Supervision

ISARA-Lyon
23 Rue Jean Baldassini
69007 Lyon
France

Wageningen University and Research
6708 PB Wageningen,
The Netherlands

Host organisation

Center for Sustainable Food Systems
3461 Ross Drive
Vancouver, BC Canada V6T 1Z4
Canada

INTEGRATED SUSTAINABILITY ASSESSMENT FOR SMALL-SCALE DIVERSIFIED FARMS: A PILOT STUDY AT THE UBC FARM, VANCOUVER, CANADA.

Master Thesis
16th Cohort Agroecology (2017-2019)

Adrien Kroese
WUR nb: 961205479130

Tutor ISARA: Dr. Alexander Wezel
Tutor WUR: Dr. Ir. Tjeerd-Jan Stomph
External Tutor: Dr. Hannah Wittman

Defended September 25th, 2019.

Integrated sustainability assessment for small-scale diversified farms: a pilot study at the UBC Farm

Adrien H. Kroese[abc]*, Alexander Wezel[a], Tjeer-Jan Stomph[b], Hannah Wittman[c], Mollie McDowell[c], Jules Savès.

[a] *ISARA-Lyon*

[b] *Wageningen University and Research*

[c] *The University of British Columbia, Center for Sustainable Food Systems, Vancouver, Canada.*

Adrien Kroese is a master’s student at WUR and ISARA-Lyon in Organic Agriculture. His research interests include sustainable agriculture and how components of the food system interact creating synergies and trade-offs.

Dr. Alexander Wezel

Dr. Tjeerd-jan Stomph

Dr. Hannah Wittman

Mollie McDowell

The author states that there is no conflict of interest arising from this research.

Data from this research is available at:

Keywords: sustainability assessment; organic agriculture; farm level; systemic; RISE; SAFA; IDEA; MESMIS.

Integrated sustainability assessment for small-scale diversified farms: a pilot study at the UBC Farm

Abstract

The present research took place at the UBC Farm, a diversified small-scale vegetable farm. It aimed at (i) comparing the applicability and validity of different sustainability assessment tools applied to small-scale diversified farms and (ii) to assess the degree of sustainability of the case-study farm. The appropriate method should be the best compromise between feasibility and scientific accuracy. Five (5) sustainability assessment frameworks were selected and applied to the case-study farm, to cover both the three pillars of sustainability (Environmental, Social and Economic) and the properties of sustainability (Productivity, Adaptability, Resilience & reliability, Equity, and Self-reliance). The results were that (i) the farm showed positive to best sustainability results in all but economic viability. Second, (ii) common assessment methods require meticulous data collection and are not accessible to all small-scale growers with the current data infrastructure. Third (iii), assessment frameworks give a broad description of the farming system through numerous indicators, covering all three pillars of sustainability. In addition, insights are valuable for researchers in a context of mapping and assessment. However, the issues reflected in the indicators with the lowest scores were already known to the farmers. Finally, (iv) the frameworks have different benchmarks and interpretations for the same indicator, causing important variations in results, and making indicator validation and benchmarking a crucial step.

Acknowledgements

As an MSc student, I would like to thank the Centre for Sustainable Food Systems for entrusting me with the development of a topic that matters to their research. I would like to sincerely thank all the people; professors, researchers, technicians, students, staff and farmers who have made this thesis possible. I would like to thank my supervisors with inspiring systems thinking and pedagogy, who have guided me along the way, providing (in)valuable input and support. The people who collaborated on this work also deserve my sincere gratitude. Finally, I would like to thank all the community that provided publications, knowledge or software giving much matter to this research.

Navigation

Contents

Abbreviations

Development FAO: Food and Agriculture Organisation of the United Nations SDG: Sustainable Development Goal
Sustainability assessment SA: Sustainability Assessment EIA: Environmental Impact Assessment ISA: Integrated Sustainability Assessment FSA: Full sustainability assessment RSA: Rapid Sustainability Assessment
IFrameworks DEA: *Indicateurs de Durabilité des Exploitations agricoles* [Sustainability indicators of Farming Operations]

MESMIS: *Marco para Evaluación de Sistemas de Manejo de Recursos Naturales Incorporando Indicadores de Sustentabilidad* [Framework for the Evaluation of Natural Resource Management Systems Incorporating indicators of Sustainability]. RISE: Response-inducing Sustainability Evaluation, SAFA: Sustainability Assessment of Food and Agriculture
Other MO: Organic Matter SWOT: Strengths, Weaknesses, Opportunities and Threats analysis

Important definitions

Integration: the application and interaction of several scientific disciplines in a study (Rothman 1998 cited in Bezlepkina 2010) (e.g. social studies, economy, agronomy, ecology and impact assessment in sustainability assessment).

Systems view: consideration that the object or system under study comprises sub-systems and components that interact to create a complex outcome. The system cannot be solely explained by the disaggregation of its components (Gray, 2018).

Sustainability assessment/evaluation: measurement, estimation or appraisal of the level of sustainability of a defined system.

Strong sustainability: school of thought assuming that natural capital is a heritage irreplaceable by human-made capital.

Weak sustainability: conversely, school of thought assuming full substitutability between natural and human-made capital as long as equilibria are not broken and that the total capital and its productivity does not decrease (Bond and Morrison-Saunders 2011).

Outer sustainability: property of a system that does not produce negative externalities; it does not decrease its environment's (social, economic and natural) resource base.

Inner sustainability: capacity of a system to function perennially; to produce a reliable output in the long-term.

1. Introduction

1.1. Objectives

Many agricultural sustainability assessment frameworks have been developed, as recalled by de Olde, Sautier and Whitehead (2018), with varying degrees of integration and adapted to different farming systems. Fewer frameworks target small-scale diversified farms, especially vegetable growers of post-industrial contexts. The present research aims at determining an appropriate means of assessing sustainability for the latter production systems, with the UBC Farm as a case-study. UBC Farm is a diversified pluri-functional organic farm conducting research activities, producing food and connecting communities through social programmes (Indigenous food sovereignty group, farmer's market, annual events) on the traditional, ancestral and unceded territory of the Musqueam people. The UBC Farm researches "the impacts of food production, transformation and consumption on environmental and community health" (UBC accessed 2019); it seeks to be a living laboratory for food sovereignty and strives for sustainable food and services production. To build on the projects assessing components of sustainable food systems (e.g. greenhouse gasses, value-chain fairness, local dynamics), there was a need for a full assessment of the farm's sustainability, particularly its performance on several indicators and their interactions. The UBC Farm being a complex system, we build on the assumption that a tool able to assess the UBC Farm will be transposable to other production systems.

The research questions are summarised as:

- *How sustainable is the UBC Farm?*
- *How can the sustainability of diversified production systems be effectively measured, in regard to feasibility and scientific accuracy?*
- *How do different frameworks define and measure sustainability, and how do they vary in their results?*

For this purpose, five (5) sustainability assessment tools were selected and used with the data from the UBC Farm from 2018. The first 4 frameworks enabled to assess results consistency. The addition of the 5th serves to determine the feasibility of each method.

1.2. Operational definition of sustainability

Central to sustainability assessment is the need for an operational definition of sustainable agriculture (Binder, Feola, and Steinberger 2010; Grenz et al. 2016). The sense of urgency (Gold 2007), the implication of long-term reliability and the consideration of social, economic and environmental dimensions (triple bottom-line) are widely agreed upon in sustainability science (Bond and Morrison-Saunders 2011). However, how these principles are operationalised into practices and assessed in studies remains highly debated or unclear (Weaver and Rotmans 2006; Pope & al. 2017; Rigby & Caceres 2001). As a result, assessment methods differ per context. Based on our literature review, we take a dual approach to sustainability, described by Zahm et al. 2019, based on both the properties of sustainable systems (*how does a systems with defined boundaries function sustainably?*) and the mandate of sustainable agriculture (*how does the system contribute to larger-scale sustainability?*).

Firstly, sustainable systems are defined by López-Ridaura et al. (2005) by 5 properties or attributes, namely *productivity, stability, reliability, resilience* and *adaptability* (referred to here as *inner sustainability*). Desirable levels of these properties enable the system to function on the long run and withstand disturbances. Although contribution to sustainability beyond the farm gates is implicit in the authors' and others' case studies (López-Ridaura et al. 2002 & 2005; Ripoll-Bosch et al. 2012), depending on the scale of assessment (e.g. at farm gates or value-chain), a system might function perennially while producing externalities. Moreover, these principles lie in the field of weak sustainability as they assume substitutability between natural and human-made capital (see George 1999) as long as the system is not threatened. Therefore, and second, the attributes are consolidated with the contribution of the system to *economic viability*, *social and human integrity*, and *environmental preservation* (referred to as *outer sustainability*). However, we consider that economic sustainability (inner and outer) is not an end per se but a way of assessing the farm's viability and its contribution to livelihoods.

Finally, we take the position of Briquel et al. 2001 and Zahm et al. 2008 that sustainability is a non-compensatory measure, or in assessment terms that a system is as sustainable as its least sustainable essential component. If a farm shows a good performance on the environmental domain but its economic viability is at threat, it cannot be considered fully sustainable and the environmental performance will not balance the economic vulnerability. Underlying indicators, however, can compensate one-another within a domain.

1.3. Integrated sustainability assessment of agriculture

1.3.1. General use, applicability

Agricultural production has led to environmental and social pressure and transition towards more sustainable production systems is required (Sydorovych & Wossink, 2007). Despite the consensus on the need for more agricultural sustainability, translating concepts into practices remains challenging. ISA aims exactly at filling this think-do gap (van der Werf & Petit, 2002).

ISA can help identify and implement sustainable practices. As such, ISA is an important tool at the policy level, enabling a comparison of the expected outcomes of different scenarios (Bertocchi, Demartini, & Marescotti, 2016). Sustainability assessments are regularly implemented in private sector value chains and on large agricultural operations (FSA accessed 2019). However, we see fewer on-farm uses, especially small-scale growers, despite the ability to reveal and favour best practices (Pope et al. 2017, Triste et al. 2014).

ISA has emerged from a tradition of impact assessment (IE). IEs focus on one specific sustainability domain (Binder, Feola, & Steinberger, 2010) while ISA advocates for integration of all pillars. Recent papers have seen an increase in epistemological considerations for ISA after calls to integrate systems views and trade-off analysis order to give a more accurate measure of sustainability in (de Olde, Sautier and Whitehead 2018, Ikerd 1993). In the face of increasing complexity however, de Ode, Sautier and Whitehead (2018) have highlighted the methodological trade-offs between simpler frameworks that are easy to implement and others that are more comprehensive. Critics of the increasing call for holism affirm that this can undermine efficiency and that the consideration of different dimensions can distract users from their original environmental-protection goal (Morrison-Saunders & Fischer, 2006). Assessment of systems properties, such as interactions and feedback, is however necessary to achieve the most pertinent and impactful choices towards sustainability. A full understanding (or to a sufficient degree) of the system is thus required. The holism challenge for this research is to provide a monitoring plan that is in agreement with these scientific considerations on ISA (systems analysis, pluri-encompassing), while also being able to deliver practical information to farmers and researchers without shortcomings on the sustainability result (e.g. inaccurate approximations or non-recognition of externalities).

1.3.2. Functioning, outputs and results applications

Agricultural sustainability assessment tools usually function in a hierarchical aggregation structure, where raw data is transformed into indicators, aggregated into scores and into themes. Two different types of frameworks can be differentiated; **systems based** where the processes and their outcomes are assessed and **content-based** that monitor key indicators. The first is most relevant to inner sustainability assessment (López-Ridaura et al. 2005) and the second to outer sustainability or objective-driven impact assessment (Van Cauwenbergh et al. 2006). There is important variation in vocabulary selection for the processes and hierarchical levels (see figure 1). Here we define the following: raw data, indicator, (possiblysub-themes,) themes and domains or pillars.

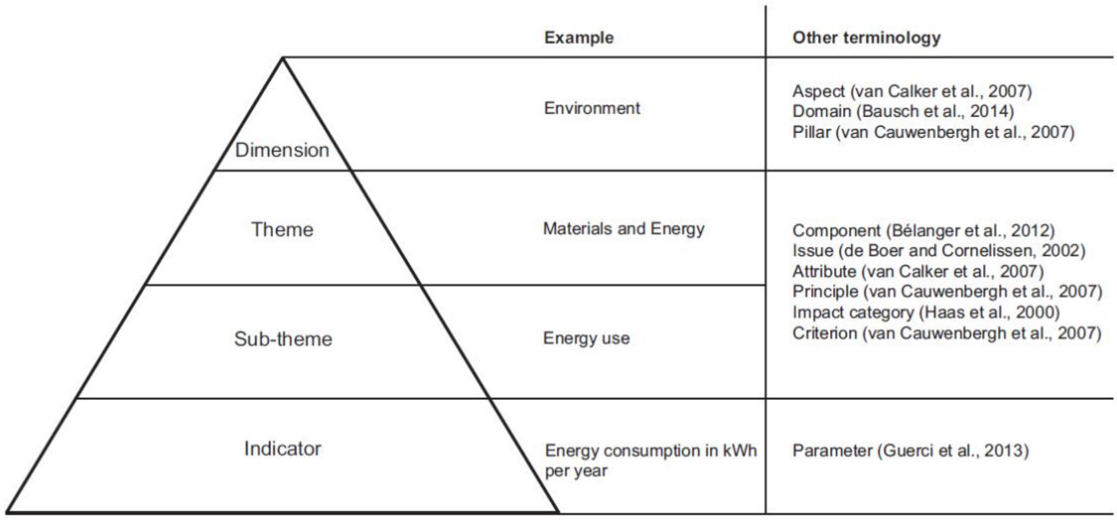

Figure 1 Hierarchy and terminology of frameworks' aggregation levels (from de Olde et al. 2016)

2. Methods

2.1. Site

The UBC Farm is part of the Centre for Sustainable Food Systems at the University of British Columbia (Vancouver, Canada). It manages 24 ha including 3.5 ha (8.6 acres) for commercial organic fruit and vegetable production, 12 ha (30 acres) of secondary forest as well as several educational gardens and recreational areas. The high number of different cultivated crops (52 crops harvested in 2018), the multiple activities and important biodiversity (both productive, functional and spontaneous) (CSFS 2018) lead to the consideration of the Farm as a **diversified production, educational, research and social environment**. The Farm is located on the university endowment lands, part of the Musqueam People's Territory and neighbouring an area under rapid urbanisation. This has in the past led to important land pressure and uncertainty on the Farm's future. The farmland is now secured by a Green Academic Zoning agreement, preventing a change in land use (UBC Land Use Plan, 2018).

The farm is managed by 2 directors and 15 permanent staff. In 2018, 43 students participated in farm activities through work-learn appointments, internships, volunteering or research. Twenty-eight (28) associates conduct research on the Farm. In addition, the farm offers a 1-year practicum, where 14 students get training in farm management.

2.2. Frameworks selection

The original proposal was the construction of a new assessment tool adapted to small-scale diversified farms. However, many frameworks, developed by research teams or certification bodies already exist, have proven wide application and have evolved from experience (de Olde, Sautier and Whitehead 2018). This thesis examined what tools exist already and their applicability, in order to identify gaps and future opportunities for farmer-led on-farm assessment tools.

