BRILL

RESEARCH DATA JOURNAL FOR THE HUMANITIES AND SOCIAL SCIENCES 6 (2021) 1–22

BRILL OPEN

brill.com/rdj

Sustainability Performance of Certified and Non-certified Food

Social and Economic History

Valentin Bellassen (*corresponding author*) | ORCID: 0000-0001-8581-2814
CESAER & AgroSup Dijon & INRAE & Université Bourgogne Franche-Comté, Besançon, France
valentin.bellassen@inrae.fr

Filippo Arfini | ORCID: 0000-0002-5179-2541
University of Parma, Parma, Italy
filippo.arfini@unipr.it

Federico Antonioli | ORCID: 0000-0002-1201-2629
University of Parma, Parma, Italy
federico.antonioli1@gmail.com

Antonio Bodini
University of Parma, Parma, Italy
antonio.bodini@unipr.it

Michael Boehm
ECOZEPT, Montpellier, France
boehm@ecozept.de

Ružica Brečić
Department of Agricultural Marketing, Faculty of Agriculture, University of Zagreb, Zagreb, Croatia
rbrecic@efzg.hr

Sara Chiussi
University of Parma, Parma, Italy
sara.chiussi@gmail.com

Peter Csillag
Department of Agribusiness, Corvinus University of Budapest, Budapest, Hungary
csillag.peter@ecosensus.hu

Michele Donati
University of Parma, Parma, Italy
michele.donati@unipr.it

Liesbeth Dries | ORCID: 0000-0002-1061-1441
Department of Social Sciences, Wageningen University, Wageningen, The Netherlands
liesbeth.dries@wur.nl

Marion Drut
CESAER & AgroSup Dijon & INRAE & Université Bourgogne Franche-Comté, Besançon, France
marion.drut@agrosupdijon.fr

Matthieu Duboys de Labarre | ORCID: 0000-0002-0131-3872
CESAER & AgroSup Dijon & INRAE & Université Bourgogne Franche-Comté, Besançon, France
matthieu.duboys-de-labarre@inrae.fr

Hugo Ferrer
CREDA, Centre for Agro-Food Economy & Development, Catalonia Polytechnic University, Barcelona, Spain
hugo.ferrer@upc.edu

Jelena Filipović
Faculty of Economics and Business, University of Belgrade, Belgrade, Serbia
jfilipovic@ekof.bg.ac.rs

Lisa Gauvrit
ECOZEPT, Montpellier, France
gauvrit@ecozept.com

José M. Gil
CREDA, Centre for Agro-Food Economy & Development,
Catalonia Polytechnic University, Barcelona, Spain
chema.gil@upc.edu

Matthew Gorton | ORCID: 0000-0002-4579-5092
University of Newcastle, Newcastle, United Kingdom
matthew.gorton@ncl.ac.uk

Viet Hoàng
University of Economics Ho Chi Minh City (UEH), Ho Chi Minh,
Vietnam
viet.hoang@ueh.edu.vn

Mohamed Hilal | ORCID: 0000-0001-7679-7026
CESAER & AgroSup Dijon & INRAE & Université Bourgogne
Franche-Comté, Besançon, France
mohamed.hilal@inra.fr

Kamilla Knutsen Steinnes
Consumption Research Norway (SIFO), Oslo Metropolitan University,
Oslo, Norway
kamilla-knutsen.steinnes@sifo.hioa.no

Apichaya Lilavanichakul
Department of Agricultural and Resource Economics,
Faculty of Economics, Kasetsart University, Bangkok, Thailand
apichaya.l@ku.ac.th

Agata Malak-Rawlikowska | ORCID: 0000-0002-1484-0989
Institute of Economics and Finance / SGGW, Warsaw University
of Life Sciences, Warsaw, Poland
agata_malak_rawlikowska@sggw.pl

Edward Majewski | ORCID: 0000-0003-0886-6645
Institute of Economics and Finance / SGGW, Warsaw University
of Life Sciences, Warsaw, Poland
edward_majewski@sggw.pl

Sylvette Monier-Dilhan
US ODR, INRAE, Castanet-Tolosan, France
sylvette.monier@inrae.fr

Paul Muller
BETA, University of Lorraine & University of Strasbourg & CNRS, Nancy, France
paul.muller@univ-lorraine.fr

