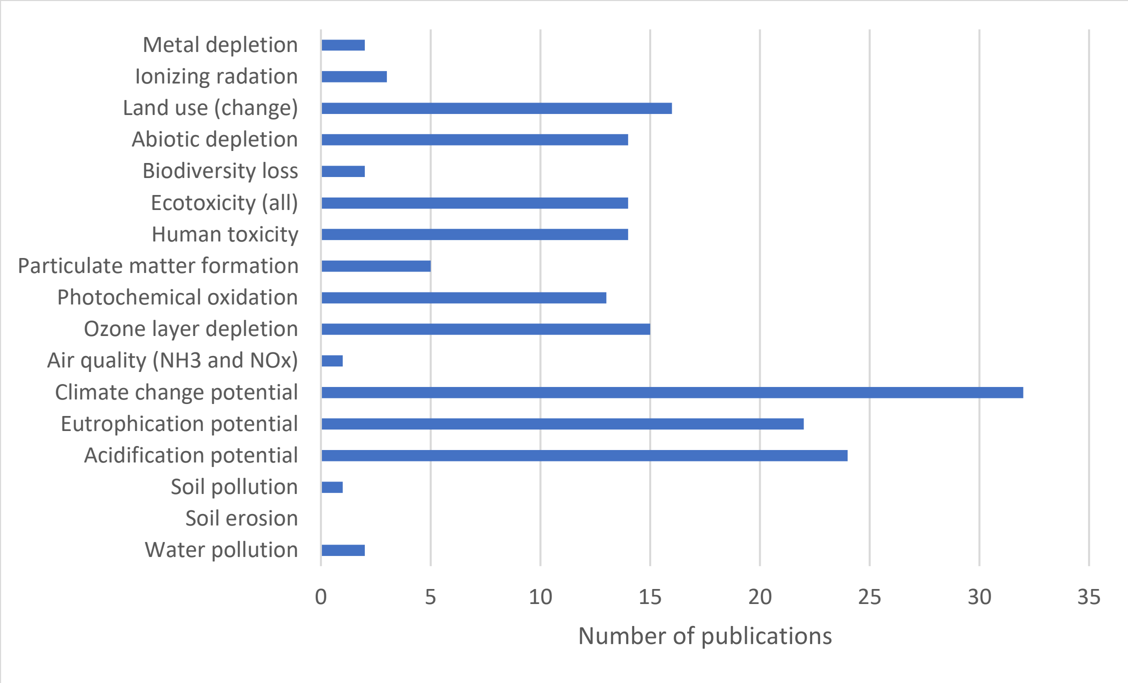


Figure 7: Overview of impact categories included in LCAs reviewed. Note: ecotoxicity is a combination of freshwater-, terrestrial- and marine-ecotoxicity

Contribution, sensitivity, and uncertainty analyses

Khatri and Jain (2017) regarded the interpretation of *edible oil* LCA results as too shallow: 67% of the studies analysed by them included a breakdown of the contributions made on impacts by differing life cycle stages (also known as a contribution analysis), 48% of the studies included in their sample applied a sensitivity analysis, and 15% of the publications included an uncertainty analysis. Indeed, evaluation methods seem to be applied somewhat haphazardly. Therefore, researchers recommend standardization of sensitivity analyses within FSSAs to reduce the inherent variability currently present in the body of assessment literature (e.g. Petit et al., 2018).

In my evaluation of *edible oil* LCAs, I also found that evaluation methods were often left out (Appendix 3): 21 out of 33 studies included a contribution analysis, 13 a sensitivity analysis and only 5 an uncertainty analysis.

3.3. Consumer FSSAPs

This section attempts to answer the research question 'How do current providers of consumer FSSAPs assess sustainability?' To answer this question, three existing consumer FSSAPs were evaluated. Each of these organizations was evaluated on the same aspects. These aspects form the structure of this section and are summarized in Table 2.

Table 2: Overview of main differences and similarities between Tatanka, Gaia and Endymion

	Tatanka	Gaia	Endymion
Organization/project aim	To provide sustainable food choices in any restaurant and food market	To make sustainable food choices as easy as possible for consumers	To facilitate a new form of consumerism. This involves a literate decision making for the individual consumer
View on sustainability	Measurable: communication should be effective and transparent	Practical: enables transparent and practical information	Social construction: sustainability is in the eye of the beholder
Socio-economic dimensions	Excluded	Excluded	Dependent on labels present on products
Impact categories / indicators	Climate score, water scarcity footprint, tropical forest conservation, animal welfare	Use of fossil resources, climate change contribution, land use and waters stress	Main clusters include environmental, social, health and quality
System boundaries	Production, handling, transportation, storage, and disposal	Cultivation, transportation, and packaging	Platform does not use LCA data
Seasonal produce	Included	Included	Not mentioned
Local produce	Included: best rating obtained when product travels <200 km	Partial: through water stress and transport distance	Partial: through transport distance
Organic produce	Included	Excluded	Included through labels
Data collection	LCI inventories, calculations, labels and surveys	ReCiPe 2.0: LCI inventories	Compilation of already existing data from various sources
Data aggregation	Separate score for each impact category	Impact categories fully aggregated in 1-5 environmental score	Sustainability criteria ranked by consumer are aggregated in 1-10 sustainability score
Validation	Expert validation	Expert validation	Expert validation

3.3.1. Tatanka

Company description and focus

Tatanka is an organization that attempts to measure environmental footprints of food products in an exact and efficient way. According to Tatanka's website, their goal is to provide sustainable food choices in any restaurant or food market. The sustainable food choices are facilitated through environmental footprints, which are assessed in cooperation with food producers and restaurants. Besides the provisioning of these environmental footprints, Tatanka aims to make all underlying data on CO_2 publicly available. I interviewed the CEO of Tatanka, who shared that within their aim, the relation between food production and climate change is stressed more than other impact categories. The representative explained that other impacts are important too, but not as pressing:

> *"According to a World Bank study, the displacement and destruction [induced by climate change] is worse than all human wars in history combined. So, you can pretty clearly imagine that this is a humanitarian catastrophe that is out of the scope of other indicators, plastic and what not. I am not saying that these other points are not important, they need to be regarded, and if you want to do things right you try to do all things right. But from a utilitarian perspective there is kind of a 100-fold magnitude between issues that are discussed in the media right now and issues that really matter."*

View on sustainability

During our interview, I probed Tatanka's representative by stating that sustainability is a social construction. The interviewee disagreed with my statement. Their believe was that sustainability can be measured and estimated: *"It should be communicated transparently and effectively with the most important impacts being the main focus – especially climate change induced by GHG emissions."* This is in line with a statement made in supporting material published by the company – the publication emphasized Tatanka's strong focus on the link between food and climate change.

Socio-economic dimension

The representative also explained that the socio-economic dimensions are not directly within the scope of their focus. The representative reasoned that the socio-economic dimensions are addressed indirectly in the long-term: by reducing climate change impacts, the socio-economic consequences are also somewhat mitigated.

Indicators, impact categories and life cycle stages

Tatanka calculates climate scores (based on GHG emissions), other environmental footprints, and health scores. The environmental footprints are developed in collaboration with a group of scientists. This group of scientists also determined the nine most essential indicators to assess agricultural impact in a research project. The determined essential indicators include the carbon footprint, water use, land use, eutrophication, ecotoxicity, soil fertility, biodiversity, conservation of tropical rainforests and animal welfare. Of these nine essential indicators, four are incorporated in the organization's platform: the carbon footprint, water use, conservation of tropical forests and animal welfare. The other indicators are excluded from Tatanka's platform.

The reason for in- and exclusion of the essential indicators varied somewhat. Obviously, the carbon footprint was incorporated because of the company's focus. The reason for incorporation differs for water use, tropical forest conservation and animal welfare – these indicators conflict with the carbon footprint of food products. This conflicting nature emerges in food scores: when a food product is scored 'good' on water use, conservation of tropical forests and animal welfare, the product occasionally is scored 'bad' on carbon footprint. This conflicting nature can be depicted in a practical example of an on-farm measure: the extension of cattle's lifetime. This measure is more ethical towards the animal, but also increases the amount of CH_4 emitted by the animal over its lifespan, resulting in higher carbon footprints. Tatanka wants to be transparent about these conflicting natures; therefore, Tatanka intentionally adds these conflicting indicators to their platform.

The five other essential indicators identified in Tatanka's research project were eutrophication potential, land use, soil fertility, biodiversity loss and ecotoxicity. These additional essential indicators were excluded from Tatanka's platform for two reasons. First, eutrophication and land use were excluded from the Tatanka's platform since these indicators behave synergistically with carbon footprints – meaning that if a consumer is able to reduce its food carbon footprint, it probably also reduces the impact on land use and eutrophication. An example of an on-farm measure that depicts such synergistic behaviour is the reduction of overfertilizing: this measure not only reduces the amount of nutrients being washed out and therewith eutrophication potential, but also reduces the residue amount of nitrogen in soils that is potentially transformed to N_2O. Thus, reducing the carbon footprint. Second, three other essential indicators -soil fertility, biodiversity loss and ecotoxicity- were excluded from Tatanka's platform since no scientific agreement exists on how to assess food production's contribution to these three indicators.

The four indicators that are included in Tatanka's platform are assessed through different means. First, the climate score is calculated through CO_2eq values estimated in LCAs. The LCAs include the following lifecycle stages: production, handling, transportation, storage, and disposal. Second, water use is assessed through the water scarcity footprint. This water footprint constitutes of two elements: the surface- and groundwater use of a product in a particular region and the relative water scarcity in this same region. Third, the indicators for tropical rainforest conservation and animal welfare are both assessed by means of certification schemes and labels. In Tatanka's platform, the rainforest conservation and animal welfare scores automatically receive the highest rating when soy, palm oil and animal products are not part of a consumer food product.

Seasonal and local produce

Next to the most critical indicators, seasonal production and local produce are incorporated through local and seasonal scores in Tatanka's platform. A maximum local score is awarded when an ingredient travels on average less than 200 kilometres distance. No information was obtained on the compilation of Tatanka's seasonal scores.

Organic produce

Tatanka carried out a large scientific research project on organic produce. The project's conclusion was complex: organic produce improves animal welfare and reduces the amount of tropical deforestation, but at the same time worsens the carbon footprint per kg of meat product. Interestingly, although true for meat products, these findings do not apply to plant-based products per se. In addition, the assessment of organic produce is even more so complicated because different countries have different organic labelling requirements and national regulations on what constitutes organic production.

Moreover, Tatanka promotes assessment of complete meals rather than single products. Ironically, the research project carried out by Tatanka showed opposite results for complete meals – meals that constitute solely of organic products mostly have better carbon footprints than meals that constitute of non-organic products. Hence, Tatanka's research report concludes that it is more important to select products with a low carbon footprint in general than if it was produced organically or not, even though substantial differences between organic and non-organic product environmental footprints exists.

Nonetheless, a distinction between organic agriculture and non-organic is included in the organization's carbon footprint calculations. However, this distinction is not explicitly mentioned in their final environmental scores.

Data collection and trustworthiness

Different forms of data collection are applied in Tatanka's platform. Tatanka's representative explained: *"[the very core of data collection] is the engine that we have programmed and the data we have collected from the Life Cycle Inventories and at the other side that we have systematic calculations and matchings of all these ingredients."* Besides LCI data from databases, Tatanka also offers more tailored food scores in collaboration with producers. In such a collaboration, Tatanka sends a questionnaire to the producers which in turn fill the questionnaire out and return them to Tatanka. Based on the producer's answers, precise GHG emissions are estimated by Tatanka. According to Tatanka's representative, collecting data through producer questionnaires at the one hand and collecting data from Life Cycle Inventories at the other hand represent two sides of a spectrum. Within this spectrum, a middle ground also exists: *"[in this middle ground] we try to gather as much possible digital data about the products and then do the calculation."*

Tatanka's representative also elaborated on the trustworthiness of their data. Unfortunately, I had to leave out this information because I lost this part of the interview recording due to technical issues with my voice recorder.

Data aggregation

Tatanka's representative clearly showed dissatisfaction with totally aggregated sustainability scores. The interviewee explained that by totally aggregating sustainability scores, certain actors can put pressure on the importance of some aspects, resulting in reduced weighted importance of other aspects. The pressure of actors can politicize overarching sustainability scores. The representative explained that to avoid this politicization, Tatanka shows individual rates and scores for certain impact categories in their methodology.

A summary of the built of Tatanka's platform's is depicted in Figure 8. The figure includes the data collection methods (bottom line), environmental footprints (circles) and final presentation of the environmental footprints to consumers (stars and circles).

Validation

Tatanka works in close cooperation with researchers and European universities to validate scores and assessment methods. No information was found on the validation of assessments' results.

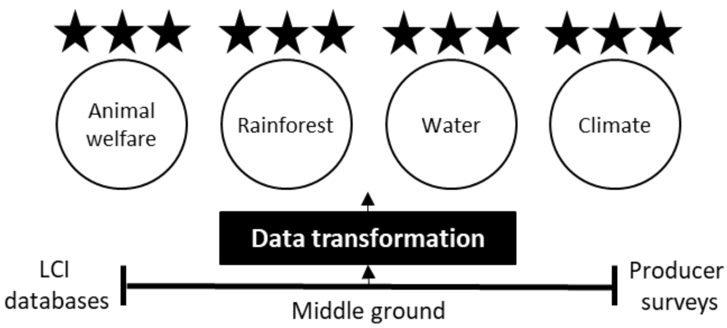


Figure 8: Conceptual drawing of the built of Tatanka's environmental footprints

3.3.2. Gaia

Company description and focus

Gaia is an independent foundation that provides consumers with practical and reliable information on environmental questions. One of their projects is a smartphone application that shows the environmental footprints of fruits and vegetables. These footprints go beyond GHG emissions only. I spoke with a senior advisor who represented the organization. The representative explained that the project's aim is to make sustainable food choices as easy as possible for consumers.