Based on the objective of assessing at farm-gates and scientific literature, we gathered a set of criteria for frameworks selection presented in table 1. These frameworks were researched in Google Scholar and Web of Science using the following keywords and snowballing:

Sustainability (assessment OR evaluation) framework (agriculture OR farm)

Integrated sustainability (assessment OR evaluation) (agriculture OR farm)

A first pre-selection was done to identify twenty one (21) potential assessment frameworks. Then, a final selection of 4 were made for application (table 2). Based on the conclusion that a full assessment at farms gate is tedious and might exclude some impacts, a 5th framework was tried, one that rapidly assesses a multiplicity of production entities on a territory, through a limited set of key indicators.

2.2.1. Preselection

For frameworks selection, Gasparatos and Scolobig (2012) identified several families of criteria: namely assessment perspective (conceptual focus), features (practical elements), objective (particularly decision-making) and values. We prioritised the objective of sustainability appraisal; that is, assessing the degree of sustainability at a certain point in time without the results serving a decision-making purpose. Then, the perspective of the assessment should include socio-economic wellness and environmental protection. The features should include a set of indicators, aggregated in a way that assess both the contribution to sustainability (outer sustainability) and the sustainability of the system itself (inner sustainability) described in 1.12. Finally, values were not a criterion as they might bias the assessment and there were no diplomatic needs. Availability, language and visibility constrained the selection.

Twenty-one (21) frameworks were originally proposed, from search engines or snowballing. Selection criteria included criteria of de Olde et al. (2016) who conducted a similar experiment aimed at comparing the applicability of assessment frameworks to dairy farms. Other criteria correspond to the expectations of the host research institution. These criteria are listed in table 1 in on-hierarchical order). There were no expectations regarding the functioning of the tool.

Table 1 Preliminary selection criteria

Criteria	Source
Aimed at sustainability assessment	De Olde et al. 2016
Enable sustainability appraisal without a practical objective.	Gasparatos and Scolobig 2012
Published in a peer-reviewed scientific journal or report, describing a clear and repeatable methodology.	De Olde et al. 2016, modified by authors
The assessment must cover both economic, environmental and social dimensions of sustainability.	De Olde et al. 2016
Designed for agricultural operations.	Proposed by researchers
Assessing sustainability at the farm gates.	Proposed by researchers
Available in English, French or Spanish.	Proposed by researchers

2.2.2. *Final selection*

From the preselection we selected four (4) frameworks (RISE v3, SAFA, IDEA v3 and MESMIS) based on the criteria below in table 2. Most assessment tools, as well as scientific literature, stress the importance of considering the three pillars of sustainability (Slätmo, Fischer and Röös 2017), representing the contribution to global sustainability by local action. However, we are also interested in knowing the sustainability of the system itself, represented by properties (e.g. productivity, stability, resilience and self-reliance in López-Ridaura, Masera, and Astier 2002). It is thus important that these two aspects of sustainability evaluation are covered, either simultaneously in the same tool or in separate ones. The frameworks differ in the level of integration and range from Full Sustainability Assessments (Marchand et al. 2014) with RISE to a set of indicators corresponding to the farmers' perceived threats.

Table 2 Final selection criteria

Criteria
• The tools should not be designed for specific productions systems, • The scope of the tools should be global, • Experience of tools use should be proven, and the tools (re)adapted based on feedback, • The tools should allow for long-term comparison of systems and for sensitivity analysis (even indirectly by exporting results), • The assessments should be repeatable across production systems and places, • The definition of sustainability that the tools are based on should consider resource preservation, equity and viability. • The final selection should include both thematic (3 pillars) and systemic (properties of sustainable systems) assessments.

All frameworks proved time-consuming with the current data infrastructure of the UBC Farm. Therefore, a 5th framework was added based on the criteria of (i) time requirement and (ii) that it gathers evidence on the impact of practices without an analysis of the underlying mechanisms. The framework used is the Multi-dimensional assessment of Agroecology (FAO upcoming).

2.3. Indicators selection

RISE and IDEA provide a standardised set of indicators. MESMIS does not give indicators but a method to derive them. The indicators result from workshops (See 2.3.1). SAFA gives a set of indicators but allows to remove sub-themes (2nd level of aggregation) or add indicators (1st of aggregation) to a sub-theme. Irrelevant sub-themes (e.g. targeting large corporations, developing countries or commercial enterprises) were removed). Additional indicators (obtained from the workshops) were added to SAFA when a relevant theme existed.

2.3.1. Link with sustainability

The frameworks providing sets of indicators justify their indicator choices in regard to the developers' definitions of sustainability (Grenz et al. 2016, Zahm et al. 2008, FAO 2014). The relevance of the indicators was recapitulated in Appendix B and linked to a supporting publication. The authors also linked the indicators with a Sustainable Development Goal (SDG).

The indicators obtained in MESMIS workshops (2.4.1) reflect the farmers' perception of their system, in regard to their priorities and conception of sustainability. This is a justification per-se.

2.4. Data collection

The UBC Farm monitors field and management data since 2005 for decision making. Several research projects throughout the years have made occasional data available for non-management variables. Monitoring data includes field operations, inputs use (except water (only monitored since 2019) and self-produced compost), soil chemical quality, land use and financial results. Although not systematic, this data is comparable to that of an accounting scheme.

Biodiversity indicators for MESMIS and SAFA were collected according to the respective methodologies of the SAFA guidelines (FAO 2014) and of various frameworks for each indicator for MESMIS (see appendix C). Land cover was assessed through remote-sensing imaging.

2.4.1. MESMIS workshop

The MEMIS framework is the only selected one that does not contain a pre-defined set of indicators. Instead, it proposes to organise a workshop with farmers to make a SWOT analysis of their farm in regard to sustainability (see methodology in López-Ridaura, Masera and Astier 2002 and Ripoll-Bosch et al. 2012). 3 workshops were conducted: 2 with 7 farm staff each and one with 15 researchers. In farmer staff workshops, participants were asked 3 questions:

- *What is sustainable agriculture?*
- *What makes the UBC Farm sustainable?*
- *What are the threats to the sustainability of the UBC Farm and where is it unsustainable?*

For each question, there was a 5 minutes reflection time, then participants shared and debated their insights. Researchers were asked to list down variables to measure for sustainability assessment, on separate sticky notes. The notes were then gathered, the occurrences of each indicator counted, and the indicators debated.

The resulting ideas were organised hierarchically by degree of conceptuality into indicators (practical raw data. E.g. Green House Gas emissions), and overarching criteria (e.g. Air quality) and critical points (e.g. environmental integrity). Indicators were derived from literature if there were none for a conceptual criterion.

Following the MESMIS V2 method from López-Ridaura et al. 2005, criteria were also linked with properties of sustainable systems. Aggregation into higher conceptual levels is done by weighted sum. For selected indicators, a critical threshold was defined, equal to the level of the variable that gives 33 points in the RISE framework. If an indicator falls below it, the criteria takes the value of this indicator, following the logic that the system is as sustainable as its least sustainable component (Briquel et al. 2001).

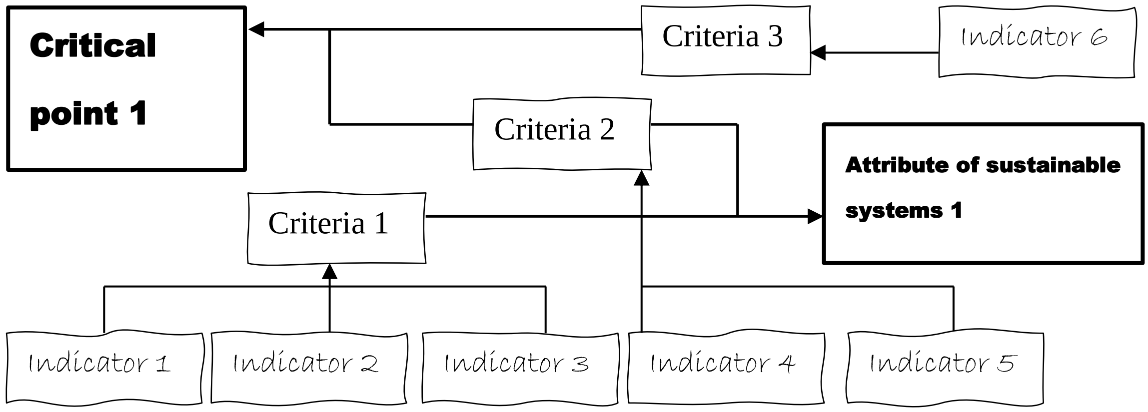

Figure 2. Aggregation structure of the MESMIS Framework.

2.4.2. Feedback sessions

After the results from all frameworks were obtained, they were presented to the same farm staff and researchers as in the 1st round of workshops. This step is important to gain feedback on the relevance of the indicators (Mascarenhas 2010) and maintain interest from the farmers (Jakku and Thorburn 2010). It is also an opportunity to get an interpretation and explanation of the results that was not perceived by the authors (Fraser et al. 2006).

The overall results in regard to sustainability was first presented to a group of 14. Then, the hierarchical and aggregative structure of the frameworks were explained. Finally, intermediate scores per theme were shown upon request.

Participants were then split in groups of 5 or 4 where they discussed the results with two questions from De Olde et al. (2016):

- *Are the outcomes presented in a way that is easy to understand?*
- *Do you consider that these results accurately reflect the UBC Farm?*

Finally, the groups shared feedback on the tools and the results. Feedback was recorded by two assistants.

2.5. Frameworks application and validation

The main step of the sustainability assessment is the tool use; inputting data and obtaining an output.

2.5.1. Operational structures of the frameworks

All frameworks are organised hierarchically; raw data is inputted and compared to a benchmark to form a score or indicator. Scores on related themes are possibly weighted and eventually aggregated into a score for that theme. Results presentation is done at different levels of aggregation. IDEA and MESMIS aggregate the themes to give 3 scores, on both 3 pillars and 5 properties of sustainability. RISE and SAFA give scores on 10 and 21 themes respectively.

RISE and IDEA function in a pre-defined way; data from accounting, practices, remote sensing and interviews are inputted in a calculator. SAFA also comes with a calculation software but requires a more in-depth interview with the farm manager. It also requires indicators with specific data collection methods. Once the data is inputted, RISE and IDEA proceed to comparing it to benchmarks to transform it into scores. Scores are then weighted and aggregated (weighted sum). In SAFA scoring is left to the user's discretion. Quantitative indicators have pre-set thresholds, whereas some practice-based indicators sometimes require making a subjective choice of score. The calculator then proceeds to weighted summing.

MESMIS does not come with a supporting calculator, which was thus developed by the researchers following the aggregation method described in López-Ridaura Masera and Astier 2002;

- Critical points and criteria are obtained from the workshop in addition to justification of how they are relevant to the farm's sustainability.
- These justifications are grouped into the properties of sustainable systems they correspond to. A subjective choice is sometimes made by the assessor.
- The critical points and criteria are integrated into the different properties based on their justification. They are also linked with a pillar of sustainability (economic, environmental or social) based of the assessor's perception.

Binder et al. (2010) developed a framework to assess sustainability assessment frameworks. They identify 3 major dimensions: normative, systemic and procedural. De Mey et al. (2011) and Marchand et al. (2014) have added critical success factors for framework implementation. The five (5) frameworks were compared using this perspective.

Normative reflects the goal of the assessment, the sustainability concept employed and resulting benchmarking, as well as the possible underlying value-judgements. The Systemic dimension refers to the boundaries of the studied entity, the way it is fragmented and the extent to which it is covered. Finally, the procedural dimension refers to how the assessment is carried-out, the results aggregated presented.

2.5.2. Divergence analysis

Once frameworks were scored, they present an output under the form of a sustainability diagram and a table detailing the subjacent level of aggregation. (E.g. Sub-themes and indicators visualised under a theme in SAFA (Figure 3). Or themes visualised under pillars for IDEA (Figure 4)

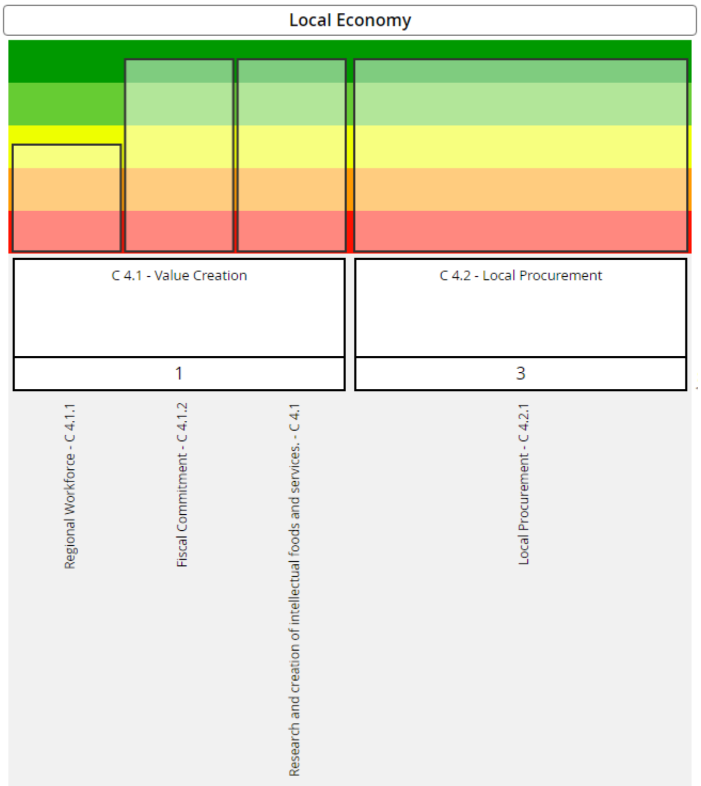

Figure 3 SAFA Local economy theme with underlying sub-themes and indicators

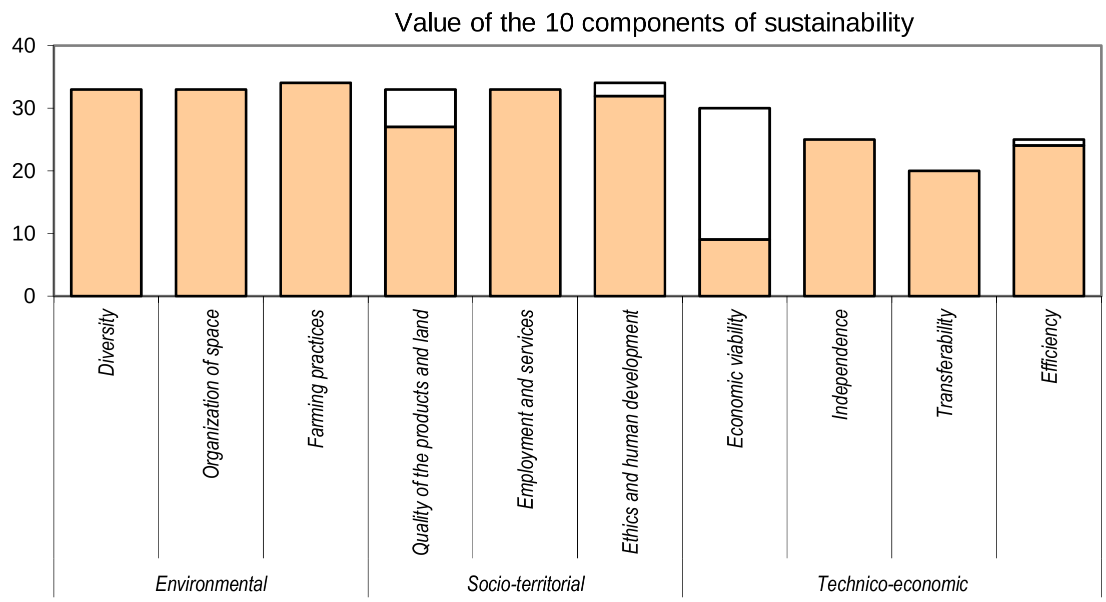

Figure 4 IDEA scores compared to maximum with overarching pillar below

As the same themes or indicators differed in score between frameworks, variation was systematically calculated; they were centred (indicator/max value) and subtracted (Highest scoring indicator – Lowest). Each variation higher than 10% (arbitrary threshold) was investigated qualitatively with the researchers' knowledge of the system or informal interviews with farmers, before concluding on the most accurate or adapted indicator.

