Star4bbs

D1.4

Report on existing monitoring schemes, with recommendations for new system

Funded by the European Union

www.star4bbs.eu
info@star4bbs.eu

@STAR4BBS

D1.4

Report on existing monitoring schemes, with recommendations for new system

DELIVERABLE TYPE

Report

MONTH AND DATE OF DELIVERABLE

M12, 08/31/2023

WORK PACKAGE

WP 1

LEADER

TUB

DISSEMINATION LEVEL

Public

AUTHORS

Nikola Matovic, Luana Ladu

Programme	Grant Agreement	Start	Duration
HORIZON EUROPE	101060588	Sept.2022	36 Months

Funded by the European Union

Contributors

NAME	ORGANISATION
Nikola Matovic	TUB
Luana Ladu	TUB

Peer Reviews

NAME	ORGANISATION
Tilman Denkler	BAM
Kristin Komives	ISEAL
Enrica Imbert	UNITELMA

Revision History

VERSION	DATE	REVIEWER	MODIFICATIONS
D1.4_ draft1	17/07/2023	Nikola Matovic	First draft
D1.4_ draft2	20/07/2023	Luana Ladu	Review and editing
D1.4_ draft3	25/07/2023	Tilman Denkler	General remarks, results improvement, editing
D1.4_draft4	09/08/2023	Kristin Komives	Review and suggestions for improvement
D1.4_draft5	28/08/2023	Enrica Imbert	Review and suggestions for improvement
D1.4_final_draft	29/08/2023	Luana Ladu	Final remarks
D1.4_revised	08/07/2024	Luana Ladu, Nikola Matovic	Updated version; Changes made according to reviewers comments

Index of Contents

Index of Tables

Index of Figures

Partners short names

TUB	Technische Universitat Berlin
UNITELMA	Università degli studi Unitelma di Roma
UNI	Ente Italiano di Normazione
AUA	Geoponiko Panepistimion Athinon
USC	Universidad de Santiago de Compostela
APRE	Agenzia per la Promozione della Ricerca Europea
NOVA	Institut für politische und Ökologische Innovation GMBH
BB	Better Biomass
BAM	Bundesanstalt für Materialforschung und Prüfung
RSB	Roundtable on Sustainable Biomaterials Association
ISEAL	Iseal Alliance

Abbrevations

ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
ASC	Aquaculture Stewardship Council
B2B	Business-to-Business
BAP	Best Aquaculture Practices
CAT	Certification Assessment Tool
CCRF	Code of Conduct for Responsible Fisheries
CEPI	Confederation of European Paper Industries
CFMB	Corporate Fiber and Materials Benchmark
CFRN	Canadian Fisheries Research Network
CGF	Consumer Goods Forum
DoC	Degree of Criticality
DoO	Degree of Obligation
EU	European Union
FAO	Food and Agriculture Organization of the United Nations
FEFAC	European Feed Manufacturers' Federation
FSC	Forest Stewardship Council
FSI	Floriculture Sustainability Initiative
GAP	Good Agricultural Practices
GIZ	Deutsche Gesellschaft für Internationale Zusammenarbeit
GLR	Grey Literature Review
GSCP	Global Social Compliance Program
GSSI	Global Sustainable Seafood Initiative
IDH	Sustainable Trade Initiative
ISCC	International Sustainability and Carbon Certification
ISO	International Organization for Standardization
ITC	International Trade Centre
LCA	Life-cycle Assessment

MANOVA	Multivariate Analysis of Variance
MAVT	Multi-Attribute Value Theory
MCDA	Multi-Criteria Decision Analysis
MCI	Material Change Index
MNE	Multinational Enterprises
MS	Monitoring System
MSC	Marine Stewardship Council
NAP	German National Action Plan on Business and Human Rights
NGO	Non-Governmental Organisation
OECD	Organisation for Economic Co-operation and Development
PAA	Arrangement Analysis Approach
PEFC	Programme for the Endorsement of Forest Certification
QMS	Quality Management Systems
RED	Renewable Energy Directive
REML	Restricted Maximum Likelihood Approach
RSB	Roundtable on Sustainable Biomaterials
SAFA	Sustainability Assessment of Food and Agriculture Systems
SCS	Sustainability Certification Schemes
SCT	Sustainability Certification Tools
SDGs	Sustainable Development Goals
SEM	Standard Error of the Mean
SIFAV	Sustainability Initiative Fruit and Vegetables
SLR	Systematic Literature Review
SME	Small and Medium-sized Enterprises
SSCI	Sustainable Supply Chain Initiative
SSCT	Sustainability Standards Comparison Tool
STAR4BBS	Sustainability Transition Assessment Rules for Bio-Based Systems
Star ProBio	Sustainability Transition Assessment and Research of Bio-based Products
T4SD	Trade for Sustainable Development
UNFSS	United Nations Forum on Sustainability Standards
UNGP	United Nations Guiding Principles on Business and Human Rights
VSS	Voluntary Sustainability Standards
WBA	World Benchmarking Alliance
WP	Work Package
WWF	World Wide Fund for Nature

Definitions of important terms

- **Certification** = A procedure by which a third party (independent certification body) gives written assurance that a product, process or service is in conformity with certain standards. (ISO, 1996)
- **Certification Label** = A certification label is a label or symbol indicating that compliance with standards has been verified. (FAO, 2003)
- **Certification Scheme (Program)** = Certification system related to specified products, to which the same specified requirements, specific rules and procedures apply (ISO/IEC 17065:2012; RSB Glossary of Terms)
- **Certification System** = Rules, procedures and management for carrying out certification. (Source: ISO/IEC 17000:2004)
- **Standards** = Documented agreements containing technical specifications or other precise criteria to be used consistently as rules, guidelines or definitions, to ensure that materials, products, processes and services are fit for their purpose. (FAO, 2003)
- **Voluntary Sustainability Standards (VSS)** = Voluntary guidelines used by producers, manufacturers, traders, retailers, and service providers to demonstrate their commitment to good environmental, social, ethical, and food safety practices. (UNFSS, 2012)

Executive Summary

This deliverable contains a review of existing systems for monitoring sustainability certification schemes (SCS) and labels, and, more specifically, a thorough analysis of different monitoring tools used to assess and compare those schemes. The purpose is to understand the range of different monitoring systems and methodologies, identify best practices and assess the appropriateness of the selected systems' methodological approaches for the goals of the STAR4BBS project - evaluating the robustness and effectiveness of certification systems in achieving sustainability goals.

The findings of an initial systematic literature review (SLR) of peer-reviewed papers indicate the use of various methodological approaches in evaluating sustainability standards and certification schemes. To account for recognized variety, and after identifying similarities, we clustered the studies into three assessment groups: benchmark, comparison, and exploring a relationship. By analyzing the methodologies used in each assessment group, we aimed to identify the strengths and limitations of different approaches at 3 key stages and gain insights into the factors that may influence the results of each type of study: study design & data collection methods, data analysis, representation and interpretation.

The research conducted within SLR highlights the importance of comprehensive data analysis and emphasises that no single method is inherently superior but that different methods are appropriate for different purposes. Rigorous and systematic data collection and analysis are essential for providing accurate and reliable information in the monitoring system. Understanding different methodological approaches helps make informed decisions in designing a new monitoring system effectively, especially in its conceptualization phase (see STAR4BBS deliverable *D4.1. Concept of the monitoring system*). In particular, developing a new monitoring system necessitates careful consideration of the scope, audience, and contextual factors in choosing an appropriate methodological approach.

The initial SLR was complemented with the review of grey literature (GLR) and existing relevant databases. A selection of a few best-fitted tools and their in-depth analysis enriched these reviews and served to obtain a more detailed picture of their features and functionalities. Those tools included: the FEFAC Responsible Soy Benchmarking Tool, Siegelklarheit, ITC Standards Map, SME Compass, FSI Basket of Standards, and GSSI Global Benchmark Tool. These tools assess certification schemes and labels from different sectors, analysing their purposes, benchmarking approaches, and evaluation framework, thus shedding light on their varied applications and effectiveness.

The reviewed monitoring tools share several commonalities regarding their purpose, key features, database assessment, and data interpretation. All tools provide a structured assessment of compliance with specific criteria related to sustainable and responsible practices, whether environmental, social, or governance-related. These tools commonly rely on comprehensive databases to collect and analyze data, ensuring thorough evaluation. Environmental and social indicators are widely employed, reflecting the importance of assessing the ecological and societal impacts of certified products or feedstock. Traceability indicators are also common, particularly where operational

requirements (i.e. system-level assessments) are covered. Most of the tools are ongoing systems rather than one-off projects. The visualization of results is a key feature, with most tools presenting findings in accessible formats such as ratings, charts, or tables to facilitate comparison and decision-making. Qualitative assessments focusing on specific sustainability requirements are a recurring element.

Despite these commonalities, there are notable differences among the tools. Each tool targets specific sectors or product groups, tailoring their criteria and assessment methodologies accordingly. The tools emphasize the importance of either sustainability content performance or operational systems, or both, with rare occasions of covering outcomes of the schemes (impacts). Specific approaches among tools differ due to distinct goals and target audiences, reflected in performance depths, information access, and geographical coverage. Some tools have multiple target audiences, with assessments and information relevant to various stakeholders involved at different supply chain stages or for different sustainability practices and certification purposes. The approaches to data interpretation also differ, ranging from pass/fail results to detailed scoring systems and star ratings. The visualization methods vary, with some tools using pie charts or percentage bars, while others provide comprehensive benchmark reports or integrate results into larger mapping platforms. These distinctions reflect the unique needs and priorities of the sectors each tool serves.

Drawing from the data and analysis of the selected six monitoring tools, it is important to consider the clarity of criteria, the ability to address regional issues, the implementation of standards, data availability and accessibility, regular revisions, and comprehensive scheme guidance when designing a new monitoring system. Finally, the report generated recommendations for important features that should be included in the new monitoring system to be developed for the bio-based industry. Those include leveraging data from existing sources (e.g. ITC Standards Map), incorporating the multidimensional evaluation approach of the Siegelklarheit methodology, engaging stakeholders, and considering methodological approaches, which include scoring and setting minimum thresholds. Continuous improvement and periodic revisions are crucial for the system's effectiveness. In the table below key components of the proposed Framework for continuous improvement of the system and its regular revisions are summarized:

Component	Objective	Activities	Frequency and Timing
Testing and Validation	Test and validate the monitoring system by assessing pilot SCS and labels	- Test the applicability of the monitoring system by assessing SCS and labels - Validate changes in the structure of the system and its indicators	At least two rounds of pilot testing should be conducted

Feedback Mechanism	Collect and integrate feedback from stakeholders	- Internal feedback through team meetings, surveys, and designated channels - External feedback through surveys, workshops, and stakeholder platform	Feedback should be collected for the initial tool development, for the testing rounds, and for each revision and update
Regular Reviews and Updates	Systematically evaluate and update the monitoring system	- Update of the indicators - Revision and update of the thresholds and metrics - Communication of reviews and revisions with stakeholders	Biannual (comprehensive) Periodical (as needed, based on feedback)
Revisions	Implement feedback and ensure revisions are effective	- Development of revision proposals - Validation and testing - Release of updates	Following iterative process of reviews and updates

The concept for the monitoring system (deliverable D4.1 of the STAR4BBS) was structured based on the findings of the literature review, in-depth analysis of the 6 monitoring systems, and following the steps recommended in the ISEAL's Sustainability Benchmarking Good Practice Guide (ISEAL, 2020). The conceptualization was carried out following three non-sequential but highly connected phases: determine the audience and purpose, which elements to include, and the evaluation structure.

Besides informing the conceptualization and design of a novel monitoring system, these insights will contribute to the overall objective of the project, which is to promote sustainability and improve the environmental, social, and economic performance of the bio-based industry.

1 Introduction

The development of sustainable bioeconomy has been identified as a promising pathway to foster sustainable economic growth, providing opportunities for reducing carbon emissions and decreasing dependence on non-renewable resources (EU

Bioeconomy Strategy[1], EU Sustainable Carbon Cycle Communication[2]). This transition towards a circular bioeconomy has been extensively studied and existing barriers have been identified (Gottinger et al. 2020), in particular, concerns about the sustainability of bio-based value chains. Sustainability certification schemes (SCS) and labels are relevant instruments to assess the sustainability of bio-based products monitoring, reducing and overcoming environmental and socio-economic risks (Ladu & Blind, 2017, Ladu & Morone, 2021). The role of SCS and labels in accelerating the development of circular bioeconomy and in achieving several SDGs is also highlighted in EU policies, including the EU Bioeconomy Strategy, the Circular Economy Action Plan[3] and the European Ecodesign Directive[4]. Notably, the EU Strategy on Standardisation[5] highlights that the digital and green transition of EU industries should rely on a well-functioning and resilient standardisation system that adequately reflects EU policy priority.

However, the rapid proliferation of the certification schemes and labels led to questioning their effectiveness and robustness and to what extent they deliver important outcomes, in line with sustainability policy priorities. This also creates confusion among consumers and a reduction of trust in sustainability certification systems. In addition, existing schemes and labels tend to overlook various crucial aspects when assessing sustainability, including end-of-life options and circularity (Ladu & Morone, 2021); the risk of indirect land use changes; and social and economic impacts (Majer et al. 2018; Villamil and Hallstedt, 2020). Higher transparency, robustness and harmonization of SCS and labels are therefore needed, providing a holistic and accurate level of assurance and sustainability coverage. Trade-offs between the rigor and complexity of the tools/assessment methodologies and the principles of feasibility and effectiveness (structure as simple as possible) should also be considered.

To assess the transparency, credibility and effectiveness of the SCS and labels, monitoring systems are being increasingly developed as evaluation tools. By providing reliable and accessible information about the environmental, social and economic impacts of different certification schemes and labels, they could support sustainability achievements and improvements in the bio-based industry (Majer et al. 2018).

In this regard, the STAR4BBS project aims to develop indicators and a monitoring system to assess the effectiveness and robustness of existing sustainability certification schemes (SCS) and labels applicable to biological feedstock and bio-based materials and products. The project's overall goal is to maximize the potential of these certification schemes and labels to support a successful transition to a sustainable bio-based economy while aligning with EU policy objectives and the Sustainable Development Goals (SDGs). A fit-for-purpose monitoring system will be one of the main achievements of this project, and it will be used to assess existing SCS and labels against a range of indicators associated with the effectiveness of these market-based instruments in supporting improved sustainability outcomes. As indicated in the Box 1, within the HORIZON-CL6-2021-ZEROPOLLUTION call, a Joint Monitoring System will be developed by the three sister projects.

[1] https://op.europa.eu/en/publication-detail/-/publication/edace3e3-e189-11e8-b690-01aa75ed71a1/language-en/format-PDF/source-149755478
[2] https://climate.ec.europa.eu/system/files/2021-12/com_2021_800_en_0.pdf
[3] https://environment.ec.europa.eu/strategy/circular-economy-action-plan_en
[4] https://faolex.fao.org/docs/pdf/eur90694.pdf
[5] https://ec.europa.eu/docsroom/documents/48598

Box 1

The **STAR4BBS** project (Sustainability Transition Assessment Rules for Bio-Based Systems) is a Coordination and Support action, which addressed the Horizon Europe call HORIZON-CL6-2021-ZEROPOLLUTION-01-07: *International and EU sustainability certification schemes for bio-based systems*. Two other projects – **HARMONITOR** (Harmonisation and monitoring platform for certification schemes and labels to advance the sustainability of bio-based systems) led by the SQ Consult, and **SUSTCERT4BIOBASED** (Sustainability Certification for Biobased Systems) led by the Stichting Wageningen Research, were also awarded to address the same call. The three sister projects work together in the implementation of different joint activities. One of them is the development of a Joint Monitoring System, of which the initial proposal, accepted by the EU officials in June 2023, is included in the Annex of *D4.1 Concept of the monitoring system*.

The goal of the three sister projects working together to develop a Joint Monitoring System (JMS) is to reduce confusion, divergences, and mistrust among stakeholders by creating a harmonized, overarching system. This would bring coherence to the space and clarity for policymakers driving the transition to a bioeconomy in the EU. Working together would allow the projects to build on each other's knowledge and experience, subjecting the JMS to a higher level of scrutiny, and maximizing the effective use of resources. The JMS would streamline stakeholder consultations and reduce fatigue while eliminating competition among the three projects and maximizing the synergies and impacts of the results. The creation of a JMS will require greater coordination, but it is believed to be feasible and worthwhile to work together to provide a more comprehensive and detailed tool, covering a wide range of bio-based sectors and products.

In this deliverable, we will refer to the term "Monitoring System" because it provides recommendations generated within the implementation of STAR4BBS project activities. The deliverable will provide inputs to the development of the Joint Monitoring System.

The monitoring system will incorporate a comprehensive set of indicators aimed at gathering crucial information on the effectiveness and robustness of certification schemes and labels. These indicators will be structured into three levels:

I) System Level: These indicators will focus on the characteristics of the certification scheme, including its governance and the development process of standards or labels.

II) Content Level: These indicators will specify the requirements of the certification scheme concerning various EU environmental, social, economic, and circularity priorities and targets. Minimum requirements that all SCS and labels should adhere to will be defined with a life cycle perspective.