Orachos Napasintuwong
Department of Agricultural and Resource Economics, Faculty of Economics, Kasetsart University, Bangkok, Thailand
orachos.n@ku.ac.th

Kalliroi Nikolaou
Department of Agricultural Economics, Aristotle University of Thessaloniki, Thessaloniki, Greece
kallirro@agro.auth.gr

Mai Nguyen | ORCID: 0000-0001-8013-2705
UEH, Ho Chi Minh, Vietnam
mainguyen1012@gmail.com

An Nguyễn Quỳnh | ORCID: 0000-0001-7842-4637
UEH, Ho Chi Minh, Vietnam
quynhan1995@gmail.com

Ioannis Papadopoulos
Department of Agricultural Economics, Aristotle University of Thessaloniki, Thessaloniki, Greece
ioannig@agro.auth.gr

Jack Peerlings
Department of Social Sciences, Wageningen University, Wageningen, The Netherlands
jack.peerlings@wur.nl

Aron Török | ORCID: 0000-0001-6769-7103
Department of Agribusiness, Corvinus University of Budapest, Budapest, Hungary
aaron.torok@gmail.com

Thomas Poméon | ORCID: 0000-0002-3352-9117
US ODR, INRAE, Castanet-Tolosan, France
thomas.pomeon@inrae.fr

Bojan Ristic
Faculty of Economics and Business, University of Belgrade, Belgrade, Serbia
risticbj@gmail.com

Burkhard Schaer
ECOZEPT, Montpellier, France
schaer@ecozept.com

Zaklina Stojanovic
Faculty of Economics and Business, University of Belgrade, Belgrade, Serbia
zaklina@ekof.bg.ac.rs

Barbara Tocco | ORCID: 0000-0003-0072-559X
University of Newcastle, Newcastle, United Kingdom
barbara.tocco@newcastle.ac.uk

Marina Tomic Maksan
Department of Agricultural Marketing, Faculty of Agriculture, University of Zagreb, Zagreb, Croatia
matomic@agr.hr

Mario Veneziani | ORCID: 0000-0001-6228-9514
University of Parma, Parma, Italy
mario.veneziani@unipr.it

Gunnar Vitterso
SIFO, Oslo Metropolitan University, Oslo, Norway
gunnar.vitterso@sifo.hioa.no

Abstract

The dataset *Sustainability performance of certified and non-certified food* (https://www.doi.org/10.15454/OP51SJ) contains 25 indicators of economic, environmental,

and social performance, estimated for 27 certified food value chains and their 27 conventional reference products. The indicators are estimated at different levels of the value chain: farm level, processing level, and retail level. It also contains the raw data based on which the indicators are estimated, its source, and the completed spreadsheet calculators for the following indicators: carbon footprint and food miles. This article describes the common method and indicators used to collect data for the twenty-seven certified products and their conventional counterparts. It presents the assumptions and choices, the process of data collection, and the indicator estimation methods designed to assess the three sustainability dimensions within a reasonable time constraint. That is: three person-months for each food quality scheme and its non-certified reference product. Several prioritisations were set regarding data collection (indicator, variable, value chain level) together with a level of representativeness specific to each variable and product type (country and sector). Technical details on how relatively common variables (e.g., number of animals per hectare) are combined into indicators (e.g., carbon footprint) are provided in the full documentation of the dataset.

Keywords

sustainability performance – economic performance – environmental performance – social performance – certified food – protected designation of origin – protected geographical indication – organic farming

Online publication date: 13-12-2021

– Related data set "Sustainability performance of certified and non-certified food" with DOI www.doi.org/10.15454/OP51SJ in repository "Data INRAE"

1. Introduction and Research Problem

EU and national food quality policies have witnessed recent reforms. In 2007, the EU agreed on a new Council Regulation (Council Regulation (EC) No. 834/2007) setting out the principles, aims, and overarching rules of organic production and defining how organic products were to be labelled. In 2012, the Quality Package (Regulation (EU) No. 1151/2012) was passed to improve and promote the operation of schemes to protect Geographical Indications (GIs) for agri-food products. The Regulation details the rationale for establishing/

promoting GIs as a means to generate a fair return for farmers and producers for the qualities of particular goods and to enable consumers to make better-informed purchasing decisions through effective labelling. The diversity and quality of EU agricultural and fisheries production are one of its main strengths in both domestic and international markets. Supporting Food Quality Schemes (FQSs) – here understood as Protected Designation of Origin (PDO), Protected Geographical Indications (PGI) and organic products – is thus regarded as consistent with Europe 2020 policy priorities for 'sustainable and inclusive growth', which seek to achieve competitive and high employment economies (economically sustainable) delivering social and territorial cohesion (socially sustainable), while paying attention to the burden placed on the environment and natural resources (environmentally sustainable). But are FQSs really more sustainable than other food products?