3. Results

3.1. UBC Farm results

The results are presented in radar diagrams, with each axis representing a theme of sustainability. The diagrams also show three zones represented by green, orange/yellow and red/orange. The upper zone represents a sustainable situation (based on the framework's criteria), the middle states that there is an imbalance and the red area that a threat exists. The scores of the farm are marked on each axis and joined so that the area under the diagram represents overall performance. However, despite giving a visual impression, the total score does not correspond to a tangible as pillars do not compensate one-another reality (López-Ridaura, Masera and Astier 2002; Zahm et al.

2019).

Overall, the Farm showed sustainable scores across all relevant dimensions (here, social and environmental). With RISE, the Farm scored in the green (satisfactory) area in all but "Economic viability" (Figure 5), the latter being limited by the profitability and indebtedness scores. The lower economic score is explained by the farm not making a profit, as it is not a commercial enterprise, and not using credit to create a leverage effect.

SAFA showed only *good* to *best* scores (Figure 6). Similarly, IDEA shows almost maximum scores in all themes but *economic viability* and *quality of the products and the land*. The economic is limited for the same reason as in RISE and is thus consistent. For the lower *quality* score, this is first due to the fact that a single-use plastic is needed for vegetable crops and that there is no formalised reuse plan. Second, this theme also captures labels associated with traditional products, which are important in Europe where the tool was developed but less relevant to Canada. Last, the farm does not have the opportunity to maintain heritage buildings or landscapes according to the criteria in IDEA.

Finally, MESMIS also resulted in all scores being in the desirable zone, with the lowest score being on productivity (as an enterprise), consistent with the RISE result that the farm does not produce a commercial profit, added to slightly lower yields than conventional production (90%), assessed from the reference yields in RISE. This represents a trade-off in sustainability between public and marketable goods.

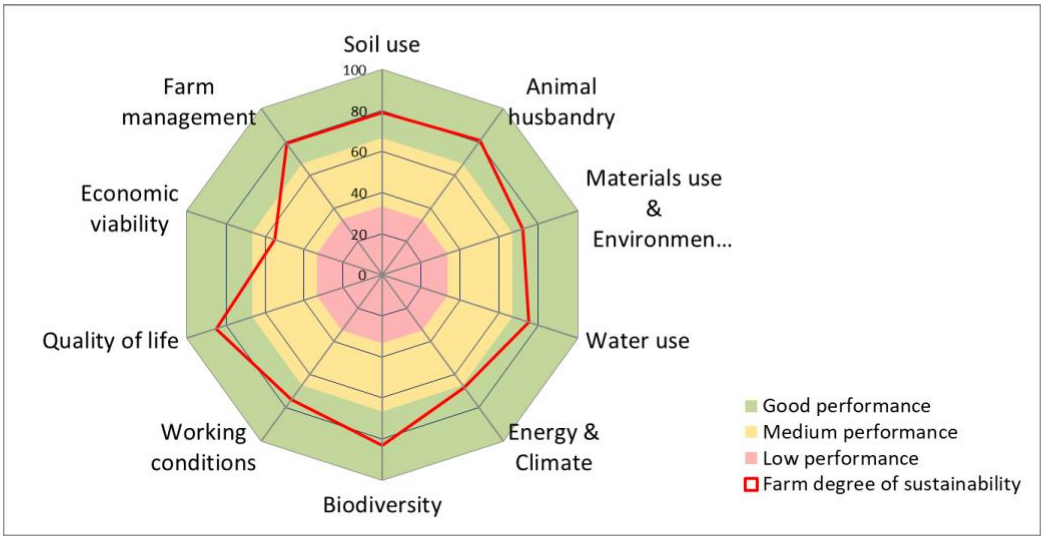

Figure 5 RISE sustainability polygon, UBC Farm 2018

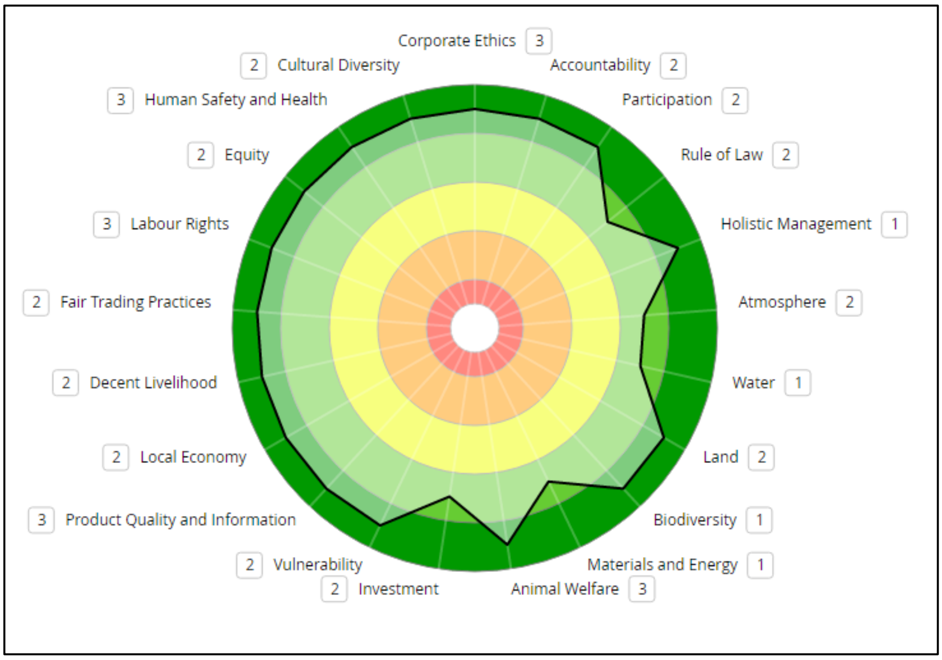

Figure 6 SAFA sustainability Polygon, UBC Farm 2018

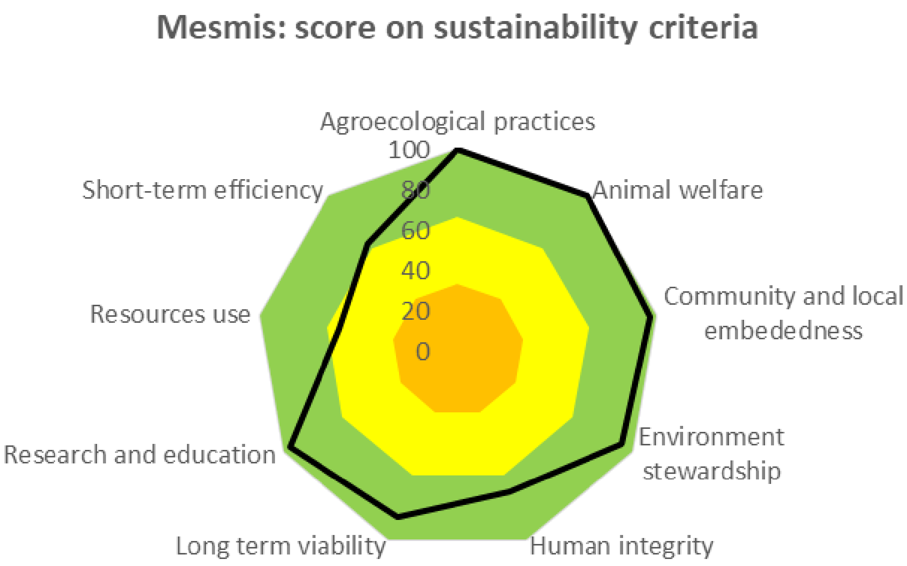

Figure 7 MESMIS sustainability Polygon, UBC Farm 2018. Critical points/criteria

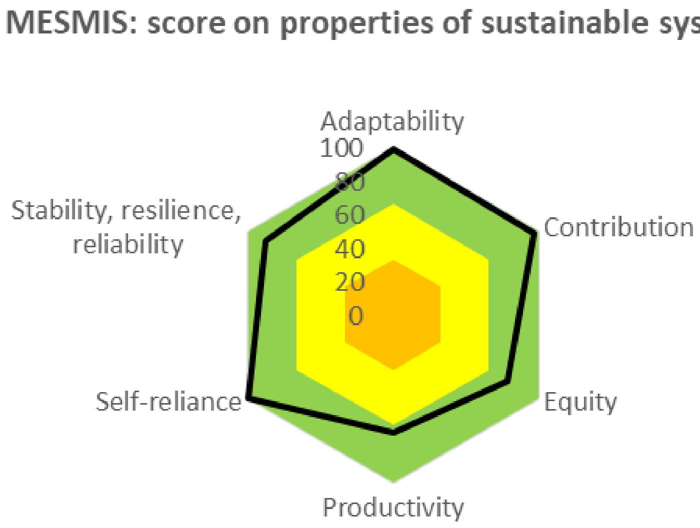

Figure 8 MESMIS sustainability Polygon, UBC Farm 2018. Properties of sustainable systems.

Domain	Issue	Indicators		Farm score	Evaluation score	*Maximum possible*
Agro-ecology	Diversity	Diversity of annual and temporary crops	A1	14	14	*14*
		Diversity of perennial crops	A2	14	14	*14*
		Animal diversity	A3	5	5	*14*
		Use and conservation of the genetic heritage	A4	3	3	*6*
		Diversity	*Subtotal:*		*33*	*33*
	Organization of space	Crop rotation	A5	8	8	*8*
		Plot size	A6	6	6	*6*
		Components management	A7	5	5	*5*
		Ecological Regulation Area (ERA)	A8	10	10	*12*
		Contribution to local environmental stakes and issues	A9	2	2	*4*
		Optimum use of space	A10	2	2	*5*
		Management of grazing surfaces	A11	2	2	*3*
		Organization of space	*Subtotal:*		*33*	*33*
	Farming practices	Fertilization	A12	0	0	*8*
		Liquid organic manures	A13	3	3	*3*
		Pesticides	A14	13	13	*13*
		Veterinary treatments	A15	3	3	*3*
		Soil protection	A16	5	5	*5*
		Water management	A17	4	4	*4*
		Energy dependency	A18	8	8	*10*
		Farming practices	*Subtotal:*		*34*	*34*
			Total:		**100**	**100**
Socio-territorial	Quality of the products and land	Quality management	B1	7	7	*10*
		Use/enhancement of built heritage and landscape	B2	7	7	*8*
		Non organic waste management	B3	2	2	*5*
		Public access to farm	B4	5	5	*5*
		Social involvement	B5	6	6	*6*
		Quality of the products and land	*Subtotal:*		*27*	*33*
	Employment and services	Valorization by/in short chains	B6	7	7	*7*
		Autonomy and valorization/enhancement of local resources	B7	8	8	*10*
		Services, multi activity	B8	5	5	*5*
		Contribution to employment	B9	6	6	*6*
		Work sharing	B10	5	5	*5*
		Foreseeable sustainability	B11	3	3	*3*
		Employment and services	*Subtotal:*		*33*	*33*
	Ethics and human development	Contribution to world food balance	B12	6	6	*10*
		Animal welfare	B13	3	3	*3*
		Training	B14	6	6	*6*
		Intensity of work	B15	4.9	4.9	*7*
		Quality of life	B16	6	6	*6*
		Isolation	B17	3	3	*3*
		Board facilities, hygiene and safety standards	B18	3	3	*4*
		Ethics and human development	*Subtotal:*		*31.9*	*34*
			Total:		**91.9**	**100**
Economy	Economic viability	Economic viability	C1	0	0	*20*
		Rate of economic specialization	C2	9	9	*10*
		Economic viability	*Subtotal:*		*9*	*30*
	Independence	Financial /dependency	C3	15	15	*15*
		Dependence on subsidies	C4	10	10	*10*
		Independence	*Subtotal:*		*25*	*25*
	Transmissibility	Economic transmissibility	C5	20	20	*20*
		Transmissibility	*Subtotal:*		20	*20*
	Efficiency	Efficiency of production process	C6	24	24	*25*
		Efficiency	*Subtotal:*		24	*25*
			Total:		**78**	**100**

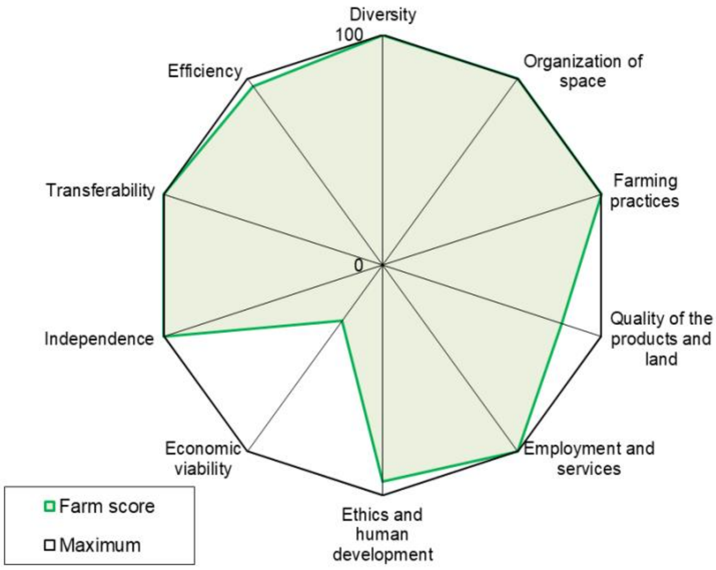

Figure 9a & 9b IDEA sustainability table and Polygon, UBC Farm 2018

3.1.1. Conclusion on UBC Farm's sustainability

RISE, the most comprehensive assessment, takes the WCED (1987) definition of sustainable development ("*development that meets the need of the present without compromising the ability of future generations to meet their own needs*"). As such, a score above 66 reflects a situation where the farm supposedly does not exceed the carrying capacity of ecosystem and does not deplete the capitals it relies on. It is important to note that RISE does not lie in the field of strong sustainability; it adopts a "Sensible sustainability" approach, stating that natural capital is partially substitutable by human-made capital within some defined boundaries.

However, the assessment at farms gates remains limited and a broader perspective should be adopted, to assess whether the farm's value-chains have a net positive value creation in a full-cost accounting perspective.