III) Outcome Level: The outcome indicators will measure the impact generated by the certification schemes and labels. They will encompass life-cycle assessment comparison indicators and continual improvement indicators.

This deliverable identifies and reviews existing monitoring systems that assess and compare sustainability certification schemes (SCS) and labels. It intends to provide a comprehensive understanding of existing monitoring systems and evaluate the methodologies currently used for benchmarking and comparing certification systems, to determine whether they can be effectively applied to the development of a novel monitoring system. We focused our search on monitoring and benchmarking systems in general, not necessarily strictly related to bio-based systems.

The reminder of this deliverable is structured as follows:

Chapter 2 - describes the methodology used for the assessment of existing monitoring systems.

Chapter 3 - presents the results from the systematic literature review (SLR), particularly main findings and key lessons learned.

Chapter 4 - presents the results from the grey literature review (GLR), including in-depth analysis and key lessons learned.

Chapter 5 - outlines recommendations for important features that should be included in the new monitoring system developed for bio-based industry.

Chapter 6 - draws important conclusions and summarizes the findings from the previous chapters, reflecting on results and lessons learned, highlighting best practices.

2 Methodology

For the assessment of existing monitoring systems, we adopted a four-step approach in methodology: (1) identification of existing monitoring systems; (2) review and qualitative synthesis of identified systems (3) selection of monitoring tools, and (4) in-depth analysis of the selected tools (Fig. 1).

Figure 1 Four-step methodology approach

Step 1: Identification of existing monitoring systems

As indicated in Figure 1, the identification of existing monitoring systems and related methodological approaches started with an extensive desktop research of the various categories of documents, including: peer-reviewed publications, grey literature and existing tools and relevant databases.

We firstly started with the **systematic literature review (SLR)**, targeting scientific articles which cover the task topic, following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, to ensure the consistency of our work (Moher et al., 2009). These guidelines are a set of evidence-based recommendations for reporting systematic reviews and meta-analyses in scientific publications. Figure 1 of Annex A3 presents the four-step process we applied, and the number of studies considered at each stage.

For the identification of records, we used the interdisciplinary academic database - Web of Science (WoS). To reduce the risk of selection biases, we chose a systematic approach, which enabled us to identify a large number of publications that represent the research field as precisely as possible, while avoiding individual decisions to include or exclude a certain publication.

As the starting point for selection, we used the following keywords and search strings:

- monitoring system AND certification schemes: yielded 85 results
- monitoring system AND labels: 10.763 results
- benchmarking AND certification schemes: 47 results
- benchmarking AND labels: 5.873 results
- comparing AND sustainability certification schemes: 66 results

As two keyword strings alone yielded more than 16.630 results, we decided to substitute the word “labels” with “B2B labels”, and tie it with “benchmarking” and “monitoring”, i.e.:

- benchmarking AND B2B labels: yielded 2 results.
- monitoring AND B2B labels: yielded 1 result.

This process led to an output of 201 publications. After removing duplicates, a list of 142 studies remained.

In order to further select and include only relevant literature in the final sample, the publication had to meet three criteria:

(I) the study explores monitoring, benchmarking or assessment of sustainability certification schemes and labels;

(II) the study compares at least two different sustainability certification systems;

(III) the research questions aim at assessing the contribution of SCS and labels to sustainability outcomes.

After excluding literature that does not address previously mentioned criteria, by reviewing titles and abstracts of 142 marked studies, a selection of 31 studies remained. Following the content screening process, a total of 26 studies were selected for the purpose of conducting a qualitative synthesis (Fig. 1 of Annex A3).

At the same time, we performed **search of the grey literature**, using keywords and strings in search engines and databases.

For the identification of relevant grey literature papers and databases, we paired anchor keywords with search strings. The following keywords and search terms were used for search engines and databases:

- benchmarking tool
- monitoring certification schemes
- monitoring systems

The above keywords were combined with the following search strings:

- sustainability
- sustainability certification schemes
- labels
- B2B labels

This was followed by existing tools and relevant databases search.

The preliminary list of identified tools was further integrated with key inputs and sources from project partners (e.g. ISEAL) and external stakeholders. In particular, we interacted with a relevant existing benchmarking initiative (ITC Standards Map), from which we could gain insights on benchmarking systems to be considered in our analysis.

The identification of existing databases and grey literature yielded 23 results, which were generated as a preliminary list and grouped together in an Excel document (reported in Annex A2). Of those, we discarded 4 as not eligible (e.g. they were classified as a certification scheme or are used for benchmarking responsible sourcing practices in supply chains etc). In total 19 monitoring tools were selected for the review.

Step 2: Review and qualitative synthesis of identified systems

Step 2 started with the review of the identified 26 academic papers resulting from the SLR (see Annex A1) and of the 19 selected monitoring tools resulting from the GLR. Due to the specificities of the findings (i.e. differences in the methodologies and study scopes), different approaches for conducting reviews and qualitative synthesis were applied.

Within the matrix, we categorized GLR findings into 11 main parameters, while SLR categorization consisted of 7 parameters (see Annex A4). The GLR matrix was the basis for further qualitative synthesis of the identified monitoring tools and an in-depth analysis of the selected tools.

The finalized categorization was presented to the STAR4BBS Consortium, the respective members of the other two projects under the same call (HORIZON-CL6-2021-ZEROPOLLUTION-01-07)—HARMONITOR and SUSTCERT4BIOBASED—and the Joint Advisory Board members during the second STAR4BBS co-creation workshop to collect their feedback and input on the review task.

Step 3: Selection criteria development and evaluation of grey literature records

Based on the review step and collated list of 19 tools identified through the review of grey literature and existing databases, a selection of the best-fitting systems was based on the following carefully designed selection criteria, developed by TUB and ISEAL (Table 1).

Table 1 Selection criteria with their definitions and interpretation for the identified monitoring tools

Selection criteria	Definition	Interpretation
Coverage	This criterion refers to the coverage of certification systems and labels that the monitoring systems are evaluating. It can encompass both the breadth and depth of coverage, including different stages of value chains, the number of schemes and labels covered, as well as the types of criteria and indicators used to evaluate them. Additional value: SCS analyzed and selected in WP1.	The MS would be suitable for selection if it covers a range of value chain stages, various schemes and/or labels, and criteria for their assessment. It would be scored higher if it evaluates SCS selected within WP1 of STAR4BBS.
Comprehensiveness	This criterion refers to the degree to which the monitoring systems analyzes all three levels of indicators, i.e. system, content and outcomes level. It can also include the degree to which the systems are able to capture the relevant dimensions and aspects of sustainability, including environmental, social, circular and economic dimensions, as well as issues related to governance, transparency, and stakeholder engagement.	The MS would be suitable if it analyzes all database indicators (system, content and outcome level) and 4 pillars (economic, social, environmental and circular), set as a goal within the STAR4BBS project.
Interpretation Approach	It includes if and how MSs interpret the results (Evaluation and Visualisation), and the extent to which the monitoring systems are user-friendly and accessible.	The MS is considered suitable if it uses appropriate evaluation methods and representative visualisation- it is user-friendly, and results are accessible to the public.
Audience	Refers to having different or multiple audiences, with some covering policy makers and SCLs as their primary audiences (identified as STAR4BBS target audience).	The well-suited MS would cover one or both intended users of the STAR4BBS monitoring system (policy makers and SCS owners).
Operationality	This refers to the operationality of the monitoring systems, whether they are ongoing or a one-off initiative, and whether they were developed recently (from 2015 onwards).	The MS is suitable if it was developed or revised recently (from 2015 onwards) and is still operating.

Step 4: Evaluation and in-depth analysis of the selected systems

An in-depth analysis of six monitoring systems, selected based on the selection criteria listed above, was later conducted, primarily assessing their evaluation methodologies, benchmarking purpose and evaluation structure.

As we first looked at the publicly available resources, our approach involved meticulous data compilation from websites, publications, and other sources. In pursuit of a comprehensive understanding of the methodological approaches employed by the selected tools, we proactively engaged with monitoring system owners. To further enrich our analysis, we conducted online consultation interviews in collaboration with ISEAL partners and representatives of the monitoring systems. This approach allowed us to gather supplementary information, address uncertainties, and gain insights into the latest updates pertaining to the tools.

The results of the selected monitoring tools analysis in this report are presented following a preliminary structure of the STAR4BBS monitoring system, developed for the purpose of conceptualization task (please refer to the STAR4BBS deliverable *D4.1. Concept of the monitoring system* and the Strategy of the Joint Monitoring System).

3 Results of the Systematic Literature Review (SLR)

Based on the SLR results, we observed a high variety among the identified 26 papers in terms of the study scope and methodological approach

To account for recognized variety, and after identifying similarities, we clustered the studies into three assessment groups (Table 2): i) benchmark; ii) comparison, and iii) exploring a relationship.

Table 2 Summary and description of the three assessment groups identified in the SLR

Overall assessment	
Benchmarking	Studies which perform benchmarking of the VSS or SCS and labels, against either performance indicators or reference points.
Comparing	Studies which perform comparative analyses, either between VSS or SCS themselves, or between certain aspects of those schemes, or comparing a certification tool with other rating tools
Exploring a relationship	Studies which do not fall into previous two groups, as they mainly examine the connection between certification and an outcome (e.g. impacts).

i) Benchmark

Certain studies benchmark voluntary sustainability standards (VSS) or SCS and labels, against either performance indicators or reference points. Methodologies used in benchmark studies tend to involve identifying or introducing specific criteria or indicators, collecting and analyzing data, and using statistical methods to assess performance.

For example, in one study three certification schemes were benchmarked against the (ecological, institutional, socio-economic) elements of the CFRN framework (Mussells and Stephenson, 2020).

In contrast, in the study done by Troster and Hiete (2019), the design characteristics of 17 SCS in the sector of mineral resources were assessed against the demands of important stakeholder groups.

ii) Comparison

Among identified articles, there are studies that are performing comparative analyses, either between certification standards/schemes themselves, or against certain aspects of those schemes, or even comparing a certification system with other rating tools.

For example, FSC was compared (i.e., used as a baseline) with other four certification programmes in the study done by Garzon et al. (2020). Specifically, principles, criteria and

indicators of the schemes were compared, exploring their consistency with FSC principles.

On the contrary, Kadam et al. (2021) used the relative temporal flow quantitative changes as an aspect of the SCS comparison, to track change and continuity in text and concept frames within 5 different global forest schemes and protocols and to compare framing of their standards using generic categories.

A different approach comes from Olawumi et al. (2020), where the proposed Building Sustainability Assessment Method (BSAM) scheme was compared with six existing green building rating systems, by allocating credit points to the key sustainability assessment criteria of all seven green rating schemes.

Finally, Ali et al. (2021) compared the newly developed green certification (rating) tool *Diana* with other green building rating tools, to account for differences in terms of structure, impact categories and indicators according to the building type and other contextual particularities.

iii) Exploring a relationship

There are studies which do not fall into previous two groups, as they explore different kind of relationships among certain elements, for e.g. they investigate the relationship between two or more variables or systems, or the impact of certification on environment, livelihoods, etc.

For example, Kittinger et al. (2021) explored key ecological and geographical mismatches in existing certification and ratings schemes, regarding entire production geography.

Another study, done by Sugiura and Oki (2018), performed an analysis of questionnaire survey data, to explore impacts of certification (and 2 schemes) and certification costs on the decision of enterprises for obtaining certain forest management certification.

3.1 Assessment results and findings

Of the 26 identified studies, 10 were grouped as *benchmark*, 10 as *comparison* and remaining 6 belonged to *exploring the relationship* group (Fig. 2).

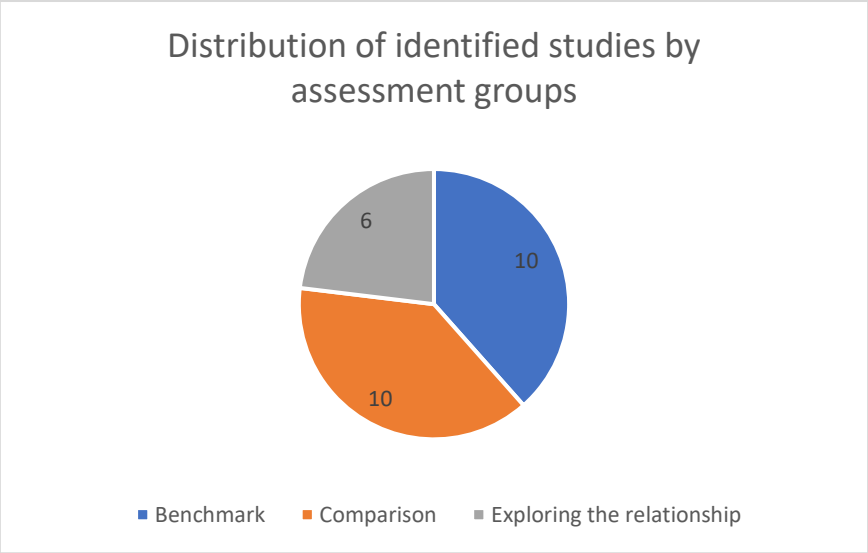


Figure 2 Distribution of identified studies by assessment group

Table 1 of the Annex B1 provides an overview of the overall assessment and assessment elements for various sectors, including forestry, construction, agriculture and floriculture, food industry, energy, mining, metals, textiles, tourism and hospitality, and chemicals (bio-based). Each sector is categorized based on the type of assessment it undergoes, which includes comparison, benchmarking, and exploring a relationship. The assessment elements covered in each sector are listed, including specific VSS, SCS, labels, programmes, regulations, and initiatives that are relevant to that sector.

Here, we would like to highlight the importance of tailored approaches in assessing sustainability performance, as different industries require unique methodologies to address their specific challenges.

By analyzing the methodologies used in each of the 3 assessment groups, we aim to identify the strengths and limitations of different approaches at 3 key stages and gain insights into the factors that may influence the results of each type of study: study design & data collection methods, data analysis and representation & interpretation. The findings are summarized in Figure 3.

Data collection	Data analysis	Representation and interpretation
• Structured and semi-structured questionnaires • Open interviews • Survey data • Semi-structured interviews • Evaluation criteria analysis • Desk-research • Factorial survey • Review of scientific articles • Benchmark comparison • Framework development • Manual coding process • LCA studies comparison • Expert questionnaire	• Descriptive comparison • MCDA • Comparative analysis • Detailed content analysis • ANCOVA and ANOVA • Two-dimensional analytical method • Standards comparison • Standardized rating scale analysis • Internal consistency assessment • Categorization analysis • Propensity score matching • REML • Random utility framework	• Tabular comparison • Scoring chart • Descriptive comparison • Tabular summary statistics • Graph representing cumulative distribution • Color-coding table • Breadth and depth scores graph • Rating system • Hexagon of criteria • Visualizing degrees of environmental performance • Relative ordinal ranking system

Figure 3 Summary of findings categorized within three key stages of the assessed studies

Detailed findings of the synthesis of the methodological approaches of each of these three assessment groups can be found in Table 2 of Annex B2.

Study design & data collection methods

In the course of the STAR4BBS monitoring system conceptualization task, as mentioned before, it is important to properly define the purpose and goal of such a monitoring system. In this regard, an important aspect is to analyze how the data are collected and processed (see last part of the *Methodology* section). The results of this analysis are relevant for collecting inputs to be used for the database layer of the new monitoring system (see STAR4BBS deliverable *D4.1. Concept of the monitoring system*). Thus, by exploring how this element is treated in the context of the identified studies, it can provide us with insights into the methods used to collect data, such as surveys, case studies, or interviews. Furthermore, this information helps to assess the validity and reliability of the data and determine whether the data are suitable for the purposes of the monitoring system.

The studies in the *benchmark* assessment group primarily analyzed and evaluated existing sustainability certification schemes and labels against specific criteria, such as ecological, governance, and socio-economic aspects. They employed a variety of methods, including desk research, online questionnaires, and comparative analysis of LCA studies. The *comparison* assessment group mainly focused on comparing different sustainability certification schemes in terms of consistency and changes in

categorization (e.g. principles, criteria, indicators). They utilized a wide range of data collection methods, such as expert surveys, structured and semi-structured questionnaires, interviews, and literature reviews. Finally, the *exploring a relationship* assessment group examined the relationships and mismatches between different certification and rating schemes, as well as the effects of certification on enterprises, using questionnaire surveys, interviews, and household surveys. Overall, the studies varied greatly in their methodological approaches, but indicating the need of careful consideration of the specific objectives, stakeholders, and contextual factors involved in the development of new monitoring system's assessment methods.

Data analysis

Another element to consider in different methodological approaches is the methods used to analyze the data, such as for example comparative analysis. This information helps determine the level of rigor and sophistication of the analysis and whether a specific method is appropriate for our purposes.

The results from the studies synthesis on this element revealed that the methods used are dependent on the research question, the type of data being collected, and the overall goals of the assessment.

The different assessment types (benchmark, comparison, exploring the relationship) all utilized a variety of data analysis techniques to evaluate sustainability standards and/or schemes and labels. These techniques included both qualitative and quantitative methods, such as categorization analysis, content analysis, statistical tests, and data visualization tools. While each study had its own unique set of research questions and objectives, they all shared a common goal of evaluating sustainability standards in a systematic way.