To answer this question and as part of the H2020 Strength2Food project, we gathered raw data on 54 food value chains spanning over 13 countries. The sampling design is paired: 27 certified – PDO, PGI, or organic – products and 27 reference products (products similar to the certified value chain but not certified). This raw data allows for the estimation of 25 performance indicators covering the three sustainability pillars: 9 economic indicators, 7 environmental indicators and 9 social indicators.

2. Methods

Disclaimer: being a summarized description of the method used to estimate sustainability indicators, this article largely draws from two existing documents from the same authors: Bellassen et al. (2019) and Bellassen et al. (2016). More technical details on the Methods are available in the data repository.

– **Sustainability Performance of Certified and Non-certified Food deposited at Data INRAE – DOI:**www.doi.org/10.15454/OP51SJ

2.1. *General Points on Indicators and Their Analysis*

2.1.1. Overview of Indicators and Minimal Systematic Comparison

The choice of indicators was made on the basis of the SAFA methodology (Sustainability Assessment of Food and Agriculture systems) developed by the Food and Agriculture Organization of the United Nations (2013) to measure the sustainability of food production. SAFA provides guidelines on how to consider each sustainability dimension, including which indicators could be

relevant and useful indications on how to implement them. SAFA, however, is primarily focused on processing firms and stops short of formulating a complete method which goes from primary data collection to indicator estimation and interpretation.

The indicators presented in this document operationalise a subset of SAFA indicators, complementing them along the following three lines:

– Most SAFA indicators cannot be directly implemented from the SAFA indicators report. They require the definition of specific data to be collected and calculation or aggregation methods which are not explicated in the report, although the report sometimes refers to existing tools for doing this. Our method defines all necessary data and variables, and provides associated calculators or aggregation methods, together with a data storage and source traceability system.
– Because they were designed to be collected for a single firm, many SAFA indicators require a substantial amount of data. This makes it difficult to cover more than a few indicators for an entire value chain within 3 person-months. Our method simplifies indicators by prioritising data collection on the key drivers of the indicators, by providing default values for many non-key but necessary variables and, where necessary, by restricting the scope of an original SAFA indicator down to the scope for which data is most accessible. As a result, it is possible in most cases to estimate 25 sustainability indicators across the three sustainable development pillars for both a specific product produced by several firms and a generic reference product in 3 person-months.
– Finally, several SAFA indicators rely only on the subjective views of specific stakeholders. Where stakeholder views are a necessary part of the indicator (e.g. bargaining power distribution), our indicators combine stakeholder views with objective data.

To make the collection of information and the subsequent analysis on the twenty-seven case studies efficient, operational choices were made concerning the type of indicators and their management. One of the most important choices is the distinction between "systematic indicators" which should be computed on all case studies and "complementary indicators" which concern only a subset of case studies, often based on data availability. There was a total of 13 systematic indicators (four economic; four environmental; five social), and a total of eleven complementary indicators (five economic; three environmental; four social). Around 150 variables were collected and refined into the 25 indicators (see Table 1).

TABLE 1 List of indicators for sustainability assessment

	Sustainability pillar	Indicator type	Indicator sub-type (code)	Level of analysis along the value chain
Systematic	Economic	Price premium	Price premium (Ec1.1, EUR kg^{-1})	One value per level of the value chain
		Profitability and value added distribution	Gross Operating Margin (Ec1.3, % of turnover)	
		Trade	Share of value exported within Europe (Ec1.5, % of turnover)	Single value for the whole value chain
		Local multiplier	Local multiplier (Ec2.1, no unit)	
	Environmental	Carbon footprint	Carbon footprint per unit of product (En1.1, kgCO_2e ton^{-1})	Single value for the whole value chain
		Foodmiles	Distance travelled per unit of product (En2.1, ton.km ton^{-1})	
		Water footprint	Blue water footprint (ground water consumption, En 3.3, m^3 kg^{-1}) Grey water footprint (water pollution by nitrates, En 3.2, m^3 kg^{-1})	
	Social	Employment	Labour to production ratio (So 1.1, AWU ton^{-1})	One value per level of the value chain