SAFA draws a very positive picture, IDEA similarly gives almost maximum score. These positive results provide little educational value according to farmers but reflect that key practices have been implemented and, in most case, good results reached.

Environmental sustainability

The UBC Farm has a mandate to preserve biodiversity and aims at being a model of sustainable farming. As such, environmental-preserving and socially just practices are put in place, reflected in the performance indicators of SAFA. For instance, greenhouse gas emissions are monitored, and the farm relies on pest-suppressive designs. It also offers opportunities for cultural preservation and gender-based empowerment. However, there is no formalised plan for environmental action, undermining the score in SAFA.

Soil organic matter (OM) indicators in MESMIS and RISE are irrelevant as data was not available to calculate organic matter balances, the soil OM has consistently varied around 11.36% ($\sigma = 2.05$) in the last 9 years. Although studies are needed to trace the evolution of organic matter, stable levels are observed with the regular input of external compost.

The farm has a lower productivity (10% less yields) than compared to yield references given by RISE. However, the limited use of resources and the important manual work (unassessed) make it function at a high level of eco-efficiency assessed in IDEA.

Socio-economic sustainability

In the considered assessment frameworks, social sustainability encompasses working conditions, social relations and contribution to local wellness and dynamics. IDEA stresses the contribution to the local territory the most, in terms of food provisioning, employment and landscape. SAFA focuses on the limitation of nuisances to the territory. Both frameworks give high to highest scores.

The UBC Farm is compliant with national regulations in regard to wage levels, workers' rights and workplace safety. As this theme may not be relevant internationally, here it does mean satisfactory labour protection and livelihoods. Although wages are higher than minimum for all positions, the important cost of living in British Columbia (WorkBC.ca, accessed 2019) undermine the *wage and income* indicators.

Working conditions indicators also entail non-mandatory workplace wellness such as social relations, gender equity and empowerment and training opportunities which are captured differently by each framework.

The farm is sustainable in a social perspective as it can deliver positive human impact while functioning as a sustainable enterprise. On an inner-sustainability perspective, good working conditions may provide attractive jobs, increasing labour availability. This was not assessed.

Properties of sustainable systems

The MESMIS framework also links indicators to properties of sustainable systems (inner sustainability). The farm also showed desirable results in sustainable properties (Figure 8).

The high adaptability (99/100) of the farm is attained by strengthened by a high number of activities and skills present on the farm. The farm has close bounds with its local community to which it provides fresh food and access to a recreational space. Conversely, the community provides clients and labour availability. The diverse expectations and services show the farm’s adaptability to its socio-economic environment.

Productivity (70/100) is affected by the negative net income. Yields are lower than their potential for the same crops and the food production (Kcal/ha) is low compared to grain or tubers. The farm relies on important labour inputs, reducing its efficiency ($\frac{Total\ income - labour\ costs}{Total\ income}$). This same labour, however, decreases the input-use intensity.

The farm is part of a public organisation with a research and education mandate, ensuring its stability (88/100) and the reliability of its output. Research projects however require grants, and food production and extra-curricular education need to be self-funding. The numerous activities contribute again to the stability.

Because the inputs produced on the farm, it was not possible to fully assess the farm’s self-reliance. As it imports most of its compost and seeds, it is not autonomous, but its diverse suppliers provide a base for resilience in case of shocks along the value-chain.

Finally, the *contribution* property (97/100) was added by the authors to encompass all elements that do not correspond to the properties given by MESMIS. They do not threaten or enhance the farm but create externalities, either positive or negative (e.g. greenhouse gas emissions, community services, public goods, plastic waste).

The main limitation of this MESMIS application is that it builds on assumptions (on how indicators link to properties, as these have not been assessed in this context) and on perceptions. Ripoll-Bosch et al. (2012) prefer considering the MESMIS framework as a SWOT analysis.

3.1.2. Consistencies and dissimilarities between results and frameworks

To appraise the accuracy of the results, the scores on similar indicators were compared between the different frameworks. There are variations in both the definitions, benchmarking and calculations of the indicators. With variations in the definition or scope of similarly named indicators, this comparison also enables to check which framework had the most appropriate indicator to the case study. A arbitrary threshold of 10% was defined, below which differences were not investigated.

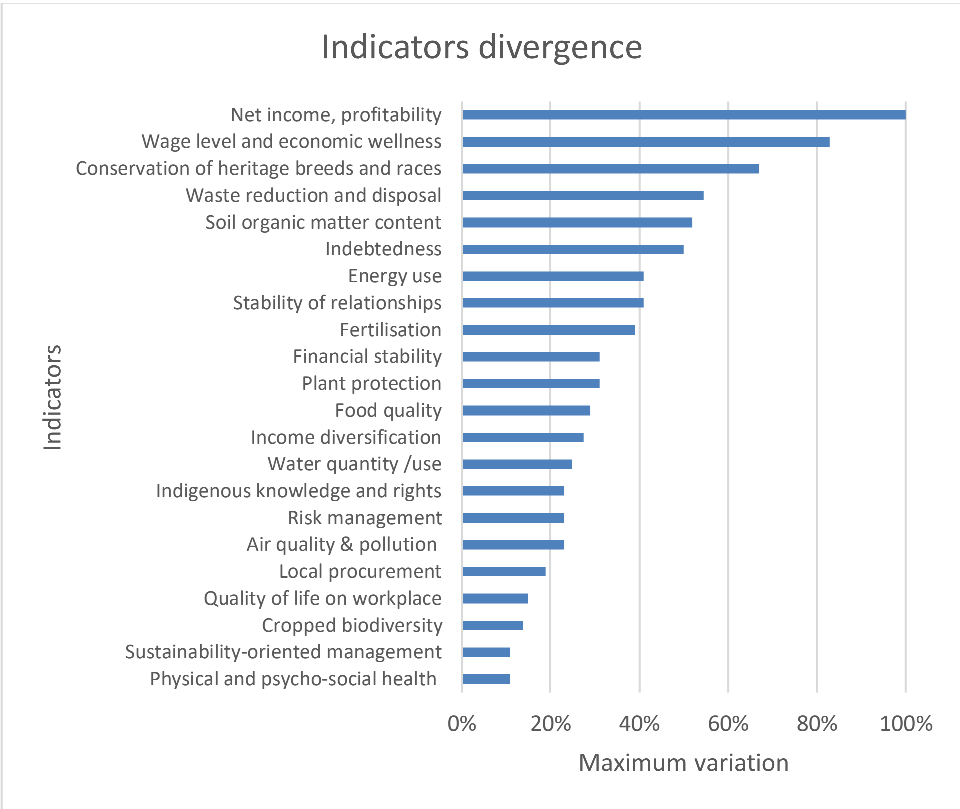

Figure 10 Divergence between indicators across frameworks

Economic

Net income is the most divergent indicator. In RISE and IDEA it scores 0% based on the value of the total net income, which is negative; the farm is not making a profit. SAFA interprets income as the ability to pay all wages and purchase all essential inputs, which is fully satisfied. For the UBC Farm, the interpretation from SAFA is more relevant, whereas net income is a common indicator for commercial farms. The financial stability and indebtedness indicators follow the same logic; the debt repayment (0$) is ratioed to cashflow in RISE (negative) and to total equity (0% debt) in MESMIS and IDEA. The RISE calculation is not appropriate for null debt levels.

The wage level and economic wellness is benchmarked on the cost of living from the local government for RISE and MESMIS, pulling it to low levels. In SAFA, the indicator is qualitative, asking if "*all employees are paid a living wage*". Although contradicting the latter justification, it cannot be argued that farm employees cannot fare for their livelihoods.

Finally, income-sources diversification differs as RISE and MESMIS consider the different economic activities whereas IDEA focuses on the crops, with the assumption that a failing crop can be compensated by another. In the case of a highly diversified farm, a hybrid model is needed, assessing both the diversity of crop groups in regard to their sensitivities and the diversity of income-generating activities.

Environmental

The environmental domain also shows important dissymmetry, linked either with differences in thresholds, indicators or interpretation. For the environmental pillar, SAFA asks about targets, practices and results. While results and practices are desirable, there are no formalised and public plans for environmental integrity. This undermines the value of indicators such as *atmosphere* and *water pollution*. RISE also asks for monitoring practices; the absence of monitoring of energy and water use leads RISE to estimate that important amounts are use, explaining lower scores and divergence.

The theme *conservation of heritage breeds and varieties* scores 100% in SAFA with one single animal variety, which is insufficient in IDEA and RISE (50% and 33% respective scores. This measure is not compensated by the 100% score of conservation of crop varieties. This is be less relevant to small-scale growers who might not have the opportunity to preserve many animal breeds.

Soil organic matter is repeatedly measured above 10% (11.36%) giving a high score in SAFA and IDEA; frameworks requiring the raw value. Oppositely, RISE calculates the organic matter balance, which is more relevant but likely inaccurate because compost use was estimated. High values of SOM should be checked against soil respiration to determine if the organic matter is being renewed.

Social

The indicators *quality of life on workplace* and *Sustainability-oriented management* have different interpretation, causing the results to vary. In RISE, this is done through a business strategy perspective, making it less relevant to the current context. The farm also shows desirable social scores. However, the educational output, recreational ecosystem services and public goods are not captured by most frameworks.

3.2. Frameworks comparison

Here, we applied the framework of Binder et al. (2010) to assess the normative, systemic and procedural dimensions of these frameworks. To this frameworks, critical success factors were added from De Mey et al. (2011) and Marchand et al. (2014). To recall section 2.5.1: normative aspect reflects the goal of the assessment, the sustainability concept employed, the benchmarking, as well as the possible underlying value-judgements. The Systemic dimension refers to the boundaries of the studied entity and the way it is fragmented. Systems-based frameworks that assess the properties of the system resulting from interactions and dynamics are here opposed to content-based frameworks which monitor key variables (von Wirén-Lehr 2001). MESMIS is the only purely systems-based framework. The version 4 of IDEA (currently under development) will include both a systems and contents approach. Moreover, a framework like RISE assesses the economic vulnerability of the system.

The Normative and Systemic dimensions are recapitulated in table 3. Finally, the procedural dimension is the *how-to*; it corresponds to the data collection and aggregation structure, and the results presentation.

Table 3 Selected frameworks presentation

Framework	IDEA	MESMIS	RISE	SAFA
Full name	Indicateurs de durabilité des exploitations Agricoles	Marco de Evaluación de Sistemas de Manejo Incorporando Indicadores de Sustentabilidad	Response-inducing sustainability assessment	Sustainability Assessment of Farming and Agriculture
Aim	Quantify strengths and weaknesses in regard to sustainability to identify courses of improvements and foster education.	Identify critical sustainability points of a system for evaluation and monitoring of its performance.	Provide a holistic sustainability assessment for the “dissemination and consolidation” of sustainable agriculture.	Integrated all aspects of sustainability and provide a universal assessment tool.
Scale	Farm-gates	Farm or agricultural community	Farm-gates	Agricultural production, processing and distribution plants or value-chain.

Leading organisation	IRSTEA	Departamento de Energía y Recursos Naturales, Instituto de Ecología, Universidad Nacional Autónoma de México	HAFL (School of Agricultural, Forest and Food Sciences), Bern university of applied sciences	Natural Resources Management and Environment Department, FAO
Origin	France	South America	Switzerland	Italy, multiple countries
Updated	019	2005	2016	2014
Uses	>1 500 farms (Zahm et al. 2008)	≥77 Case studies (multiple farms per CS)	>2 300 farms (de Olde et al. 2016)	>8 600 farms (de Olde et al. 2016)
Countries	>5 (de Olde et al. 2016)	≥12	>51 (de Olde et al. 2016)	>30 (de Olde et al. 2016)

Table 4 Normative and Procedural dimensions of selected frameworks

	RISE	SAFA	IDEA	MESMIS
Normative dimension				
Sustainability concept	WCED (1987) definition: "development which meets the needs of the present without compromising the ability of future generations to meet their own needs"	"Management and conservation of the natural resource base, and the orientation of technological and institutional change [for the continued] satisfaction of human needs for present and future generations. [Sustainability] conserves land, water, plant and animal genetic resources, is environmentally non-degrading, technically appropriate, economically viable and socially acceptable."	Sustainability is represented both by the contribution on social, environmental and economic dimension and by properties of sustainable systems (Responsibility; territory-embeddedness; stability; productivity & reproducibility; and autonomy)	A sustainable system is productive, stable, reliable, resilient, adaptable, equitable and self-reliant.
Way of goal setting	Pre-selected indicators (defined by top-down and participatory methods) must attain a minimum level defined as sustainable, based on local benchmarks.	Set of preselected indicators from which the assessor or interviewee can add or remove elements.	Top-down approach with a standardised set of indicators.	Interviews with farmers and farm-managers to determine indicators relevant to their context.
Scoring and aggregation method	Each indicator is scored between 0 and 100 points. Indicators get 66 points when they meet the sustainable benchmarks, 33 when below a critical limit.	Indicator scores are weighted and averaged into sub-categories, then these are averaged together.	Weighted aggregation of indicators within components, weighted aggregation of the latter into (a) pillars and (b) properties of sustainable systems	Ratio of each indicator with a benchmark.
Original function	Agriculture is, and will be further, facing pressures in economic, social and environmental domains. RISE assess performance in all of them and supports decision-making.	Provide a universally applicable framework to easily assess the sustainability and map sustainability around the globe.	Tool for sustainable agriculture education, extended to research, decision-making and extension.	Integrate sustainability evaluation in the decision-making process to improve the likelihood of success in the design of alternative development projects.

	Justification of thresholds[1]	High: corresponding to situations with tolerable pressure on the resource-base	Medium: varies per indicators. Some are based on desired results rather than resource-thresholds.		Low: value-judgements.
	Influence of value-judgement[1]	Low: estimations based on practices and context.	High: influence of the assessor's discretion while	Medium: redundancy in some indicators. Arbitrary thresholds.	High: arbitrary thresholds.
Systemic dimension					
	Boundaries	Farm gates.	From farm to value chain.	Farm gates	An assessment is only valid for a specific system, location and time period.
	Parsimony (simplicity)	Implicit goal by providing a clear operational structure and user-friendly interface.	Interview-based evaluation with selection of a score.	Scores are easily retrieved from routine data and entered in a user-friendly interface.	Possible redundancies; indicators might represent several critical features.
	Sufficiency (complexity)	High: Holistic assessment tool that aims at covering all elements of a farming system. Integrates calculation of supra-indicators (e.g. GHG based on crops, practices and climate)	High: extensive documentation and methodology for each indicator. Numerous indicators.	Medium: Numerous indicators to describe a farming system in detail. The assessment covers key features of all 3 pillars of sustainability.	Medium: enables to determine a number of indicators high enough to cover the all critical features. Aggregation is done to represent the different critical points.
	Systems analysis[2]	Medium: implicit in aggregation, not present in reporting.	Medium: separation in management, drivers and impacts of the system.	High: implicit in aggregation explicit in reporting.	Low: limited insight on how variables interact
	Coverage of sustainable systems' propertie[2]	Medium: representation of the system through a subset of indicators.	Low: set of indicators	High: linkage of indicators to properties of sustainable systems	High: linkage of indicators to properties of sustainable systems.