For *benchmark* assessments, the studies used analytic methods such as categorization analysis, multi-criteria decision analysis, and detailed content analysis. They also used various statistical methods such as ANCOVA, ANOVA, MANOVA, and restricted maximum likelihood approach (REML). For *comparison* assessments, studies utilized qualitative analysis, descriptive comparisons, and comparative cross-case analysis, to name a few. They also used tools such as Delphi technique, Analytic Hierarchy Process, and credit points (score-weighting). Finally, for *exploring the relationship* assessments, studies employed descriptive comparison, propensity score matching, and case study approaches.

To develop a new monitoring system, it is important to consider a variety of data analysis techniques, including both qualitative and quantitative methods, and to previously carefully design the scope. Additionally, it is important to consider the context in which the monitoring system will be implemented, including the goals and objectives, the available data, and the resources required to carry out the study. By taking a comprehensive and thoughtful approach to data analysis, it is possible to develop an effective and reliable monitoring system that can help to promote sustainable goals.

Representation & interpretation

Data interpretation indicates how the data are presented and interpreted, which makes it possible to assess the clarity and transparency of the results and whether they are appropriate for our intended audience. In addition, it can include visualisation of results, which shows how information can be accessed and is communicated to the users.

From the *benchmarking* assessment type, it is clear that different methods were used to evaluate the representation and interpretation element. Benchmark assessments tend to use direct rating methods, color-coding tables, scoring charts, and relative ordinal ranking systems to compare the degree to which different schemes consider specific criteria or principles (Fig. 4).

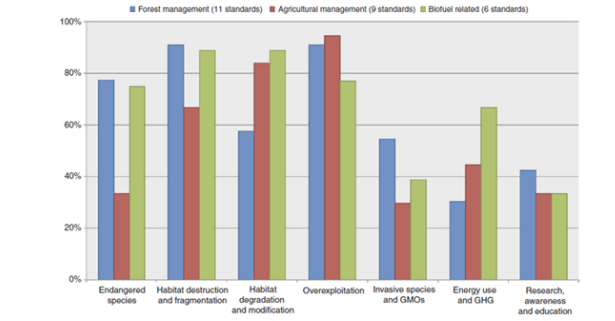
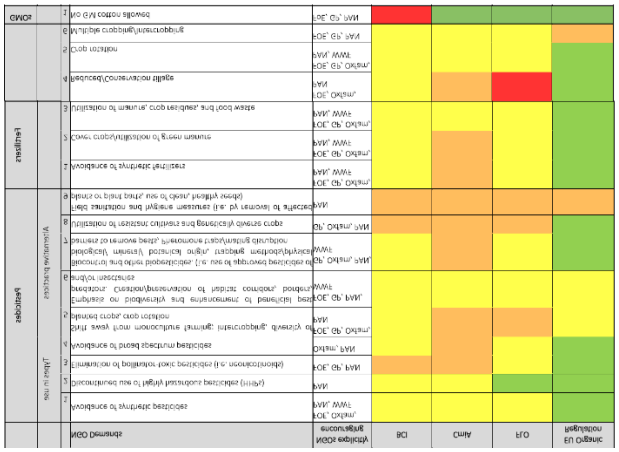


FIGURE 2 | Percentage of benchmark criteria complied with on a principle level for the three standard types, respectively. Well considered principles are those with >80% criteria complied with in total; Poorly considered principles are those with <50% criteria complied with in total.

Englund & Berndes, 2015



Partzsch et al. 2019

CFRN Framework Element	Forest Stewardship Council	Aquaculture Stewardship Council	Marine Stewardship Council
Obligations to laws and indigenous peoples			
Good governance structure			
Effective decision-making processes			

Mussells & Stephenson, 2020

Figure 4 Some examples of the benchmark assessment interpretation elements

These methods allow for a clear comparison between different schemes and the identification of their strengths and weaknesses.

In contrast, *comparison* assessment studies covered this element in less breadth, and used graphical representations (such as Sankey diagrams, radar diagrams, or pie charts), ranking of criteria, and classification of sub-attributes to compare the overall sustainability performance of different schemes (Fig. 5).

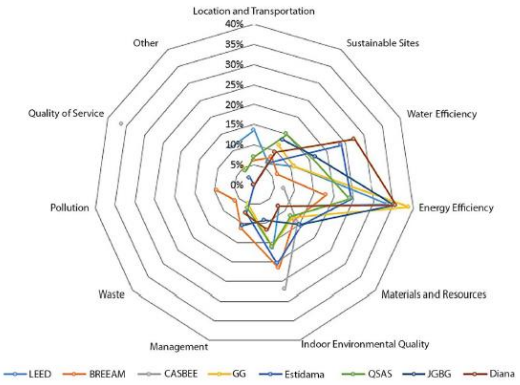


Fig. 14. Comparing category assigned weights for Diana and other green building rating tools.

Figure 5 An example of a diagram from the study of Ali et al. 2021

These methods focus more on identifying the contribution of different indicators to the overall sustainability performance and provide a clear picture how schemes perform relative to each other. Finally, *exploring the relationship* assessments were more descriptive and focus on summarizing and comparing the results of different studies. They used tabular comparison, descriptive comparison, and summary statistics to present findings on the relationship between different variables and what effects sustainability certification schemes can have (Fig. 6).

Table 3. *Average treatment effects for household expenditure and poverty*

	Nearest neighbor matching		Kernel matching		Γ
	ATT	S.E.	ATT	S.E.	
All certified vs. non-certified					
Per capita expenditure (UGX)	478.99**	191.88	369.44**	180.24	1.9
Poverty headcount index	−0.08	0.05	−0.05	0.05	1.5
Poverty gap index	0.01	0.04	0.01	0.04	1.3
Fairtrade vs. non-certified					
Per capita expenditure (UGX)	1028.58***	239.84	871.27***	229.69	1.6
Poverty headcount index	−0.15**	0.06	−0.13**	0.05	2.0
Poverty gap index	−0.09*	0.04	−0.11**	0.04	1.5
UTZ vs. non-certified					
Per capita expenditure (UGX)	−51.70	269.70	36.72	254.52	1.2
Poverty headcount index	−0.02	0.08	−0.03	0.07	1.3
Poverty gap index	0.05	0.07	0.05	0.07	1.1
Organic vs. non-certified					
Per capita expenditure (UGX)	242.42	286.99	0.55	252.84	1.3
Poverty headcount index	−0.04	0.08	0.02	0.07	1.4
Poverty gap index	0.06	0.05	0.07	0.06	1.1
Fairtrade vs. UTZ					
Per capita expenditure (UGX)	984.83***	318.74	850.20***	286.93	1.8
Poverty headcount index	−0.07	0.07	−0.07	0.07	1.4
Poverty gap index	−0.21**	0.06	−0.22***	0.06	2.3
Fairtrade vs. Organic					
Per capita expenditure (UGX)	619.75*	334.15	484.8	331.01	1.4
Poverty headcount index	−0.08	0.08	−0.07	0.08	1.1
Poverty gap index	−0.19**	0.08	−0.24**	0.1	2.3
UTZ vs. Organic					
Per capita expenditure (UGX)	97.53	405.28	−106.55	343.34	1.2
Poverty headcount index	0.15	0.11	0.13	0.09	1.1
Poverty gap index	−0.17	0.18	0.03	0.13	1.1

Notes: ATT: average treatment effect on the treated; S.E.: bootstrapped standard errors; Γ: Rosenbaum bounds (critical levels for hidden bias).
*$p < 0.1$.
**$p < 0.05$.
***$p < 0.01$.

Figure 6 An example of exploring the relationship assessment element, from Chiputwa et al. 2015.

During the initial conceptualization of the new monitoring system, it has been discussed that the assessment of the schemes can have a goal of identifying schemes' strengths and weaknesses and improving their robustness and effectiveness. In this regard, methods used by the benchmarking assessment type (rating methods, color-coding tables, scoring charts, relative ordinal ranking systems, etc.) can in fact be taken into account for results interpretation and visualization of the new monitoring system.

3.2 Key lessons learned

Overall, our findings suggest that there are various methodological approaches employed in reviewed studies regarding their design, data collection, data analysis, representation and interpretation of results. It can be derived that the development of a new monitoring system requires careful consideration of the specific scope, audience, and purpose to consider a choice of a specific methodological approach.

As explained above, the studies reviewed in the *benchmark*, *comparison*, and *exploring the relationship* assessment groups used a variety of data collection and analysis techniques, both qualitative and quantitative, as well as interpretation techniques, to evaluate sustainability standards and/or schemes and labels. Effective and reliable monitoring systems should be developed by taking a comprehensive and thoughtful approach to the mentioned elements, especially concerning data analysis.

In this part of the study, we did not highlight or recommend any specific methodological approach and put it above others. As ISEAL clearly states it in its practical guide for sustainability systems (ISEAL, 2020), no one method is necessarily better than the other. It can be tentatively concluded from the analysis of the results in accordance with the study design, data analysis, and representation methods, that the benchmarking group exhibited the most comprehensive and thorough assessment of the elements which can be used for the core purpose of the new monitoring system.

The complexity is also reflected in variability, but since the ultimate goal of a method is to inform some decision, the specific needs and purposes will command the choice of method.

It is clear that by selecting appropriate data collection methods and analyzing the data using rigorous and systematic techniques, the monitoring system can be developed so that it provides accurate and trustworthy information on sustainability standards and certification schemes and labels.

Finally, an insight into different methodological approaches of identified studies can provide a systematic way to assess their strengths and limitations and help us decide which methodologies to use in our monitoring system and how to design and implement them effectively.

4 Results of the Grey Literature Review (GLR)

The extensive search of existing databases, grey literature review and stakeholders' input provided us with 23 results, of which we analysed 19, as explained in the Methodology section above. The results were compiled in a matrix table, which is available in Annex A2. As an output of an internal consultation with Consortium members, 11 different parameters (see *Box 2*) were selected, as a basis for a further review and qualitative synthesis of the identified and selected monitoring systems.

Box 2

The 11 parameters for categorization of findings are as follows:

- **Name of the monitoring system:** specific name or title of the monitoring system being analyzed or evaluated.
- **Development entity:** refers to the organization or group responsible for developing the monitoring system.
- **Scope:** specific area or domain that the monitoring system covers and the intended purpose or goals of the monitoring activities.
- **Operationality (Ongoing/One-off project):** refers to whether the monitoring system is an ongoing initiative or a one-time project.
- **Structure:** refers to the internal organization of the monitoring system, including its implementation and content.
    - **Implementation:** practical aspects of how the monitoring system is put into action, including its procedures, methods, and tools.
    - **Content:** specific areas or topics covered by the monitoring system, including analysed standards and metrics.
- **SCS and labels benchmarked:** specific sustainability certification schemes (SCS) and labels being compared or evaluated by the monitoring system.
- **Database Indicators** (identifying and reviewing sustainability indicators in the existing monitoring systems)
    - **System level** (e.g., certification schemes procedures, rules, accreditation, transparency, audit system, standard settings).
    - **Content level** (e.g. social, environmental, economic and circular indicators)
    - **Outcome level** (impact on the environment, health and safety of the workers, etc.)
- **Description:** a brief summary of the monitoring system, including its purpose, aim and usage.
- **Link:** a web link or other resource where the monitoring system can be accessed or further information can be obtained.
- **Interpretation of results:** refers to the analysis and interpretation of the monitoring system's findings, including any ratings or recommendations drawn.
- **User interface and Audience (Users):**
    - **Tool (type):** specific type of tool or platform used for monitoring or benchmarking, such as a website tool or a report. A comprehensive monitoring system should be tailored to the needs of its audience, providing relevant and actionable information that can be used to inform decision-making.
    - **Audience:** intended users or target audience for the monitoring system, such as consumers, producers, or policymakers.

In this sub-section, we will present an overview of the main findings (of all 19 reviewed monitoring tools) and summarize results that can inform the development of a new monitoring system to assess the effectiveness and robustness of certification schemes and labels. An overview of the commonalities and differences of the analysed monitoring tools is provided in Table 3.

Table 3 A summary of the commonalities and differences among reviewed systems according to the most important aspects

Aspect	Commonalities	Differences
Purpose	Assessing different schemes and labels for transparency, credibility, and sustainability in various industries.	Tools approach assessment from different sectors and target specific product groups, tailoring their criteria and methodologies accordingly.
Key Features	Use of comprehensive databases to collect and analyze data; rely on structured criteria related to sustainable and responsible practices; mostly ongoing tools ensuring continuous monitoring.	The depth and focus of benchmarks vary. Various assessments are used to measure schemes' performances and provide a basis for comparison. There are distinct goals and target audiences.
Database Assessment	Employment of mostly environmental, social (content level) and traceability indicators (system level).	Tools emphasize sustainability performance, operational systems, or both, with rare coverage of impacts (outcomes).
Data Interpretation	Facilitating criteria comparison between two or more schemes/standards; use of matrices or matrix-like frameworks; predominance of qualitative assessments.	Interpretation methods range from pass/fail results to detailed scoring systems and star ratings; varied visualization methods like pie charts, percentage bars, and benchmark reports.

The synthesis here comprises exclusively the findings pertaining to specific parameters associated with the expected structure of the new monitoring system (described in STAR4BBS deliverable D4.1. Concept of the monitoring system).

Scope

Certain monitoring tools focus on simple assessments of sustainability certification schemes, while others provide more complex benchmarking platforms for identifying common denominators between different certification schemes. There are tools that only evaluate labels, but also systems for providing recommendations on responsible supply chains.

Some of the tools can help companies identify and choose certification schemes that align with their sustainability goals and values, as is the case of the Textile Exchange CFMB tool. Others highlight the schemes that have fulfilled a certain set of criteria used

as a threshold (e.g. Soy Benchmarking Tool, SIFAV basket of standards), or show how a scheme performs against a range of criteria, in order to drive improvement or inform user decision-making (e.g. WWF CAT, FAO's Comparative Matrix and the SME Standards Compass).

Moreover, there are benchmarking processes and indices that assess the sustainability performance of certification schemes on different levels, as well as the companies that implement them. FEFAC Soy Benchmarking Tool evaluates the performance of soy supply chains, while ITC Standards Map does not score or rate schemes but documents the presence (or absence) of related criteria within the compared standards. All these various approaches can help to increase transparency and accountability, and drive improvements in sustainability practices.

Reviewed monitoring systems also vary according to the specific sector they cover. These include metals and minerals (OECD Due Diligence Guidance for Responsible Supply Chains of Minerals), fishery products (GSSI, Seafood Stewardship Index), textiles (ITC Standards Map, The Textile Exchange CFMB), food and feed (Labelinfo.ch, SOJA), fuels (WWF CAT, ADVANCEFUEL) and more. The scope of the STAR4BBS monitoring system will target biological feedstock and bio-based materials and products, and few of the identified systems cover a similar scope (CEPI, STAR ProBio SCT, ITC Standards Map, Siegelklarheit, Blue Angel Comparison of Certification Systems, SME Compass, Textile Exchange CFMB).

In regard to their function, certain identified monitoring systems serve to aid economic operators to demonstrate their sustainability compliance (ADVANCEFUEL), while others provide reliable advice to customers and companies involved in the products trade (CEPI) or provide value chain actors with clear guidance on which certification schemes cover key sustainability requirements and apply relevant governance and verification practices as third-party auditing (SSCI).

The differences that were identified in the scope of the systems during the review stage vary strongly in terms of performance, depth, information access and geographical range. For instance, CEPI lacks on providing content of forestry performance standards, and is dependent upon information provided by SCS owners. On the other hand, the proposed STAR ProBio SCT's Benchmarking Platform (BP) requires the voluntary participation of existing certification schemes and becomes fully dynamical instrument once the SCS engage in the process of its establishment. IDH SIFAV basket of standards system lacks full access to information that is publicly available to interested parties. Finally, Labelinfo.ch limited its geographical scope nationally, focusing on labels present on the Swiss food market.

Most of the tools were developed to promote and ensure the credibility, consistency, and effectiveness of sustainability certification schemes or standards. Their goals include building trust between stakeholders, including consumers, companies, and governments, and facilitating the transition towards more sustainable and responsible production and consumption practices.

Operationality

Overall, the monitoring systems listed have varying levels of operationality, with some being one-off projects (or two-times initiatives) and others being ongoing initiatives. Some systems were developed in the past and are now updated periodically, while others were established more recently. Most of them are ongoing systems, as it can be seen from the Figure 7.

One-off projects include the comparative matrix CEPI developed in 2000, SOJA as a two-times benchmarking exercise developed in 2017 and then repeated (using CAT WWF criteria) in 2020, and a WWF CAT itself which was developed as a project report in 2012-2013. The STAR ProBio SCT was developed in 2020 as a project deliverable, similar to ADVANCEFUEL developed in 2017. Blue Angel Comparison of Certification Systems was another one-off project developed in 2019, as well as Labelinfo.ch published as a rating guide in 2015.

Ongoing initiatives include the ITC Standards Map as a dynamic tool for benchmarking of standards and indicators launched in 2011, the SSCI benchmarking initiative established in 2022, the GSSI drafted in 2014 but being updated as a benchmarking process, and the SIFAV Basket of standards updated recently, etc. Finally, there are some systems that were developed in the past, but are still ongoing initiatives, regularly updated over time. Such is the case with the Siegelklarheit which had its last revision in 2021.