TABLE 1 List of indicators for sustainability assessment (*cont.*)

	Sustainability pillar	Indicator type	Indicator sub-type (code)	Level of analysis along the value chain
		Governance	Bargaining power distribution (S02.1, no unit)	Single value for the whole value chain
		Social capital	Educational attainment (S03.1, no unit) Generational change (S05.1, no unit) Gender equality (S05.2, no unit)	One value per level of the value chain
Complementary	Economic	Profitability and value added distribution	Gross Value-added (Ec1.2, % of turnover)	One value per level of the value chain
		Profitability and value added distribution	Net result (Ec1.4, % of turnover)	
		Trade	Share of value exported outside Europe (Ec1.6, % of turnover) Share of volume exported within Europe (Ec1.7, % of production) Share of volume exported outside Europe (Ec1.8, % of production)	Single value for the whole value chain

TABLE 1 List of indicators for sustainability assessment (*cont.*)

	Sustainability pillar	Indicator type	Indicator sub-type (code)	Level of analysis along the value chain
	Environmental	Carbon footprint	Carbon footprint per hectare (En1.2, $kgCO_2e\ ha^{-1}$)	One value per level of the value chain
		Foodmiles	Emissions from transportation per unit of product (En2.2, $kgCO_2e\ ton^{-1}$)	
		Water footprint	Green water footprint (total – rain + groundwater – water requirements, En3.1, $m^3\ kg^{-1}$)	
	Social	Employment	Turnover to labour ratio (So1.2, EUR AWU^{-1})	One value per level of the value chain
		Governance	Stability of the value chain level (So2.2, no unit)	
		Social capital	Wage level (So3.2, EUR AWU^{-1}) Gender equality index (So5.3, no unit)	

2.1.2. Relative Difference and Value Chain Averages

Indicators are estimated at each level of the value chain (farm level, processing level and, where relevant, retail level). To control for country and product specificities, we analyse relative differences between the FQS and its reference product rather than absolute values.

Equation (1) is used, where rel_diffj is the relative difference for an indicator at level j of the value chain, and $indic_{FQS,j}$ and $indic_{FQS,j}$ are the indicator value at level j of the value chain for the FQS and the reference product respectively.

$$rel\_diff_j = \frac{indic_{FQS,j} - indic_{REF,j}}{indic_{REF,j}} \tag{1}$$

For environmental indicators and for bargaining power distribution, the opposite of the relative difference is used in the analysis so that a positive difference consistently indicates higher performance of FQS (e.g., more added value, more employment, lower carbon footprint).

In a second step, to assess the difference in performance for the entire value chain, we compute aggregated values or "value chain averages as shown in equation (2):

$$\begin{cases} \text{All indicators except environmental indicators and labour to production: } {}^{VC}_{average} = \dfrac{\sum_{j=1}^{n} rel\_diff_j}{n} \\ \text{Environmental indicators: } {}^{VC}_{average} = rel\_diff_n \\ \text{Labour to production ratio:} \begin{cases} {}^{VC}_{average} = \dfrac{cum_{indic_{FQS}} - cum_{indic_{REF}}}{cum_{indic_{REF}}} \\ cum_{indic_X} = \dfrac{indic_{x,farm}}{final_{prod_{ratio}} \times \left(1 + coproducts_{farm}\right)} + \dfrac{indic_{x,proc}}{\left(1 + coproducts_{proc}\right)} \end{cases} \\ \text{Value chain stability:} \begin{cases} {}^{VC}_{average} = \dfrac{cum_{indic_{FQS}} - cum_{indic_{REF}}}{cum_{indic_{REF}}} \\ cum_{indic_X} = min\left(indic_{x,farm}, indic_{x,proc}\right) \end{cases} \end{cases} \tag{2}$$

For most indicators, these aggregated values are simply averages across value chain levels for which the indicator could be estimated (farm, processing and, where relevant, retail). There are, however, two exceptions.