View on sustainability and socio-economic dimension

Gaia's representative emphasised that the foundation mainly aims to provide practical sustainability advice on daily matters. For example, Gaia offers practical consumer advice on environmentally friendly food choices, the choice of least polluting modes of transport to use in daily life and how to deal with household waste sustainably. The interviewee explained that besides the practical application of sustainability, the socio-economic side of sustainability is also important. However, Gaia does not include socio-economic side of sustainability in their approach:

> *"By solely focussing on the environmental side we don't deny the importance of other aspects. But, we are the experts in this specific field. That is why we emphasise that we only provide environmental scores. Of course, there are other aspects, but for these, other organizations exist with the expertise needed to assess these. Expertise that we have to a lesser extent."*

Indicators

Gaia's application is built on the ReCiPe 2.0 method – a method for the impact assessment in LCAs. The impact assessment is done by the conversion of emissions of pollutants and resource depletions through so-called characterisation factors (Dekker et al., 2019). The ReCiPe 2.0 method derives these characterisation factors on two levels, namely the midpoint level that includes categories like fossil resources, global warming and ionizing radiation; and, the endpoint level that includes human health, damage to ecosystems, and damage to resource availability (Huijbregts et al., 2016).

Gaia's application uses four indictors to assess vegetables and fruits by means of the ReCiPe 2.0 method. These four indicators are water stress, climate change contribution, land use and fossil resource use. Huijbregts et al. (2016) elaborated on what each of these indicators consist of in their publication on ReCiPe 2.0. In short:

- Water stress is estimated considering water use and water scarcity in a production area by means of the water stress index. When data allows, the water scarcity is approximated on regional level, if not, it is assessed on a national level;
- The climate change estimation is based on GHG emissions in a production cycle;
- Land use is characterized by relative species loss due to local land use – assuming that an area would be in a natural state if land use did not occur; and
- Fossil resource use is based on the extraction of fossil fuels.

The ReCiPe 2.0 method includes many more indicators (or midpoint categories). However, these indicators are not included in Gaia's platform. I did not identify the motivations for exclusion of these indicators during data collection, except for the exclusion of pesticide use. Gaia's representative explained that pesticide use is excluded from their platform due to the lack of reliable data.

System boundaries

Gaia's application includes the following life cycle stages: cultivation, packaging, transport. In terms of packaging, impact estimates are based on the packaging material used most frequently for a certain product. Concerning transport, emissions are estimated by combining the distance from the country of produce and the mode of transport.

Farming systems, local produce and seasonality

The distinction between different farming systems like organic and conventional farming is not made in Gaia's platform. Also local produce is not explicitly considered, although local considerations are still partially included in Gaia's platforms since water scarcity and transport distances are considered. In terms of seasonality, every vegetable and fruit its environmental score is determined by the month it is consumed.

Data collection and trustworthiness

The development of the assessment method was outsourced to a consultancy firm. This firm makes use of the ReCiPe 2.0 method to compile the environmental footprints. No other information on data collection or the trustworthiness of data was mentioned.

Data aggregation

Gaia's environmental footprints are presented in a single 'traffic light' score that includes five levels, ranking from 'A' which stands for good choice to 'E' meaning that the product should be avoided. These footprints are an aggregation of the underlying impact categories. The representative explained that the rationale for this aggregation is *"to make it as easy as possible for consumers, so they can make a choice in the blink of an eye."*

Figure 9 summarizes the data collection (black box), indicators used (middle circles) and final aggregation (upper circle) of Gaia's fruits and vegetables application.

Validation

Gaia has a scientific board of advisors that validates the methods used. In addition, the external party that oversaw the platform's technical development has its own expert validation methods. I did not find information on the validation of assessments' results.

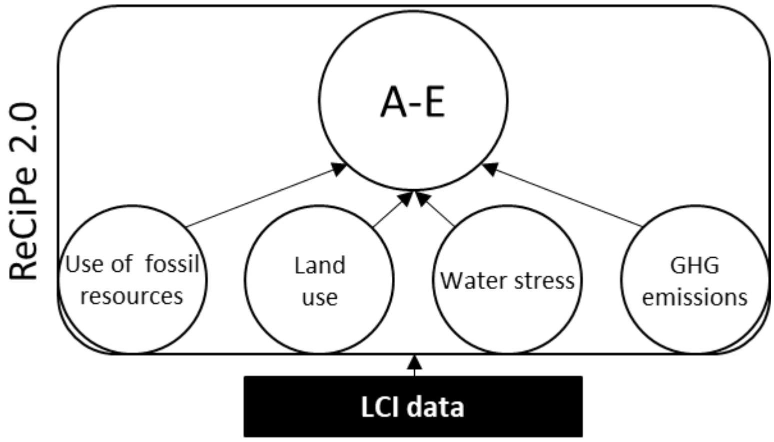


Figure 9: Conceptual drawing of the aggregation of Gaia's environmental score

3.3.3. Endymion

Company description and focus

Endymion is in its essence a longitudinal European consumer study. The research project aims to facilitate a new form of consumerism, which involves informed decision making for the individual consumer. I interviewed the coordinator of Endymion on their research project. The representative emphasised that the project was only partially focussing on the provisioning of sustainability information, since this was needed to fulfil a larger goal. This larger goal was a research hypothesis. namely, to confirm if consumers behave differently once they have access to sustainability scores for food products:

> *"What we did is not to compare standards, in a sense of rates, to each other. Our contribution is whether, once you have them, once they are developed, does it make sense to have them? Does it make sense to have another dispute on standards? Does it make sense to have… to let's say invest in another 5 years of 10 people research work on soil quality while in the end nobody will care, and it will not help anybody? And in order to address this specification… so what if… we had better quality? What if – we can make a statement…. do people change? That's what we address and that's the impact."*

During the interview, the representative emphasised multiple times that the presentation of sustainability data was merely a prerequisite to answer their research hypothesis:

> *"In the end our hypothesis was that "if users have a better view on the products, how do they change their behaviour." What we in the end did not tackle is whether what you provide them is fully trustworthy and objective, but let's assume it is trustworthy and objective… we had to provide at least a useable level of information. To be able, or to continue with the whole project, to address the hypothesis that we were aiming for and this was just the prerequisite for this. I understand that in the meantime there are other projects who dedicate millions with a large consortium to dedicate years of research on these prerequisites… for us it was just prerequisites. And we didn't even have so much money. So, the best was a trade-off."*

The sustainability data was presented to the supermarket customers through a smartphone application. This application collected data on consumer behaviour. The products included in the application were those that had a package and barcode, so, fresh produce was excluded. The field tests lasted for a few months. However, consumers that participated in the field experiment were tracked for two years beforehand so that actual behavioural change could be observed.

View on sustainability and consumer behaviour

Although the collected consumer behaviour data is still analysed, the representative could confirm that people behaved more sustainably. Important here is the project's definition of sustainability, as the representative explained:

> *"You could argue what is more sustainable, but there is a certain, definition and setting we took from various backgrounds... According to this people behaved more sustainable than before… [The application itself] did not include a "standard setting of sustainable." So, sustainable is in the eye of the beholder. You say what it is. And once you set what it is for yourself, then we go along with this. So that's kind of one of the main points on sustainability, because it's always a matter of how far you reach out for the context."*

Endymion included 25 sustainability criteria in their application to provide consumers the possibility to interpret sustainability in their own way. Consumers themselves could then select which criteria were important for him/her. In doing so, consumers had the possibility to construct their own version of what sustainability is. The representative explained that this approach towards sustainability was not always received wholeheartedly: *"I had repeated discussions with representatives of the United Nations and they were always saying 'you have to tell people what is sustainable" and I always said "you don't have to, [.] rather not."'* Endymion's vision on sustainability was underscored after concluding their field experiment: *"I am rather lucky that in the end this hypothesis turned out to be correct to some extent."* The representative could confirm that change in consumer behaviour was observed to a certain degree, although not much information could be disclosed since the data collected in the project was still being analysed.

Socio-economic dimension

Endymion's project did not standardize methods to assess the economic impacts of food produce. However, social criteria were part of Endymion's platform. The social criteria were included as part of already existing labels such as the 'fair trade' label.

Indicators

Endymion's platform covered 25 sustainability criteria that were clustered under four main categories, being: environmental, social, health and quality. The environmental cluster included aspects like travelling distances of food products, pesticides, and disposability of packaging. Examples of the social dimension criteria were fair trade, working conditions, production transparency and animal rights. Health included additives; protein, sugar and fat contents; ordinary properties; lactose free or not; and whether a product was vegetarian or vegan. A criterion for quality was freshness, among other things. The representative explained that these main categories were heavily limited by data availability and the time of development:

> *"Back then… I mean now things are different, actually, I am kind of, let's say trembling or negotiating with myself whether to reinitiate such a project again. Because both publicity, results and data quality have significantly increased. So, I mean, you might see, there are others now that strongly go into that direction. [But] we have a project that already finished this with a social science study, and it started four years ago. Which means, it was quite pioneering, and consequently, the results have to be considered on the pioneering stage. Given the situation that you have now, it is not fair to compare us, from 2015, with a project from 2020 – there is a huge gap. So is the data quality, data availability, interfacing, and motivation of retailers to participate."*

Data collection, trustworthiness, and disputability

Endymion's representative saw the platform as a communication tool rather than an assessment tool: *"we do not label products; we do not obtain or certify anything. We take information that is there, more or less… so the things are only as good as [what is there]."* So, Endymion's platform compiled existing data and based the sustainability categories on these. Data sources included existing labels, data from online stores, crowdsourced databases, open source databases, product data digitized by online retailers, commercial product databases and feedback from an advisory board that included amongst others multiple NGOs. Endymion's representative elaborated that this was not ideal, but that it was the way to go at that time:

> *"We classified data on highly trustable and low trustable versions and those low trustable versions ended up less in the calculation scheme. The highly trustable entered to a larger extend… At some point we had to allow trade-offs to say well in certain fields we go as well with the low value, low trustable data and in some cases we did not, just in order to make sure that we could still raise a point. But, of course, you can negotiate between, you can*

always claim that the numbers behind the comma, or the three numbers behind the comma are rounded or whatever. And it is all subjective to some extent."

The level of subjectivity in data sources was a recurring topic during the interview. The representative explained that to the core, every indicator is somewhat subjective, and therefore disputable. The disputability of indicators resulted in difficulties with the inclusion of certain criteria due to internal disagreement:

"[Even for rating products vegetarian or vegan], we had high disputes within the consortium because we had a consumer organization involved and they for example were reluctant to say that an apple is vegan, because you don't know what chemicals or what processing is involved. [During] harvesting, or even in the production process. And even such argumentation like 'well let's at least say it is 95% vegan', they were reluctant to approve [.] Difficult. So, you can only rely on the products that call themselves vegan. And then… yeah, so, as I said, we did this to some extent. Again, we had to accept the trade-offs. But in the end, also here you would say we are 95% sure that what we are doing is appropriate, and in some cases, it is common sense. But common sense is heavily disputable. What is common sense?"

Organic, local and seasonal produce

Organic produce is considered in Endymion's platform through existing labels. The representative explained that the legal requirements of organic labels differ per country of produce and it would be interesting to include these differences in a consumer FSSAP. However, that requires manpower, which was lacking in Endymion's project. Regarding seasonal produce, I did not gain any insights during the interview and analysis of documents. In terms of local produce, travel distances were considered in Endymion's platform.

Data aggregation

Endymion's application included both a totally aggregated score and scores from the underlying criteria, each score ranking from 1-10. Every underlying sustainability criterion was rated on importance by the consumer before doing groceries. This 'importance' rating and the 'actual' sustainability score of the corresponding criterion were then combined to a totally aggregated sustainability score. By combining consumer preferences and sustainability scores of products, the consumers themselves could partially decide how sustainable a product was.

Figure 10 summarizes the data collection and consumer inputs (bottom), the rated sustainability criteria (middle circles) and the totally aggregated sustainability score (top circle) that were part of Endymion's platform at the time it was still running.

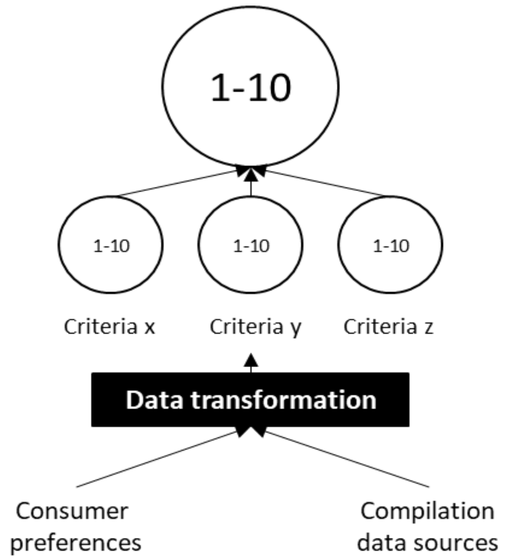


Figure 10: Conceptual drawing of the aggregation of Endymion's sustainability score(s)

Validation

The assessment methodology was screened by an expert panel. I did not find information on the validation of assessments' results.

3.4. Synthesis: the ideal FSSA, edible oil FSSAs and consumer FSSAPs

I outlined an ideal FSSA in Section 3.1. Subsequently, I compared both edible oil FSSAs and consumer FSSAPs (Section 3.2 and 3.3, respectively) with this ideal FSSA – both the edible oil FSSAs and consumer FSSAPs all differed substantially from the ideal FSSA. The most remarkable differences are described in this section.

Edible oil FSSAs seldomly included effects on socio-economic conditions. Especially, socio-economic data on a local level were often lacking. A logical explanation for this deficiency is the scope of most studies reviewed. They were LCAs identified overall environmental impacts. Similar deficiencies were diagnosed for consumer FSSAPs. Socio-economic factors were seldomly included apart from some labels (e.g. fair trade).

Ideally, FSSAs consider the effects of production systems on site specific aspects. However, I found that edible oil FSSAs often neglect these specific spatial units of analysis. Consumer FSSAPs also rarely considered these more local units of analysis apart from water stress adjusted to regional water scarcity and transport distance.