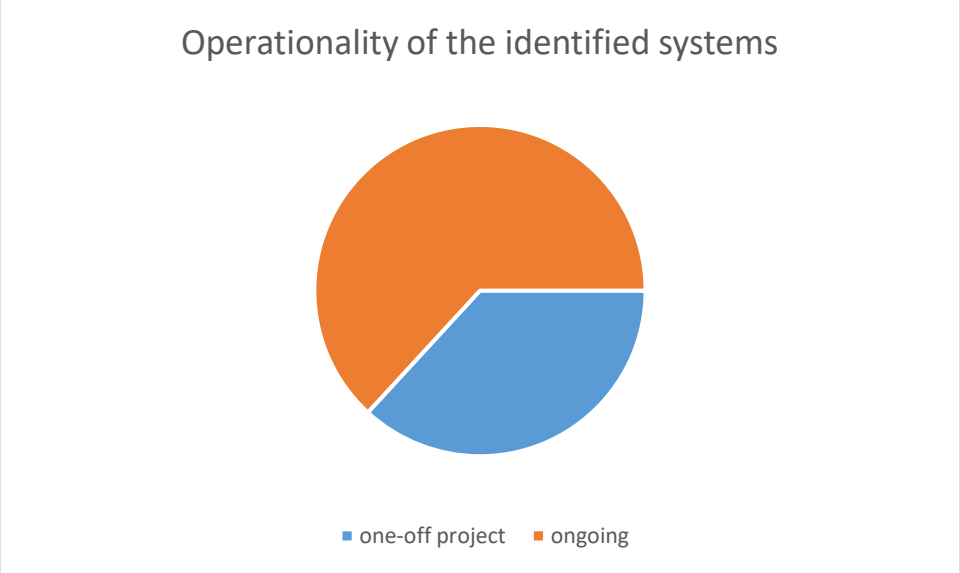


Figure 7 Proportion of the operationality sub-categories within identified monitoring tools

Database Indicators

The identified monitoring systems use a variety of **database indicators** to assess the sustainability performance of certified products or feedstock. These indicators were categorized in our matrix into three levels: system, content and outcome level indicators (following the intended STAR4BBS monitoring system structure, as mentioned before).

At the system level, indicators focus on aspects such as certification governance and scheme management, performance, legal compliance, transparency, and traceability.

These indicators assess the overall framework and procedures governing the certification schemes, including the rules and procedures for implementation and assessment. Additionally, some systems evaluate the credibility and reliability of certification systems, ensuring adherence to systemic requirements and monitoring the chain of custody. Supply chain management and due diligence obligations are also considered.

Content-level indicators primarily assess social and environmental performance. The environmental indicators cover topics such as forest protection, soil quality, water resources, biodiversity conservation, and climate change. Some systems specifically evaluate the performance of farming practices, animal welfare, and seafood quality. Social indicators encompass human rights, labor rights, community well-being, supply chain responsibilities, and more. Economic indicators are included in a few systems and may include aspects like investment, techno-economic performance, risk assessment and vulnerability. Circular economy indicators focus on measuring the extent to which certified products or feedstock contribute to circularity, product renewability and resource efficiency.

At the outcome level, the identified monitoring systems utilize a smaller set of indicators to assess the sustainability outcomes or impacts achieved by the certified products or feedstock. Stakeholder engagement and interviews play a significant role in assessing the outcomes of certification schemes. The perspectives and feedback of various stakeholders (such as for e.g. actors in the supply chain, like buyers, retailers etc.) are considered to understand the effectiveness and impact of the certified products or feedstock production on the environment, society, and other relevant areas. Data collection requirements are of utmost importance in this regard, for measuring and tracking sustainability outcomes accurately.

While there are some commonalities in the types of indicators used across the monitoring systems, there are also differences based on the specific focus and objectives of each system. Environmental and social indicators appear to be widely employed, reflecting the importance of assessing the ecological and societal impacts of certified products or feedstock. Traceability is another common indicator at the system level, highlighting the significance of tracking and verifying the origin and flow of materials throughout the supply chain.

Interpretation of Results

The analyzed monitoring systems exhibited both commonalities and differences in the ways they interpreted the findings of their assessments, showing various practices and approaches.

Several systems, such as the SAFA Tool and Siegelklarheit employ rating systems or scales to provide a straightforward assessment of the sustainability along value chains. By assigning ratings or stars based on predefined criteria, these systems simplify the interpretation process and allow stakeholders to quickly grasp the sustainability performance of certification schemes, standards or supply chains.

Some identified monitoring systems use a point system or a joint consideration of individual dimension ratings to provide an overall rating that ranges from "Not assessed"

(requirements not fulfilled) to "Good choice!" or "Very good choice!" (GIZ Sustainability Standards Comparison Tool (SSCT) within Siegelklarheit) or to rate labels as Excellent, Highly recommended, Recommended, and Conditionally recommendable (e.g. Labelinfo.ch).

The Seafood Stewardship Index and the Textile Exchange CFMB provide scores or percentages to evaluate companies' performance against key indicators. This allows companies to compare their results to industry averages or sector peers and identify areas for improvement.

The ITC Standards Map, SME Compass and the FEFAC Responsible Soy Benchmarking Tool provide comprehensive and in-depth information about existing standards and schemes. They enable users to independently gain insight into the sustainability performance of various certification schemes and standards. In addition, they facilitate criteria comparison between two or more schemes/standards and they assess compliance and governance of the schemes. The ITC Standards Map generates graphs and statistics on criteria coverage and proportion for better visualisation of results.

Despite these commonalities, there are also differences in how the interpretation of results is conducted. Some systems focus on assessing compliance with specific criteria or requirements, while others take a broader perspective by evaluating social and environmental practices. Additionally, certain systems emphasize internal use for self-evaluation and communication (e.g. SAFA), while others provide publicly available information and encourage dialogue and collaboration within the industry.

Overall, the identified monitoring systems strive to present their findings in a clear, transparent, and user-friendly manner. The emphasis on transparency and continuous improvement is evident across the systems, fostering a greater understanding of sustainability performance and encouraging collaboration among stakeholders. However, specific approaches may vary, reflecting the diverse goals, scopes, and target audiences of the different systems.

Structure[6]

The *Implementation* sub-category of the **Structure** category demonstrated that the identified monitoring systems use different methodologies to assess the robustness and effectiveness of standards and certification schemes. The benchmarking tools evaluate the performance of supply chains, companies, or products against a set of sustainability criteria clustered under different pillars or dimensions. These systems utilize matrix-based assessments, qualitative evaluations, benchmarking processes, and systematic frameworks to evaluate and compare different schemes.

The *Coverage* sub-category of the Structure category in the identified monitoring systems employs various methodologies to assess the credibility, environmental impact, and social responsibility of certification schemes and sustainability standards.

[6] This category does not follow the planned monitoring system structure but is included in the matrix and subsequently here because it highlights the diverse methodologies utilized by the identified monitoring systems to assess the robustness and effectiveness of SCS and labels. The use of matrices, qualitative assessments, and benchmarking processes by identified tools allows for a valuable insight into comprehensive evaluation and comparison of different schemes.

Some of the identified systems, such as the SSCI, evaluate social and environmental sustainability in supply chains using a step-by-step approach that assesses labor rights, environmental protection, and business integrity. Other systems, such as the Blue Angel Comparison of Certification Systems, use a set of test criteria to evaluate the degrees of fulfilment of certification systems' sustainability criteria.

For a systematic evaluation against a set of predefined criteria, the GSSI Global Benchmark Tool is divided into three parts, including a Framework, Process, and Result, which provide information on the GSSI Essential and Supplementary Components of a scheme, steps the scheme goes through to be recognized by GSSI, and the statement of GSSI recognition and benchmark report.

The ITC Standards Map allows the comparison of several standards and has a comprehensive coverage of system characteristics and sustainability topics through 1650 indicators. On the other hand, Soy Benchmarking Tool uses the FEFAC Soy Sourcing Guidelines providing 73 criteria clustered under six pillars. WWF CAT uses a detailed set of questions and criteria, applying a point system to assess the strategic, structural, social and ecological strengths and weaknesses of standards and certification schemes against WWF's requirements for a sustainable environmental and social standard.

One common approach found across several monitoring systems is the use of matrices or matrices-like frameworks. These matrices align certification schemes with specific criteria and indicators to assess their credibility and adherence to sustainability standards. For instance, the CEPI Forestry Committee's agreed-upon criteria serve as a basis for evaluating different schemes. The matrices facilitate a structured evaluation process and enable comparisons between various certification schemes.

Another recurring element is the inclusion of qualitative assessments. These assessments focus on specific sustainability requirements, such as greenhouse gas emissions, sustainable forest management, biodiversity, and water sustainability. The qualitative evaluations consider the unique challenges associated with different product groups and establish criteria that align with these specific challenges. Additionally, the assessments often incorporate internationally recognized standards to benchmark and ensure compliance with labor rights, environmental protection, and economic resilience.

The benchmarking processes employed by several monitoring systems involve comprehensive evaluations across multiple dimensions. This includes assessing sustainability issues (content performance, i.e. *Content level*), but also issues related to operational system (or *System level*, see *Methodology* section) such as scheme governance, operational management, etc. The benchmarking exercises use various assessments to measure scheme performance and provide a basis for comparison.

For instance, the identified monitoring tools may emphasize the importance of either sustainability content performance or operational systems, or both. Evaluations of sustainability content encompass various aspects, such as ecosystems, land use, circularity, responsible working conditions, business integrity. Evaluations of operational systems included legal requirements, chain of custody, procedures, rules, assurance, claims, to name a few. The assessments consider the transparency and accessibility of standard setting organizations to ensure credibility. Additionally, some systems employ a basket approach, tailoring social and environmental standards to different risk levels,

practices and geographical areas (e.g. FSI Basket of Standards, IDH SIFAV basket of standards).

In conclusion, identified monitoring tools utilize diverse methodologies to evaluate and compare certification schemes and sustainability standards. All these different types of assessments and benchmarking processes consider specific categories (pillars, principles, criteria and indicators) as common elements across these systems. These approaches facilitate assessments of environmental impact, socio-economic responsibility, and circularity, enabling stakeholders to make informed decisions.

SCS and Labels benchmarked

It is worth noting that the selection of benchmarked SCS and labels varied widely across the monitoring systems, reflecting their specific scope and objectives. While some systems benchmark a wide range of standards and labels across sectors, others have a narrower scope or focus on specific industries. The evaluated schemes and labels address sustainability concerns related mostly to environmental protection, social responsibility, economic development, circularity aspects.

Among the commonly benchmarked schemes and labels are those related to forestry certification. The Forest Stewardship Council (FSC) and Programme for the Endorsement of Forest Certification (PEFC) are frequently assessed for their commitment to sustainable forestry practices and responsible sourcing of wood and forest products.

In the bio-based products sector, several certification schemes have been compared. The Roundtable on Sustainable Biomaterials (RSB), International Sustainability and Carbon Certification (ISCC), and REDcert are among the commonly evaluated schemes. These certification systems ensure that bio-based products, but also biofuels, are produced using sustainable practices and meet specific environmental and other criteria.

The seafood industry also sees significant attention from monitoring systems, with seafood certification schemes being frequently benchmarked. Schemes such as the Marine Stewardship Council (MSC), Best Aquaculture Practices (BAP), and Aquaculture Stewardship Council (ASC) are evaluated to assess their contributions to sustainable fishing and aquaculture practices.

While sustainability certification schemes and labels are the primary focus for most monitoring systems, there are cases where no specific SCS or labels are compared. Instead, certain systems track and evaluate sustainability performance across supply chains without specifically benchmarking SCS or labels. These evaluations often assess the progress of industries towards more sustainable practices, such as responsible mineral supply chains (e.g. OECD MNE Guidelines benchmarking) or the sourcing of sustainable materials in the apparel, footwear, and home textile sectors (e.g. Textile Exchange Corporate Fiber & Materials Benchmark (CFMB) and MCI).

User Interface and Audience (Users)

Based on the type of tool used for benchmarking, our findings can be categorized as:

1. Information tools (publications, methodology documents, feasibility studies, reports, desk-based exercises)
2. Websites and web-based tools
3. Spreadsheet-based tools (multi-stakeholder platforms, etc.)

Most common were the web-based tools, such as ITC Standards Map, SAFA Tool (which include various online platforms, databases, and tools accessed through websites) and information tools, such as CEPI, Star ProBio SCT, WWF CAT (which encompass reports, publications, and other documents that present information, guidelines, and other reference materials).

The intended users or target audience for the monitoring systems varied according to the stakeholders' categories to which they belong:

1. Companies: A certain number of monitoring systems aimed to provide sustainability assessments to business entities or even benchmark the companies themselves. These systems targeted companies involved in various sectors, including consumer goods, textiles, seafood, biofuels, and agriculture. The assessments were designed to help companies evaluate their sustainability performance, identify areas for improvement, and set targets for sustainability.
2. Consumers: Some monitoring systems focused on providing information and transparency to consumers. These systems aimed to empower consumers to make informed choices by assessing and comparing sustainability certification schemes and labels associated with different products.
3. Policy makers and regulatory bodies: Several monitoring systems targeted policy makers, regulatory bodies, and governments. These systems provided insights and assessments to inform policy development and regulatory decisions related to sustainability certification schemes and labels.
4. Industry and supply chain actors: Certain monitoring systems aimed to serve economic operators, SMEs and industry associations. These systems provided assessments and benchmarks that could be used to guide investment decisions, evaluate sustainability performance within industries, or support industry-wide initiatives for sustainability.
5. Civil societies and NGOs: Some monitoring systems had a focus on civil society and non-governmental organizations. They aimed to provide information and assessments that could be used by NGOs, civil society groups, and other similar stakeholders to advocate for improved sustainability practices and drive positive change.
6. Academia, researchers and consulting experts: Certain monitoring systems targeted academia and consulting firms and similar institutions. These systems provided resources and tools that could be used for research, analysis, and consulting purposes in the field of sustainability certification and labeling.
7. Other stakeholders: The target audience for monitoring systems also included in smaller depth certification scheme and label owners, auditors, standards organizations, and various stakeholders interested in improving the sustainability of certification systems, specific industries or sectors.

It is worth noting that some monitoring systems had multiple target audiences, as their assessments and information were relevant to various stakeholders involved at different supply chain stages, or for different sustainability practices and certification purposes.

4.1 In-depth analysis of selected monitoring systems

4.1.1 Selection results

The results of the selection process for the best-fitting systems are presented in the Table 1 of Annex C1. As it can be retrieved from the table, the best supported tools were ITC Standards Map and Siegelklarheit (GIZ Sustainability Standards Comparison Tool (SSCT)) receiving the highest scores of 8 points each. Close behind, with scores of 7 each, were the GSSI Global Benchmark Tool, FEFAC Responsible Soy Benchmarking Tool, SME Compass, and FSI Basket of Standards. For this reason, we decided to analyse these 6 monitoring systems further.

The chosen systems demonstrated strong alignment with the selection criteria, making them the most promising sources of information for the subsequent evaluation and providing valuable insights into their strengths, limitations, and overall effectiveness in assessing sustainability performance.

4.1.2 Overview of the selected monitoring systems and their features

In Annex A5, we present factsheets of the 6 selected monitoring systems, with their key features. Each system offers unique characteristics and approaches to assessing sustainability performance across different industries. By examining the background, scope, purposes, and applicability of these systems, we aim to provide an understanding of their intended use and goals.

4.1.3 Evaluation of each monitoring system's adherence to ISEAL's principles for credible benchmarking

ISEAL Alliance, as a global membership organization that works with sustainability standards and certification systems, created the ISEAL Credibility Principles[7], which outline fundamental criteria and requirements that sustainability standards and certification systems must adhere to in order to be considered credible and trustworthy. The ISEAL Credibility Principles cover various aspects of sustainability standards, including governance, assurance, impacts, and continual improvement. These principles are designed to ensure that standards and certification systems are transparent, impartial, reliable, and truthful in achieving their intended sustainability goals (Fig. 8).

[7] https://www.isealalliance.org/defining-credible-practice/iseal-credibility-principles

For this reason, we checked whether the selected monitoring tools are addressing the ISEAL Credibility Principles.

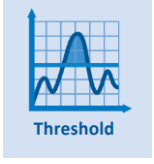
Figure 8 ISEAL Credibility Principles (source: isealalliance.org)

Furthermore, we wanted to see to what extent the selected systems can be structured according to the ISEAL's Sustainability Benchmarking Good Practice Guide (ISEAL, 2020).

ISEAL identifies four basic benchmarking models that are commonly applied in benchmarking initiatives - *Threshold*, *Improvement*, *Ranking* and *Peer comparison*. The benchmark itself must reflect the minimum set of performance and operational requirements that are acceptable to the convenor. The benchmarking process must be rigorous enough to ascertain whether the entities being benchmarked meet this performance threshold.

Out of the 6 selected benchmarking tools, 4 fall under the Threshold model, while the other 2 can be categorized as the Improvement model (Table 4).