The first exception concerns indicators expressed on a *per ton* basis, that is the environmental indicators and the labour to production ratio. Because these indicators follow a life cycle assessment logic, and in particular because they use a functional unit (one ton of product), aggregated values over the value chain must be calculated cumulatively. If one ton of cheese requires 10 tons of milk, the aggregated indicator sums the footprint of 10 tons of milk at farm level and 1 ton of cheese at processing level rather than averaging the footprints of one ton of milk and one ton of cheese. This cumulative process also allocates the footprint to all products (e.g., milk and meat at farm level) based on their relative economic value. For environmental indicators, this is already done in the estimation of the indicator. For labour to production ratio, the formula is provided in equation (2).

The second exception concerns the indicator on value chain stability for which the aggregated value is the minimum across value chain levels. In equation (2), the $VC_{average}$ is the aggregated performance difference for the entire value chain, rel_diffj is the relative difference in performance at level j of the value chain (see equation (1)), n is the lowest level of the value chain where the indicator could be estimated (most often the processing level), cum_indicX is the cumulative indicator over different value chain levels for product X (either FQS or reference), $indic_{X,farm}$ and $indic_{X,proc}$ are the indicator value for product X at the farm and processing levels respectively, final_prod_ratio is the amount of raw product at farm level (e.g., milk) necessary for one ton of final product (e.g., cheese), and coproducts_farm and coproducts_proc are the value of coproducts (e.g., meat) expressed as a percentage of the value of the main product (e.g., milk) at farm and processing levels respectively.

2.2. *Reference, Data Collection and Metadata Documentation*

2.2.1. Selection of a Reference Product/Case: Elements of Guidance

To provide a basis for comparison, each sustainability indicator has been estimated for the same product category (for example cheese) in two different value chains: specific quality (organic or geographical indication) and generic quality (reference product). To define the reference, the following guidance, composed of two objectives and three constraints, was applied. The two objectives are:

– comparability of contexts: the two cases (food quality scheme and its standard reference) should be produced in territorial contexts (in terms of location) as similar as possible;

– comparability of the products: the two products/basket of products (food quality scheme and their standard reference) should be as comparable as possible.

These objectives should be sought until one of the three following constraints are met:

– data resolution limit: data for the reference are only available at a larger scale than for the case studied;
– confusion of the case and its reference: for example, for an apple under geographical indication (GI), the reference would ideally be the production of "standard" apples in the same area. Nevertheless, if almost all the apple production of that area is under GI, a reference should be chosen at a larger scale (regional or even national scale);
– the case studied is the only one of its type: with the example of an apple under GI, the ideal reference would be a standard apple of the same variety. Nevertheless, as mentioned for geographic scale, data may be scarce at this detailed level (variety), or even all the apples of this variety may be sold under GI. In this case, a suitable reference would be one, or a mix of, the main varieties.

In practice, the choice of a relevant reference by case study conductors will strongly depend on data availability, so that a national average can be used if a more suited reference cannot be documented. Moreover, a mix of specific references and national averages can be used. For example, looking at the Comté cheese, some variables (e.g., price of milk, price of cheese, …) may be specific to Emmental, a non-certified ripened, hard, cow-milk based cheese, while national averages are used for other variables (e.g., quantity of mineral fertilizer per hectare, share of exports over total production, …) for which Emmental-specific data are not readily available.

Note that the use of the reference is primarily to interpret the results from the case, so even if the reference presents some peculiarities, this can be accounted for in the discussion of results. An extreme case of such peculiarities is Sjenica cheese in Serbia. Because it is almost the only sheep cheese produced in the country, the reference product is a conventional cow cheese. But as a result, many differences between Sjenica cheese and its reference are better explained by the difference between sheep and cow than by the technical specifications or the terroir of Sjenica cheese. For this reason, Sjenica should be excluded from most cross-comparisons. To the contrary of many performance assessments, we thus opted for real *relative references* as opposed

to *normative references*, that is references which correspond to fictive cases or to targets to be reached (Acosta-Alba & Van der Werf, 2011).

2.2.2. Data Collection

2.2.2.1. *Value Chain Diagram*

The first step in data collection was to identify the firms which belong to the value chain (see the box text for the criteria) and to classify them into different levels (e.g., farm level, processing level, retail level). This first step resulted in a value chain diagram which is inserted in the second sheet of each data file and provides the code of each value chain level.