3.2.1. Normative dimension

The consideration of 3 pillars of sustainability, the goals of the frameworks and the systemic aspect were original selection criteria. As such, these are very close. The scoring and aggregation method are also similar and appear to be usual in sustainability assessment (van Cauwenbergh et al. 2007; Slätmo, Fischer and Röös 2017).

There is variation in the influence of value judgement. In IDEA, being certified organic gives points in the environmental domains, which farmers have criticised for being inaccurate (de Olde et al. 2016). In SAFA, it is the assessor's judgement that can influence the results, as the choice between two scores is left to their discretion.

[1] Obtained from feedback sessions

[2] Proposed by authors

RISE, SAFA and IDEA use a similar sustainability concept; environmentally sound, socially just and economically viable and MESMIS an internal stability definition. Together, they deliver a broad picture of inner and outer sustainability. RISE and MESIMS reached the sustainability goal through both stakeholder workshops and literature review, whereas SAFA and IDEA rely purely on a top-down approach (De Olde et al. 2016).

3.2.2. Systemic dimension

Binder et al. (2010) define the systemic aspect as the way a complex system is represented through selected indicators. They also oppose sufficiency (for accurate representation) to parsimony (simplicity). In MESMIS, the degree to which the system is represented depends on the assessor's and interviewees choice. RISE and IDEA are designed in a way that makes them applicable with routine farm data where they originate. SAFA is made to require only a short interview and specific data collection. None of the tools give the possibility to directly assess interactions (not without exporting the results), even though MESMIS does state the importance of considering them to understand the system.

3.2.3. Procedural dimension

De Olde, Sautier and Whitehead (2018) as well as triste et al. (2014) observe limited adoption of sustainability assessment tools, despite the multitude of available ones. They state a lack of stakeholder involvement in the development process and resulting mismatch in value-judgement between developers and users. They also remind that tools are only adopted when they provide valuable information for the farm's management. From similar observations, Rose et al., (2016) drew a list of success factors for tool adoption, including core factors (e.g. performance, accuracy and ease of use), enabling and driving factors (e.g. level of communication). In order to identify the framework with the best chances of being successfully applied, the frameworks' procedural dimensions were compared. The framework of Binder et al. (2010) was completed with "critical success-factors for implementation" from De Mey et al. (2011) and "characteristics" from Marchand et al. 2010). They are recapitulated in table 5

Table 5 Procedural dimension of selected frameworks

		RISE	SAFA	IDEA	MESMIS
Procedural dimension					
	Attitude of the users towards sustainability[3]	Sustainable agriculture is a primary objective of the organisation			
Preparatory phase		Plan interviews and centralise traceability data.	Plan interviews, centralise data, define the system boundaries.	Centralise traceability and research data.	Organise workshops and feedback sessions., centralise traceability and research data.
Indicator selection		NA, fixed set	Possibility to add indicators in sub-themes or to delete sub-themes.	NA, fixed set	Indicators are determined by a participatory approach
Measurement phase	Data availability [3]	Low: routine data needs to be derived, aggregated and transformed.	High: known by managers, gathered routinely or retrievable	High: retrievable from farm management.	Medium: retrievable from management, remote sensing
	Data source [4]	Remote sensing, accounting, HR and traceability documents	Interviews, accountings, HR and traceability documents.	Remote sensing, accounting, HR and traceability documents	Remote sensing, accounting, HR and traceability documents, testing, transects, interviews, workshops.
	Data correctness [1]	Medium-High: requires quantitative monitoring data. Influence of perception for some indicators (e.g. investment capability).	Moderate-low: scoring depends on the assessor's judgement and the value of each score can be changed by the user. A lot of the data required is not collected.	Moderate-high: Requires quantitative monitoring data, yet the scientific underpinning of the scores is unclear and indicators can be compensated within a them.	Moderate: relies on farm-logs and scientific benchmarks, however aggregation is arbitrary, and elements can compensate.
	User-friendliness [3]	High: clear and intuitive software	High: 5 levels scoring with clear formulation of questions. Very extensive data gathering procedures.	High: clear scoring procedure on a single sheet.	Low: no existing infrastructure. Requires several tools from different publishers
	Compatibility [3]	Low: data intensive; needs derivation from several logs. Limited compatibility with diversified systems.	High: data from interviews and routinely procedures are easy to gather.	Medium: requires data and accounting in a form that is partly used by UBC Farm. Measurement/consideration of systems properties.	High: the indicators are derived to fit the context.

[3] Proposed by De Mey et al. 2011

[4] Proposed by authors

	Assessment phase	Transparency	Medium: calculations are explained in a detailed manual but not intuitive in the software.	Low: aggregation procedures were not accessible or through extensive documents.	High: Excel sheet detailing all the calculations.	High: simple ratio to benchmarks and averaged weighting.
		Benchmarking complexity[5]	Low: benchmarks were available through local government and research	Medium: benchmarks are necessary for some indicators but available in scientific research.	NA: levels are pre-defined in the tool.	High: benchmarks for site-specific indicators were not always available.
		Scientific accuracy of aggregation [6]	Medium-high: visual integration for the differences. Final score is an average but there is insight in disaggregated variables.	Medium: high set of metrics that are not aggregated when too different	Medium: the system is as sustainable as the least sustainable pillar. Value-judgements in weighting.	Low: value-judgements in weighting.
		Time requirement[7]	High: 2 weeks	Low: 3 days	Low: 4 days	Medium: 1 week
	Applicability of results	Output accuracy [8]	High: rigorous calculations based on scientific tools	Low: averaging with influence of assessor's discretion.	Medium: scores are scientifically underpinned but may sometimes reflect value judgement and/or another context.	Low: indicators have varying degrees of scientific underpinning.
		Output complexity [5]	Medium: complex underlying calculation but user-friendly process to understand the makeup of each score.	High: high number of indicators on the final diagram using unusual semantic.	Low: presentation in a simple radar diagram without scale.	Medium: presentation in radar diagrams with access to underlying variables.
		Communication aid [5]	High: Clear radar diagram showing the scores of indicators on a theme with the overall score and a scale.	Low: Complex visual polygon.	Low: Very basic visual polygon.	Medium: Visual polygon with scale.
		Effectiveness [5]	High: ease of understanding and linkage with concrete farm operations.	Low: difficulty in linking scores to farm operations.	Low: the farm was not recognised in the scores.	Medium: recognition of the farm processes but limited insight for change.
		Comparability [9]	High: standardised metrics and data collection structure. Software embedded comparison tool.	High: possibility to plot several case-studies on the same graph	Medium: standardised metrics but absence of benchmarking.	Low: an assessment is only valid for a defined system at a given time.

5 Proposed by De Mey et al. 2011

6 Obtained from feedback sessions

7 Proposed by de Olde et al. 2016

8 Proposed by Marchand et al. 2014

9 Proposed by authors

Data sharing[10]	High: the data is centralised on a server.	Low: data sharing server is retired.	Medium	Low
Insight in aggregation processes and variables[11]	High: provision of an interactive output table to look into the desired level of detail.	Low: calculations are not transparent, and some information is available in lengthy documentation.	Medium: calculations are available in separate documentation.	Medium

10 Proposed by authors

11 Obtained from feedback sessions

Sydorovich (2007) distinguishes practice-based and result-based assessments. The first ones base their judgement of sustainability on a set of practices considered sustainable. This is of limited accuracy because of the variable impacts of the same practice in different contexts. The second type evaluates the impact of such practices on end-level indicators of sustainability. There is an increasing call for this second type of assessments, more appropriate to impact assessment, yet complicated studies. Here, IDEA and MESMIS are the most practice based, translating practices into scores. When results can be measured, SAFA assesses them and gives performance indicators a higher weight in aggregation. Finally, RISE inputs practices and crosses them with the context to estimate the impact.

. All frameworks require both interviews and retrieving farm data, in addition, workshops are organised for RISE. It is thus possible to combine the steps of several frameworks to make the application less time consuming. All frameworks also present results in the form of a radar diagrams of the scores in each indicator

3.3. Applicability to the diversified context

3.3.1. Ease of use (data intensity, collection methods, understanding…)

The application of the frameworks revealed that the data infrastructure at the UBC Farm was inappropriate. RISE and IDEA are tailored to use European traceability and/or accounting data, which may not be available in small scale farms, even less in Canada. Although an important share of the data was part of the routine collection plan, specific data points were unavailable. Similarly, farmers mentioned criteria such as plastic use and self-inputs self-sufficiency which could not be assessed because not monitored.

SAFA was the easiest framework to apply as it required a 1.5 h interview and data collection by the assessor. However, there is important subjectivity emanating from the interviewee's perception and the assessor's judgement.

3.3.2. Applicability of results

Sustainability assessment of diversified systems can give farmers a decision-making support and researchers a tool to compare and map sustainability of different systems. New insights were limited, as farmers were already aware of the issues resulting in the lowest scores, particularly nitrogen balance and economic loss. This is consistent with the findings of De Olde et al. (2016), who add that the problems were known but it was hard for farmers to remediate without creating trade-offs. However, having an overall result of sustainability was an appreciated feature and serves easy communication.

For researchers, the systematic application of frameworks in a region can enable to identify recurring issues, priority research needs and a larger-scale response. Furthermore, systematic assessments can allow to compare the sustainability of different practices. This is especially true with the FAO's Agroecology framework; it first characterises the "level" of agroecology before monitoring key indicators, enabling to link ways of organising to results on key metrics.

3.3.3. Multi-dimension assessment of Agroecology.

The trial of the frameworks delivers the paradoxical results that, despite being designed for farmers, they bring limited new insights (de Olde et atl. 2016) or levels of adoption, which is consistent with other findings (Triste et al. 2014, Paracchini et al.2012). On the other hand, they are too data intensive for research purposes.

The FAO is developing a *Multi-dimensional assessment of Agroecology* tool that aims at gathering the evidence on the impacts of agroecology. It uses a first characterisation step; through a short interview, the *level* of agroecology applied is assessed. Then, 10 to 13 key indicators corresponding to objectives of sustainable systems are measured without aggregation. This rapid assessment on a multitude of farms will enable to gather empirical evidence of the impact of agroecology. This framework adds the flexibility of assessing individual farms or communities.

Table 6 Indicators of the FAO agroecology framework

Indicators
1. Soil organic matter 1 bis. Soil health 2. Cropped and occurring biodiversity 3. Dietary diversity, 3 bis. Food insecurity, 4. Exposure to pesticides, 5. Women empowerment, 6. Youth employment, 7. Net income, 7 bis. Income stability 8. Income distribution, 9.Productivity, 9 bis. Stability of production 10. Secure land tenure

Multi-dimensional assessment of Agroecology	
Normative dimension	
Sustainability concept	Implicit: agroecological systems are sustainable through diversity, co-creation, synergies, efficiency, resilience and recycling. Their responsible governance respects human values, preserves traditions and promotes a circular economy.
Way of goal setting	Set-of indicators selected in top-down approach and based on existing frameworks.
Scoring and aggregation method	Comparison with a desirable benchmark. No aggregation
Original function	Utilising existing research to centralise and compare key metrics.
Justification of thresholds	*Desirable level for sustainable systems.*
Influence of value-judgement[1]	Low: standardised procedures for each indicator.
Systemic dimension	
Boundaries	Farm, territory or value-chain. Most adapted at territory level.
Parsimony (simplicity)	High:
Sufficiency (complexity)	Low: reduced number of variables without aggregation give a limited insight.
Systems analysis	Low: limited analysis (no quantification) of the system properties.
Coverage of sustainable systems' properties[1]	Medium: covered in an interview-based scoring.

Procedural dimension		
Preparatory phase		Gather relevant monitoring data from existing research projects.
Indicator selection		NA
Measurement phase	Data availability	High: limited indicator set.
	Data source	Interviews, testing, accounting, existing datasets.
	Data correctness	High: limited set of indicators with replicable and accurately described evaluation methods.
	User-friendliness	Medium: no existing infrastructure but each indicator is well-described.
	Compatibility	Low: limited relevance at farm-gates and in post-industrial contexts.
Assessment phase	Transparency	High: no aggregation, clear methodology.

	Benchmarking complexity	Medium: variability of available data depending on indicator and region.
	Scientific accuracy of aggregation	NA
	Time requirement	Low: 1 day.
Applicability of results	Output accuracy	High: describes the indicators directly.
	Complexity	Low: limited set of indicators without aggregation.
	Communication aid	Medium: radar graph with different scale per branch.
	Effectiveness	Medium: recognition of the farm processes but limited insight for change.
	Comparability	High: standardised data collection. Method aimed at comparing systems.
	Data sharing	High: data centralising platform under development.
	Insight in aggregation processes and variables	NA

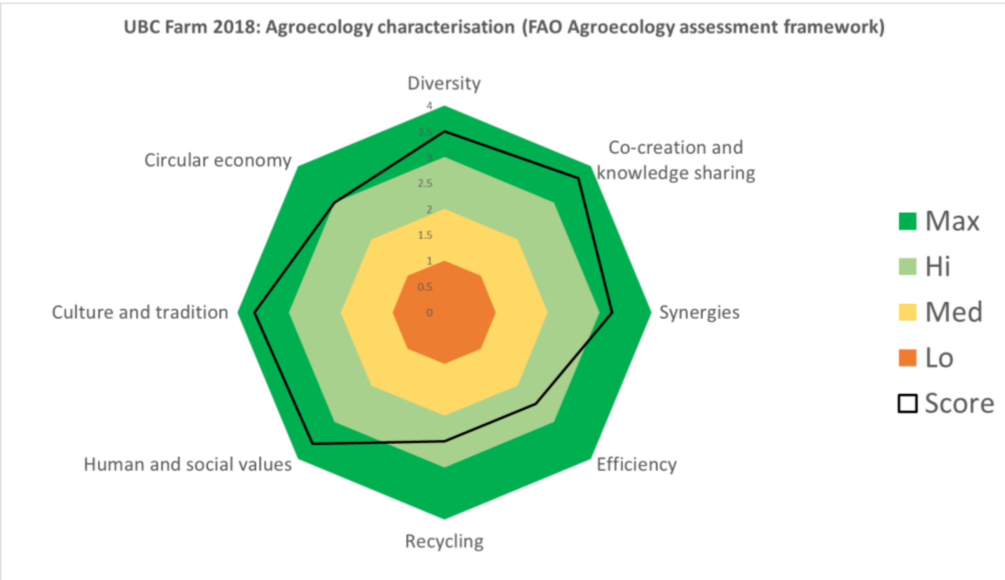

Figure 11. Agroecological characterisation of the UBC Farm based on the Multi-dimensional assessment of Agroecology's criteria

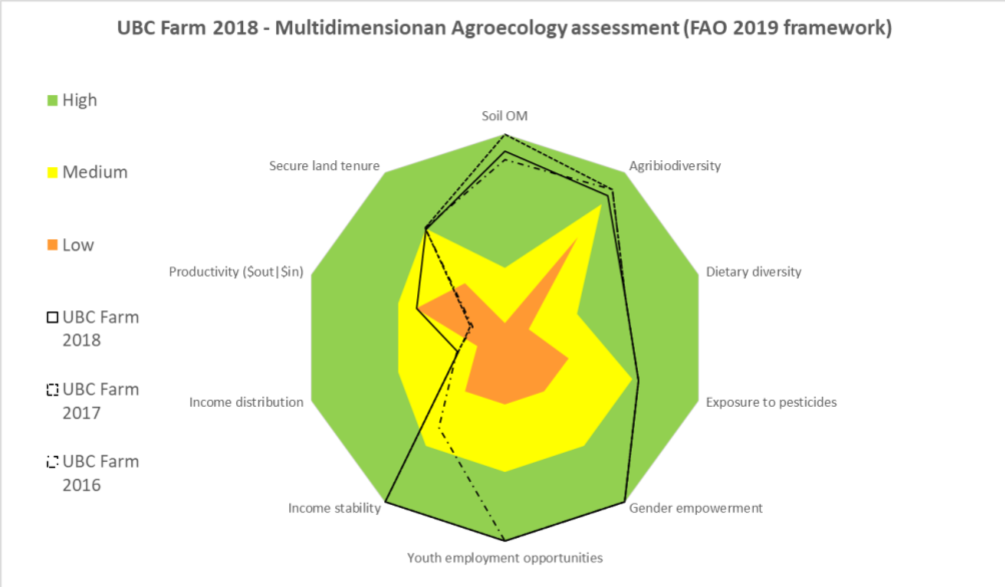

Figure 12. Score on indicators, Multi-dimensional assessment of agroecology.