Table 4 Four Benchmarking models proposed by ISEAL (ISEAL, 2020) and selected benchmarking tools that correspond to the models

Model	Purpose	Key features	Benchmarking tool
	To qualify entities that meet or exceed a common baseline or threshold. Can have different incentives: often used for recognition.	Performance bar set at level of acceptable practice.	**FEFAC Responsible Soy Benchmarking Tool** **Siegelklarheit** **ITC Standards Map** **GSSI (partly)** **FSI Basket of Standards**

	To encourage improved practices by showing progress toward good practice.	Aspirational performance bar set beyond current practice to provide direction and incentive.	**FEFAC (partly)** **SME Compass** **GSSI (partly)**
	To compare performance of similar entities through a ranked evaluation.	Entities are scored against performance topics and compared.	**Siegelklarheit (partly)**
	To conduct an internal comparison of an entity's own performance against its peers	The reference benchmark is the practices of the benchmarking entity itself. Examples are not readily shared.	**ITC Standards Map (partly)**

Our findings indicate that most of the selected monitoring tools have used ISEAL's guidance to inform the process of developing their systems and majority of them incorporated ISEAL's credibility principles in defining criteria.

Siegelklarheit, in particular, collaborated with ISEAL to develop their assessment methodology, which incorporates credibility criteria based on ISEAL Credibility Principles, including transparency, truthfulness, accessibility, and rigour. Similarly, ITC Standards Map upholds key features and objectives such as impartiality, quality, transparency, and neutrality in its tool.

FEFAC Responsible Soy Benchmarking Tool used ISEAL definitions for benchmarking in its FEFAC Soy Sourcing Guidelines 2021.

SME Compass worked with ISEAL, Shift, and other organizations to develop their methodology and scoring system. ISEAL, with ITC and WBA established the system of scoring and weighting as part of two consultation workshops held with the tool owner - Helpdesk on Business & Human Rights.

GSSI utilized ISEAL benchmarking guidances for their Benchmark Processes and QMS system, making sure all recommendations made by ISEAL are implemented. ISEAL codes or principles were used to select criteria for the benchmark, but in a limited fashion. FAO CCRF and Guidelines and ISO were the main references for criteria selection. The criteria used in the benchmark referred to clauses in the ISEAL codes or ISEAL principles under version 1, but they were removed for version 2 of the tool. A challenge for GSSI was that only some GSSI-recognized Scheme Owners were ISEAL members (only ISEAL members can be formally recognized as being Code compliant), so equal treatment of all scheme owners undergoing the benchmark process needed to be ensured.

Majority of the monitoring systems have shown efforts to align with ISEAL principles and guidance in their methodologies to declare credibility in their benchmarking.

4.1.4 Comparative analysis of the selected monitoring systems

The selected monitoring systems, namely the ITC Standards Map, Siegelklarheit, GSSI Global Benchmark Tool, FEFAC Responsible Soy Benchmarking Tool, SME Compass, and FSI Basket of Standards, each serve distinct purposes and offer unique features in comparing and evaluating sustainability standards and certification schemes and labels (Table 5).

Table 5 An overview of the main features of selected monitoring tools

Monitoring tool	Purpose	What does the benchmark do?	Database assessment	What is the result? (visualization)
FEFAC Responsible Soy Benchmarking Tool	To identify soy certification schemes that meet European feed industry requirements for responsible soy production. Comparing performance with the industry's requirements → making improvements	Assesses certification schemes against the 73 criteria in FEFAC's Soy Sourcing Guidelines. Self-assessment module or official benchmark against the FEFAC Soy Sourcing Guidelines 2021.	Content Level: Assesses legal compliance, responsible working conditions, environmental responsibility, and community relations in soy certification standards. Outcome Level: Evaluates presence of criteria but does not explicitly measure continuous improvement over time.	Schemes that meet the criteria are shown in the *ITC Standards Map* as 'FEFAC benchmarked'. (Pass/fail. Schemes that don't pass this threshold are not shown).
Siegelklarheit	To provide clarity and enable users to compare different standards within product groups.	Assesses labels for certain product groups, using the SSCT (Sustainability Standards Comparison Tool) - evaluation and comparison methodology. It uses ITC Standards Map for data collection but applies its own assessment methodology.	System Level: Evaluates credibility, transparency in decision-making, and effectiveness of the control system. Content Level: Analyzes environmental and social issues such as chemical use, forced labor, and discrimination.	Labels' performance is shown as the 3-star rating system to consumers (B2C) (Scoring included – 0-2 points scale)
ITC Standards Map	To provide information on the criteria and requirements of each standard, as well as the certification process. It also provides foundation for T4SD customized tools.	It offers a comprehensive overview of more than 300 sustainability standards and certification schemes, and coverage of system characteristics and sustainability topics through 1650 indicators.	System Level: Documents criteria presence or absence, provides information on certification process and standards.	It documents the presence (or absence) of related criteria within the standard. (Results are shown in the form of pie charts or table lists)
SME Compass	To offer guidance to companies to identify and manage their sustainability risks. Performance in regard to	Benchmarking certification and participation standards against 40 due diligence criteria (based on UNGP).	System Level: Guides companies through strategy and policy, risk analysis, measures, monitoring and communication, grievance	Standards Tool compares max 4 standards at the same time and visualize the results in form of percentage bars.

	achieving requirements.	Data is sourced from ITC Standards Map.	and remedy processes.	ITC Data Entry Tool as an Excel-based table contains the criteria grid that serves as the basis for the assessment methodology in the Standards Tool.
FSI Basket of Standards	To showcase standards/schemes with responsible sourcing of flowers and plants. It helps traders and buyers to identify suppliers with responsible practices and products.	Assessing compliance of standards and schemes against Good Agricultural Practices (GAP), Environmental and/or Social (GSCP/SSCI) basic requirements.	System Level: Uses independent benchmarking bodies for Good Agricultural Practices (GAP), Environmental, and Social basic requirements. Content Level: Evaluates specific criteria and requirements related to GAP, environmental responsibility, and social practices of each standard.	Standards that successfully pass benchmarking are added to the FSI Basket (recognized for responsible sourcing). Standards are published and publicly available for consultation and comparison of their scope and criteria in the ITC Standards Map online portal.
GSSI	To provide formal recognition of seafood certification schemes that successfully complete benchmark process.	Benchmarking against performance areas covering scheme governance, operational management (including chain of custody) and standards for aquaculture and fisheries certification.	System Level: Assesses governmental and operational criteria for seafood certification schemes. Content Level: Focuses on environmental aspects but excludes social issues. Outcome Level: Provides formal recognition for alignment with FAO standards; includes outcome level indicators through benchmark reports.	The Global Benchmark Framework is an Excel spreadsheet containing components and guidance text with FAO reference. GSSI recognized certification schemes are publicly announced on the website, with a benchmark report included.

Purpose

The selected monitoring systems exhibit a diverse range of purposes. The FEFAC Responsible Soy Benchmarking Tool is geared towards identifying soy certification schemes that meet European feed industry requirements for responsible soy production. Siegelklarheit aims to provide clarity and enable users to make informed decisions for sustainable consumption based on credible labels. ITC Standards Map serves as an extensive information platform, offering details about sustainability standards and certification schemes. SME Compass is designed to guide companies in identifying and managing sustainability risks. The FSI Basket of Standards showcases standards and schemes with responsible sourcing of flowers and plants. Lastly, GSSI Global Benchmark Tool provides formal recognition to seafood certification schemes that successfully complete the benchmark process.

Benchmark Function

Regarding benchmark functions, each monitoring system employs unique methodologies. The FEFAC Responsible Soy Benchmarking Tool assesses certification schemes against specific criteria outlined in FEFAC's Soy Sourcing Guidelines. Siegelklarheit uses the SSCT methodology to evaluate labels for product groups, employing a 3-star rating system with scoring. ITC Standards Map offers a comprehensive overview of sustainability standards and certification schemes, covering various sustainability topics through 1650 indicators. SME Compass benchmarks certification and participation standards against 40 due diligence criteria based on UNGP. The FSI Basket of Standards assesses compliance against Good Agricultural Practices (GAP), Environmental, and/or Social basic requirements. GSSI Global Benchmark Tool evaluates certification schemes against performance areas encompassing scheme governance, operational management, and standards for aquaculture and fisheries certification.

Result Visualization

The visualization of results also varies across the selected monitoring systems. The FEFAC Responsible Soy Benchmarking Tool displays schemes that meet the criteria as 'FEFAC benchmarked' within the ITC Standards Map, adopting a straightforward pass/fail system. Siegelklarheit empowers consumers by presenting label performance through a 3-star rating system with scoring, enabling them to discern between various labels and their sustainability performance. ITC Standards Map documents the presence or absence of related criteria within each standard, presenting the information using pie charts or table lists. SME Compass compares up to 4 standards simultaneously, visualizing the results as percentage bars using the ITC Data Entry Tool. The FSI Basket of Standards acknowledges successfully benchmarked standards, adding them to the FSI Basket, and making them publicly available for consultation and comparison. GSSI Global Benchmark Tool utilizes an Excel-based Global Benchmark Framework, publicly announcing recognized certification schemes on their website along with a benchmark report.

Database Indicators Assessment

There are three different levels for database indicators identified through STAR4BBS project, as mentioned in the *Methodology* section.

By using these three levels of indicators, the monitoring systems can provide a comprehensive assessment of the sustainability standards and certification schemes, covering both their design and implementation aspects as well as their real-world impact on environmental, social, and economic dimensions. This approach allows stakeholders to gain a deeper understanding of the strengths and weaknesses of each system.

Below, a brief overview of each identified monitoring tool's database assessment is presented.

FEFAC Responsible Soy Benchmarking Tool

The FEFAC Responsible Soy Benchmarking Tool focuses on soy certification schemes and their compliance with the FEFAC Soy Sourcing Guidelines. These guidelines consist of criteria covering areas such as legal compliance, responsible working conditions, environmental responsibility, and community relations. The tool primarily assesses the content level indicators, evaluating the specific criteria and requirements of the sustainability standards related to responsible soy production. It provides a detailed evaluation of each criterion's presence and considers 54 essential (obligatory) and 19 desired (optional) criteria.

Siegelklarheit

Siegelklarheit evaluates sustainability standards and labels in three dimensions: Credibility, Environmental, and Social. It covers a wide range of social and environmental standards from various consumer goods sectors. The tool assesses both the system level and content level indicators. In terms of system level indicators, it evaluates the credibility of the standard-setting organization, transparency in decision-making, and effectiveness of the control system. As for content level indicators, it analyzes how well the standards address relevant environmental and social issues, such as chemical use, forced labor, and discrimination.

SME Standards Compass

The SME Standards Compass focuses on guiding companies through the due diligence process to identify and manage their sustainability risks. It covers five phases of the due diligence process, which include Strategy and Policy, Risk analysis, Measures, Monitoring and Communication, and Grievance and remedy. The tool primarily addresses system level indicators, providing guidance and support to companies in implementing sustainability measures throughout their supply chains.

FSI Basket of Standards

The FSI Basket of Standards currently includes 16 Voluntary Sustainable Standards and Schemes, which are transparent and comparable through independent benchmarking. The tool focuses on compliance with Good Agricultural Practices (GAP), Environmental, and Social basic requirements. For the benchmarking process, different independent

benchmarking bodies are used for each category. For GAP benchmarking, GLOBALG.A.P., an internationally recognized standard for farm production, is used. The environmental benchmarking is carried out by FOOD EXPERTS, an independent third-party, while the social benchmarking is performed by the Global Social Compliance Program (GSCP). The tool primarily addresses content level indicators, evaluating the specific criteria and requirements related to GAP, environmental responsibility, and social practices of each standard in the FSI Basket (Fig. 9).

Figure 9 The FSI Basket of Standards with benchmarked standards and schemes (source: https://www.fsi2025.com/basket/)

GSSI

The GSSI Global Benchmark Tool assesses seafood certification schemes against 175 Essential Components, aiming to be aligned with FAO standards and guidelines. It does not include a ranking or declaration of schemes as "sustainable." The benchmark process focuses on environmental, governmental, and operational criteria and does not include social issues. The outcome level indicators are addressed through the benchmark report for each "GSSI recognized" scheme. The tool provides formal recognition to schemes that are "in alignment" with the Essential Components and related guidelines, ensuring responsible fisheries and aquaculture practices.

ITC Standards Map

The ITC Standards Map serves as an online platform documenting the presence or absence of related criteria within sustainability standards and certification schemes. The tool does not score or rate schemes, and it does not assess the effectiveness of the schemes. It provides information on the criteria and requirements of each standard, certification process, and other relevant details. The database is populated through document reviews and consultations with the schemes. The tool mainly focuses on system level indicators, providing stakeholders with comprehensive information about the various sustainability standards available without ranking or rating their performance (Fig. 10).

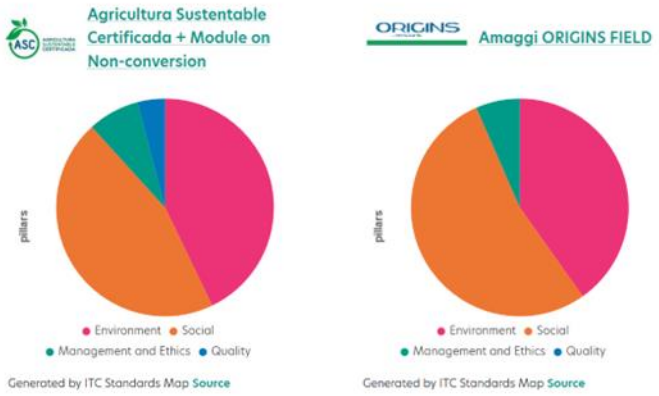


Figure 10 An example of the certification schemes' comparison on the ITC Standards Map platform (source: https://standardsmap.org/en/compare)

Evaluation Structure - Interpretation of the results, including threshold, ranking and other metrics

The ways in which results are analyzed and interpreted vary among different monitoring systems. Some systems set specific threshold values or benchmarks against which each scheme's or standard's performance is measured. This could involve setting minimum requirements that a scheme/label must meet to be considered effective or robust. Certain benchmarking tools also use ranking systems or scoring mechanisms to compare and quantify the performance of the schemes objectively.

The ***FEFAC Responsible Soy Benchmarking Tool*** adopts a comprehensive evaluation structure based on the FEFAC Soy Sourcing Guidelines. It consists of 73 criteria categorized into six pillars, with 54 being essential (obligatory) and 19 desired (optional). To pass the benchmarking exercise, a scheme must meet at least 8 out of the 19 desired criteria, in addition to fulfilling the verification requirements. This tool uses a pass/fail rating system, indicating whether a scheme meets the established benchmarking criteria or not.

In case of the FEFAC Benchmarking Tool, responsible soy scheme owners voluntarily apply to have their respective (certification) standard or program benchmarked against the criteria in the Guidelines. This benchmarking exercise is independently facilitated and executed by ITC (the International Trade Centre).

Siegelklarheit, in contrast, employs a multidimensional evaluation approach to assess labels' sustainability performance. It evaluates labels based on three dimensions: Credibility, Environmental, and Social, using comprehensive grids of criteria. Within these criteria grids we distinguish between minimum requirements (of the German Government[8]) and additional requirements, making it altogether more than 300 requirements. The credibility dimension examines the transparency and decision-making structures of standard-setting organizations, while the environmental and social dimensions analyze relevant issues such as the use of chemicals and labor practices.

First, it is examined whether a label meets the minimum requirements by the German Federal Government in the dimension of credibility. If not, it is not assessed further. If yes,

[8] https://www.bundesregierung.de/breg-en/issues/sustainability

the entire criteria grid for the corresponding product group and for the credibility dimension is then checked in detail for the second step (Fig. 11).

In a second step, it is checked whether the label meets or even exceeds the minimum requirements, which ultimately leads to the statement "Good choice" or "Very good choice". The overall rating ("Good Choice" or "Very Good Choice") is determined based on the joint consideration of individual dimension ratings. A label achieves the overall rating "Good Choice" if at least 2 stars have been achieved in the area of credibility and in the area of environment **or** social issues. A label receives the overall rating "Very Good Choice" if it has achieved 3 stars in the area of credibility and 3 stars in at least one other area.

The percentage of the achieved score (incl. basic & advanced) of the total criteria grid per area is calculated to determine whether the label meets the threshold of 60% (requirement for three stars). If the percentage is below 60%, the label remains at 2 stars in this area. If the percentage is equal to or higher than 60%, the label receives 3 stars for the relevant area.

The evaluation approach for the "Very good choice" rating involves meeting minimum requirements in at least two dimensions and considering the **total number of points achieved**. The system employs an absolute approach, distinguishing between basic and advanced (higher) levels of intensity for various requirements. Each fulfilled basic requirement receives 1 point, while advanced requirements receive 2 points. The scores are then totalled to calculate the label's percentage of the maximum score, determining the star rating. Some requirements are evaluated with 1 point when fulfilled without further subdivision.