Box. Criteria Used to Identify Which Firms Belong to the Value Chain

When firms are making only part of their turnover from the FQS product (– e.g., a freezing plant which is freezing and packaging all kind of fruits, including the FQS (organic raspberries) – criteria are needed to determine whether they belong to the FQS value chain. The key recommended criterion is that the firm makes at least 50% of its turnover from the FQS product. As such, most firms at retail level will be excluded. However, a few systematic or ad hoc exceptions are made:)

- The retail level is included for two economic indicators, namely price premium and export;
- A firm/value chain level can be retained on an ad hoc basis when its impact on an indicator is substantial (e.g., impact of freezing on the carbon footprint of frozen raspberries);
- A firm/value chain level can be retained on an ad hoc basis when stakeholders consider it as part of the value chain despite it making less than 50% of its turnover from the product.

2.2.2.2. *Two Angles of Prioritisation*

Two distinctions were made to convey a sense of priority for data collection:

- *Systematic versus complementary indicators*: systematic indicators were to be computed for all case studies while complementary ones could be restricted to a subset of particularly interesting cases;
- *Key versus secondary variables*: a reasonable approximation of the indicator can be obtained from key variables data while obtaining values for secondary variables would create even more precise estimates.

In other words, most of the data collection/gathering effort should be spent on *key variables which contribute to systematic indicators*, while the rest should only be provided if data is readily available, and should not be the object of a dedicated data collection effort.

2.2.2.3. *Relying on Existing Sources of Information*

In general, given the resource and time constraints, most variables were designed to be common enough to be obtained from existing studies, reports and databases. A good strategy for a comprehensive overview of existing sources may be to conduct a few (3–5) interviews with key stakeholders in the chosen case study's value chain.

2.2.2.4. *Default Values*

In parallel to case-by-case data collection, an effort was made to obtain national average values for as many variables as possible, and cover all the sectors studied (dairy, meat products, seafood/fish, cereals, fruits & vegetables). These values do not refer to specific products but to larger product categories which can be identified in systematic surveys. For this purpose, databases with pan-European coverage, such as the Farm Accountancy Data Network (FADN) and different surveys and datasets available via Eurostat database (i.e., Farm Structure Survey, Structural Business Statistics, Labour Force Survey, etc.) have been explored.

These default values were used in three different manners:

- to check that the collected data for the case and/or its reference is of a reasonable order of magnitude;
- to estimate indicators for a "national average" reference product;
- to save time on data collection when there is evidence (e.g., expert judgement) that a given variable is not significantly different from the national average.

This last option was infrequently used and, in all cases, data sources for each variable and product are transparently documented in the data repository.

2.2.3. Quality Checks in Data Collection and Indicator Estimation

2.2.3.1. *Principles*

Considering the scale and the complexity of data collection (measuring the sustainability level of 54 products using 25 indicators referring to the environmental, economic and social dimensions of sustainability), an organizational model was developed. This thorough quality check procedure was

implemented to limit the risk of misreporting data. The three key aspects of this procedure were 1) to record all data, their date and source in a shared spreadsheet, 2) to separate the person who collected data from the person who estimated the indicator, and 3) to come up with a written and consensual interpretation of the results between these people.

The most important principle of the procedure for data collection and indicator estimation is an early and repeated interaction between the case study conductor and the indicator coordinator (see Figure 1). The case study conductor is responsible for collecting the data and ensuring its traceability, which implies creating a repository with all source files and intermediary calculations. The indicator coordinator is responsible for the quality check of the data provided (e.g., verifying, together with the case study conductor, the source when an order of magnitude seems wrong, etc.) and for providing the case study conductor with the estimated indicator(s). Both are responsible for interpreting the results. Results are considered valid only when both the case study conductor and the indicator coordinator agree that the estimated resulting indicators are plausible and, should a large difference occur between the certified product and its conventional reference, that they have plausible ways of explaining this large difference.