3.3.4. Lite-Farm to easily collect the data

The Centre for Sustainable Food Systems at the UBC farm is developing a decision-support application: Lite Farm. This application was initially designed to monitor management data and enable to track costs of production per crop for small-scale farms. It will be possible to input the data collected into sustainability assessment tools, particularly the more demanding RISE, shortening the time needed for data collection. The input data that Lite Farm collects was identified in Annex D. The compatibility of Lite-Farm was also determined based on the number of these covered (table 9).

Table 7. Compatibility of Lite Farm with the assessment frameworks

	RISE	SAFA	IDEA	MESMIS	Agroecology
Compatibility with Lite-Farm	High: output data from Lite farm can be inputted in RISE 8/10 themes are covered in Lite Farm, the rest is retrievable from interviews or management documents.	Low: different data form. 8/21 themes covered.	Medium: output data from Lite farm can be used for some indicators. 28/53 indicators covered	Medium: can be tailored to match Lite Farm. Depends on chosen indicators.	Medium: direct data for 5/10 indicators.

4. Discussion and conclusion

4.1. Conclusion on sustainability and applicability

Despite variability in indicator scores and uncertainty on the results, the frameworks consistently showed sustainable scores. Particularly, the Farm has excellent environmental management practices with biodiversity enhancement particularly, soil preservation and negative GHG balance. The Farm also delivers positive social impact with indigenous cultural preservation, gender equality and education to food literacy. The pitfalls lie in the economic domain, where the farm does not make a commercial profit. This domain is however less relevant to this case study but should be considered for commercial farms.

Indicator value comparison showed important divergence between frameworks This divergence is linked with different benchmarks, data needs, interpretations and calculations. For the comparison in 3.1.2, indicators were grouped together when the authors considered that they corresponded to the same concrete process. Yet, depending on the framework, the breadth of indicators and themes varies. This is a methodological issue of the present study. However, frameworks have different benchmarks, corresponding to *what is deemed* sustainable. While for some indicators such as greenhouse gas emissions or erosion it is easy to define a balance (with a sustainable situation corresponding to a net zero), for other indicators, especially regarding heritage preservation, benchmarks do not correspond to *sensible* processes and the interpretation varies per author. Consequently, framework and indicator selection, require indicator validation. This can be done by determining the environment's carrying capacity; *for what level of specific variables can a system not function perennially?* Selecting the indicator of sustainability and its benchmark value should be done in regard to what the environment can withstand.

RISE overcomes this challenge by requiring a benchmarking phase where acceptable levels of variables is defined. However, as the scores rely on comparison of the data to a benchmark, RISE is also the most sensitive to input errors and should not be used with estimations. This can also be overcome in a true-cost accounting perspective, where the value of elements such as heritage breed sis estimated.

The Frameworks are branded as easy to apply (Grenz et al. 2016, Zahm et al. 2008). While SAFA needed a 1.5 hours interview and additional data collection taking 2 full days, gathering data for MESMIS, RISE and IDEA proved to tedious for a systematic application with the current data infrastructure (3 weeks for data gathering). RISE and IDEA are designed for use with European accounting schemes data but should be compatible in British Columbia with a similar data collection method.

4.2. Strengths, weaknesses and methodological discussion

The frameworks give a multi-encompassing portrait of the farm's sustainability, incorporating all three pillars, inner, and outer sustainability. They vary in degree of accuracy towards real phenomena. RISE is the most precise but suffers from inaccurate inputs. Although the application of five (5) frameworks enables to discern where results converge and help form a better understanding, this is also a timely process, not realistic for farmers.

The results of SAFA and IDEA show are very positive, making the farmers doubt their accuracy. This is consistent with the findings of De Olde (2016) where farmers found the output of IDEA and SAFA "too positive". These frameworks should thus be reserved, as IDEA's supporting publication states, to learning (Zahm et al. 2008). As IDEA is not benchmarked for the region of the current application, it is unsure whether the scores given correspond to what the environment can withstand. It still gives a valuable summary of the good practices implemented on the farm. To overcome this issue, the methodology used by the authors for indicators selection should be applied to the British Columbian context.

Finally, in general debate, sustainable agriculture is opposed a reliable food supply (Jacobsen et al. 2013). Agriculture cannot be deemed sustainable if the human labour input (throughout the value-chain) requires more food than is produced on the farm (for the staff involved in food production activities, not considering education and research). It is of interest, before conducting an in-depth assessment at farms gates, to assess the performance on a limited number of key metrics; particularly food balance, economic viability and greenhouse gas emissions.

A similar assessment on the services dimension is also not made by the framework. The farm produces food but also research, education, public good and awareness raising. It is of interest to assess whether the immaterial output equates or surpasses the financial input.

4.3. Future research needs

4.3.1. Trade-offs and synergies (between indicators and properties).

Sustainable agriculture is faced with multiple trade-offs (Klapwijk et al. 2014). These are process trade-offs between competing resource allocation patterns or substantive trade-offs defined as "gains at the expense of adverse outcomes" (Morrison-Saunders and Pope 2013). Klapwijk et al. (2014) add the trade-offs between resource consumption and resource allocation.

Empirical knowledge about trade-offs can strengthen decision making, by knowing the adverse effects associated with a decision. Conversely, knowing synergies between indicators would enable to maximise capital gain. However, adverse outcomes are weighted differently depended on rationale (Weston 2000) and trade-offs assessment should require a true-cost accounting method. Empirical sustainability evaluation and sensitivity analysis

The intensive data collection schemes *revealed* the need for less data intensive techniques if researchers want to systematically monitor agricultural sustainability. Moreover, researcher workshops underlined the lack of clear relation between the indicators and properties of sustainable systems, particularly resilience and human integrity. It would be of crucial interest to monitor both the current indicators, exterior shocks and resulting responses on a set of farms in the same region. This would first enable to check for correlation between indicators, possibly reducing the indicator sets by keeping uncorrelated factors. Second, this would enable to empirically *regress* measures of resilience and stability to shocks and determine which levels of the indicators are, in practice, the most sustainable.

Finally, for outer-sustainability assessment, there is a need for clear benchmarking of what disturbances the local environment can withstand. This can be done in a true-cost accounting perspective, by evaluating the value of natural, social and economic capital, enable to easily compare the gains or losses from certain activities. It is then easy to conclude, in a weak sustainability *perspective* which practices are sustainable; lead to an increase in total capital (True Price Foundation 2009).

5. References

Anderson, Weston, and Timothy Johnson. 'Global Cost of Land Degradation'. In *Economics of Land Degradation and Improvement – A Global Assessment for Sustainable Development*, 85–116. Cham: Springer International Publishing, 2016. https://doi.org/10.1007/978-3-319-19168-3.

Bertocchi, Mattia, Eugenio Demartini, and Maria Elena Marescotti. 'Ranking Farms Using Quantitative Indicators of Sustainability: The 4Agro Method'. *Procedia - Social and Behavioral Sciences* 223 (June 2016): 726–32. https://doi.org/10.1016/j.sbspro.2016.05.249.

Bezlepkina, Irina, Pytrik Reidsma, Stefan Sieber, and Katharina Helming. 'Integrated Assessment of Sustainability of Agricultural Systems and Land Use: Methods, Tools and Applications'. *Agricultural Systems* 104, no. 2 (February 2011): 105–9. https://doi.org/10.1016/j.agsy.2010.11.002.

Binder, Claudia R., Giuseppe Feola, and Julia K. Steinberger. 'Considering the Normative, Systemic and Procedural Dimensions in Indicator-Based Sustainability Assessments in Agriculture'. *Environmental Impact Assessment Review* 30, no. 2 (February 2010): 71–81. https://doi.org/10.1016/j.eiar.2009.06.002.

Bond, Alan J., and Angus Morrison-Saunders. 'Re-Evaluating Sustainability Assessment: Aligning the Vision and the Practice'. *Environmental Impact Assessment Review* 31, no. 1 (January 2011): 1–7. https://doi.org/10.1016/j.eiar.2010.01.007.

Briquel, Vincent, Lionel Vilain, Jean-Louis Bourdais, Philippe Girardin, Christian Mouchet, and Philippe Viaux. 'La méthode IDEA (indicateurs de durabilité des exploitations agricoles) : une démarche pédagogique [THE IDEA (INDICATORS OF FARM SUSTAINABILITY) METHOD: A PEDAGOGIC APPROACH]'. *Ingénieries* 25, no. 1 (2001): 29–39.

FAO. 'SAFA Sustainability Assessment of Food and Agricultural Systems', 2014. http://www.fao.org/3/a-i3957e.pdf.

Fraser, Evan D.G., Andrew J. Dougill, Warren E. Mabee, Mark Reed, and Patrick McAlpine. 'Bottom up and Top down: Analysis of Participatory Processes for Sustainability Indicator Identification as a Pathway to Community Empowerment and Sustainable Environmental Management'. *Journal of*

Environmental Management 78, no. 2 (January 2006): 114–27. https://doi.org/10.1016/j.jenvman.2005.04.009.

Gasparatos, A., and A. Scolobig. ‘Choosing the Most Appropriate Sustainability Assessment Tool’. *Ecological Economics* 80 (August 2012): 1–7. https://doi.org/10.1016/j.ecolecon.2012.05.005.

George, Clive. ‘Testing for Sustainable Development through Environmental Assessment’. *Environmental Impact Assessment Review* 19, no. 2 (March 1999): 175–200. https://doi.org/10.1016/S0195-9255(98)00038-9.

Gold, Maty. ‘Sustainable Agriculture: The Basics’. In *Sustainable Agriculture*. Special Reference Briefs Series, SRB 99-02. National Agricultural Library, USDA, 2007.

Gray, Steven. ‘Measuring Systems Thinking’. *Nature Sustainability* 1, no. 8 (August 2018): 388–89. https://doi.org/10.1038/s41893-018-0121-1.

Grenz, Jan, Raphael Mainiero, Michael Schoch, Firesenai Sereke, Susanne Stadler, Christian Thalmann, and Rebekka Wyss. ‘RISE 3.0 Manual’. Bern University of Applied Sciences, 2016. www.rise.hafl.bfh.ch.

Häni, Fritz, A Stampfli, and Th Keller. *RISE [Response-Inducing Sustainability Evaluation]: A Tool for a Holistic Sustainability Assessment at the Farm Level.* Zollikofen (Switzerland): Berner Fachhochschule, 2001. https://www.bfh.ch/en/research/reference-projects/rise/.

Ikerd, John E. ‘The Need for a System Approach to Sustainable Agriculture’. *Agriculture, Ecosystems & Environment* 46, no. 1–4 (September 1993): 147–60. https://doi.org/10.1016/0167-8809(93)90020-P.

Jacobsen, Sven-Erik, Marten Sørensen, Søren Marcus Pedersen, and Jacob Weiner. ‘Feeding the World: Genetically Modified Crops versus Agricultural Biodiversity’. *Agronomy for Sustainable Development* 33, no. 4 (October 2013): 651–62. https://doi.org/10.1007/s13593-013-0138-9.

Jakku, E., and P.J. Thorburn. ‘A Conceptual Framework for Guiding the Participatory Development of Agricultural Decision Support Systems’. *Agricultural Systems* 103, no. 9 (November 2010): 675–82. https://doi.org/10.1016/j.agsy.2010.08.007.

Klapwijk, Cj, Mt van Wijk, Ts Rosenstock, Pja van Asten, Pk Thornton, and Ke Giller. ‘Analysis of Trade-Offs in Agricultural Systems: Current Status and Way

Forward'. *Current Opinion in Environmental Sustainability* 6 (February 2014): 110–15. https://doi.org/10.1016/j.cosust.2013.11.012.

López-Ridaura, S. 'Evaluating the Sustainability of Complex Socio-Environmental Systems. the MESMIS Framework'. *Ecological Indicators* 2, no. 1–2 (November 2002): 135–48. https://doi.org/10.1016/S1470-160X(02)00043-2.

López-ridaura, S., H. Van. Keulen, M. K. van. Ittersum, and P. A. Leffelaar. 'Multiscale Methodological Framework to Derive Criteria and Indicators for Sustainability Evaluation of Peasant Natural Resource Management Systems'. *Environment, Development and Sustainability* 7, no. 1 (January 2005): 51–69. https://doi.org/10.1007/s10668-003-6976-x.

Marchand, Fleur, Lies Debruyne, Laure Triste, Catherine Gerrard, Susanne Padel, and Ludwig Lauwers. 'Key Characteristics for Tool Choice in Indicator-Based Sustainability Assessment at Farm Level'. *Ecology and Society* 19, no. 3 (2014): art46. https://doi.org/10.5751/ES-06876-190346.

Mascarenhas, André, Pedro Coelho, Eduarda Subtil, and Tomás B. Ramos. 'The Role of Common Local Indicators in Regional Sustainability Assessment'. *Ecological Indicators* 10, no. 3 (May 2010): 646–56. https://doi.org/10.1016/j.ecolind.2009.11.003.

Morrison-Saunders, Angus, and Thomas B Fischer. 'WHAT IS WRONG WITH EIA AND SEA ANYWAY? A SCEPTIC'S PERSPECTIVE ON SUSTAINABILITY ASSESSMENT'. *Journal of Environmental Assessment Policy and Management* 8, no. 1 (n.d.): 19–39.

Olde, Evelien M. de, Frank W. Oudshoorn, Claus A.G. Sørensen, Eddie A.M. Bokkers, and Imke J.M. de Boer. 'Assessing Sustainability at Farm-Level: Lessons Learned from a Comparison of Tools in Practice'. *Ecological Indicators* 66 (July 2016): 391–404. https://doi.org/10.1016/j.ecolind.2016.01.047.

Olde, Evelien M. de, Marion Sautier, and Jay Whitehead. 'Comprehensiveness or Implementation: Challenges in Translating Farm-Level Sustainability Assessments into Action for Sustainable Development'. *Ecological Indicators* 85 (February 2018): 1107–12. https://doi.org/10.1016/j.ecolind.2017.11.058.