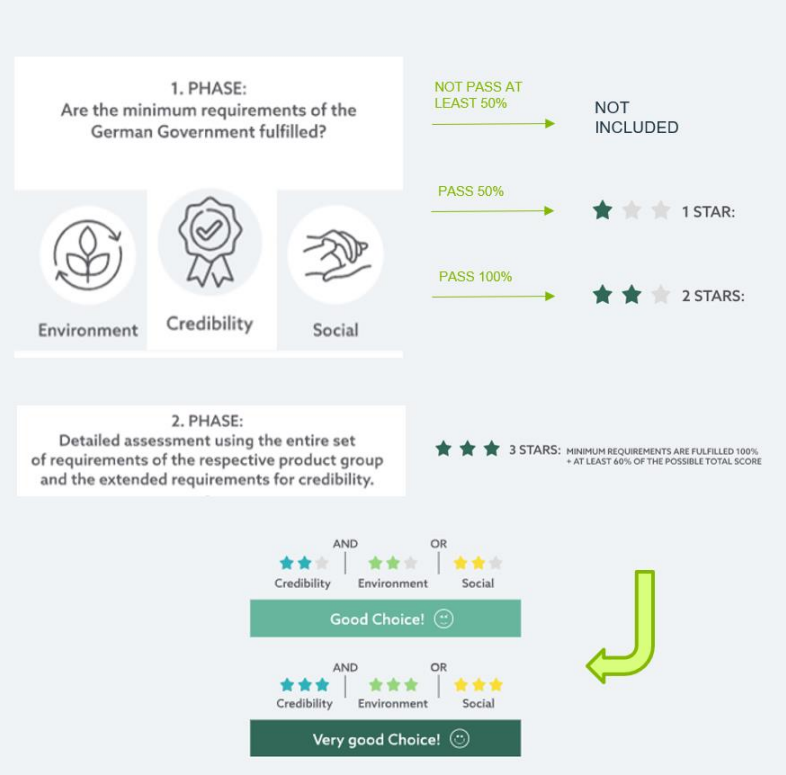


Figure 11 Assessment methodology of the Siegelklarheit tool (adapted from: https://www.siegelklarheit.de/en/assessment-methodology)

Within the framework of the Siegelklarheit, the assessment of labels based on the criteria sets is carried out by independent experts from the International Trade Centre (ITC).

SME Compass is designed to assist small and medium-sized enterprises in making informed decisions about the opportunities and limitations of standards in their due diligence process and choose the right standards for their needs. Tool's scoring is based on two data points from the ITC Standards Map - Degree of Obligation (DoO) and Degree of Criticality (DoC) metrics (Fig. 12). DoO specifies the deadline set by the standard for indicator fulfillment, classified as immediate action, transition period, or recommendation. The DoC defines the consequences of failing to meet an indicator, ranging from "Deal Breaker" to "Optional Compliance".

"Deal Breaker": If the indicator is classified as a Deal Breaker, failure to satisfy the indicator will lead immediately to de-certification or withdrawal of membership.

Major non-compliance: If the company fails to satisfy the indicator, it must immediately submit a Corrective Action Plan that will be reviewed within a very short space of time.

Minor non-compliance: If the company fails to satisfy the indicator, it must submit a Corrective Action Plan that will be reviewed as part of its regular audits.

Optional compliance: Just like with DoO, it is recommended that companies satisfy this indicator, but failure to satisfy it will not lead to de-certification or withdrawal of membership.

SME Compass doesn't employ a rating system or set thresholds but offers a qualitative assessment of how standards align with the SME's needs and due diligence process.

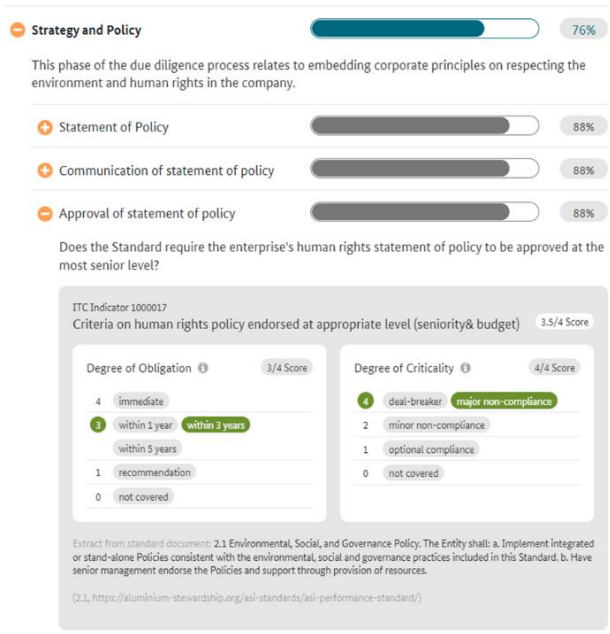


Figure 12 Assessment methodology of the SME Compass tool (source: https://kompass.wirtschaft-entwicklung.de/en/standards-compass/standards-tool#/standards)

In SME Compass, scoring and weighting of standards are based on the criteria and accompanying ITC indicators. Each indicator is awarded scores according to the level of requirement set by the standard. The Degree of Obligation (DoO) and Degree of Criticality (DoC) are integrated equally, each contributing 50% to the overall assessment of an indicator. If a criterion is based on multiple ITC indicators, the average score is calculated to ensure fair and unbiased scoring, avoiding any skewing of the results. This approach helps SMEs gain insight into opportunities and limitations of standards in their due diligence process and choose the most suitable standards for their needs.

The ***FSI Basket of Standards*** employs independent benchmarking based on Good Agricultural Practices (GAP), Environmental, and Social basic requirements. Standards are categorized as required, optional, or recommended, enhancing transparency and comparability. However, no other information is available on their website on evaluation structure, and currently the tool which is hosted on the ITC Standards Map is not available for further analysis.

GSSI Benchmark Tool uses Essential Components, aligned with the FAO Code of Conduct for Responsible Fisheries and Ecolabelling Guidelines, as a compulsory benchmark for all applicant schemes. Additionally, Supplementary Components demonstrate schemes' diverse approaches to sustainable seafood. GSSI-recognized Scheme Owners must align with all Essential Components applicable to them. Although no specific rating system is mentioned, GSSI recognition provides formal acknowledgment of schemes meeting the Essential Components.

The GSSI Tool incorporates a pre-assessment process aimed at driving improvements in non-recognized seafood certification schemes. This process offers valuable insights into how these schemes align with the Global Benchmark Tool, encouraging them to meet the required standards and potentially achieve GSSI recognition.

GSSI-recognized Scheme Owners demonstrate full alignment with the Essential Components in addition to their respective aquaculture or fisheries standards. This alignment ensures compliance with specific sections (A, B, C, and/or D) based on the type of standard they hold, without the need for Supplementary Components.

ITC Standards Map serves as a comprehensive database documenting the presence or absence of related criteria within standards without scoring or rating schemes. However, it generates graphs and statistics on standards' performance, allowing for comparison based on specific sustainability topics. It offers valuable insights into the scope and coverage of each standard's criteria and requirements. It is part of the back-end T4SD database (Trade for Sustainable Development), programme of ITC providing comprehensive, verified and transparent information on sustainability standards.

The experts at ITC have been specially trained in the requirements of the tools they are hosting. The final quality assurance is conducted by the ITC experts before the raw data are imported into the back-end of the customized tool and the assessment methodology is applied.

This tool assesses in detail the scope and coverage of sustainability topics by each standard, the degree of obligation, compliance level of criticality, explicitness, full text and source of information (Fig. 13).

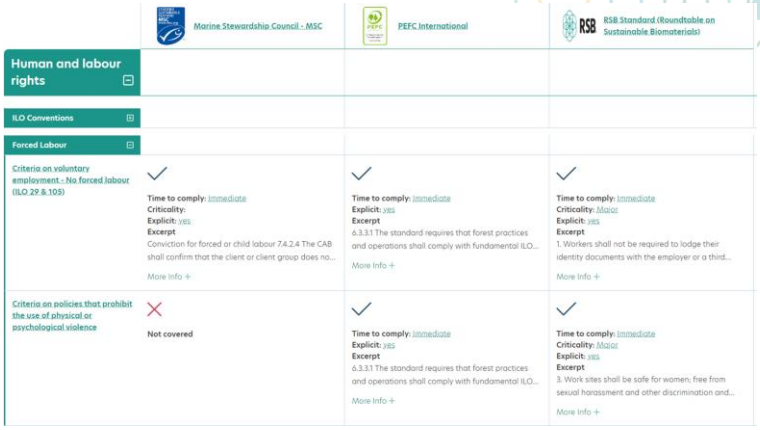


Figure 13 Example of the comparison of the full criteria list on the ITC Standards Map platform (source: https://standardsmap.org/en/compare)

Key lessons learnt

The data and gap analysis of existing benchmarking methodologies and associated monitoring tools is a crucial step in conceptualizing and developing new monitoring systems to ensure proper testing of the effectiveness and relevance of existing schemes and labels. Based on the information gathered from various selected existing tools, several important considerations and potential obstacles have emerged that should be addressed when designing new monitoring tools for assessment and benchmarking purposes.

One key aspect to consider is the clarity of criteria included in the monitoring tools. Ensuring that the requirements are well-defined, easy to be used and transparent is essential to facilitate the assessment and/or benchmarking process and gain stakeholder trust. Additionally, the complex assessment framework (e.g., heavy criteria assessment and extensive requests during the assessment/benchmarking process) may lead to challenges for both the benchmarking body and the schemes owners being assessed. Striking a balance between comprehensive evaluations and streamlined assessments will be critical for the successful implementation and application of the new monitoring systems, considering a certain level of simplicity and feasibility.

Another important consideration is the ability of the assessment tools and their benchmarking methodologies to assess the extent to which a standard addresses specific regional issues. Individual assessments might be required to understand how certifications impact the unique situations of sourcing or production sites. This aspect connects to the outcome or impact level, and integrating regional considerations will enhance the relevance and applicability of the new monitoring systems.

Additionally, it would be important for new monitoring tools/benchmarking methodologies to provide evidence, when possible, of the achieved impact of adopted practices, for example, in terms of sustainability outcomes. Evaluating how well standards are implemented and their effectiveness is crucial to identify gaps and drive continuous improvement in sustainability practices. Aligning with internationally

recognized criteria, such as the OECD Guidelines[9], can further enhance the robustness and credibility of the monitoring systems.

Furthermore, the availability and accessibility of data play a significant role in the effectiveness of benchmarking methodologies. Ensuring that tools and portals hosting the data are well-maintained and regularly updated will provide reliable information for analysis and comparison. For instance, some existing tools faced limitations due to the unavailability of data on the ITC portal (where they are hosted), highlighting the need for continuous data management and maintenance.

The experience with existing benchmarking methodologies, such as the GSSI Global Benchmark Tool, demonstrates the importance of regular revisions and updates. Refining the framework criteria and reducing redundancy ensures the benchmarking process remains relevant and aligned with evolving sustainability priorities.

To drive improvement in the schemes, the new monitoring systems should identify gaps and relevant requirements against various criteria. However, it is essential to recognize that not all criteria may be applicable to every certification scheme and labels. Providing schemes with clear guidance on identifying gaps and areas for improvement will facilitate their sustainability journey.

Moreover, to ensure that the new monitoring system remains effective and responsive to stakeholders' needs, it should include a comprehensive framework for continuous improvement and regular revisions. Based on our findings, we propose a framework consisting of four key components (Figure 14).

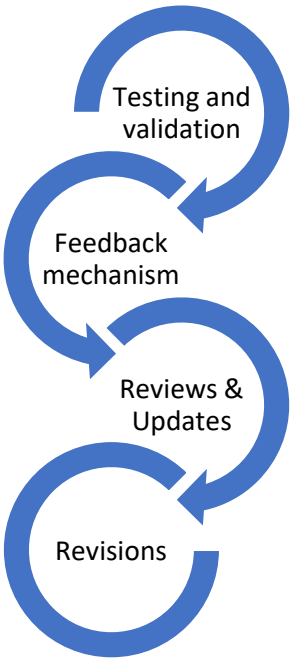


Figure 14 Framework loop for continuous improvement and viability of the monitoring system

[9] https://www.oecd.org/daf/inv/mne/48004323.pdf

1. Testing and Validation

Objective: Assess SCS and labels for bio-based feedstock and products and validate the feasibility of the monitoring system

Timeframe: At least two rounds of pilot testing should be conducted.

Testing Process:

- Assess pilot SCS and labels against the indicators included in the monitoring tool.
- Track the soundness and applicability
- Allocate reference score based on the proposed scales.
- Involve SCS owners to reflect on preliminary results and recommendations derived from the testing.

- **Validation Process:**
  - Validate changes proposed by testers and scheme owners and reiterate indicators accordingly.
  - Assess the feasibility of the pre-set minimum requirements for the three levels of the monitoring system.
  - Implement necessary changes to the system's evaluation structure.

2. Feedback Mechanism

Objective: Collect and integrate regular feedback from stakeholders.

Timeframe: Feedback should be collected for the initial tool development, for the testing rounds, and for each revision and update.

- **Internal Feedback:**
  - Project consortium (regular team meetings, emails)
  - Sister projects partners (live meetings or online calls, designated events)
- **External Feedback:**
  - Co-Creation Workshops - engaging stakeholders in the co-design and iterative improvement of the monitoring system
    - Participants: Key internal and external stakeholders, including certification scheme owners, industry experts, and decision-makers.
  - Via different feedback channels, including surveys and online exercises (e.g. *Mentimeter*).
  - Feedback sessions during conferences and other dissemination events.
  - Online platforms for broader input.

3. Regular Reviews and Updates

Objective: Systematically evaluate and update the monitoring system.

Timeframe: Biannual (comprehensive) and periodical (as needed, based on feedback).

- **Review Process:**
    - Update of the indicators.
    - Revision and update of the thresholds and metrics.
    - Communication of reviews and revisions with stakeholders.

4. Integration of Feedback and Revisions

Objective: Implement feedback and ensure revisions are effective and aligned with stakeholders' needs.

Timeframe: Following iterative process of reviews and updates.

- **Implementation Process:**
    - Development of revision proposals based on feedback.
    - Validation and testing of proposed changes.
    - Release of updates with noted changes.

In conclusion, based on the data and gap analysis of the six monitoring tools in the design phase of the STAR4BBS monitoring tool, **the clarity of criteria, the ability to address regional issues, the implementation of standards, data availability and accessibility, regular revisions, and effective guidance for schemes** should be considered. By addressing these aspects, the new monitoring systems can enhance their effectiveness, credibility, and impact in promoting sustainability across various industries and value chains.

5 Recommendations

Based on the research conducted in this deliverable and the in-depth analysis of selected monitoring tools, we formulated six key recommendations for developing a new monitoring system for assessing the effectiveness and robustness of SCS and labels.

1. **Potential to utilize ITC Standards Map data**: As many of the selected monitoring systems use the database of the ITC Standards Map, especially for data collection, the new monitoring system can also benefit from collaborating with this platform or being hosted by it. The ITC Standards Map provides comprehensive and verified information on sustainability standards, which can be used to gather information

on the SCS and labels to be tested by the monitoring system. Additionally, the new monitoring system can inspire ITC to develop new indicators not already available in the database, enhancing the scope and coverage of the platform's overall structure.

2. **Applying evaluation structure elements (rating system) based on the Siegelklarheit methodology**: The new STAR4BBS monitoring system should consider applying the multidimensional evaluation approach of the Siegelklarheit and its SSCT in its methodology development. This method allows for a clear and user-friendly (star) rating system, distinguishing between labels that meet minimum requirements and those that exceed them, and taking into account the special features of the labels.

3. **Enhancing the *connectivity* between three monitoring system levels**: The GSSI benchmarking tool's approach separates its sections on *Governance and Operational Management of Seafood Certification Schemes* from *Fisheries and Aquaculture Certification Standards*. To improve the benchmarking process, the new monitoring system should consider separately outcome- and system-level criteria for all schemes and labels (potentially with compulsory requirements for those two sections), while also ensuring that there is connectivity between the three levels (see also STAR4BBS deliverable *D1.2 Report on existing international and EU SCS and B2B labels for feedstock and bio-based materials & products*). This will allow for a more holistic and comprehensive evaluation of the schemes.

4. **Introduction of new indicators**: Building on the experience of certain tools (e.g. SME Compass tool), the new monitoring system should have a mechanism for introducing new indicators for each of the identified levels (system, content and outcome levels) based on emerging sustainability issues and evolving industry requirements. This will enable the monitoring system to stay up-to-date with the latest developments and trends in the bio-based products sector, ensuring its relevance and effectiveness in promoting sustainability.

5. **Reliability of Methodological Approach**: A new monitoring system should adopt a flexible and adaptive approach to its methodological design, ensuring that it aligns with the objectives, goals, and context of the bio-based products industry. It should also involve assessments against performance areas that cover not only sustainability indicators, but also governance and operational management, as well as the impact of the schemes and labels. In addition, potential reliance on ITC could help increase reliability, as the ITC follows a consistent and systematic process to obtain and check information about each indicator in its database. This comprehensive approach will provide a more well-rounded evaluation of the bio-based products' sustainability performance.