With regards to the four main environmental spheres, the atmosphere and hydrosphere were mostly considered in edible oil FSSAs. The biosphere, rhizosphere and geosphere were not. Biodiversity loss was rarely included (2 out of 33 papers) due to the inability of methods to measure it (De Gennaro et al., 2012). The same holds for soil erosion and soil fertility (0 out of 33 papers). Similar results were found in case of consumer FSSAPs: the representatives of Tatanka and Gaia explained that their platforms could not deal with the effects of food production on soil quality and biodiversity because of the scientific lack of consensus.

Ideally, system boundaries should include all phases from cradle to grave or farm to fork (Andersson, 2000). In edible oil FSSAs, farm-to-fork analysis were seldom conducted. The majority of the LCAs reviewed by me only focussed on the cultivation (31 out of 33) and manufacturing phases (20), and some on the packaging phase (10). Other phases like warehousing (2) and final disposal (2) of products were mostly not included by authors. Consumer FSSAPs differed from each other in this respect: Tatanka included farm-to-fork boundaries; Gaia included cultivation, packaging and transportation; and Endymion did not make use of LCA-based assessments.

Different farming systems are considered in the ideal FSSA (e.g. conventional and organic cultivation). However, different farming system were not always considered in edible oil FSSAs – only half of the LCA cases reviewed by me reported the analysed farming system. Consumer FSSAPs differed with this regard: Tatanka differentiates between organic and non-organic systems in its calculations, so did Endymion by using organic product labels. Interestingly, the representatives of both Tatanka and Endymion raised questions if differentiating between organic and non-organic contributes to the accuracy of assessments. Tatanka's representative explained that focussing on organic produce might convey the wrong message: rather than focussing on eating an organic product or not, a consumer should focus on eating the right products instead. Endymion's representative explained that distinguishing between organic and non-organic produce is difficult due to different national regulations and labelling requirements.

The ideal FSSA requires scientific evaluation of results. In case of the reviewed edible oil LCAs, results were evaluated through contribution analysis by the majority, but not by all studies (21 out of 33), roughly one third did a sensitivity analysis (13) and only very few did an uncertainty analysis (5). With regards to consumer FSSAPs, the methods were validated through scientific panels in every consumer FSSAPs analysed. However, no information was found on the evaluation of their results.

Discussion

4. Discussion

The discussion chapter puts my findings in comparison with other publications. Subsequently, I raise some thoughts for contemplation with regards to the use of consumer FSSAPs. As a wrap-up, I outline further avenues for research and the limitations of my thesis in this chapter.

4.1. Literature review: parallels and differences with other authors

The body of edible oil literature is thin regarding two aspects. First, I observed that not many edible oil LCAs exist, especially in case of rape-, sunflower- and peanut-oil. The small number of published LCA cases is also observed for agricultural products other than edible oils (Poore & Nemecek, 2018). Second, most studies I sampled focussed on biofuel systems rather than edible oil systems. This was also found in other publications (e.g. Muñoz et al., 2014; Williams et al., 2010).

Additionally, I identified the main methodological issues in the edible oil FSSA literature. Many of my findings were in line with earlier publications. First, I confirmed that little is known about the socio-economic sustainability of 1) edible oil production and 2) agri-food LCAs in general (Fleskens et al., 2009; Jeswani et al., 2010; Menozzi, 2014). This is particularly undesirable because socio-economic impacts are considered highly relevant by many consumers (Notarnicola et al., 2017).

Second, my review confirmed the strong incline towards the agricultural and manufacturing phases in edible oil LCAs. An explanation for this inclination is that most of the GHGs are emitted in the agricultural and manufacturing phases (Khanali et al., 2018; Khatri & Jain, 2017). This inclination varies somewhat in LCAs on other agri-food products. Studies on corn for example, mainly focus on the agricultural phase; studies on rice largely vary in their system boundaries; and studies on products derived from wheat mostly incorporate all stages of production and sometimes even waste disposal (Renzulli et al., 2015). In these examples on different cereals, the most applied cut-off point was still cradle-to-farm-gate regardless of the different system boundaries (Renzulli et al., 2015). Excluding production phases could be justified if it does not substantially change the overall conclusion of a study (Salomone et al., 2015). However, the implications of excluding production phases were rarely mentioned in the studies I reviewed. Consequently, these LCAs could lead to misinterpretations and inaccurate conclusions by third parties.

Third, I also confirmed the absence of temporal system boundaries in most edible oil LCAs (Khatri & Jain, 2017). In other agri-food LCAs, temporal system boundaries are also applied haphazardly (Notarnicola et al., 2017). The exclusion of temporal boundaries might be of lesser importance with non-food products LCAs; but, in agri-food LCAs they can determine studies' results due to seasonality. Therefore, LCA results of exactly the same product could vary within a single year (Notarnicola et al., 2017) – emphasizing why the exclusion of temporal boundaries should not be the standard.

Fourth, my findings confirmed the large heterogeneity in allocation methods reported in 1) edible oil LCAs and 2) other agri-food LCAs (Khatri & Jain, 2017; Poore & Nemecek, 2018). Ultimately, using a different allocation method will lead to different results and conclusions. Therefore, the standardization of allocation criteria for specific agri-food products likely helps to improve interpretation and comparability of different studies' results (Notarnicola et al., 2017).

Fifth, the neglection of farming systems in edible oil LCAs found in other reviews (e.g. Basset-Mens et al., 2010; Salomone et al., 2015), was also confirmed by my literature review. This issue also occurred in LCAs on other plant-based food systems. Renzulli et al. (2015) reported that farming systems were oftentimes erroneously assumed in the LCA literature on cereals; and Cerutti et al. (2015) found that the farming system type was frequently left out in the LCA literature on fruits.

However, these findings must be interpreted carefully. By including standardized inventory sets for each farming system type, inaccuracies can be introduced to FSSA's results. This is because a specific farming system type (e.g. organic or conventional farming) involves a wide spectrum of practices. The application of these practices differ from farm to farm, even though they might all belong to the same farming system type (Notarnicola et al., 2017). Therefore, agri-food LCAs are more accurate when all underlying farming practices rather than predefined 'sets' of farming system types are included.

Sixth, my literature review confirmed the scarce application of evaluation methods like contribution-, sensitivity- and uncertainty-analyses (Khatri & Jain, 2017; Petit et al., 2018). Similarly, evaluation methods were not the standard within LCAs on wine and cereal systems (Petti et al., 2015; Renzulli et al., 2015). The use of uncertainty and sensitivity analysis is essential for LCAs to filter out the results' dependence on methodological choices and uncertain parameters (Roma et al., 2015). This is especially the case in literature bodies with large methodological heterogeneity – such as the edible oil literature.

Seventh, in line with Basset-Mens et al. (2010), my results showed that the lack of specific methods and data to estimate field N_2O emissions indeed exists. The consequences of this issue are large. The method applied most frequently is the IPCC method (see Eggleston et al., 2006 for explanation). The IPCC method is criticized for having a too large uncertainty range for field emissions and for not being robust, especially in the lowest tier(s) (Figueiredo et al., 2017). This uncertainty can be so large that it dominates uncertainty in edible oil LCAs (Figueiredo et al., 2017; Williams et al., 2010). This issue is also present in other agri-food LCAs; for example, a large range of variation was found in estimated field emissions in cereal and livestock system LCAs (de Vries & de Boer, 2010; Renzulli et al., 2015).

Eighth, my literature review also confirms the conclusion by Banias et al. (2017) that the use of pesticides and fertilizer is generally included in olive oil LCA inventories. An important remark is that these inventory items are often only estimated in agri-food LCAs, due to a lack of available data on production and diffusion of fertilizers and pesticides (Notarnicola et al., 2015). The choice of fertilizer and pesticide type has strong implications on environmental impacts of a production system (Queiros et al., 2015). As such, the inclusion of these inventory items can introduce additional uncertainty in studies' results.

Last, the findings in my literature review were mostly in line with the review on olive oil by Banias et al. (2017) and the review on multiple edible oils by Khatri and Jain (2017) except for two small aspects. The first aspect is the inclusion of fuel use in FSSAs. Banias et al. (2017) claimed that assessments of olive oil production chains only marginally included fossil fuel use. I found the opposite in my review: 29 of the 33 articles examined included fossil fuel use. The difference in fuel use can be explained by 1) the edible oils included – my review focussed on four vegetable oils whereas Banias et al. (2017) focussed solely on olive oil; and 2) the scope of both reviews: Banias et al. (2017) included multiple assessment tools in their review whereas my literature review only focussed on LCAs. The second aspect is a small addition to the study of Khatri and Jain (2017). My addition is the inclusion of abiotic depletion as an impact category. This was ignored in the study of Khatri and Jain (2017). My results showed that abiotic depletion was assessed in 14 out of 33 edible oil LCAs reviewed.

4.2. Neglection of essential impact categories

A recurring topic in both edible oil FSSAs and consumer FSSAPs was the inability of LCAs to deal with multiple essential impact categories, namely: land use change, changes in soil quality and biodiversity loss. Land use change is seen as one of the most substantial impact categories in agricultural production (Mattsson et al., 2000). Land use change even has the potential to dominate GHG emissions in LCA results (Schmidt, 2015). However, the GHG emissions induced by land use change are often not accounted for in LCAs; since the link between land use and deforestation is not well described, nor is there a scientific consensus on how to establish this link (Schmidt, 2015). The potentially large impact of land use change combined with the difficulty to assess the impact category introduces high uncertainty to LCA results (Figueiredo et al., 2017). In terms of the evaluated consumer FSSAPs, I found that one platform assessed land use change by means of LCI data, the second platform included land use change by using labels, and the third did not include land use change at all. Also, only 16 out of the 33 edible oil LCAs reviewed by me included land use change. I should note that this number of studies is somewhat biased, as my sample included a high amount of olive oil publications (16). Olive trees are known for living very long (> 100 years); hence, the emissions from land preparation in the past are often deemed insignificant (e.g. El Hanandeh & Gharaibeh, 2016).

In addition, decreased soil quality can be seen as a substantial impact on the long-term sustainability of a food production system (De Luca et al., 2018). Indeed, soil quality and its underlying midpoint indicators (e.g. change in SOM, erosion, salinization, etc.) play a central role in agriculture (Garrigues et al., 2012). As such, the inclusion of impacts on soil quality is essential in food system LCAs. However, this rarely happens because of multiple reasons (Salomone et al., 2015). First, soil quality inventory items are site-specific and therefore locally unique (Garrigues et al., 2012). Second, soil characteristics vary spatially and temporarily (Reap et al., 2008). Third, the soil system includes many complex interactions (Garrigues et al., 2012). The frequent omission of changes in soil quality as an impact category is confirmed in my thesis: none of the consumer FSSAPs assessed changes in soil quality; and only 1 out of the 33 edible oil LCAs reviewed included a form of changes in soil quality as an impact category. In addition, soils were rarely included in the LCA inventories I reviewed: only seven papers considered soil types, and only one paper applied an actual soil analysis. This large neglection of soil quality and soil characteristics in FSSAs could lead to management and governance that jeopardizes food production on the long-term due to soil degradation (Notarnicola et al., 2017).

Also, direct biodiversity loss can only be partially assessed by current methodologies (De Gennaro et al., 2012). This was clearly visible in the results of this thesis: none of the consumer FSSAPs included biodiversity loss; and only 2 out of the 33 edible oil LCAs reviewed included biodiversity loss as an impact category (see Mattsson et al., 2000; Schmidt, 2010). The implication of this large neglection is that LCAs can be seriously misinterpreted since decision makers do not take this important impact category into account. As such, this can lead to support for less favourable policies with regards to sustainability (Notarnicola et al., 2017); or in the case of consumer FSSAPs, it can lead to misinformed consumer choices. This possible misinterpretation becomes especially evident when organic and conventional agriculture are compared. Conventional agriculture is often seen as favourable in food LCAs because of its lower impact per unit of produce. However, impacts on soil quality and biodiversity -on which organic agriculture has a more positive impact- are mostly not considered in these LCAs (Notarnicola et al., 2017).

4.3. Recurring assessment issues

In my literature review, I found several additional aspects that other authors also deemed important for FSSAs, but I did not include them in my results. This because I only learned about them towards the end of my analysis.

First, multiple studies highlighted that the nursery and planting of cultivars are seldom included in the system boundaries of both edible oil LCAs and LCAs on other agri-food products, despite the potential impacts induced in these phases (e.g. De Gennaro et al., 2012; Petti et al., 2015; Salomone et al., 2015). Second, food system LCAs struggle with the definition and use of functional units (Salomone et al., 2015). A simple example is the comparison of impacts induced by olive oil and virgin olive oil production. This might sound like logical comparison to make. However, 1L of olive oil as a FU is not comparable to 1L of virgin olive oil because these oils differ a lot in terms of quality properties (Salomone et al., 2015). Problems with the definition of FUs have a large consequence: carrying out a LCA on a product system with different functional units can result in different or even contrasting LCA results (Notarnicola et al., 2015). Third, differing soil and residue management strategies included in LCA inventories substantially affect environmental emissions of the overarching farming systems. For example, using pruned residues of olive trees as fertilizer results in a large environmental gain in GWP (Iraldo et al., 2014). However, this example cannot be generalized to other edible oil systems since residue management strategies are not fully evaluated for each oil production system (Mousavi-Avval et al., 2017). As a result, management implications cannot always be derived. Fourth, the inclusion of local practices, microclimates and physical elements in FSSAs was deemed important by several authors (e.g. Bélanger et al., 2012; Fleskens et al., 2009; Petti et al., 2015). Site specific conditions can partially determine the environmental impacts of (and on) a production system; hence, these should always be part of an assessment (Notarnicola et al., 2015). However, the LCA framework is known for not making any distinction between spatial units of analysis, even if an impact is a regional or a local one (Jeswani et al., 2010). Nonetheless, the exclusion of these site-specific aspects probably leads to misinterpretation of LCA results and can therefore not be simply ignored (Notarnicola et al., 2017).