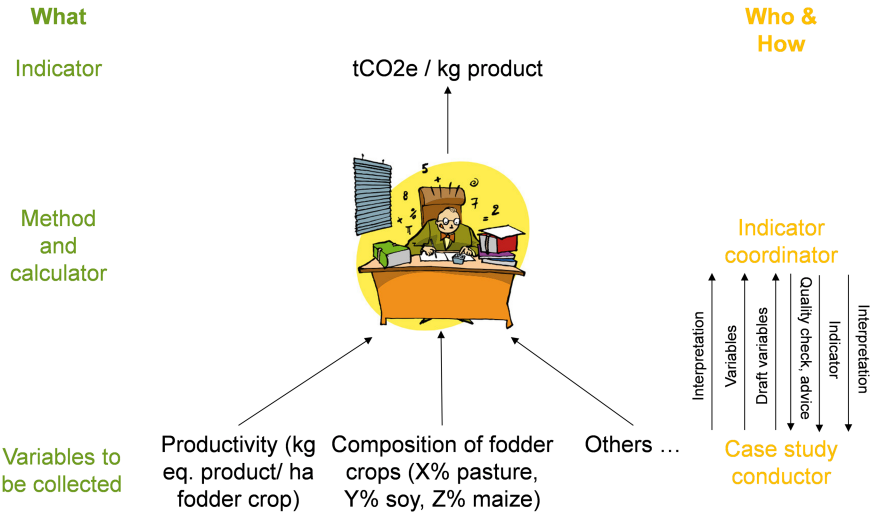

FIGURE 1 Organization of data collection and indicator estimation and interpretation

Those conducting case studies received initial guidance and tips for data collection, and regular online meetings were organized to share data collection practices and problems and thus ensure consistency across case studies.

2.2.4. Metadata Documentation

For each variable value, two metadata items were documented:

- the source/reference for the values (e.g., "Dupond et al., 2010");
- to which time period the variable's values correspond. Time periods should be as recent as possible, and to the extent possible, similar between different variables. When relevant and available, time-series and/or multi-year averages can be used.

In addition, all original documents from which the data are sourced and the intermediary calculations (e.g., excel or word documents) have been stored in an online repository, so that both the case study conductor and the indicator coordinator can go back to them easily to double-check some values or interpret the results.

2.3. *Description of Indicators, Their Purpose and Their Estimation Method*

The exhaustive list of the raw data collected and the technical details of the method to estimate the 25 performance indicators based on this raw data are provided in the full documentation posted in the data repository.

Table 2 sets out the sample characteristics where the sectors are highlighted (red, green and blue lines for animal, vegetal and seafood respectively). The indicated turnover is either at processing or farm level, whichever is higher. Arfini & Bellassen (2019) provides a detailed description of each value chain, its structure, its governance and its sustainability performance.

As a result of the quality check procedure described in 2.2.3, the applicant-PGI Sjenica sheep cheese was removed from the sample: its reference product is a cow cheese, and the difference between cow and sheep was identified as the main drivers for the differences in performance. The procedure also resulted in the exclusion of employment indicators at processing level for PGI Doi Chaang coffee and PGI TKR Hom Mali rice, for which differences between certified food and its reference were both high and unexplained.

TABLE 2 Sample characteristics with sectors highlighted

Product name	Country	Product description	Type of FQS	Processed?	Turnover of FQS (M€ yr-1)	Reference product
Dalmatian prosciutto	Croatia	Dry pork ham	PGI	Yes	4.20	Local non-PGI firm
PDO olive oil	Croatia	Olive oil	PDO	Yes	0.25	National average
Comte cheese	France	Hard pressed cooked cheese from cow milk	PDO	Yes	504.19	Similar uncertified cheese (Emmental) or national average (cow cheese)
Organic flour	France	Wheat flour	Organic	Yes	34.80	National average
Saint-Michel bay bouchot mussels	France	Mussels produced on "bouchots"	PDO	No	25.45	National average (TSG Bouchot mussels)
Organic rice	France	Rice	Organic	Yes	17.64	Non-organic rice (mostly PGI)
Organic pork	Germany	Raw meat	Organic	Yes	69.00	National average
Organic yoghurt	Germany	Organic yoghurt from cow milk	Organic	Yes	387.00	National average
Zagora apples	Greece	Apple	PDO	No	10.11	Kissavos apples (non-GI apples from another region)
Kastoria apples	Greece	Apple	PGI	No	7.50	Kissavos apples (non-GI apples from another region)
Gyulai sausage	Hungary	Sausage	PGI	Yes	55.00	Non-PGI Hungarian sausage
Kalocsai paprika powder	Hungary	Paprika powder	PDO	Yes	10.75	Imported Chinese pepper milled in Hungary
Parmigiano Reggiano cheese	Italy	Hard pressed cooked cheese from cow milk	PDO	Yes	1,009.94	Biraghi cheese (similar non-PDO cheese)

TABLE 2 Sample characteristics with sectors highlighted (*cont.*)