6. **Stakeholder engagement**: The majority of the analysed monitoring tools have extensively used stakeholder participation and involvement for the development or improvement of their methodologies. To ensure the relevance and effectiveness of the new monitoring system, it is essential to engage stakeholders throughout the development process. Stakeholder engagement can take the form of workshops or consultations, where input and feedback from various stakeholders, including industry experts, certification bodies, and civil society organizations, can be gathered. This will help in identifying key priorities, refining

the methodology, and ensuring that the system meets the needs and expectations of all stakeholders.

7. **Continuous improvement and regular revisions**: To ensure system's effectiveness, a framework for monitoring system's continuous improvement and revisions was created, encompassing key components: testing and validation, feedback mechanism, regular reviews and updates, and integration of feedback and revisions. This will contribute to establish a proper baseline for long-term sustainability of the monitoring system.
8. **Specific action plan steps**: Learnings from this deliverable are being implemented into the new monitoring system's framework in D4.1 *Concept of the monitoring system*. Coupled with the recommendations from the Sustainability Benchmarking Good Practice Guide (ISEAL, 2020), three phases are being considered in such conceptualization:
    - Phase 1 - Determine the audience and the purpose
    - Phase 2 - Determine which elements to include
    - Phase 3 - Determine the evaluation structure

The conceptualization was informed by the following core principles defined in ISEAL's Sustainability Benchmarking Good Practice Guide:

- Rigour: Benchmarking exercises and programmes are structured and implemented in a way that is sufficient to produce quality outcomes.
- Accessibility: Benchmarking exercises and programmes avoid structures that create unnecessary barriers to participation and seek to minimize the reporting and engagement burden for entities being benchmarked.
- Efficiency: Benchmarking exercises and programmes are structured as simply as possible and avoid redundancies.
- Improvement: Benchmarking exercises and programmes are structured to incentivize better practices in the entities that they cover.

By incorporating these recommendations into the development of the monitoring system for bio-based products, this can become a valuable and credible tool for promoting sustainability in the bio-based industry. The system will provide valuable insights and guidance to policy makers, certification bodies, and other stakeholders, driving the adoption of responsible and environmentally friendly practices throughout the bio-based products supply chain.

6 Conclusions

The report provides a comprehensive analysis and evaluation of the existing monitoring tools used in the context of sustainability standards, certification schemes, and labels for various industries. The tools examined in-depth include the FEFAC Responsible Soy Benchmarking Tool, Siegelklarheit, ITC Standards Map, SME Compass, FSI Basket of Standards, and GSSI Global Benchmark Tool. Through an exploration of their purpose, benchmark functions, and result visualizations, we gained valuable insights into their

applicability and effectiveness in promoting sustainable practices across different sectors.

Each tool serves a unique purpose, ranging from identifying responsible soy production schemes to providing clarity for consumers on choosing sustainable products. The benchmark functions of these tools involve assessing and comparing certification schemes and labels against specific criteria and indicators, often drawing data from the extensive ITC Standards Map database.

The result visualizations varied, ranging from pass/fail indicators to 3-star rating systems, pie charts, and percentage bars. These visualizations aim to offer stakeholders clear and easily interpretable information on the sustainability performance of various schemes and labels, enabling them to make informed decisions and encourage improvements in the supply chain.

The evaluation structure of each tool revealed diverse approaches to assessing sustainability performance. Some tools incorporate comprehensive grids of criteria across multiple dimensions, while others focus on aspects such as governance, operational management, and due diligence. Scoring and weighting methods vary, with considerations for degree of obligation and criticality to determine compliance levels.

In developing a novel monitoring system, several important features should be considered. Leveraging data from the ITC Standards Map can enhance the system's information base, while stakeholder engagement can provide valuable insights and feedback. Moreover, careful consideration of the specific scope, audience, and contextual factors is essential when selecting methodological approaches for data collection, analysis, and interpretation.

The report also highlights the need for continuous improvement and refinement of monitoring systems. Feedback from stakeholders and periodic revisions ensure that the systems remain up-to-date, relevant, and effective in achieving their intended goals.

The lessons learned from the evaluation and comparison of existing monitoring tools provide valuable guidance for the conceptualization and development of the novel STAR4BBS monitoring system for bio-based products. By incorporating best practices, robust evaluation structures, and stakeholder engagement, the new system can become a powerful tool in advancing sustainability practices and responsible sourcing within the bio-based products sector.

7 References

General References:

Gottinger, A.; Ladu, L.; Quitzow, R. (2020): Studying the Transition towards a Circular Bioeconomy—A Systematic Literature Review on Transition Studies and Existing Barriers. In *Sustainability* 12 (21), 8990. https://doi.org/10.3390/su12218990

ISEAL (2020): Sustainability Benchmarking Good Practice Guide. Version 1.1, September 2020.

ISEAL (2021): ISEAL Credibility Principles. Version 2, June 2021.

Ladu, L.; Blind, K. (2017): Overview of policies, standards and certifications supporting the European bio-based economy. In *Current Opinion in Green and Sustainable Chemistry* 8, pp. 30–35. DOI: 10.1016/j.cogsc.2017.09.002

Ladu, L; Morone, P. (2021): Holistic approach in the evaluation of the sustainability of bio-based products: An Integrated Assessment Tool. In *Sustainable Production and Consumption* 28, pp. 911-924er. https://doi.org/10.1016/j.spc.2021.07.006.

Majer, S.; Wurster, S.; Moosmann, D.; Ladu, L.; Sumfleth, B.; Thrän, D. (2018): Gaps and Research Demand for Sustainability Certification and Standardisation in a Sustainable Bio-Based Economy in the EU. In *Sustainability* 10, p. 2455. https://doi.org/10.3390/su10072455

Moher, D.; Liberati, A.; Tetzla, J.; Altman, D.G. (2009): Preferred reporting items for systematic re-views and meta-analyses. The PRISMA statement. In *BMJ* 2009, 339, b2535.

Villamil, C.; Hallstedt, S.I. (2020): Sustainability integration in product portfolio for sustainable development: Findings from the industry. In *Bus. Strat. Environ.* 30 (1), pp. 388-403. https://doi.org/10.1002/bse.2627

References from the systematic literature review:

Ali, Hikmat H.; Barakat, Diana K.; Sharif, Ahlam A. (2021): Establishing a Green Building Certification Scheme and Standards for Multifamily Residential Buildings: Case of Jordan. In *J. Archit. Eng.* 27 (2), Article 05021003. DOI: 10.1061/(ASCE)AE.1943-5568.0000468.

Barnett, Brad (2016): An evaluation of the UK's use of SFM standards to procure solid woody biomass for electricity generation using sustainable bioenergy criteria. In *Biofuels* 7 (1), pp. 1–11. DOI: 10.1080/17597269.2015.1118775.

Chaplin-Kramer, Rebecca; Jonell, Malin; Guerry, Anne; Lambin, Eric F.; Morgan, Alexis J.; Pennington, Derric et al. (2015): Ecosystem service information to benefit sustainability

standards for commodity supply chains. In *Annals of the New York Academy of Sciences* 1355, pp. 77–97. DOI: 10.1111/nyas.12961.

Chiputwa, Brian; Spielman, David J.; Qaim, Matin (2015): Food Standards, Certification, and Poverty among Coffee Farmers in Uganda. In *World Development* 66, pp. 400–412. DOI: 10.1016/j.worlddev.2014.09.006.

Chkanikova, Olga; Kogg, Beatrice (2018): Sustainability governance service providers: the role of third-party product certification in facilitating corporate life cycle management. In *Int J Life Cycle Assess* 23 (7), pp. 1383–1395. DOI: 10.1007/s11367-015-0865-z.

Englund, Oskar; Berndes, Göran (2015): How do sustainability standards consider biodiversity? In *WIREs Energy Environ.* 4 (1), pp. 26–50. DOI: 10.1002/wene.118.

Gutierrez Garzon, Alba Rocio; Bettinger, Pete; Siry, Jacek; Abrams, Jesse; Cieszewski, Chris; Boston, Kevin et al. (2020): A Comparative Analysis of Five Forest Certification Programs. In *Forests* 11 (8), p. 863. DOI: 10.3390/f11080863.

Hazen, Lucie; Le Cornu, Elodie; Zerbe, Amanda; Martone, Rebecca; Erickson, Ashley L.; Crowder, Larry B. (2016): Translating sustainable seafood frameworks to assess the implementation of ecosystem-based fisheries management. In *Fisheries Research* 182, pp. 149–157. DOI: 10.1016/j.fishres.2015.11.019.

Kittinger, John N.; Bernard, Miranda; Finkbeiner, Elena; Murphy, Erin; Obregon, Pablo; Klinger, Dane H. et al. (2021): Applying a jurisdictional approach to support sustainable seafood. In *Conservat Sci and Prac* 3 (5). DOI: 10.1111/csp2.386.

Lee, W. L. (2013): A comprehensive review of metrics of building environmental assessment schemes. In *Energy and Buildings* 62, pp. 403–413. DOI: 10.1016/j.enbuild.2013.03.014.

Mai-Moulin, Thuy; Armstrong, Simon; van Dam, Jinke; Junginger, Martin (2019): Toward a harmonization of national sustainability requirements and criteria for solid biomass. In *Biofuels, Bioprod. Bioref.* 13 (2), pp. 405–421. DOI: 10.1002/bbb.1822.

Mengistie, Belay T.; Mol, Arthur P. J.; Oosterveer, Peter (2017): Governance of agro-pesticide through private environmental and social standards in the global cut flower chain from Ethiopia. In *Ambio* 46 (7), pp. 797–811. DOI: 10.1007/s13280-017-0914-x.

Merger, Eduard; Dutschke, Michael; Verchot, Louis (2011): Options for REDD+ Voluntary Certification to Ensure Net GHG Benefits, Poverty Alleviation, Sustainable Management of Forests and Biodiversity Conservation. In *Forests* 2 (2), pp. 550–577. DOI: 10.3390/f2020550.

Meyer, Markus A.; Priess, Joerg A. (2014): Indicators of bioenergy-related certification schemes – An analysis of the quality and comprehensiveness for assessing local/regional environmental impacts. In *Biomass and Bioenergy* 65, pp. 151–169. DOI: 10.1016/j.biombioe.2014.03.041.

Montazeri, Mahdokht; Zaimes, George G.; Khanna, Vikas; Eckelman, Matthew J. (2016): Meta-Analysis of Life Cycle Energy and Greenhouse Gas Emissions for Priority Biobased Chemicals. In *ACS Sustainable Chem. Eng.* 4 (12), pp. 6443–6454. DOI: 10.1021/acssuschemeng.6b01217.

Mussells, Claire; Stephenson, Robert L. (2020): A comparison of sustainability objectives: how well does the Canadian Fisheries Research Network framework compare with fisheries, forestry, and aquaculture certification schemes? In *E&S* 25 (1). DOI: 10.5751/ES-11368-250117.

Partzsch, Lena; Zander, Macy; Robinson, Hannah (2019): Cotton certification in Sub-Saharan Africa: Promotion of environmental sustainability or greenwashing? In *Global Environmental Change* 57, p. 101924. DOI: 10.1016/j.gloenvcha.2019.05.008.

Schut, Marc; Florin, Madeleine J. (2015): The policy and practice of sustainable biofuels: Between global frameworks and local heterogeneity. The case of food security in Mozambique. In *Biomass and Bioenergy* 72, pp. 123–135. DOI: 10.1016/j.biombioe.2014.11.009.

Sikkema, Richard; Junginger, Martin; van Dam, Jinke; Stegeman, Gerben; Durrant, David; Faaij, Andre (2014): Legal Harvesting, Sustainable Sourcing and Cascaded Use of Wood for Bioenergy: Their Coverage through Existing Certification Frameworks for Sustainable Forest Management. In *Forests* 5 (9), pp. 2163–2211. DOI: 10.3390/f5092163.

Sugiura, Katsuaki; Oki, Yoshinao (2018): Reasons for Choosing Forest Stewardship Council (FSC) and Sustainable Green Ecosystem Council (SGEC) Schemes and the Effects of Certification Acquisition by Forestry Enterprises in Japan. In *Forests* 9 (4), p. 173. DOI: 10.3390/f9040173.

Tlusty, Michael F.; Thompson, Matthew; Tausig, Heather (2016): Statistical tools to assess the breadth and depth of shrimp aquaculture certification schemes. In *Fisheries Research* 182, pp. 172–176. DOI: 10.1016/j.fishres.2015.10.008.

Tröster, Rasmus; Hiete, Michael (2019): Do voluntary sustainability certification schemes in the sector of mineral resources meet stakeholder demands? A multi-criteria decision analysis. In *Resources Policy* 63, p. 101432. DOI: 10.1016/j.resourpol.2019.101432.

Young, S. B.; Zhe, Y.; Dias, G. (2014): Prospects for sustainability certification of metals. In *Metall. Res. Technol.* 111 (3), pp. 131–136. DOI: 10.1051/metal/2014008.

Zielinski, Seweryn; Botero, Camilo (2015): Are eco-labels sustainable? Beach certification schemes in Latin America and the Caribbean. In *Journal of Sustainable Tourism* 23 (10), pp. 1550–1572. DOI: 10.1080/09669582.2015.1047376.

8 Annexes

A1. Full list of identified scientific publications

Excel file STAR4BBS SLR Matrix (additional material)

A2. Full list of identified monitoring tools from GLR

Excel file STAR4BBS Existing Monitoring Tools Review Matrix (additional material)

A3. SLR Methodology

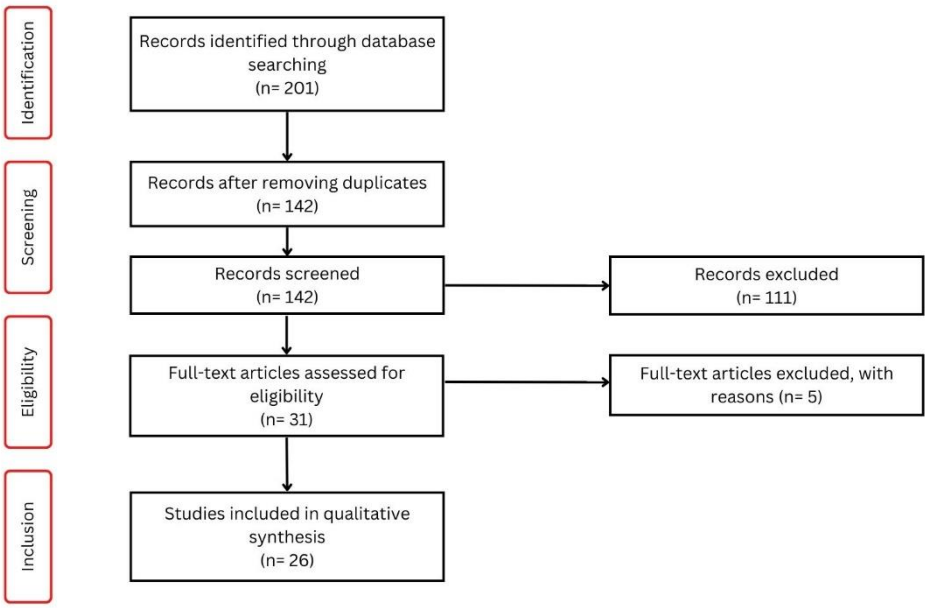


Figure 1 The PRISMA flow diagram for the review process linked to literature search on existing sustainability moni-toring systems. Source: adapted from Moher et al., 2009.

A4. List of parameters

The identified articles from the ***systematic literature review*** were categorized according to the following parameters:

- **Article Title:** title of the academic article or research paper being discussed in the matrix.
- **Authors:** individuals who wrote the article.
- **Source Title:** title of the journal or publication in which the article was published.
- **Description:** a summary of the article and its findings, and more precisely the outline of the methodology used for comparing various certifications schemes and/or labels.
- **Standards:** specific standards or certification schemes/labels being compared in the study.
- **Methodological Approach:** specific approach or methods used to conduct the research.
- **Sector:** the industry or sector being studied.

A5. Factsheets of the six selected monitoring systems

ITC Standards Map

Background

The ITC Standards Map was developed by the International Trade Center - a joint agency of the World Trade Organization and the United Nations.
It provides detailed information on the criteria, certification process, and benefits of each standard. This web-based tool allows stakeholders to discover trends on voluntary sustainability standards and it centralizes information such as the standards' content requirements, the related verification and certification mechanisms, product scope, market outreach and standards-related research results.

Scope and purpose

The Standards Map is an online platform offering a comprehensive overview of more than 230 sustainability standards and certification schemes, on environmental protection, worker and labour rights, economic development, quality and food safety, as well as business ethics. It is comprised of 1.650

baseline indicators, which allow the comparison of standards.

By encompassing a wide geographical scope, this platform promotes international collaboration and knowledge exchange in the field of sustainability. It facilitates the understanding of diverse sustainability initiatives across different regions, promotes harmonization and alignment of standards, and encourages the adoption of best practices globally.

Applicability

The Standards Map has a broad applicability across various stakeholders involved in sustainable production and trade. It serves as a valuable resource for businesses, policymakers, industry associations, certification bodies, and other actors along the value chains. The platform covers standards and schemes from both developed and developing countries, ensuring a comprehensive representation of sustainability initiatives on a global scale.