Pope, Jenny, David Annandale, and Angus Morrison-Saunders. 'Conceptualising Sustainability Assessment'. *Environmental Impact Assessment Review* 24, no. 6 (August 2004): 595–616. https://doi.org/10.1016/j.eiar.2004.03.001.

Rigby, Dan, Phil Woodhouse, Trevor Young, and Michael Burton. 'Constructing a Farm Level Indicator of Sustainable Agricultural Practice'. *Ecological Economics* 39, no. 3 (December 2001): 463–78. https://doi.org/10.1016/S0921-8009(01)00245-2.

Ripoll-Bosch, R., B. Díez-Unquera, R. Ruiz, D. Villalba, E. Molina, M. Joy, A. Olaizola, and A. Bernués. 'An Integrated Sustainability Assessment of Mediterranean Sheep Farms with Different Degrees of Intensification'. *Agricultural Systems* 105, no. 1 (January 2012): 46–56. https://doi.org/10.1016/j.agsy.2011.10.003.

Rose, David C., William J. Sutherland, Caroline Parker, Matt Lobley, Michael Winter, Carol Morris, Susan Twining, Charles Ffoulkes, Tatsuya Amano, and Lynn V. Dicks. 'Decision Support Tools for Agriculture: Towards Effective Design and Delivery'. *Agricultural Systems* 149 (November 2016): 165–74. https://doi.org/10.1016/j.agsy.2016.09.009.

Sachs, Jeffrey, Roseline Remans, Sean Smukler, Leigh Winowiecki, Sandy J. Andelman, Kenneth G. Cassman, David Castle, et al. 'Monitoring the World's Agriculture'. Nature 466, no. 7306 (July 2010): 558–60. https://doi.org/10.1038/466558a.

Slätmo, Elin, Klara Fischer, and Elin Röös. 'The Framing of Sustainability in Sustainability Assessment Frameworks for Agriculture: Framing of Sustainability'. *Sociologia Ruralis* 57, no. 3 (July 2017): 378–95. https://doi.org/10.1111/soru.12156.

Sydorovych, Olha & Wossink, Ada. (2007). Assessing Sustainability of Agricultural Systems: Evidence from a Conjoint Choice Survey. Southern Agricultural Economics Association, 2007 Annual Meeting, February 4-7, 2007, Mobile, Alabama.

Triste, Laure, Fleur Marchand, Lies Debruyne, Marijke Meul, and Ludwig Lauwers. 'Reflection on the Development Process of a Sustainability Assessment Tool: Learning from a Flemish Case'. *Ecology and Society* 19, no. 3 (2014): art47. https://doi.org/10.5751/ES-06789-190347.

True Price Foundation, Deloitte, EY, and pwc. 'The Business Case for True Pricing', 2014. https://trueprice.org/wp-content/uploads/2015/02/True-Price-Report-The-Business-Case-for-True-Pricing.pdf.

Weaver, Paul M., and Jan Rotmans. 'Integrated Sustainability Assessment: What Is It, Why Do It and How?' *International Journal of Innovation and Sustainable Development* 1, no. 4 (2006): 284. https://doi.org/10.1504/IJISD.2006.013732.

Werf, Hayo M.G van der, and Jean Petit. 'Evaluation of the Environmental Impact of Agriculture at the Farm Level: A Comparison and Analysis of 12 Indicator-Based Methods'. *Agriculture, Ecosystems & Environment* 93, no. 1–3 (December 2002): 131–45. https://doi.org/10.1016/S0167-8809(01)00354-1.

Wirén-Lehr, S von. 'Sustainability in Agriculture — an Evaluation of Principal Goal-Oriented Concepts to Close the Gap between Theory and Practice'. *Agriculture, Ecosystems & Environment* 84, no. 2 (April 2001): 115–29. https://doi.org/10.1016/S0167-8809(00)00197-3.

Zahm, Frédéric, Adeline Alonso Ugaglia, Jean-Marc Barbier, Héloïse Boureau, Bernard Del'homme, Mohamed Gafsi, Pierre Gasselin, et al. 'Évaluer La Durabilité Des Exploitations Agricoles. La Méthode IDEA v4, Un Cadre Conceptuel Combinant Dimensions et Propriétés de La Durabilité'. *Cahiers Agricultures* 28 (2019): 5. https://doi.org/10.1051/cagri/2019004.

Zahm, Frédéric, Philippe Viaux, Lionel Vilain, Philippe Girardin, and Christian Mouchet. 'Assessing Farm Sustainability with the IDEA Method - from the Concept of Agriculture Sustainability to Case Studies on Farms'. *Sustainable Development* 16, no. 4 (July 2008): 271–81. https://doi.org/10.1002/sd.380.

Zahm, Frédéric, Mohamed Gafsi, Laurence Guichard, Bernard Del'homme, and Vincent Manneville. *La Méthode IDEA*, 2008. https://idea.chlorofil.fr/index.php.

6. Tables

7. Figures

8. Appendices

8.1. *Appendix A: Pre-selected frameworks and properties*

Table 8 Pre-selected frameworks

Nb	(Acronym)	Long name	Organisation	Publication	Pilot application	Farm Type	Target region	Comment
1		Appropriate method to assess the Sustainability of Organic farming systems	UMR Dynamiques Rurales	Gafsi and Favreau 2010	South-western France	Multiple, organic	France	Stress the importance of region-adapt.
2		A tool for the sustainability assessment of farms.	VESPA, University of Milan	Gaviglio, Bertocchi and Demartini 2017	Northern Italy		Italy	Those were developed for a specific context and are sets of indicators. They are of use for weighing and discriminating.
3		Empirical evaluation of agricultural sustainability using composite indicators.	IFAPA	Gómez-Limón	Spain	Multiple	Spain	
4		Methodological approach to assess and compare the sustainability level of agricultural plant production systems	Department of environmental sciences, University of Hill	Dantsis et al. 2010	Greece	Plant production	Post-industrial world	
5		Farm-level indicator of sustainable agricultural practice.	CAFRE, University of Manchester	Rigby et al. 2001	UK	Multiple	Worldwide	Compares organic with non-organic
6		Analytical framework for the multi-dimensional assessment of agroecology.	FAO	*Upcoming*	*Upcoming*	Multiple	Global south, Worldwide	
7		Participatory method for the design and integrated assessment of crop-livestock systems	INRA	Moraine et al. 2014	France	Crop-livestock systems.	France	
8		Sustainable intensification assessment framework	US AID	Musumba et al. 2017	US	Peasant farming	Developing world	
9		Systemic framework for sustainability assessment	Joint research-center for environment and sustainability, European Commission	Sala, Ciuffo and Nijkamp 2015	Europe	All	None	Framework to make a framework.
10	4Agro		Dipartimento di scienze veterinarie per la salute, la produzione animale e la sicurezza alimentaria.	Bertocchi, Demartini and Marescotti 2016)	Northern Italy	Multiple livestock systems.	Worldwide	
11	APOIA-NovoRural	System of weighted environmental impact assessment of rural activities.	Embarpa Labex Europe.	Rodrigues et al. 2010	Europe	All	Europe	
12	DELTA	Agri-environmental indicators to assess dairy farm sustainability in Quebec	Faculté des sciences de l'agriculture, Laval University	Bélanger et al. 2015	Québec, Canada	Dairy	Canada	
13	**IDEA**	**Indicateurs de durbailité des Exploitaitons agricoles**	**IRSTEA**	**Zahm et al. 2080**	**France**	**All except vegetables and wine**	**France, World**	
14	**MESMIS**	**Marco de Evaluación de Sistemas de Manejo Incorporando Indicadores de Sustentabilidad**	**PPS, Wageningen University**	**López-Ridaura, Masera and Astier 2002**	**Latin america**	**Peasant, all**	**Latin-america, worldwide**	**No indicators but method to derive them.**

15	MOTIFS	Monitoring Tool for Integrated Farm Sustainability		Meul et al. 2008	Belgium	Dairy	Flanders, worldwide	
16	PG	Public goods		Gerrard et al. 2012	UK	All	Worldwide	No direct mention of sustainability
17	**RISE**	**Response-inducing Sustainability Evaluation**	**HAFL, Bern UNiversity of Applied sciences**	**Grenz et al. 2008**	**Worldwide**	**All**	**Worldwide**	
18	**SAFA**	**Sustainability Appraisal of Farming and Agriculture**	**FAO**	**FAO 2014**	**Worldwide**	**All**	**Worldwide**	
19	SAFE	Sustainability Assessment of Farming and the Environment		Van Cauwenbergh et a. 2007		Multiple	Post-industrial contexts.	
20	SHARP	Self-evaluation and holistic assessment of climate resilience of farmers and pastoralists	FAO	Choptiany et al. 2015	Developing world	Small-scale peasant farming	Developing world	
21	SOSTARE		Regione Lombardia	Paracchini et al. 2015	Italy	Multiple	Northern Italy	Developed for regional purposes

8.2. Appendix B Indicators and relevance to sustainability

Table 9 Indicators and relevance to sustainability

Pillar	Indicators	Indicators	Appraised sustainability attribute	Causality	Reference	SDG	SDG
Environemental	Atmosphere	Greenhouse gases	Stability	GHG lead to climate change in turn affecting the productivity of agro-ecosystems in the long-run	FAO 2014	13	
		Air quality & pollution	Human integrity	*See human health*	FAO 2014	13	
	Freshwater	Water quantity /use	Stability	Water availability affects crops productivity in the long run	FAO 2014	6	14
		Water quality/pollution	Externalities and contribution	*See environmental protection*		6	14
		Water use intensity	Adaptability	Low water-use intensity provides better adaptability under changing rainfall patterns	Grenz et al. 2016	6	2
		Irrigation practices		*See water use*		2	6
	Soils	Soil organic matter content	Stability	Soil organic matter increases yield and reduces erosion	FAO 2001, 2014	2	
		Soil physical structure	Productivity	Soil structure affects root-growth (limiting water and nutrient availability) and erosion.	FAO 2001, 2014	2	
		Soil chemical quality	Productivity	Soil chemical quality (pH & CeC particularly) affect productivity.	FAO 2001	2	
		Soil erosion	Stability	Soil erosion reduces productivity in the long run.	FAO 2001, 2014	2	
		Soil compaction	Reliability	Compaction affects productivity in the short-term.	FAO 2001	2	
		Land use efficiency		Contribution to global food supply	Zahm et al. 2019	2	
		Plots size					
		Land management degradation and desertification	Adaptability	Degradation reduces productivity and cropping options	Grenz et al. 2016	2	
	Biodiversity	Habitat diversity	Stability	Hosts functional biodiversity	Costanza et al. 1997 & FAO 2014	15	2
		Ecosystem diversity	Stability	*See habitat diversity*		15	
		Ecosystem integrity	Stability	Ability of ecosystems to function and provide services.		15	
		Wild biodiversity	Reliability	Provides ecosystem services such as pollination and pest regulation	FAO 2014	15	
		Agricultural biodiversity	Resilience	Reduced risk of failures, compensation potential and habitat for functional biodiversity	FAO 2014	2	15
		Connectivity in agroecological landscape	Stability	Improves the conservation and services potential of habitats	FAO 2014	15	
		Biodiversity enhanching practices		*See biodiversity*		15	
		Conservation of heritage breeds and races	Externalities and contribution	Supports species preservation	FAO 2014	15	
	Material and energy	Non-renewable resources	Adaptability	The limited use of non-renewable energy will allow systems to function in the long-run as the cost of fossil-fuels increases	Grenz et al. 2016	13	12

	Energy use	Adaptability	*See resource use*	Grenz et al. 2016	13	12
	Eco-efficiency	Stability	High efficiency of natural resources use allows to preserve them and contribute to global food supply.	FAO 2014	13	12
	Resource conservation	Externalities and contribution			12	
	Exposure to pesticides	Human integrity	Pesticides may cause health issues.	Gilden, Huffing and Sattler	3	15
	Pesticide use	Externalities and contribution	*See exposure to pesticides*			
	Recycling, flows and closed cycles	Externalities and contribution	Circular material flow (reuse and absence of waste) reduces waste at value-chain by a more efficient use of resources.	UNEP 2011	12	
	Externalisation of energy use	Externalities and contribution	Agricultural systems may rely on energy-intensive inputs produced outside the farm	FAO 2014	13	12
	Food miles	Externalities and contribution	Food and inputs transportation use non-renewable energy		13	12
	Fertilisation	Reliability	Fertilisation enables to maintain high yields and, in some case, prevent soil depletion. Fertilisation can also cause externalities in production and leaching, causing eutrophication.	Grenz et al. 2016	2	
	Organic matter management					
	Plant protection	Stability	Plant protection can be done with agroecological means but also rely on pesticides which can be detrimental to the environment and health	Grenz et al. 2016	2	6
	Other inputs use					
	Waste reduction and disposal	Externalities and contribution	Waste represents a sub-optimal resource-use in regard to conservation. Decomposing of waste produces GHG.	FAO 2014	12	
	Livestock units					
	Livestock feed					
Animal management and welfare	Freedom from stress, welfare	Adaptability	Production systems with low welfare conditions are associated with the most detrimental practices (e.g. reliance deforested land, antibiotics residues, GHG emissions)	FAO 2014	15	
	Animal health	Reliability	*See animal welfare and productivity*		15	
	Animal productivity	Productivity	Fulfilling of productive objective and contribution to food security	López-Ridaura et al. 2005	2	
	Species-appropriate conditions	Ethics	Non-welfare systems are unsustainable because of public outcry. Welfare participates in the products' quality.	Broom 2010	15	
Agroecological practices	Crop-livestock integration	Stability	Waste reduction and on-site inputs production.		2	
	Agricultural intensity	Productivity	Too low intensity is detrimental to food production and efficiency	Musumba et al. 2017	2	

Economic and management		Access to agroecological knowledge	Stability	Sustainable practices are dependent on the access to agroecological knowledge and awareness.	Grenz et al. 2016	9	
		Agroforestry	Stability	Agroforestry offers ecosystem services while being a viable land-use option.	Jose 2009	2	
		Soil cover	Stability	Soil cover limits erosion.	Grenz et al. 2016	2	
		IPM		IMP reduces the reliance on external PPP	Lewis et al. 1997	2	
		Crop productivity	Productivity	The system meets the needs of its stewards and provides for as many as possible	López-Ridaura et al. 2005	2	
	Production	Access to decision-making information	Stability	The managers have access to information on sustainability, enabling them to take decisions accordingly.	Grenz et al. 2016	9	
		Internal investment	Reliability	Internal investment allows the system to adapt, strengthen and diversify. Investment increases the probability of progress.	FAO 2014	8	
	Investment	Community investment	Stability	Creation of dynamics and public goods at the local level.	FAO 2014	11	17
		Efficiency	Productivity	Fulfilling of the productive objective.	Grenz et al. 2016	8	
		Long-ranging/term investment	Stability	Assurance of long-term stability	FAO 2014	8	
		Financial stability	Adaptability	Enables to system to provide for livelihoods in the long-term.	Zorn et al. 2018	8	
		Stability of relationships	Stability	Variations in the market affect income	FAO 2014	8	
		Liquidity	Reliability	Ensures that livelihoods can be provided for and that the system is economically viable.	Grenz et al. 2016	8	
		Employment	Reliability	Employs enables to fare for livelihoods, a direct goal of social sustainability		8	
		Secure land tenure/ competing land claims		The farm cannot function in the long-term if the land its needs is a risk.		2	
	Vulnerability and resilience	Income diversification	Resilience	A diversified income base makes livelihoods less sensitive to shocks and more resilient in recovery	Zahm et al. 2008; FAO 2014	8	
		Sensitivity to subsidies	Reliability	Subsidy reliance shows that the system is not productive enough or not adapted to the market. It is sensitive to external decisions.	Zahm et al. 2008	8	
		Risk management	Reliability	A risk management strategy helps the system to predict and withstand variations	Zahm et al. 2008	8	
		Demand variability				8	
		Diverse client base	Resilience	Reduces dependency on single economic actors	Joyce Nohria and Robertson 2003	8	
		Stability of production	Stability	A system is sustainable if it is able to reliably provide for the livelihoods	Grenz et al. 2016	2	