4.4. Consumer FSSAPs: translating sustainability

The three consumer FSSAPs I evaluated show similarities, but also large differences (see Table 2). The three organizations' aims are almost identical, they all focus on assisting consumers in making informed sustainable food choices. Although Gaia has its own accent in this aim: to make a choice as easy as possible. Despite these similar aims, the organizations' approaches towards sustainability assessments are less aligned. The approaches of Tatanka and Gaia are both rooted in LCA thinking, in which sustainability is measurable and could be communicated transparently. Endymion differs with this regard: in their opinion, sustainability is as a social construction that is in the eye of the beholder. So, the only social institution that can present a sustainability score to the end consumer — is the end consumer himself. This notion exemplifies a main finding of my thesis, namely that an organization's view on sustainability translates throughout their platforms all the way to the end consumer.

This can also be observed in the organizations' varying views on data aggregation. Endymion's totally aggregated sustainability scores are constructed by consumers. Endymion's platform did this by combining: 1) the importance of sustainability aspects ranked by an individual consumer, and 2) sustainability data collected from various sources like databases and existing labels. Interestingly, the other two more LCA rooted organizations also differ with their view on data aggregation. Gaia aggregates all data and impact categories in a single end score in order to be as practical as possible, whereas Tatanka's representative rejected the idea of aggregating environmental footprints in one sustainability score. Instead, Tatanka presents their environmental scores separately, one score for each impact category. The differences in the presentation and aggregation of environmental scores by each consumer FSSAP potentially has large influence on the final decision making of consumers.

In addition, the three organizations all made use of different methods to collect data. The methods used are all firmly rooted in science but differ in levels of preciseness. The representatives admitted themselves that their data collection methods come with flaws. The organizations are facing similar difficulties as the ones I diagnosed in the academic literature on edible oil FSSAs. These similarities include difficulties dealing with the social and economic sides of sustainability; the inclusion of important indicators like pesticide use; the handling of essential impact categories such as soil erosion, land use change and biodiversity loss; data availability; and the inclusion of different production systems. Noteworthy is that the consumer FSSAPs are dealing with these flaws in creative ways. One way was simply to be transparent about missing indicators; for example, Gaia's desire to include pesticides in their platform. Another way was the use of existing sustainability labels (e.g. fair trade) in cases were impact categories could not be directly assessed. This was done by Tatanka and Endymion. Additionally, Tatanka puts impact categories that lack in their platform in perspective; for example, the organization's main focus on GHG emissions was justified by its view that indirect consequences of climate change will outnumber many other 'smaller' impacts of food production. Last, Tatanka justified the exclusion of land use and eutrophication by its view that a decrease of carbon footprints most likely concurs with the reduction of land use and eutrophication potential.

4.5. Total aggregation of sustainability scores: go or no-go?

A strong claim was made in this thesis's problem statement (p. 4):

> *"More clarity on the wide range of negative production impacts is urgently needed… Transparency on how these impacts are aggregated in current consumer applications and, ultimately, a totally aggregated sustainability score that combines both this clarity and transparency in an easily understandable way, must be developed."*

I based this claim on findings presented by Evocco (2019) and Kalnikaitė et al. (2013). Evocco (2019) found that consumers perceived tangible factors (e.g. waste reduction) as the most impacting aspects of the food production chain, even though science has proven this is not the case. In addition, Kalnikaitė et al. (2013) found that consumers only have a very limited decision making process in a supermarket. As such, overloading consumers with sustainability information would be futile. I deduced a possible solution from these findings – totally aggregated sustainability scores for each supermarket product based on expert insights.

However, the results presented in my thesis undermine the claim I made in my problem statement to a certain extent. Only one of the consumer FSSAPs (Gaia) evaluated in this thesis included a single totally aggregated sustainability score that was solely based on expert knowledge. The flaws of this method became obvious during the evaluation of the platform. For example, Gaia's aggregation included land use, an indicator that is somewhat disputed in the academic literature. However, due to the total aggregation it was impossible to trace how this indicator affected the final totally aggregated score. As such, a 'very good' product score probably conceals large ranges of uncertainty in the underlying data layers. Furthermore, Endymion also aggregated sustainability scores as part of their research project, but these scores were weighted by consumers. The draft findings of Endymion's research project show that the 'social construction' of sustainability scores did motivate consumers to behave more sustainably. Or, at least more sustainably according to the consumers' own social construction of sustainability. The danger here is that consumers likely think that they consume sustainably but only make minor improvements in terms of avoided environmental impacts. So, this type of aggregated sustainability scores can form biased or even wrong ideas regarding 'sustainable' products. This possible bias was also emphasized by Tatanka's representative, who disputed aggregated scores because these can become politicized depending on *who* aggregates them.

Academics also vary in their view on total aggregation. Schmidt (2009), for example, is strongly against it. He argues that creating an aggregated one-dimensional ‘environmental indicator’ is impossible due to the variety of material and non-material interactions in a production system. In addition, he added that reliable consumer purchase decisions cannot be made based upon aggregated scores due to the accumulated uncertainty ranges. At the other hand, complete bodies of literature attempt to develop aggregated sustainability scores for food products – examples include composite indicators (see Gómez-Limón & Sanchez-Fernandez, 2010) and LCA-based endpoint impact categories (see Huijbregts et al., 2016). Although arguably, these authors aimed to develop aggregated sustainability scores for decisionmakers rather than consumers. Therefore, the question arises: can reliable sustainability decisions based on total aggregated scores be made by consumers without any background knowledge? Sargant (2014) doubts this: consumers weigh up sustainable aspects against other considerations that are context dependent. As such, only giving consumers the responsibility for sustainable consumption through total aggregated sustainability scores is not fair. A shared responsibility between consumers and actors higher upon the retail chain might be more effective and ethical.

4.6. Future avenues of research

To my knowledge, my thesis is the second contribution in which an overview is given of methodological issues regarding multiple edible oil LCAs (after Khatri & Jain, 2017) and the third in case of olive oil FSSAs (after Banias et al., 2017; Salomone et al., 2015). Additionally, my thesis is the first research that evaluated and compared multiple consumer FSSAPs.

First, the results of these analyses raised an interesting thought for contemplation. The main issues in food system LCA methodologies are simultaneously the ones that supposedly assess aspects that cause large environmental impacts. Examples include the inability of LCAs to deal with essential impact categories like biodiversity loss, changes in soil quality and land use change; the lack of accurate N_2O field emission models; and, the lack of data on the production and dispersion of pesticides and fertilizers. Further research on these issues, generation of data and methods, and scientific consensus, are essential to adequately address several main environmental impacts and to reduce the large extent of uncertainty present in LCAs. Second, the issue of integrating socio-, economic- and environmental-sustainability is rarely addressed in edible oil LCAs. Future research could expand on this gap. Third, edible oil LCAs seldom include LCA phases beyond the agricultural and manufacturing phases. Further research on phases like nursery, planting, warehousing and disposal could confirm if ignoring these phases is justified. Fourth, local practices, physical elements and microclimates partially determine the impact of (and on) production systems. However, LCAs have difficulties to deal with these local factors. The development of LCAs that can deal with site-specific elements could tackle this issue. Fifth, the body of edible rapeseed-, sunflower- and peanut-oil LCAs is very thin. Therefore, more LCAs on these edible oils are desirable for comparability’s sake. These LCAs should then focus on the inclusion of aforementioned aspects that are now often lacking (see point 2, 3 and 4). If authors still wish to leave out these and other aspects, they should be clearer about the implications of not including them. Sixth, consumer FSSAPs had difficulties with assessing different farming systems. Renewing methods to address different farming systems and the underlying practices could help to improve accuracy of assessments. Last, the FSSAPs reviewed presented their sustainability scores and underlying data aggregation in different ways. Each method was built on a vision that substantially influences the decisions of end-users. However, the actual preferences of these end-users were not highlighted. Future research could explore consumer preferences on the presentation of sustainability scores.

4.7. Limitations

I should note that my thesis unintendedly focussed on two subjects, namely olive oil and LCAs. These two focusses are not a limitation by itself, but as a reader you probably wonder how these came to be. When I started this project, I wanted to compare olive-, rapeseed-, sunflower- and peanut-oil by using academic studies selected through predefined search terms. However, the academic literature did not include any publications on rapeseed-, sunflower- and peanut-oils apart from LCA-based publications. As an outcome, most of the publications used in my literature review were focussing on olive oil. Consequently, Section 3.2 has a very strong olive oil bias. Additionally, the limited body of literature on rapeseed-, sunflower- and peanut-oils resulted in an unintended focus on LCAs. Also, most of the organizations I planned to interview made use of LCA-based methodologies. Hence, the focus of my thesis gradually shifted from the 'Sustainability Indicator' literature –on which the ideal FSSA is primarily based—towards LCA-based literature. I tried to align this 'renewed' focus on LCAs by adapting the ideal FSSA towards key aspects that both methods incorporate.

My thesis is also subject to multiple limitations of a methodological nature. First, the evaluation sheet I used to review edible oil FSSAs was incomplete. The sheet was based on publications by Khatri and Jain (2017) and Banias et al. (2017) appended with some additional items found in the literature. However, while reviewing the edible oil literature by means of the evaluation sheet, it became clear that some additional aspects were also deemed highly important by other authors. Some examples of these aspects are the ambiguity around functional units, some missing LCA inventory items and the neglection of nursery and planting phases in LCA system boundaries. By not incorporating these items into my evaluation sheet, I was not able to assess the extent of these issues in the FSSA literature. If these and other aspects would have been included in my evaluation sheet, this might have guided my 1) consumer FSSAPs interviews; 2) results; and 3) conclusions in a different direction. Based on the results presented by Khatri and Jain (2017) and Schmidt (2004), I believe that I could have had some additional main findings if I evaluated the implications of 1) the ambiguous use functional units; and 2) the traditional attributional LCA approach rather than consequential LCA approach.

Second, an issue with system boundaries arose during my review of edible oil LCAs. Transport and distribution were originally included during data collection. However, transport and distribution were reported in various ways because they take place between multiple phases of product chains. I did not consider this distinction between multiple transport 'moments' when I commenced with my literature review. As such, distinguishing between these different moments of transport became increasingly difficult overtime. Consequently, I omitted transport and distribution from my results, making my evaluation less complete. This could have been avoided if I subdivided different moments of transport from the start rather than putting everything under the same header.

Third and foremost, my analysis of consumer FSSAPs could have been deeper. The methods I chose were document analyses and interviews. These methods were appropriate for the exploratory nature of this research. However, this also limited my analysis. For example, I gained limited insights on how data was collected by consumer FSSAPs. Therefore, a more in-depth analysis in which I would have had access to the consumer organizations' methodologies could have resulted in a better understanding of the organizations' assessment frameworks. However, this type of in-depth analysis was hindered by time and access.

Fourth, the use of interviews might have potentially introduced some form of subjectivity to the findings of my thesis because the used interview quotes represent the interviewees’ view rather than the organizations’ view. I aimed to reduce this subjectivity as much as possible by triangulating the quotes through documents published by the organizations. Conducting interviews with other company stakeholders or an organizational ethnography could have reduced this subjectivity even more. Again, the use of more in-depth methods was hindered by time and access limitations.

Last, I only analysed a limited amount of consumer platforms. The notion of consumer FSSAPs is rather new and as such not many exist. I listed all consumer platforms I could find before approaching them. This list consisted of merely nine organizations, in the end three replied to my e-mail. However, this is only a small limitation since this thesis project has a more explorative nature, in which a reasonable novel concept was analysed. In the end, the three organizations analysed differed substantially – nicely depicting the wide spectrum of choices and issues organizations must deal with.

Conclusion

5. Conclusions

This thesis commenced by briefly touching upon the impacts of the global food system and the effects of consumer choices. Subsequently, this thesis narrowed to the theory behind the assessment of food systems. In doing so, three research questions were introduced, which were analysed by the development of an ideal Food System Sustainability Assessment (FSSA), a review of the literature on edible oil FSSAs and an analysis of three consumer Food System Sustainability Assessment Platforms (FSSAPs). In this chapter I will conclude this thesis by answering these research questions.

RQ1: What would a complete and accurate (i.e. ideal) FSSA look like?

This thesis outlined an ideal FSSA. This ideal FSSA was based on the bodies of literature on Sustainability Indicators and Life Cycle Assessment (LCA). By doing so, I found that an ideal FSSA does not exist, although conceptually some conclusions could be drawn. An ideal FSSA would be holistic and include all sustainability dimensions; all earth spheres; complete production chains; a condensed set of indicators that covers both site specific and globally generic factors; a form of time integration; local stakeholders, external experts and researchers; and, some form of scientific validation and evaluation. However, researchers have to make choices because of time, money and access constrains, making the ideal FSSA an impossibility.

Considering the remainder of the thesis, I argue that it could have been better to base the ideal FSSA only on LCA literature. However, by also basing the ideal FSSA on the literature on Sustainability Indicators, a somewhat broader few of FSSAs was presented. This gave insights into factors that are rarely mentioned in the LCA literature such as stakeholder inclusion and site specific conditions.

RQ2: What methodological problems must be diagnosed when an ideal FSSA is compared to current scientific sustainability assessments on edible oil systems?

I reviewed the scientific body on edible oil FSSAs to test if the ideal FSSA was realistic. I found found that most edible oil LCAs diverged from the ideal FSSA on several aspects. The main aspects were the neglection of the social and economic dimension; farming systems being overlooked; the exclusion of product chain phases other than the agricultural and manufacturing phase; the neglection of time boundaries; and, the failure to execute evaluation methods. Excluding these aspects might be justified by the narrow scope of most studies reviewed, although many authors failed to report on the possible implications for their conclusions. As a result, I found it difficult to assess if these narrow scopes were indeed legitimized.

Furthermore, I found that the results of edible oil LCAs are diverse, mainly due to the heterogenous choices made by authors. Most heterogenous choices were made in the allocation methods applied. The variety in researchers' approaches can largely be explained by the lack of a standardized methodology for agri-food system LCAs. Consequently, the comparability of LCA-based FSSAs is jeopardized. As such, these agri-food systems require specifically tailored LCA guidelines—preferably for specific food systems—which yet have to be developed.