Product name	Country	Product description	Type of FQS	Processed?	Turnover of FQS (M€ yr-1)	Reference product
Lofoten stockfish	Norway	Dried fish	PGI	No	71.24	Clipfish (cod)
Organic salmon	Norway	Salmon	Organic	Yes	144.71	Conventional salmon
Organic pasta	Poland	Pasta	Organic	Yes	0.52	Simulated conventional farms with sample characteristics
Kaszubska strawberries	Poland	Strawberry	PGI	No	0.64	National average
Sjenica cheese	Serbia	Sheep cheese	PGI	Yes	1.21	National average (cow cheese)
Organic raspberries	Serbia	Frozen rapsberries	Organic	Yes	4.37	National average
Sobrasada Porc Negre	Spain	Raw, cure saussage from pork meat	PGI	Yes	1.80	National average
Ternasco de Aragon	Spain	Unprocessed lamb meat	PGI	No	16.97	Non-PGI lamb in the same region (Aragon)
Thung Kula Rong-Hai Hom Mali rice	Thailand	Rice	PGI	No	300.74	Non certified rice from the same region (90% of GI rice is organic as well)
Doi Chaang coffee	Thailand	Coffee	PGI	Yes	756.00	Non-PGI coffee from the same province
Buon Ma Thuot coffee	Vietnam	Coffee	PGI	Yes	89.58	Non-PGI coffee from Dak Lak province in Vietnam

Note: Red, green and blue shading denotes animal, vegetal and seafood sectors respectively. The indicated turnover is either at processing or farm level, whichever is higher.

3. Reuse Potential

We anticipate two major avenues for reuse of the dataset *Sustainability performance of certified and non-certified food*. First, the analysis we have made so far and submitted to academic journals is not exhaustive. Sector-specific or standard-specific analysis are only sketched and have not been synthesized. Other cross-comparisons could also be envisaged, based on geographical or cultural proximity for example, and more systematic sensibility studies could be performed.

Second, other food value chains may be willing to assess their sustainability using the same method. Having access to our detailed dataset will allow them to better understand the method – seeing it applied to a broad set of examples – and to undertake detailed quality checks (e.g., identifying outliers in both raw data and indicator values). Finally, should such an assessment be conducted, we encourage these fellow researchers to enrich the dataset by sending us their results. Such a virtuous cycle could, over time, lead to further interesting analyses, such as intertemporal comparisons or testing past results on a higher sample.

Acknowledgements

This research was funded by the European Commission under the H2020 Strength2Food project, grant number 678024.

References

Acosta-Alba, I., & Van der Werf, H. (2011). The use of reference values in indicator-based methods for the environmental assessment of agricultural systems. *Sustainability*, 3(2), 424–442. https://www.doi.org/10.3390/su3020424.

Arfini, F., & Bellassen, V. (Eds.). (2019). *Sustainability of European food quality schemes: Multi-performance, structure, and governance of PDO, PGI, and organic agri-food systems*. Springer International Publishing. https://www.doi.org/10.1007/978-3-030-27508-2.

Bellassen, V., Antonioli, F., Bodini, A., Donati, M., Drut, M., Duboys de Labarre, M., Hilal, M., Monier-Dilhan, S., Muller, P., Poméon, T. & Veneziani, M. (2019). Common methods and sustainability indicators. In F. Arfini & V. Bellassen (Eds.). *Sustainability of European food quality schemes: Multi-performance, structure,*

and governance of PDO, PGI, and organic agri-food systems (pp. 23–45). Springer International Publishing. https://www.doi.org/10.1007/978-3-030-27508-2_2.

Bellassen, V., Giraud, G., Hilal, M., Arfini, F., Barczak, A., Bodini, A., Brennan, M., Drut, M., Duboys de Labarre, M., Gorton, M., Hartmann, M., Majewski, E., Muller, P., Monier-Dilhan, S., Poméon, T., Tocco, B., Tregear, A., Veneziani, M., Vergote, M.-H., Vitterso G., Wavresky, P. & Wilkinson, A. (2016). *Methods and indicators for measuring the social, environmental and economic impacts of food quality schemes, short food supply chains and varying public sector food procurement policies on agri-food chain participants and rural territories* (Strength2Food project, deliverable 3.2). INRA.

Food and Agriculture Organization of the United Nations. (2013). *SAFA indicators*.

Sustainability performance of certified and non-certified food (2021). [Dataset]. https://data.inrae.fr/dataset.xhtml?persistentId=doi:10.15454/OP51SJ.