Social	Product safety and quality	Product information and traceability	Reliability	To make sustainable choices, producers and consumers need to access information about their purchases.	FAO 2014	12	
		Food safety	Human integrity	Food safety is a basic human right, stated in the sustainable development goals.	FAO/WHO 2003	3	
		Food quality	Human integrity	Right to quality food	FAO/WHO 2003	3	2
	Local economy	Direct sales	Externalities and contribution	Component of food sovereignty	Wittman, Desmarais, and Wiebe 2010	9	11
		Local sales	Stability	Local sales reduce food miles and strengthens the territory's food sovereignty		11	8
		Value creation	Productivity	Value creation is central to economic sustainability byt ensuring that the farm provides desirable goods and services.	FAO 2014	8	11
		Local procurement	Stability	Local procurement contributes to make the economy more dynamic.	FAO 2014	8	11
	Management practices	Income sources	Stability	*See income diversification*			
		Personnel management and cost	Productivity				
		Sustainability-oriented management	Stability	Sustainability results are" heavily dependent on the approach and quality of the farm management". A good management will ensure long-term viability.	Grenz et al. 2016	8	15
	Autonomy	Labour availability	Reliability	Obtained from workshops		8	
		Labour turnover	Stability	Obtained from workshops		8	
		Net income, profitability	Productivity	Income assures that livelihoods are supported and that the enterprise is viable.	FAO 2014	1	8
		Indebtedness	Stability	Indebtedness shows a limited capacity of internal investment.	Grenz et al. 2016	8	
		Reliance on externally produced inputs	Reliability	Sensitivity to inputs availability	FAO 2014	8	
	Decent livelihood	Wage level and economic wellness	Human integrity	Decent wages ensure that livelihoods are provided for and dignity respected	FAO 2014	1	3
		Fair trade practices	Human integrity	Respect of human rights, dignity and integrity	FAO 2014	10	8
		Quality of life on workplace	Human integrity	Respect of human dignity and integrity	FAO 2014	3	8
		Livelihood security	Human integrity	*See economic wellness*		1	
		Capacity building	Resilience	Capacity building enables development and satisfaction linked with *quality of life*	FAO 2014	4	9
	Labour right	Employment relations and connectedness	Human integrity		FAO 2014	17	8
		Forced labour	Human integrity	Respect of human rights, dignity and integrity	FAO 2014	8	
		Child labour	Human integrity	Respect of human rights, dignity and integrity	FAO 2014	8	
		Freedom of association and bargaining	Human integrity			8	1

	Working hours	Human integrity	Long hours cause non-wellbeing and leads to sickness .	Ala-Mursula et al. 2008	8	3
Equity	Non-discrimination	Human integrity		FAO 2014	10	
	Gender equality	Human integrity		FAO 2014	5	
	Women empowerment	Human integrity	*See gender equity*		5	
	Income distribution	Human integrity	Feedback loop between inequality and unsustainability.	Neumayer 2013	10	
	Support to vulnerable people	Human integrity	Respect of human integrity	FAO 2014	10	
Human health and safety	Physical and psycho-social health	Human integrity		FAO 2014	3	
	Public health	Stability		FAO 2014	3	
	Safety at work	Human integrity		FAO 2014	3	8
	Provision of healthy food	Human integrity		FAO 2014	3	2
	Food security	Human integrity		FAO 2014	2	
Cultural diversity	Indigenous knowledge and rights	Human integrity	Core component of food sovereignty	Wittman, Desmarais, and Wiebe 2010	10	
	Cultural preservations	Productivity	Cultural appropriate measures are a success factor for development operations	UNSECO 2008	10	
	Freedom and values	Human integrity			10	
	Food sovereignty	Stability	Promotes sustainable agriculture in its 3 dimensions.	Wittman, Desmarais, and Wiebe 2010	Multiple	
Platforms and sharing	Food literacy, education and awareness raising	Stability	Contribution to knowledge dissemination and food sovereignty	Wittman, Desmarais, and Wiebe 2010	4	
	Knwoledge creation				9	8
	Knowledge creation and sharing	Productivity	Research and innovation as productive activity		4	9
	Total					

8.3. *Appendix C: Indicators and reference for measurement for MESMIS framework.*

Table 10 Indicators and reference measurement methods for MESMIS frameowrk

Criteria	Indicator	Method reference
Financial stability	Variation in Gross Operating Surplus in last 5 years	
Financial stability	debt/total equity	
Subsidy reliance	Subsidies/Gross operating surplus	Zahm 2008
Competing land use	Land lost in 5 years	Niewöhner et al. 2016
Competing land use	Competing use	
Competing land use	Frozen long-term land use agreement	
Competing land use	Share of rented land	
Parent organisation	Part of an organisation	
Soil preservation	Humus balance (calc) or SOM balance (lab)	Hénin and Dupuys, cited in Flury et al. 2015
Soil preservation	pH	Hijbeek et al. 2018
Soil preservation	Compaction (% of land)	
Soil preservation	Erosion (t/ha/year)	
Labour availability	Applications/opening	
Labour turnover	% of employees with more than 5 years	Zahm 2008
Activities diversification	% of turnover generated by the main activity	Zahm 2008
Goals diversification	Nb of stated goals	Zahm 2008
Customer base	Nb customers buying more than 25%	IDEA v3
Sustainability-oriented management	Mission statement	
Too high demand	CSA waiting list	
Adaptation to demand	% of unsold crops	
Water access security	% of year with water restrictions	
Inputs availability	% input (in t or MJ equ.) not available in less than 100 km	
Land use efficiency	$ revenue/hectare	
Waste	% of land not harvested	
Labour efficiency	(Total income-labour)/Total income	Zahm 2008
Crop productivity	Yield/base yield	Musumba et al. 2017
Animal productivity	Yield/base yield	
Income	GOS (with Labour except 1 manager)/average GOS	FAO upcoming
Soil quality	Water reserve	
Non-renewable resource use	In MJ equivalent (on farm use)	
Input-use efficiency	(Output-input)/output	Zahm 2008
water usage	Irrigation practices	Zahm 2008
water usage	water use/ rainfall	
Reliance on external inputs	Imported inputs/total inputs	Zahm 2008
Reliance on external inputs	Imported seeds/total seeds	Zahm 2008
Closing cycles	On-farm inputs % of total value	
Food production	Calories/ha	Zahm 2008
Job well-being	Perception	Forbes
Wages	Wage/minimum wage	Jobbank Gov. Of Canada Accessed 2019
Capacity building	Employee training/year	Zahm 2008
Respect, integrity and support	Discrimination	FAO 2014
Culture preservation	Ceremonial and traditional crops	
Indigenous rights	% of land used without agreement of indigenous	

Income inequality	Max wage/lowest wage	
Food literacy and awareness	Hours dedicated	
Use of unpaid labour	Contributing unpaid labour hours /total labour hours	
Gender equality	Managerment control (only if more than 3 managers)	
Gender equality	Gender balance in entire operation	
Contribution to higher education	Work hours in education	
Lower education	Classes received	
Formal research	nb of on going research projects	
Interns training	Nb of intern hours/total hours	
Experimentation (other)	%land under experiments	
Visibility	Farm public visibility	
Visibility	Transparency on operations	
Habitat	m² habitat	
Pollution	Use of chemicals beyond recommendation	Grenz et al. 2016
Pollution	No nitrogen in 5meters of water bodies	Grenz et al. 2016
Pollution	GHG	Hillier et al. 2011
Forested area	m² forest/total land	
Biodiversity	Nb of habitats	Dennis et al. 2012
Biodiversity	Biodiv est.	Hillier et al. 2011
Cropped diversity	Nb cropped species	Leyva and Lores 2018
Sustainability assessments	Assessment performed /year (for one indicator or more of sustainability)	
Engagement	CSA subscriptions	Garret and Feenstra 1999
Engagement	Part of a local farmers' association	Garret and Feenstra 1999
Nutrition	Nb of people fed/ha	
Nutrition	% of highly nutritious food in mass	
Protects health	Pesticide residues	
Locally adapted crops	% of land with local crops	
Heritage crops	% of land with heritage crops	Grenz et al. 2016
Customer proximity	% sales locally	
Territory embededness	Local clients and short value-chains	
Territory embededness	Locally produced inputs	
Employment	FT job openings per year (except students)	
Community engagement	Assessment on Community services	
Space	Space/animal	
Behaviour hindrance	SPCA criteria	BCSPCA Accessed 2019
IPM	Area under IPM %	
Cover crops	(duration bare soil (year) x size /total size	
Use of ecosystem services	ecosystem services cited	
Organic inputs	Organic fertilisation in % of fertilisation	
Sourcing sustainable inputs	Share of inputs are certified by a recognised organisation	

8.4. Appendix D: Scores and deviation per indicator for each framework:

Table 11 Indicators comparison with score and deviation

Indicators	Indicators	RISE	IDEA	SAFA	MESMIS	Divergence	Available through Lite Farm
Atmosphere	Greenhouse gases	1		1.00	1		Output
	Air quality & pollution	0.91		0.78	1	23%	
Freshwater	Water quantity /use	0.39	0.5	0.56	0.3	25%	Output
	Water quality/pollution	0.8		0.83		2%	
	Water use intensity	0.98					Derived
	Irrigation practices	0.98	X	X			
Soils	Soil organic matter content	0.48		1.00	1	52%	Output
	Soil physical structure	X		0.56	0.557	1%	
	Soil chemical quality	0.92		1.00		7%	Output
	Soil erosion	1	X	1.00	0.97	3%	
	Soil compaction	1		1.00	1	1%	
	Land use efficiency		X	0.00			Derived
	Plots size		X	0.00			Output
	Land management degradation and desertification	0.84		0.78		7%	
Biodiversity	Habitat diversity	1	1	X			
	Ecosystem diversity			1.00	1	1%	Output
	Ecosystem integrity	1		1.00	1	1%	
	Wild biodiversity	0.85		0.78		8%	
	Agricultural biodiversity	0.72	1	1.00	0.862	28%	Output
	Connectivity in agroecological landscape	1		1.00		1%	
	Biodiversity enhanching practices	0.93		1.00		6%	?
	Conservation of heritage breeds and races	1	0.5	0.78	1	50%	
Material and energy	Non-renewable resources	0.5		0.56		5%	Output
	Energy use	0.39	0.8	0.63	?	41%	Output
	Eco-efficiency	0.65					Derived
	Resource conservation			0.00			
	Exposure to pesticides						
	Pesticide use			0.00			Output
	Recycling, flows and closed-cycles	0.5			?		
	Externalisation of energy use			0.78			Derived
	Food miles			0.56			
	Fertilisation	0.39	0			39%	Output
	Organic matter management	X	X	0.00			Output
	Plant protection	0.69	1		?	31%	Output
	Other inputs use	X	X	0.00			Output
	Waste reduction and disposal		0.4	0.95		55%	
Animal management and welfare	Livestock units	X		0.00			
	Livestock feed	X		0.00			
	Freedom from stress, welfare	1	1	1.00		1%	
	Animal health	0.5					

	Animal productivity	0.55					
	Species-appropriate conditions	1	1	1.00	1	1%	
Agroecological practices	Crop-livestock integration						
	Agricultural intensity	0.5					Derived
	Access to agroecological knowledge						
	Agroforestry	NA					
	Soil cover	1	1	X	1		Output
	IPM		1	1.00	1	1%	
Production	Crop productivity	0.51			0.594	8%	Output
	Access to decision-making information	0.77					
Investment	Internal investment	X	0.8	0.78		3%	
	Community investment			1.00	X		
	Efficiency		0.96	1.00	0.938	5%	
	Long-ranging/term investment			NA			
Vulnerability and resilience	Financial stability	0.69			1	31%	
	Stability of relationships	0.58		1.00		41%	
	Liquidity	X		0.44			
	Employment	X	1		1		Output
	Secure land tenure/ competing land claims				0.74		
	Income diversification	0.69	0.9	X	0.625	28%	Output
	Sensitivity to subsidies		1	X	0.998	0%	
	Risk management	1		0.78		23%	
	Demand variability			0.00			
	Diverse client base						
	Stability of production			0.89			
Product safety and quality	Product information and traceability			0.70			
	Food safety			1.00	X		
	Food quality		0.7	1.00	X	29%	
Local economy	Direct sales				1		Output
	Local sales				1		Output
	Value creation			0.85			Derived
	Local procurement		0.8	1.00	?	19%	
Management practices	Income sources			0.00			Output
	Personnel management and cost	X		0.00			Output
	Land cover/land use	X	X	0.00			Output
Autonomy	Sustainability-oriented management	0.77		0.89		11%	
	Inputs availability			0.00			
	Labour availability			0.00	1		
	Labour turnover			0.00	0.5		
	Net income, profitability	0	1	1.00	0.475	100%	Output
	Indebtedness	0.5	1		1	50%	
	Reliance on externally produced inputs				?		
Decent lilelihood	Wage level and economic wellness	0.16		1.00	0.24	83%	
	Fair trade practices			1.00			

	Quality of life on workplace	0.85	1	1.00		15%	Output
	Livelihood security	NA					
	Capacity building	1	1	1.00	1	1%	
Labour right	Employment relations and connectedness	0.88			0.82	6%	
	Forced labour	1		1.00		1%	
	Child labour	1		1.00		1%	
	Freedom of association and bargaining			X			
	Working hours	0.75	0.7		0.7	5%	Output
Equity	Non-discrimination	1		1.00	1	1%	
	Gender equality	1		1.00	1	1%	Derived
	Women empowerment						
	Income distribution				0.244		
	Support to vulnerable people			1.00			
Human health and safety	Physical and psycho-social health	0.88		1.00		11%	
	Public health			1.00			
	Dietary diversity						Derived
	Safety at work		1	1.00		1%	
	Provision of healthy food				92.3		Derived
	Food security		0.6				
Cultural diversity	Indigenous knowledge and rights			0.78	1	23%	
	Cultural preservations				1		
	Freedom and values	0.75					
	Food sovereignty			1.00			
Platforms and sharing	Food literacy, education and awareness raising			1.00			
	Knowledge creation			0.00			
	Knowledge creation and sharing						
	Total	52	30	84	39		