In addition, I identified several methodological issues in my review. Interestingly, the unaddressed issues were often the ones that had potentially large environmental impact. These issues include essential impact categories like biodiversity loss, changes in soil quality and land use change, and the estimation of field N_2O emissions. The existence of these large issues can be explained by their complexity and the lack of scientific consensus on how to deal with them. However, that does not mean they should simply be ignored. Providers of LCA-based FSSAs must therefore be transparent about the incompleteness of the method, so that results can be interpreted with beforementioned issues in mind.

My thesis is the second publication that integratively reviewed the bodies of rape-, sunflower-, peanut- and olive-oil LCAs and the third publication that reviewed existing olive oil FSSAs. The evaluation sheet I used during my literature review was based on earlier literature reviews on various edible oils carried out by Banias et al. (2017) and Khatri and Jain (2017). The results showed that most of my findings were indeed in line with these authors. With the exception of the research gap on fossil fuel use identified by Banias et al. (2017) and the neglection of abiotic depletion as an impact category by Khatri and Jain (2017).

RQ3: How do current providers of consumer FSSAPs assess sustainability?

My thesis aimed to provide explorative insights into consumers FSSAPs. These insights were obtained by analysing three consumer FSSAPs. In this analysis, I analysed documents published by the organizations and carried out interviews with consumer FSSAPs' representatives. My results show that every organization attempted to assist consumers with making sustainable choices; yet, the way of facilitating these sustainable choices differed substantially. Interestingly, the organizations' aims and views on sustainability only differed slightly. However, these differences trickled down throughout each organization – I found large differences in the inclusion of indicators and impact categories, and the aggregation and presentation of sustainability scores. These differences can heavily influence consumers who want to consume more sustainably through a FSSAP. As such, the consumers' choices are not only dependent on the impacts of a production system, but also dependent on the philosophy behind the FSSAP used by that person.

The organizations I examined were dealing with severe limitations, just like the academic literature. However, whereas the academic literature often has a cut-off point and call a limitation an avenue for further research, the organizations seem to deal with the limitations in their own ways, depending on what purpose their project serves.

The findings on consumer FSSAPs are based on unique cases and do not apply to every consumer FSSAP. This was also not the aim of this analysis. Instead, the nature of the analysis was more explorative.

Synthesis

The title of my thesis included the following question: *"How can informed consumer food choices be facilitated?"* The answer to this question is double-sided. First and foremost, it is possible to facilitate informed consumer choices. Possibly through consumer FSSAPs like the ones I examined in this thesis. However, both the academic FSSAs and consumer FSSAPs come with many flaws. Flaws that need urgent addressing. This does not mean that these FSSA(P)s are of low quality, or useless – I tried to show quite the opposite. The FSSA(P)s have flaws but they can still be highly valuable as long as academics and organizations are transparent about these flaws. So that people without background knowledge can keep this in mind when using a FSSAP. That way, a consumer, a decisionmaker or a farmer can still have valuable insights into the sustainability of food products.

My thesis's title question was expanded in my research aim. The research aim added the three audiences of my thesis. These audiences are academics, a start-up that is also my commissioner and consumers. For academics, my thesis can be used as an overview of methodological issues within agri-food LCAs and as a steppingstone into consumer FSSAPs research. In case of my commissioner, my thesis elaborates on several considerations that are at the core of consumer FSSAPs in general and at the core of my commissioner's start-up. These considerations are essential when developing aggregated sustainability scores like my commissioner is doing at the moment of writing. Hence, I dedicated a last section to these considerations to wrap up my thesis (Section 6.1). In terms of consumers, my thesis gave insights into the flaws of consumer FSSAPs. I dedicated my thesis's last section (Section 6.2) to the consumer. This to provide some practical advice on how to make informed and sustainable food choices, regardless if somebody uses a consumer FSSAP or not. In doing so, I answered my research aim and provided the information I aimed for to each of my audiences.

Recommendations

6. Recommendations: consumers and consumer FSSAPs

I wrap-up this thesis by providing a set of recommendations to a start-up that is currently developing a consumer Food System Sustainability Assessment Platform (FSSAP) (6.1). This start-up, in which I partake myself, is also the commissioner of this thesis (Appendix 1). The start-up attempts to develop an LCA-based FSSAP for Dutch supermarkets. The recommendations to the start-up provide insights into matters that an infant FSSAP needs to consider. Additionally, I outline several considerations for consumers who want to consume more sustainably (6.2). These considerations cover aspects regarding the usage of consumer FSSAPs.

6.1. Start-up: recommendations for developing a FSSAP

Aligning organizational aim and sustainability

You should have a clear vision of what sustainability exactly means and how your organization aims to go about it. Have clear answer to questions like:

- What consumer pain do you want to solve (lack of information, reliability, quick decision making, etc.)?
- What do you exactly aim to offer to consumers in order to solve this pain?
- How do you define sustainability and how do you plan to translate this to your customer?

This clarity is needed to create a focal point to which your resources can be directed. This need became evident in my case studies, in which different operationalizations of sustainability lead to completely different ways of presenting sustainability scores to consumers.

Data aggregation

Assuming that you still want to aggregate impact categories into one total sustainability score (see Appendix 1), your organization should reflect on its persistence to do so. My findings show that it is not 100% fair to give consumers the complete responsibility to consume sustainably. Especially because totally aggregated sustainability scores include multiple layers of uncertainty, of which consumer have no knowledge. Indeed, the values in aggregated scores might conceal large uncertainty ranges, especially because LCA-based FSSAs include a few highly uncertain indicators.

If an aggregated score is still a goal by itself, you should also consider that the inclusion and exclusion of an impact category can be deemed political. Therefore, I suggest that aggregated sustainability scores should be developed in cooperation with an expert panel that includes both scientists and industry experts.

Additionally, the aggregation of sustainability scores requires weighing of impact categories. Favourably by the expert panel. However, you could also include consumers' views since this can motivate your users to continue to use your platform. A way to do this is to enable each user of your platform to rate the impact categories that were already weighted by the expert panel (see Endymion for an example of rating of sustainability criteria by consumers). This way, impact categories deemed important by experts are properly weighted and consumers have the possibility to include their view on sustainability.

Inclusion of the social dimension

Assuming that you continue to use LCA-based methodologies, you should know that the inclusion of the social dimension is extremely difficult. Social impacts are oftentimes locally rooted and therefore require locally adapted data and assessment tools, these are scarce. A solution might be the use of available databases or averages presented in the available literature; however, these are often expensive or too generic to cover a specific social context. Indeed, social data require a specific contextualized interpretation to arrive at meaningful conclusions – making the use of generic databases somewhat disputable. A possible way to still include the social dimension is the incorporation of existing labels (e.g. fair trade) in your assessment metric. A good example is the work done by Tatanka on tropical rainforest protection (see Section 3.1 and Appendix 8).

Lack of LCA data

Besides scarce social and economic LCA data, other aspects are also lacking in the LCA methodology. This would not be a problem if the potential environmental impact is not large. However, two aspects have large potential impact and low accuracy in LCAs, namely 1) fertilizer and pesticide production and dispersion, and 2) N_2O field emissions. To my knowledge, no widely applicable accurate method exists to estimate the impacts of these aspects. Therefore, your organization should be very transparent about these drawbacks to the users of your platform.

LCAs inability dealing with multiple essential impact categories

In my thesis, I showed that the LCA methodology has difficulties with assessing land use change, biodiversity loss and changes in soil quality. However, these impact categories are essential for ecosystems' quality and should therefore not be ignored. Again, incorporating existing labels into your platform enables you to give some insights of food products' impacts on these categories. Additionally, Tatanka dealt with these impact categories by being transparent about the synergies between these 'missing indicators' and other indicators like global warming potential. Whatever method you choose to use, be sure you are transparent about it and to root it in current scientific findings.

Possible partnerships

During my interviews, multiple existing consumer FSSAPs stated that they are very willing to collaborate with other organizations in order to attain their goals. A few suggestions made during the interviews were:

- Endymion's representative considered to reinitiate their research project in The Netherlands in collaboration with your start-up (see Appendix 7). You should consider if this can be a valuable partnership;
- Tatanka offers access to their product databases to your start-up for a small compensation (see Appendix 8). The representative emphasized that they want to help smaller organizations in order to stimulate sustainable consumption in other countries. Therefore, consider collaboration with Tatanka and subsequent steps; and
- Tatanka's representative also mentioned a Netherlands based company called Questionmark who is doing something similar as your start-up (see Appendix 8). Questionmark is searching for partners so you might want to approach them.

6.2. Consumers: considerations when using a FSSAP

Food products and the systems in which they are produced can be assessed in various ways. Different institutions (e.g. universities, labels, governments, companies) vary in their approaches. Hence, no uniform assessment method exists, although there are a few popular methods. One of the popular methods is the Life Cycle Assessment (LCA), on which this thesis mostly focussed. The LCA methodology is rooted in scientific approaches. However, that does not automatically mean it is fully able to assess the ‘level of sustainability’ of specific food products. The results of food LCAs are only as accurate and complete as their underlying datasets. As such, food LCAs’ results are impeded by limited availability of input data (e.g. fertilizer and pesticide production), large uncertainty in some parameters (e.g. N_2O field emissions), missing sustainability dimensions (e.g. the social dimension) and the inability of the method to assess certain essential impact categories (e.g. changes in soil quality and biodiversity loss). Hence, the LCA method is somewhat inaccurate and incomplete. Nothing is wrong with that per se – food system sustainability is very difficult to assess. Nonetheless, you should know that this inaccuracy and incompleteness has undesired influence on the choices you make when using a FSSAP – the FSSAP might not cover your view on sustainability.

Indeed, the concept ‘sustainability’ is contested – it is in the eye of the beholder. So, before using a FSSAP, try to crystalize for yourself what a sustainable food product is for you. Do you care about local produce? Fair trade? Maybe GHG emissions? The importance of this exercise becomes clear when looking at some specific sustainability aspects. For example, local sourcing is perceived as one of the most important environmental factors by Irish consumers (Evocco, 2019). However, a FSSAP can assess locally sourced products such as beef as potentially very ‘non-sustainable’ because of high global warming potential (Ritchie & Roser, 2020). This introduces cognitive dissonance to the Irish consumer – the question arises why they prefer locally sourced food. Is it because they want to support local farmers? Or because of smaller transport distances and therewith decreased GHG emissions? Whatever the answer may be, a FSSAP might cause friction with consumers’ beliefs because the platform only includes a limited set of sustainability aspects. A second example is the debate regarding organic and conventional produce. Based on LCA methodologies, many studies found that conventionally sourced products have less overall environmental impacts per unit of produce than organic produce (Notarnicola et al., 2017). The reason is simple: conventional produce yields more so a system’s production impacts are divided over a larger quantity. However, the current LCA methodology is incomplete and does not consider some essential aspects for long-term food production, such as biodiversity loss and decreased soil quality (Petti et al., 2015). Considering these two impact categories, organic sourcing is the more environmentally friendly pick. However, this is not considered in LCA-based FSSAPs due to the LCA’s inability to include these impact categories (Notarnicola et al., 2017). So, if you believe in the good cause of organic agriculture, know that a FSSAP might tell you that you are buying the wrong products. As such, as a consumer, you should know that every FSSAP is incomplete by default and might only be transparent with regards to what is included in the assessment – and not so much about what is not.

A last consideration when choosing a FSSAP is the presentation of sustainability scores. Different formats exist, of which a few are presented in my thesis. These formats are built upon organizations’ visions on how sustainability scores should be translated to consumers. Once again, nothing is right or wrong. However, you might want to consider what kind of presentation you prefer. Do you want to make choices as quick and easy as possible based on expert views? Or choices tailored to your sustainability preferences? Or do you desire more information depth so you could make more informed choices? Regardless of your answer, each FSSAP will produce scores that do cover sustainability aspects but are still incomplete and inaccurate because of underlying methodological flaws. Therefore, I recommend you choose a FSSAP that fits your view on sustainability, your preference on information depth and your need for customizability.

In case you are not considering using a FSSAP but still want to consume in a more climate friendly way, my recommendations are in line with some more recent publications. Rather than sticking to local and/or organic produce, focus more on what foods precisely you are eating (Poore & Nemecek, 2018), in which plant-based foods are preferred (Willett et al., 2019) and air-freighted foods should be avoided (Ritchie & Roser, 2020). In case of air-freighted products, the easiest way to recognize them is that they are highly perishable while simultaneously coming from a far distance (Ritchie & Roser, 2020). Some products that tend to be in this category are nonseasonal berries, green beans and asparagus.

References

Accorsi, R., Cascini, A., Ferrari, E., Manzini, R., Pareschi, A., & Versari, L. (2013). Life cycle assessment of an extra-virgin olive oil supply chain. *Proceedings of the 18th Summer School "Francesco Turco"—Industrial Mechanical Plants, Senigallia, Italy*, 11-13.

Andersson, K. (2000). LCA of food products and production systems. *The International Journal of Life Cycle Assessment, 5*(4), 239-248.

Avraamides, M., & Fatta, D. (2008). Resource consumption and emissions from olive oil production: a life cycle inventory case study in Cyprus. *Journal of Cleaner Production, 16*(7), 809-821.

Badey, L., Lahitte, N., Flenet, F., & Bosque, F. (2013). French environmental communication on sunflower and rapeseed oils based on life cycle assessment. *OCL, 20*(4), A401.

Banias, G., Achillas, C., Vlachokostas, C., Moussiopoulos, N., & Stefanou, M. (2017). Environmental impacts in the life cycle of olive oil: a literature review. *Journal of the Science of Food and Agriculture, 97*(6), 1686-1697.

Basset-Mens, C., Benoist, A., Bessou, C., Tran, T., Perret, S., Vayssières, J., & Wassenaar, T. (2010). *Is LCA-based eco-labelling reasonable? The issue of tropical food products*.

Beketov, M. A., Kefford, B. J., Schäfer, R. B., & Liess, M. (2013). Pesticides reduce regional biodiversity of stream invertebrates. *Proceedings of the National Academy of Sciences, 110*(27), 11039-11043.

Bélanger, V., Vanasse, A., Parent, D., Allard, G., & Pellerin, D. (2012). Development of agri-environmental indicators to assess dairy farm sustainability in Quebec, Eastern Canada. *Ecological Indicators, 23*, 421-430. doi:https://doi.org/10.1016/j.ecolind.2012.04.027

Bie, R. v. d., Hermans, B., Pierik, C., Stroucken, L., & Wobma, E. (2012). Smakelijk Weten, Trends in Voeding en Gezondheid. *Den Haag/Heerlen: Centraal Bureau voor de Statistiek*.

Boeije, H. (2009). *Analysis in qualitative research*: Sage publications.

Bossel, H. (2002). Assessing viability and sustainability: a systems-based approach for deriving comprehensive indicator sets. *Conservation ecology, 5*(2).

Brentrup, F., Küsters, J., Lammel, J., & Kuhlmann, H. (2000). Methods to estimate on-field nitrogen emissions from crop production as an input to LCA studies in the agricultural sector. *The International Journal of Life Cycle Assessment, 5*(6), 349-357.

Brundtland, G. H. (1987). Our common future—Call for action. *Environmental Conservation, 14*(4), 291-294.

Busset, G., Belaud, J.-P., Montréjaud-Vignoles, M., & Sablayrolles, C. (2014). Integration of social LCA with sustainability LCA: a case study on virgin olive oil production. *Social LCA in progress*.

Busset, G., Belaud, J.-P., Vignoles-Montrejaud, M., & Sablayrolles, C. (2013). An Integrated Approach for Designing an Agricultural Process Guided by sustainable Evaluation: Application to Olive-Oil Production. *Insight, 16*(4), 32-36.

Busset, G., Belaud, J., Clarens, F., Espi, J., Montréjaud-Vignoles, M., & Sablayrolles, C. (2012). *Life Cycle Assessment of olive oil production in France.* Paper presented at the Proc. 4th Int. Conf. Eng. Waste Biomass Valoris. A. Nizhou & F. Castro, Porto, Portugal.

Carimentrand, A., & Ballet, J. (2010). When Fair Trade increases unfairness: The case of quinoa from Bolivia. *Cahier FREE, 5*, 17.

Castellini, C., Boggia, A., Cortina, C., Dal Bosco, A., Paolotti, L., Novelli, E., & Mugnai, C. (2012). A multicriteria approach for measuring the sustainability of different poultry production systems. *Journal of Cleaner Production, 37*, 192-201.

Cavallaro, F., & Salomone, R. (2010). Interpretation of Life Cycle Assessment results using a multi-criteria tool: Application to the olive oil chain. *Proceedings of LCA Food*.

Cerutti, A. K., Beccaro, G. L., Bosco, S., De Luca, A. I., Falcone, G., Fiore, A., . . . Strano, A. (2015). Life cycle assessment in the fruit sector. In *Life Cycle Assessment in the Agri-food Sector* (pp. 333-388): Springer.

Chaudhary, A., Gustafson, D., & Mathys, A. (2018). Multi-indicator sustainability assessment of global food systems. *Nature communications, 9*(1), 848.

Cloquell-Ballester, V.-A., Cloquell-Ballester, V.-A., Monterde-Diaz, R., & Santamarina-Siurana, M.-C. (2006). Indicators validation for the improvement of environmental and social impact quantitative assessment. *Environmental Impact Assessment Review, 26*(1), 79-105.
Cossu, A., Degl'Innocenti, S., Cristani, M., Bedini, S., & Nuti, M. (2013). Assessment of the life cycle environmental impact of the olive oil extraction solid wastes in the European Union. *The Open Waste Management Journal, 6*(1).
Dantsis, T., Douma, C., Giourga, C., Loumou, A., & Polychronaki, E. A. (2010). A methodological approach to assess and compare the sustainability level of agricultural plant production systems. *Ecological Indicators, 10*(2), 256-263.
De Gennaro, B., Notarnicola, B., Roselli, L., & Tassielli, G. (2012). Innovative olive-growing models: an environmental and economic assessment. *Journal of Cleaner Production, 28*, 70-80.
De Luca, A. I., Falcone, G., Stillitano, T., Iofrida, N., Strano, A., & Gulisano, G. (2018). Evaluation of sustainable innovations in olive growing systems: A Life Cycle Sustainability Assessment case study in southern Italy. *Journal of Cleaner Production, 171*, 1187-1202.
de Vries, M., & de Boer, I. J. (2010). Comparing environmental impacts for livestock products: A review of life cycle assessments. *Livestock science, 128*(1-3), 1-11.
Dekker, E., Zijp, M. C., van de Kamp, M. E., Temme, E. H. M., & van Zelm, R. (2019). A taste of the new ReCiPe for life cycle assessment: consequences of the updated impact assessment method on food product LCAs. *The International Journal of Life Cycle Assessment*. doi:10.1007/s11367-019-01653-3
Egea, P., & y Pérez, L. P. (2016). Sustainability and multifunctionality of protected designations of origin of olive oil in Spain. *Land Use Policy, 58*, 264-275.
Eggleston, H., Buendia, L., Miwa, K., Ngara, T., & Tanabe, K. (2006). IPCC IPCC Guidelines for National Greenhouse Gas Inventories. Volume 1: General guidance and reporting. In.
El Hanandeh, A., & Gharaibeh, M. A. (2016). Environmental efficiency of olive oil production by small and micro-scale farmers in northern Jordan: Life cycle assessment. *Agricultural Systems, 148*, 169-177.
Erisman, J. W., Galloway, J. N., Seitzinger, S., Bleeker, A., Dise, N. B., Petrescu, A. R., . . . de Vries, W. (2013). Consequences of human modification of the global nitrogen cycle. *Philosophical Transactions of the Royal Society B: Biological Sciences, 368*(1621), 20130116.
Evocco. (2019). Evocco insights 2019. Retrieved from https://www.linkedin.com/posts/counting-carbon_sustainability-food-activity-6570244322141306880-PUc_
Figueiredo, F., Castanheira, É. G., & Freire, F. (2017). Life-cycle assessment of irrigated and rainfed sunflower addressing uncertainty and land use change scenarios. *Journal of Cleaner Production, 140*, 436-444.
Figueiredo, F., Geraldes Castanheira, E., & Freire, F. (2012). *LCA of sunflower oil addressing alternative land use change scenarios and practices.* Paper presented at the Proceedings of the 8th International Conference on Life Cycle Assessment in the Agri-Food Sector (LCA Food 2012).
Fleskens, L., Duarte, F., & Eicher, I. (2009). A conceptual framework for the assessment of multiple functions of agro-ecosystems: A case study of Trás-os-Montes olive groves. *Journal of Rural Studies, 25*(1), 141-155.
Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., . . . Gibbs, H. K. (2005). Global consequences of land use. *science, 309*(5734), 570-574.
Garrigues, E., Corson, M. S., Angers, D. A., van der Werf, H. M., & Walter, C. (2012). Soil quality in Life Cycle Assessment: Towards development of an indicator. *Ecological Indicators, 18*, 434-442.
Gladek, E., Kennedy, E., Exter, P. v., & Monaghan, C. (2018). EIT CLIMATE-KIC: Food value chains strategy: Review and recommendations from a systemic impact perspecitve. *EIT CLIMATE-KIC & Metabolic*.
Gomez-Limon, J. A., & Riesgo, L. (2010). *Sustainability assessment of olive grove in Andalusia: A methodological proposal*. Retrieved from

Gómez-Limón, J. A., & Sanchez-Fernandez, G. (2010). Empirical evaluation of agricultural sustainability using composite indicators. *Ecological economics, 69*(5), 1062-1075.
Goodland, R. (1995). The concept of environmental sustainability. *Annual review of ecology and systematics, 26*(1), 1-24.
Guermazi, Z., Gharsallaoui, M., Perri, E., Gabsi, S., & Benincasa, C. (2017). Integrated approach for the eco design of a new process through the life cycle analysis of olive oil: Total use of olive by-products. *European Journal of Lipid Science and Technology, 119*(9), 1700009.
Hayashi, K. (2000). Multicriteria analysis for agricultural resource management: a critical survey and future perspectives. *European Journal of Operational Research, 122*(2), 486-500.
Huijbregts, M., Steinmann, Z., Elshout, P., Stam, G., Verones, F., Vieira, M., . . . Van Zelm, R. (2016). ReCiPe 2016: A harmonized life cycle impact assessment method at midpoint and endpoint level Report I: Characterization.
IPCC. (2019). Climate Change and Land: An IPCC Special Report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems.
Iraldo, F., Testa, F., & Bartolozzi, I. (2014). An application of Life Cycle Assessment (LCA) as a green marketing tool for agricultural products: the case of extra-virgin olive oil in Val di Cornia, Italy. *Journal of Environmental Planning and Management, 57*(1), 78-103.
Jeswani, H. K., Azapagic, A., Schepelmann, P., & Ritthoff, M. (2010). Options for broadening and deepening the LCA approaches. *Journal of Cleaner Production, 18*(2), 120-127.
Jones, A. D., Hoey, L., Blesh, J., Miller, L., Green, A., & Shapiro, L. F. (2016). A systematic review of the measurement of sustainable diets. *Advances in Nutrition, 7*(4), 641-664.
Kalnikaitė, V., Bird, J., & Rogers, Y. (2013). Decision-making in the aisles: informing, overwhelming or nudging supermarket shoppers? *Personal and Ubiquitous Computing, 17*(6), 1247-1259.
Kazemi, H., Bourkheili, S. H., Kamkar, B., Soltani, A., Gharanjic, K., & Nazari, N. M. (2016). Estimation of greenhouse gas (GHG) emission and energy use efficiency (EUE) analysis in rainfed canola production (case study: Golestan province, Iran). *Energy, 116*, 694-700.
Khanali, M., Mousavi, S. A., Sharifi, M., Nasab, F. K., & Chau, K.-w. (2018). Life cycle assessment of canola edible oil production in Iran: A case study in Isfahan province. *Journal of Cleaner Production, 196*, 714-725.
Khatri, P., & Jain, S. (2017). Environmental life cycle assessment of edible oils: A review of current knowledge and future research challenges. *Journal of Cleaner Production, 152*, 63-76.
Lamichhane, J. R., Dachbrodt-Saaydeh, S., Kudsk, P., & Messéan, A. (2016). Toward a reduced reliance on conventional pesticides in European agriculture. *Plant Disease, 100*(1), 10-24.
Leip, A., Billen, G., Garnier, J., Grizzetti, B., Lassaletta, L., Reis, S., . . . Weiss, F. (2015). Impacts of European livestock production: nitrogen, sulphur, phosphorus and greenhouse gas emissions, land-use, water eutrophication and biodiversity. *Environmental Research Letters, 10*(11), 115004.
MacWilliam, S., Sanscartier, D., Lemke, R., Wismer, M., & Baron, V. (2016). Environmental benefits of canola production in 2010 compared to 1990: A life cycle perspective. *Agricultural Systems, 145*, 106-115.
Mattsson, B., Cederberg, C., & Blix, L. (2000). Agricultural land use in life cycle assessment (LCA): case studies of three vegetable oil crops. *Journal of Cleaner Production, 8*(4), 283-292.
Menozzi, D. (2014). Extra-virgin olive oil production sustainability in northern Italy: a preliminary study. *British Food Journal, 116*(12), 1942-1959. doi:10.1108/BFJ-06-2013-0141
Meul, M., Van Passel, S., Nevens, F., Dessein, J., Rogge, E., Mulier, A., & Van Hauwermeiren, A. (2008). MOTIFS: a monitoring tool for integrated farm sustainability. *Agronomy for sustainable development, 28*(2), 321-332.
Mousavi-Avval, S. H., Rafiee, S., Sharifi, M., Hosseinpour, S., Notarnicola, B., Tassielli, G., . . . Khanali, M. (2017). Use of LCA indicators to assess Iranian rapeseed production systems with different residue management practices. *Ecological Indicators, 80*, 31-39.

Muñoz, I., Schmidt, J. H., & Dalgaard, R. (2014). *Comparative life cycle assessment of five different vegetable oils.* Paper presented at the Proceedings of the 9th International Conference on Life Cycle Assessment in the Agri-Food Sector (LCA Food 2014), San Francisco, California, USA, 8-10 October, 2014.

Nikkhah, A., Khojastehpour, M., Emadi, B., Taheri-Rad, A., & Khorramdel, S. (2015). Environmental impacts of peanut production system using life cycle assessment methodology. *Journal of Cleaner Production, 92*, 84-90.

Noorhosseini, S. A., & Damalas, C. A. (2018). Environmental Impact of Peanut (Arachis hypogaea L.) Production under Different Levels of Nitrogen Fertilization. *Agriculture, 8*(7), 104.

Notarnicola, B., Sala, S., Anton, A., McLaren, S. J., Saouter, E., & Sonesson, U. (2017). The role of life cycle assessment in supporting sustainable agri-food systems: A review of the challenges. *Journal of Cleaner Production, 140*, 399-409.

Notarnicola, B., Tassielli, G., & Nicoletti, G. M. (2004). Environmental and economical analysis of the organic and conventional extra-virgin olive oil. *New Medit., 3*, 28-34.

Notarnicola, B., Tassielli, G., Renzulli, P. A., & Giudice, A. L. (2015). Life Cycle Assessment in the agri-food sector: an overview of its key aspects, international initiatives, certification, labelling schemesand methodological issues. In *Life Cycle Assessment in the Agri-food Sector* (pp. 1-56): Springer.

Nucci, B., Puccini, M., Pelagagge, L., Vitolo, S., & Nicolella, C. (2014). Improving the environmental performance of vegetable oil processing through LCA. *Journal of Cleaner Production, 64*, 310-322.

Petit, G., Sablayrolles, C., & Yannou-Le Bris, G. (2018). Combining eco-social and environmental indicators to assess the sustainability performance of a food value chain: A case study. *Journal of Cleaner Production, 191*, 135-143.

Petti, L., Arzoumanidis, I., Benedetto, G., Bosco, S., Cellura, M., De Camillis, C., . . . Raggi, A. (2015). Life cycle assessment in the wine sector. In *Life Cycle Assessment in the Agri-food Sector* (pp. 123-184): Springer.

Poore, J., & Nemecek, T. (2018). Reducing food's environmental impacts through producers and consumers. *science, 360*(6392), 987-992.

Queiros, J., Malca, J., & Freire, F. (2015). Environmental life-cycle assessment of rapeseed produced in Central Europe: addressing alternative fertilization and management practices. *Journal of Cleaner Production, 99*, 266-274.

Rajaeifar, M. A., Akram, A., Ghobadian, B., Rafiee, S., & Heidari, M. D. (2014). Energy-economic life cycle assessment (LCA) and greenhouse gas emissions analysis of olive oil production in Iran. *Energy, 66*, 139-149.

Reap, J., Roman, F., Duncan, S., & Bras, B. (2008). A survey of unresolved problems in life cycle assessment. *The International Journal of Life Cycle Assessment, 13*(5), 374.

Reed, M. S., Fraser, E. D., & Dougill, A. J. (2006). An adaptive learning process for developing and applying sustainability indicators with local communities. *Ecological economics, 59*(4), 406-418.

Renzulli, P. A., Bacenetti, J., Benedetto, G., Fusi, A., Ioppolo, G., Niero, M., . . . Supino, S. (2015). Life cycle assessment in the cereal and derived products sector. In *Life Cycle Assessment in the Agri-food Sector* (pp. 185-249): Springer.

Rinaldi, S., Barbanera, M., & Lascaro, E. (2014). Assessment of carbon footprint and energy performance of the extra virgin olive oil chain in Umbria, Italy. *Science of The Total Environment, 482*, 71-79.

Ritchie, H., & Roser, M. (2020). Environmental impacts of food production. *Our World in Data*. Retrieved from https://ourworldindata.org/environmental-impacts-of-food

Rolvink, R. (2019). Olie: de belangrijkste soorten op een rij. *Consumentenbond*. Retrieved from https://www.consumentenbond.nl/gezond-eten/olie

Roma, R., Corrado, S., De Boni, A., Forleo, M. B., Fantin, V., Moretti, M., . . . Vitali, A. (2015). Life cycle assessment in the livestock and derived edible products sector. In *Life Cycle Assessment in the Agri-food Sector* (pp. 251-332): Springer.

Roy, R., & Chan, N. W. (2012). An assessment of agricultural sustainability indicators in Bangladesh: review and synthesis. *The Environmentalist, 32*(1), 99-110.

Salomone, R., Cappelletti, G. M., Malandrino, O., Mistretta, M., Neri, E., Nicoletti, G. M., . . . Saija, G. (2015). Life cycle assessment in the olive oil sector. In *Life Cycle Assessment in the Agri-food Sector* (pp. 57-121): Springer.

Salomone, R., & Ioppolo, G. (2012). Environmental impacts of olive oil production: a Life Cycle Assessment case study in the province of Messina (Sicily). *Journal of Cleaner Production, 28*, 88-100.

Sargant, E. (2014). *Sustainable food consumption: a practice based approach* (Vol. 11): Wageningen Academic Publishers.

Schmidt. (2004). *The importance of system boundaries for LCA on large materials flows of vegetal oils.* Paper presented at the The Fourth World SETAC Congress, Portland, USA.

Schmidt. (2009). Carbon footprinting, labelling and life cycle assessment. *The International Journal of Life Cycle Assessment, 14*(1), 6-9.

Schmidt. (2010). Comparative life cycle assessment of rapeseed oil and palm oil. *The International Journal of Life Cycle Assessment, 15*(2), 183-197.

Schmidt. (2015). Life cycle assessment of five vegetable oils. *Journal of Cleaner Production, 87*, 130-138.

Shiklomanov, I. A., & Rodda, J. C. (2004). *World water resources at the beginning of the twenty-first century*: Cambridge University Press.

Smith, C. S., McDonald, G. T., & Thwaites, R. (2000). TIM: Assessing the sustainability of agricultural land management. *Journal of Environmental Management, 60*(4), 267-288.

Springmann, M., Clark, M., Mason-D'Croz, D., Wiebe, K., Bodirsky, B. L., Lassaletta, L., . . . Carlson, K. M. (2018). Options for keeping the food system within environmental limits. *Nature, 562*(7728), 519.

Sutton, M. A., Bleeker, A., Howard, C., Erisman, J., Abrol, Y., Bekunda, M., . . . Oenema, O. (2013). *Our nutrient world. The challenge to produce more food & energy with less pollution*. Retrieved from

Thomassen, M. A., Dalgaard, R., Heijungs, R., & De Boer, I. (2008). Attributional and consequential LCA of milk production. *The International Journal of Life Cycle Assessment, 13*(4), 339-349.

Tsarouhas, P., Achillas, C., Aidonis, D., Folinas, D., & Maslis, V. (2015). Life Cycle Assessment of olive oil production in Greece. *Journal of Cleaner Production, 93*, 75-83.

Van Asselt, E., Van Bussel, L., Van der Voet, H., Van der Heijden, G., Tromp, S., Rijgersberg, H., . . . Van der Fels-Klerx, H. (2014). A protocol for evaluating the sustainability of agri-food production systems—A case study on potato production in peri-urban agriculture in The Netherlands. *Ecological Indicators, 43*, 315-321.

Van Cauwenbergh, N., Biala, K., Bielders, C., Brouckaert, V., Franchois, L., Cidad, V. G., . . . Reijnders, J. (2007). SAFE—A hierarchical framework for assessing the sustainability of agricultural systems. *Agriculture, ecosystems & environment, 120*(2-4), 229-242.

Van Rossum, C., Buurma-Rethans, E., Vennemann, F., Beukers, M., Brants, H. A., De Boer, E., & Ocké, M. C. (2016). The diet of the Dutch: Results of the first two years of the Dutch National Food Consumption Survey 2012-2016. *RIVM letter report 2016-0082*.

Vermeulen, S. J., Campbell, B. M., & Ingram, J. S. (2012). Climate change and food systems. *Annual review of environment and resources, 37*.

Von Reischach, F., Dubach, E., Michahelles, F., & Schmidt, A. (2010). *An evaluation of product review modalities for mobile phones.* Paper presented at the Proceedings of the 12th international conference on Human computer interaction with mobile devices and services.

Von Wirén-Lehr, S. (2001). Sustainability in agriculture—an evaluation of principal goal-oriented concepts to close the gap between theory and practice. *Agriculture, ecosystems & environment, 84*(2), 115-129.
Westhoek, H., Rood, T., van den Berg, M., Janse, J., Nijdam, D., Reudink, M., . . . Woltjer, G. (2011). *The protein puzzle: the consumption and production of meat, dairy and fish in the European Union*: Netherlands Environmental Assessment Agency.
Willett, W., Rockström, J., Loken, B., Springmann, M., Lang, T., Vermeulen, S., . . . 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. doi:https://doi.org/10.1016/S0140-6736(18)31788-4
Williams, A. G., Audsley, E., & Sandars, D. L. (2010). Environmental burdens of producing bread wheat, oilseed rape and potatoes in England and Wales using simulation and system modelling. *The International Journal of Life Cycle Assessment, 15*(8), 855-868.
Winterman, P. (2019). Kabinet in paniek door stikstofverbod. *AD*. Retrieved from https://www.ad.nl/politiek/kabinet-in-paniek-door-stikstofverbod~a0d181f2/?referrer=https://www.google.nl/
Zhang, X., Davidson, E. A., Mauzerall, D. L., Searchinger, T. D., Dumas, P., & Shen, Y. (2015). Managing nitrogen for sustainable development. *Nature, 528*(7580), 51.

Image credits

1. Cover page: Adapted from Pexels / Pixabay
2. In memoriam image: Adapted from Sponchia / Pixabay
3. Introduction cover: Adapted from Aaron Escobar / Flickr
4. Methodology cover: Adapted from Mitchell Luo / Pexels
5. Results cover: Adapted from Edgar Hernandez / Pexels
6. Discussion cover: Adapted from Anna / Pexels
7. Conclusion cover: Adapted from Picasa / RiledupJournal
8. Take home messages cover: Adapted from Johnny Mckane / Pexels

Appendixes

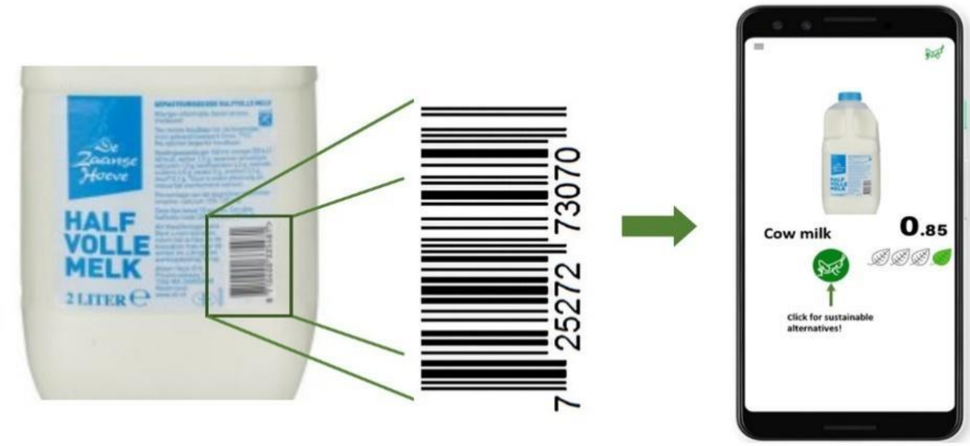
Appendix 1: Cricket

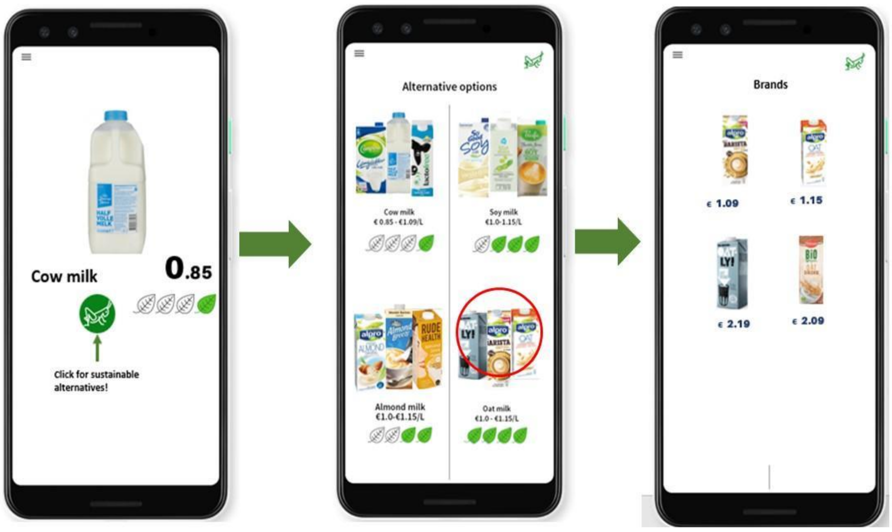
Cricket, a newly founded start-up project aims to simplify sustainability grading systems by providing supermarkets a platform, and integration of that platform in their existing digital consumer infrastructure. In doing so, it will provide their customers a real-time comparison of the shop assortment based on a user-friendly sustainability grading system. Hence, the proposed thesis project will function as a MSc thesis at the ESA group, of which the outcomes can be used as part of a feasibility study of the Cricket application.

The platform itself will show the different options available in-store, their price and their social and environmental impact grade in a single screen. The Cricket platform will be developed to be added to already existing technological infrastructure employed by supermarkets such as hand-scanners, store-apps, and store-websites to ease grocery shopping. The platform will work as follows: after selecting or scanning a product, the customer of the supermarket will see the price and sustainability grade of the chosen product on the screen and they will receive a notification advising them to see sustainable alternatives available in the store (see illustration next page). The customer will then have the option to click on the 'Cricket logo' and see, in a single screen, a list of different alternative products that serve the same dietary purpose and possess comparable qualities to the scanned product (see illustration next page). If the customer is interested in changing their choice, they can select one of the visible product options to identify which brands are available in the store they are currently at.

To conclude, we (Cricket) work on a single index that combines multiple environmental impact categories. Of course, we want to provide transparency about the underlying impact categories, but not *in situ* while this would result in information overload for the end-user. Hence, Cricket would aggregate existing impact data, from existing data bases and weigh this impact data by means of a developed assessment method so that a single index number for a large number of supermarket products can be developed. The weighing could be done by experts or consumers themselves if desired. The job Cricket has to do is finding the right partner databases, developing the assessment method, carrying it out on a continuous base and providing the end-user with a single index number *in situ* and the underlying metric *ex post.*

Developing such a platform includes on one large assumption, being data availability. As stated Negra et al. (2019) *"many environmental metrics and reporting initiatives have been put forward for the agriculture sector. However comprehensive, robust, and consistent data systems are not yet in place."* Surprisingly, organizations like Evocco and Eaternity do show that it is possible to a certain extent. This raises the question: how exactly?


HALF VOLLE MELK
2 LITER
7 25272 73070
Cow milk
0.85
Click for sustainable alternatives!


Cow milk
0.85
Click for sustainable alternatives!
Alternative options
Cow milk
€ 0.85 - €1.09/L
Soy milk
€1.0-1.15/L
Almond milk
€1.0-€1.15/L
Oat milk
€1.0 - €1.15/L
Brands
€ 1.09
€ 1.15
€ 2.19
€ 2.09

Appendix 2: Overview of categories used in edible oil LCA evaluation

Category	Question used to analyse category	Answer options	Answers explained
Sustainability dimensions	Which sustainability dimensions are considered?	SC	Social
		EC	Economic
		EV	Environmental
Goal/Scope	What was the goal of the study?	EI	Environmental impact assessment
		EN	Energy analysis
		EC	Economic analysis
		WC	Water consumption
		OT	Other
		NS	Not specified
System boundaries	Which production phases are included in the study?	FR	Farm
		MF	Manufacturing
		PK	Packaging
		WR	Warehousing
		FD	Final disposal (of packaging)
		RL	Reversed logistics
System boundaries (time)	Was a temporal system boundary included in the study?	Yes/No	
Allocation criteria	Which allocation criteria where applied?	SE	System expansion
		MA	Mass allocation
		EA	Economic allocation
		EnA	Energy allocation
		NA	Not applied (as mentioned by author)
		NS	Not specified
Inventory	Which inventory items were included in the analysis?	FR	Fertilizer
		PS	Pesticides
		OMW	Oil mill wastewater
		FL	Fuel
		WU	Water use
		EU	Energy use

Farming system	Which farming system is analysed in the case study?	CF IF TF OF OT NS	Conventional farming Integrated farming Traditional farming Organic farming Other Not specified
Soil conditions	Are soil conditions specified?	Yes/No	
Field emissions	Are field emissions assessed? If yes, how?	IPCC Brentrup NS	IPPC's method Brentrup's method Not specified
Data quality	Where any reflections on data quality made by the authors?	CA SA UA	Contribution analysis Sensitivity analysis Uncertainty analysis

Appendix 3: Overview of methodological choices made in edible oil LCAs (summarized)

Author & year	Sustainability dimensions	Scope	System boundaries	Temporal boundaries	Allocation criteria	Inventory items	Farming system	Soil conditions & management	Field N_2O emissions method	Evaluation methods used
Accorsi et al. (2013)	EV	EI	PK & WR & RL	No	NS	EU	NS	-	NS	CA
Avraamides and Fatta (2008)	EV	EI & OT	FR & MF	No	EA	FR & PS & OMWW & FL & WU & EU	OT	-	Brentrup	CA
G Busset et al. (2012)	EV, SC, EC	EI	FR & MF & PK & RL	No	NA	FR & PS & OMWW & FL & WU & EU	NS	-	NS	-

Guillaume Busset et al. (2014)	EV	EI	FR & MF & PK & RL	No	NA	FR & PS & OMWW & FL & WU & EU	TF & CF & IF & OF	Ploughing	Brentrup	CA
Cavallaro and Salomone (2010)	EV	OT	FR & MF & RL	No	NS	FR & PS & OMWW & WU	CF & OF	Ploughing	Brentrup	-
Cossu et al. (2013)	EV	OT	RL	Yes	NS	FL	NS	-	NS	CA
De Gennaro et al. (2012)	EV, EC	EI & EC	FR	Yes	NA	FR & PS & FL & WU & EU	OT	Soil preparation	Brentrup	SA & UA
De Luca et al. (2018)	EV, SC, EC	OT	FR	Yes	NS	FR & PS & FL	TF & CF & OF	-	Brentrup	SA
Guermazi et al. (2017)	EV	OT	FR & MF & RL	No	NS	FR & PS & OMWW & FL & WU	OT	-	NS	CA
El Hanandeh and Gharaibeh (2016)	EV	EI	FR & MF & PK & RL	Partial	SE	FR & PS & OMWW & FL & WU & EU	OF	Ploughing	IPCC	SA & UA
Iraldo et al. (2014)	EV	EI & OT	FR & MF	Partial	EA	FR & PS & OMWW & FL & WU & EU	OT	Ploughing	Brentrup	CA & SA
Notarnicola et al. (2004)	EV, EC	EI & EC & OT	FR & MF	No	EA	FR & PS & OMWW & FL & WU & EU	CF & OF	Ploughing	Brentrup	CA
Rajaeifar et al. (2014)	EV, EC	EI & EN & EC	FR & MF	Partial	MA	FR & PS & FL & WU & EU	NS	-	NS	CA & SA

Rinaldi et al. (2014)	EV	EI & EN	FR & MF & PK & WR & FD & RL	Partial	MA	FR & PS & OMWW & FL & WU & EU	NS	Ploughing	IPCC	CA & SA
Salomone and Ioppolo (2012)	EV	EI	FR & MF & RL	No	SE	FR & PS & OMWW & FL & WU & EU	CF & IF & OF	Ploughing	Brentrup	CA & SA
Tsarouhas et al. (2015)	EV	EI	FR & MF & PK & RL	No	NS	FR & PS & OMWW & FL & WU & EU	NS	Ploughing	NS	CA
Mousavi-Avval et al. (2017)	EV	EI	FR	Yes	EA	FR & PS & FL & EU	OT	Ploughing	IPCC	CA & SA
Khanali et al. (2018)	EV	EI & EN	FR & MF & PK	Partial	MA	FR & PS & FL & WU & EU	NS	-	Copied from other author	CA
Queiros et al. (2015)	EV	EI	FR	No	NS	FR & PS & FL & EU	OT	Soil type, SOC, ploughing and liming	IPCC	CA & SA
MacWilliam et al. (2016)	EV	EI & OT	FR	Yes	NS	FR & PS & FL & EU	NS	Different soils tested	IPCC	CA & SA
Kazemi et al. (2016)	EV	EI & EN	FR	Partial	NS	FR & PS & FL & EU	OT	Ploughing	NS	-
Nikkhah et al. (2015)	EV	EI	FR	Partial	NS	FR & FL	NS	Soil type estimated	Brentrup	-
Noorhosseini and Damalas (2018)	EV	EI & OT	FR	Yes	NS	FR & PS & FL	OT	Complete soil analysis	IPCC	-

Figueiredo et al. (2012)	EV	OT	FR	Yes	NS	FR & PS & FL & WU & EU	OT	Soil type + SOC	IPCC	SA & UA
Figueiredo et al. (2017)	EV	OT	FR & MF	No	MA & EA &EnA	FR & PS & FL & EU	OT	Soil type + SOC	NS	SA
Nucci et al. (2014)	EV	EI	FR & MF & PK & RL	Partial	OT	FR & PS & FL & WU & EU	NS	-	NS	CA
Mattsson et al. (2000)	EV	OT	FR	No	NS	-	NS	?	NS	-
Schmidt (2010)	EV	EI	FR & MF	Yes	SE & EA	FR & PS & FL & WU & EU	NS	Soil type	NS	CA & UA
Williams et al. (2010)	EV	EI	FR	Yes	NS	FR & PS & FL & EU	CF & OF	Soil type	IPCC	CA
Badey et al. (2013)	EV	EI	FR & MF & PK & FD	No	EnA	-	NS	-	NS	CA
Muñoz et al. (2014)	EV	EI	FR & MF	Yes	OT	FR & FL & WU & EU	NS	?	IPCC	SA
Schmidt (2015)	EV	EI	FR & MF	Yes	OT	FR & FL & WU & EU	NS	?	IPCC	CA & UA
Schmidt (2004)	EV	EI & OT	FR & MF	No	SE & EA	FR & FL & EU	NS	-	NS	CA

Appendix 4: Interview guide(s)

Note: the interview guides also included some questions specifically for the organizations interviewed. These are removed for anonymity's sake.

Interview questions (Dutch)

Intro

- Naam en achtergrond (culture geografie en omgevingswetenschappen)
- Uitleg thesis project
    - Vegatarier, maar zelfs dan vaak moeilijk om goede keuzes te maken in de supermarkt
    - Vandaar, is het technisch mogelijk?
    - Daarom, literatuurstudie en interviews
- Interviews zijn anoniem en vertrouwelijk
- Half uurtje uurtje
- Mogelijk om op te nemen?
- Kunt u misschien iets vertellen over uw rol binnen uw project?

Duurzaamheid/milieu definiëren

Zover ik het kon traceren, zag ik dat de sociale en economische dimensies van duurzaamheid niet opgenomen zijn in uw applicatie. Waarom niet?

Indicatoren

- Uw milieuscore is opgebouwd uit 1) gebruik van fossiele brandstoffen, 2) bijdrage aan klimaatverandering, 3) landgebruik en 4) waterstress. Waarom is er specifiek voor deze indicatoren gekozen?
    - Wegen deze indicatoren allen even zwaar mee in de eindbeoordeling?

- Ik ben veel beoordelingen tegengekomen die moeite hebben met het beoordelen van landgebruik en watergebruik.
    - Hoe gaat u om met landgebruik?
    - Hoe gaat u om met waterstress?

- In de literatuur kom ik erg vaak eutrofiëring, verzuring en toxiciteit (voor mens en ecologie) tegen. Deze zijn niet meegenomen in uw beoordeling. Waarom niet?

- Bodemkwaliteit is wordt vaak genoemd als een belangrijk indicator, toch wordt deze zelden meegenomen in voedselketen beoordelingen, zo ook in die van u. Heeft u hier een mening over?

- Hetzelfde geldt voor biodiversiteit. Wat is uw mening hierover?

- De wetenschappelijke literatuur suggereert vaak dat verschillende productie gebieden verschillende duurzaamheidsindicatoren vereisen vanwege verschillende lokale duurzaamheidsproblemen.

    - Wordt dit meegenomen in jullie applicatie?

- Waarom wel / niet?

Landbouwpraktijken

- Producten worden beoordeeld op opbrengst per hectare ipv input per product. Wat is de reden hiervoor?

- Worden verschillende landbouwvormen meegenomen in uw beoordeling?
    - Waarom (niet)?

- Hoe worden distikstofmonoxide/lachgas (N2O) veldemissies vastgesteld in uw berekening?

- Worden verschillende landbouwpraktijken zoals irrigatie, ploegen, snoeien en het gebruik van pesticide meegenomen in uw beoordeling?
    - Waarom niet?

- Milieubesparende maatregelen zoals de WKK en/of aardwarmte worden meegenomen in uw beoordeling. Hoe worden deze beoordeeld?
    - Is dit mogelijk op lokaal niveau?
    - Is dit ook toepasbaas buiten Nederland?

Verpakking

- Zover ik kan oordelen, is verpakking erg belangrijk in uw beoordeling. Is deze stap gekwantificeerd? Of is dit een kwalitatieve overweging?

LCA methodologie

- Wanneer er bij bijproducten zijn in een beoordeling (bijv. bietensuiker en pulp), worden deze meegenomen in een allocatieprocedure?
    - Hoe precies?

Algemeen

- Als u uw product/applicatie kan verbeteren, wat zou u dan als eerste willen aanpakken?

- In mijn omgeving hoor ik vaak dat transparantie in labels en beoordelingen ontbreekt. Een voorbeeld met uw applicatie is dat mijn huisgenoot laatst niet geloofde dat een avocado gelabeld werd met een B gradering, zij kon toen niet uitvogelen hoe deze gradering exact tot stand is gekomen. Wat denkt u hiervan?

Outro

- Dank u wel
- Mogelijkheid om resultaten in te zien.

Interview questions (English)

Intro

- Name and background (msc. cultural geography & Environmental sciences)
- Explanation thesis project
  - Vegeterian, but even vegetarian wholefoods are hard to trace
    - So many impact categories, LCAs in stores.
  - Own project -> to develop such metric
  - But so many considerations -> so interest translated into this thesis about food system assessments.
  - Literature review and interviews
- Confidentially
- Time (30 minutes – 1 hour)
- Possibility to record?

Warm Up

- I read some critiques on only assessing products on their CO2 scores, mainly that this might result in shifting of other environmental burdens. Why are you choosing to solely focus on CO2 in your new product?
- I did the climate-KIC journey last summer and we actually had almost the same idea for supermarkets, hence I am writing my thesis about it and I found you. One thing that keeps on bugging me though is, why does this not exist on a large scale already?
- Could you maybe elaborate a little on your role in the project?
  - What is the exact goal of the project may I ask?

Unclear (data)

- Unclear: are all supermarket products included in your platform?
- Unclear: if I understood it correctly, products are assessed on both health and sustainability indicators. Are these available for all supermarket products included?
  - What databases support your platform? Are these LCA databases?
- Is the data somehow validated?

Defining sustainability

- Are the social and economic dimensions of sustainability taken into account in your platform?
  - How exactly?

Indicators

- I noticed that products are assessed on healthy choice, CO2 emissions, water footprint, rainforest degradation, animal welfare, locality and seasonality.
  - Most of the times CO2 emissions and water footprint are included in assessments in the academic literature. However, the other ones you are including are not. May I ask, what was the reason to include:
    - Healthy choice?
    - Rainforest degradation?
    - Animal welfare
      - Is this not countering the effect of CO2 emissions because more land area is needed?

        - Locality?
            - It is often claimed that one should not focus to much on where food comes from, but rather what product one eats. What do you think about this statement?

- It seems that assessments oftentimes have difficulties with dealing with land use and water use (or water stress).
    - How are you dealing with land use?
    - What about water use?

- I noticed that eutrophication, acidification and human/ecotoxicity are often included in assessments. Are these represented in your application?
    - How?

- Soil quality is often described as an important indicator for local sustainability in the literature, yet seldom included in assessments. What is your opinion on this matter?

- Biodiversity is also becoming a larger concern (see IPCC), but also often excluded from assessments. What is your opinion on this?

- The academic literature often suggests that different production areas require different types of indicators being included in assessments due to local sustainability issues.
    - Is this considered in your application?
    - What is your opinion on this?

Farming systems

- I read in your database report that organic and traditional agriculture is taken into account. This is a topic that is often avoided in literature because of the ambiguity if 'organic' is better or not depending if the result is determined by productivity.
    - How do you deal with this within your applications?

LCA methodology

- Which other steps of the food chain are included in your assessment?
    - Why these and not others?
    - Is packaging included in the assessments included in your app?
        - How is packaging included?

General

- Is there a default 'best' option in your application for consumers who want to be sustainable but are simply not aware of all different aspects?
    - If so, are these categories equally weighted in importance?

- If you could improve your product, what would be the first step you would make?

- Purely out of **own interest** on this topic, I would like to continue to work on providing supermarket consumers with more transparent choices, also because it brings together both my scientific backgrounds. Is it possible to maybe have another chat, then not related to my thesis, on a later point in time?

- Thank you

- Insight in thesis if wanted.