Siegelklarheit

Background

Siegelklarheit is an initiative by the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) that focuses on increasing transparency and clarity in the landscape of sustainability standards and labels. It utilizes Sustainability Standards Comparison Tool (SSCT) to provide a systematic framework for analyzing and comparing the requirements, criteria, and performance of various standards across different sectors and to promote informed choices to German consumers.

Scope and purpose

This platform evaluates labels for which there are existing evaluation requirements. Such requirements currently exist for the product groups textiles, paper, laptops & co., detergents & cleaning agents, natural stones, leather and mobile phones. Siegelklarheit evaluates not only ecological and social aspects, but also the credibility of labels.

Siegelklarheit aims to enhance accessibility and understanding of sustainability standards and labels for businesses, consumers, and other stakeholders. Its purpose is to assist users in Germany in comprehending and evaluating labels, enabling them to discern which environmental and social labels possess credibility and demonstrate ambitious sustainability commitments.

Applicability

The Siegelklarheit initiative, with the support of the GIZ Sustainability Standards Comparison Tool (SSCT), is primarily applicable to German consumers. Additionally, the SSCT integrated into the Sustainability Compass (Kompass Nachhaltigkeit) extends the applicability of Siegelklarheit to public procurers in Germany.

GSSI Global Benchmark Tool

Background

The GSSI Global Benchmark Tool is an assessment framework developed by the Global Sustainable Seafood Initiative (GSSI). It evaluates the sustainability of seafood certification schemes and standards worldwide. It offers a standardized process to assess various aspects of certification schemes, ensuring they meet rigorous criteria across social, environmental, and economic dimensions. With a focus on both wild-caught and farmed seafood, the tool contributes to the long-term sustainability of the seafood industry by promoting credible and transparent certification schemes that align with global sustainability standards.

Scope and purpose

GSSI's Global Benchmark Tool provides formal recognition of seafood certification schemes that successfully complete a transparent benchmark process. The Tool's seven-step Benchmark Process is voluntary and designed to be independent, impartial and transparent. The expert-led process involves objective assessments made against performance areas covering scheme governance, operational management (including chain of custody) and standards for aquaculture and fisheries certification.

By providing a benchmarking process that is widely recognized and accepted, the GSSI aims to promote greater transparency and credibility in the seafood industry and to drive improvements in the sustainability of seafood certification schemes and standards around the world.

Applicability

This tool serves as a standardized assessment framework that enables stakeholders in the seafood industry to evaluate and compare the sustainability performance of different certification schemes and standards. By

supporting informed decision-making, it promotes the adoption of credible and transparent practices throughout the seafood supply chain. Additionally, the GSSI provides guidance and support to schemes and standards, fostering dialogue and collaboration among stakeholders to enhance sustainability, transparency, and continuous improvement in the industry.

FEFAC Responsible Soy Benchmarking Tool

Background

The FEFAC Responsible Soy Benchmarking Tool is a comprehensive tool developed by FEFAC (European Feed Manufacturers' Federation) to assess and evaluate the sustainability of soy production and sourcing practices. It allows standard setting organizations that certify responsible soy or other programmes to benchmark their performance and criteria against FEFAC minimum requirements for responsible and conversion-free soy, as outlined in the FEFAC Soy Sourcing Guidelines.

Scope and purpose

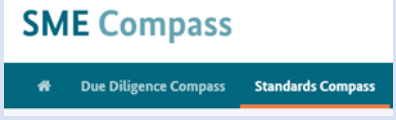
FEFAC has been supporting its members in offering animal nutrition solutions that contribute to enhancing the sustainability of livestock farming operations, considering the environmental, economic, and social aspects involved. The FEFAC tool enables stakeholders to assess key sustainability criteria related to soy production, such as deforestation, pesticide use, and social welfare, and supports sourcing soy that aligns with responsible and sustainable practices.

It is designed to promote market transparency and ensure that soy used in the European feed industry meets the required standards for responsible production.

Applicability

The FEFAC Responsible Soy Benchmarking Tool is primarily intended for feed industry stakeholders, including feed manufacturers, soy traders, and other supply chain actors involved in the sourcing and utilization of soy for animal feed production.

SME Compass



Background

SME Compass is an online platform that assists companies in enhancing the sustainability of their supply chains by establishing robust management systems. This tool is a product commissioned by the Federal Ministry for Economic Cooperation and Development, provided by the Agency for Business & Economic Development, Germany. It comprises two tools, namely Due Diligence Compass and Standards Compass. The Standards Compass utilizes up-to-date data from the Standards Map, developed by the International Trade Centre (ITC).

It offers a range of resources, instructions, and case studies to help SMEs understand and address sustainability challenges. The platform aims to empower SMEs to improve their sustainability performance, enhance competitiveness, and contribute to sustainable development.

Scope and purpose

This portal provides support to small and medium-sized enterprises to improve their due diligence processes, and to take targeted measures and comply with obligations.
Although targeted at business operations of German SMEs, this tool is equally applicable to enterprises at a larger scale, seeking to enhance their understanding of supply chain risks and fulfil their corporate due diligence responsibilities.

Applicability

The SME Compass enables small and medium-sized enterprises across different industries to identify and manage sustainability risks. The content of the SME Compass is based on the German National Action Plan on Business and Human Rights (NAP), the United Nations Guiding Principles on Business and Human Rights (UNGP) and the OECD Guidelines for Multinational Enterprises.

FSI Basket of Standards

Background

The FSI Basket of Standards, developed by the Floriculture Sustainability Initiative (FSI), is a tool designed to advance responsible sourcing practices in the floral industry, focusing on environmental and social aspects. It promotes transparency by providing clear criteria for standards and facilitates the

identification of responsible sources through independent benchmarking.

Scope and purpose

It serves as a comprehensive reference tool that allows companies and stakeholders to assess and compare the sustainability performance of different standards in areas such as environmental impact, social responsibility, and supply chain transparency. Using the FSI Basket as a reference, traders and buyers can identify suppliers with responsible practices and products. The FSI Basket of Standards aims to promote sustainability in the floral sector value chain.

Applicability

The FSI Basket of Standards is applicable to all stakeholders in the floriculture industry, including flower and plant producers, suppliers, retailers, and other relevant actors in the supply chain.

Table 1 Assessment overview according to the sectors and assessment elements

*(*VSS = Voluntary Sustainability Standards; SCS = Sustainability Certification Schemes)*

Sector	Overall assessment	Assessment elements
Forestry	- comparison - benchmark - exploring a relationship	- VSS (e.g. Plan Vivo Standard, ISO 14064:2006, CarbonFix Standard (CFS), Global Conservation Standard (GCS), etc.) - SCS (e.g. Fair Trade, FSC, etc.) - Labels (e.g. Green Gold Label S5 (GGLS5), etc.)
Construction	- comparison	- Rating systems (e.g. LEED, BREEAM, BEAM Plus, IGBC, Green Mark, Green Star, etc) - (Proposed) assessment scheme (Building Sustainability Assessment Method (BSAM))
Agriculture and floriculture	- benchmark - exploring a relationship	- SCS (e.g. Better Cotton Initiative (BCI), Fairtrade, etc.) - Labels (e.g. Green Gold Label S5 (GGLS5), etc.)

		- Programmes/Regulations (e.g. EU Organic Regulation, EHPEA-CoP, etc.)
Seafood industry	- comparison - benchmark - exploring a relationship	- SCS (e.g. MSC, Friend of the Sea, etc.) - Labels (e.g. Dolphin Safe Tuna, etc.) - Programmes (e.g. Monterey Bay Seafood Watch Program, Shrimp Improvement Program, etc.)
Energy	- comparison - benchmark - exploring a relationship	- SCS (e.g. RSB, REDcert, GGL, CSBP, IWPB, SAN, etc.) - Initiatives (Sustainable Forestry Initiative (SFI-FM), American Tree Farm System (ATFS), etc.) - Labels (e.g. Green Gold Label S5 (GGLS5), EU Organic, etc.)
Mining	- benchmark	- SCS (e.g. CTC, CFSP, etc.) - Standards (Bettercoal, Fair Stone, etc.) - Labels (Xertifix, Fairmined, etc.)

		- Programmes (Cyanide Code, iTSCi, etc.)
Metals	- comparison	- Programmes (Conflict Free Smelter Program, Green Lead Project, etc.) - Labels (Fairmined) - Standards (Fairtrade-Gold)
Textiles	- comparison	- Standards (Nordic Swan, Organic Exchange (OE), etc.)
Tourism and hospitality	- benchmark	- SCS (e.g. Bandera Azul Ecologica, Blue Flag, Premio Ecoplayas, etc.)
Chemicals **(bio-based)**	- benchmark	- SCS (ISCC, RSB, RFS2)

B2. SLR Methodological approaches

Table 1 Methodological approaches of three identified assessment groups through three key stages

Assessment type	Methodological approach	
Benchmark	Study design & data collection methods	- evaluation criteria in the standards of 3 SCS were analyzed and each requirement was matched as closely as possible to the most relevant element in the CFRN framework, based on 3 aspects (ecological, institutional, socio-economic) - design characteristics of 17 SCS assessed against the demands of important stakeholder groups; desk-research; factorial survey results served as basis for determining the criteria weights - analyzed the requirements of the 4 most dominant programs, according to the 3 criteria (pesticides and fertilizers use and GMO practices); sources: NGOs' webpages and publications available on their websites (manual coding process) - GHG emissions reductions were compared against proposed thresholds from the SCS (descriptive), if and how they met the thresholds, based on LCA studies - testing the requirements of SCSs against a set of sustainability indicators, designing questionnaire for experts for SCS effectiveness (to weight the importance of each indicator in terms of its function of measuring and enabling sustainability); online questionnaire (employing 5-point Likert scale) - a review based on scientific articles, identified through database searches; pre-assessment of the standards selected for assessment - benchmark of EU (draft) legislation with voluntary SCS frameworks - comparing SCS and assessing the indicator quality through the impact categories, using a standardized rating scale to evaluate the indicators

		- developing a framework for evaluation of the standards relative to each other using 4 substantive criteria; developing a set of criteria against which each standard's principles and criteria, the level of performance required, and the degree of overlap between standards were compared and evaluated
	Data analysis	- categorization analysis - multi-criteria decision analysis (MCDA); multi-attribute value theory (MAVT); the transformation of the performance scores was based on the quality of the design characteristics; survey data analyzed in a hierarchical multivariate linear regression analysis (using the restricted maximum likelihood approach (REML) and R package lme4) - detailed content analysis (each individual criterion of the SCS compared to the NGO demands) - comparative analysis; ANCOVA and ANOVA statistical tests to identify process and life cycle modeling factors that contributed significantly to environmental metrics - two-dimensional analytical method - breadth–depth graphical analysis: multiple analysis of variance (MANOVA) and cluster analysis (assessed how the aggregation of factors into impact areas may influence interpretation of results) + breadth–depth (B–D) analysis - standards were individually compared with the benchmark standard and for each benchmark criterion it was determined whether a specific standard was compliant or not - using a standardized rating scale to evaluate the indicators (composite scale: indicators aggregated by impact categories and arithmetic mean of all indicators per SCS is computed, per environmental impact category and the indicator subcategories respectively); the standard error of the mean (SEM) was calculated to assess the uncertainty of the arithmetic mean; internal consistency: excluding indicators that did not fall into composite scales - assessment criteria were developed on the basis of the WWF forest carbon standards assessment guide for forest carbon standards and projects, and complemented by the review of the standard schemes
	Representation & interpretation	- the schemes were classified as treating each element "completely," "partially," or "not at all" - direct rating method: the best and worst criteria values are chosen and rated with 1 and 0 respectively, results of the MCDA analysis presented using radar charts to show the performance of

		different sustainability certification schemes against the set of criteria - color-coding table depicting the extent to which each of the programs fulfilled the NGOs' demands, using the classifications: explicitly (green), addressed without obligatory implementation (yellow), only indirectly (orange), not at all addressed (red) - the breadth and depth scores for each factor were presented graphically; the 112 factors discussed in the guidelines were divided into 5 impact issues (community, environment, food safety, feed and marine resource use, and supply risk); each component factor addressed by the standard was evaluated using a seven-point scale to estimate relative rigor - scoring chart: requirements corresponding with indicators were assigned value of (1), while those that did not were assigned a value of (0) - stringency of a given standard was obtained by summing the compliance values given for all seven principles, for that standard. Standards reaching a sum of 10 or more were classified as Stringent, and standards reaching a sum of 6 or less were classified as Unstringent. - the benchmark standard was developed using 7 principles under which 26 criteria were defined and sorted; selected standards were individually compared with the benchmark standard, determining compliance for each criterion (the degree to which the standard considered the principle (0: principle disregarded; 1: principle considered in part; 2: principle fully considered) - color-coding and +/- table: green color (+) indicated considered; yellow color (+/−) indicated partly considered; orange color (−) indicated disregarded; other tables with principles, criteria and SCS standards, with marking - '•', compliance; bar charts: presenting the percentage of benchmark criteria complied with by the standard types - rating the individual indicators for feasibility in 3 requirement subcategories and for reliability in 4 requirement subcategories (feasibility: by data and qualification requirements, and clearly defined thresholds; reliability: by validity of indicators and response time) - the actual number of indicators was counted for each of the components and rated on a 3-step scale based on thirds - the certification scheme with the highest number of causal links had the best rating, i.e., 100%, and was used as a benchmark and rated as done for the indicators; analyzed ratings of certification scheme indicators were at the less aggregated level of composite scales and

		grouping them by the indicator requirements and their subcategories: visualizing different degrees of environmental performance - relative ordinal ranking system that assigns a value between '0' and '2' to each of the six criteria: the assignment of a grade to each criterion indicated the level of standards comprehensiveness relative to the other standards being evaluated, not a specific benchmark. a grade of '0' indicated that the standard *does not* provide any regulation on the respective assessment criteria, '2' indicated that the criteria was addressed comprehensively compared to other standards. The value '1' indicated that the standards only partly provided guidance on the criteria; hexagon of criteria
Comparison	Study design & data collection methods	- comparing the relative temporal flow quantitative changes in the categories of criteria and indicators in the standards of the selected systems; qualitative review of the respective systems' institutional histories; tracking change and continuity in text and concept frames within standards of the selected 5 different global forest institutions, and comparing framing of the standards using generic categories - comparing newly developed green certification tool (*Diana*) for rating the "green" practices, with existing benchmark rating tools (in terms of rating categories); questionnaire aimed at collecting informed expert feedback - comparing principles, criteria, indicators of one scheme with four other SCS (consistency with principles) - structured questionnaire survey and interview; literature review (content analysis approach); experts' consultations and surveys - comparing sustainability requirements and evaluating the comprehensiveness of national support schemes (country-wise) - combined analysis of findings from two related but separate studies - expert survey and scholarly literature review - benchmarking the strengths and characteristics of different schemes; a comprehensive review and comparison of the issues and metrics of five representative assessment schemes; assessment aspects of the *SBTool 2010* generally adopted as a common basis for the comparison

	Data analysis	- 'discursive-institutional' approach; arrangement analysis (PAA) approach; 'discourse as frames' approach; - data collection utilizing focus groups (facilitated local expert discussions); deployment of the Delphi technique and the Analytic Hierarchy Process; defining the level-wise aspects of the system using pilot study, observations, and interviews; pair-wise comparison; adaptive weighting system; AHP multilevel hierarchical structure - qualitative analysis; descriptive comparison; establishing thresholds for a minimum acceptance of a requirement - credit points (score-weighting) and determining significance of the key sustainability criteria - analysis of the sustainability approaches; comprehensive criteria adopted as the baseline for the comparison of the national sustainability requirements - comparative cross-case analysis - text analysis; comparing criteria within 3 sustainability pillars - descriptive comparison - statistical analysis: weights and ranks of weights allocated to five key assessment aspects; checking if they cover key assessment aspects with the tool (correlation between the number of criteria/sub-criteria and the number of key assessment aspects)
	Representation & interpretation	- graphical representation of the observed changes (Sankey diagrams - for graphically depicting temporal flow changes) - tabular and spider diagram - ranking of the criteria; classification of sub-attributes: "required," "optional," or "negligible"; radar diagrams and comparison charts; pie chart: percentage contribution ratio of each sustainability indicator to the overall sustainability performance; six-grade certification system - tabular comparison

Exploring a relationship	Study design & data collection methods	- identifying key mismatches in existing certification and ratings schemes - questionnaire survey (to identify the reasons for obtaining certification, the choice between two SCS, and the effects of certification on the enterprises) - structured and semi-structured questionnaires (adapted from an audit check sheet); open interviews with key informants (supplemented with personal observations) - survey data - semi-structured interviews and household surveys
	Data analysis	- reviewing the potential for jurisdictional approaches; descriptive comparison - estimating propensity score: generalized propensity score matching (with multiple treatments); average treatment effects; random utility framework - descriptive case study approach
	Representation & interpretation	- tabular comparison; scoring chart based on number of responses from the questionnaire - descriptive comparison - tabular summary statistics; graph representing cumulative distribution

C1. Selection of the best-fitted monitoring tools for an in-depth analysis

Table 1 Selection assessment of the monitoring tools

Monitoring Systems	Coverage	Comprehensiveness	Interpretation Variety	Audience Variety	Operationality
Comparative Matrix of Forest Certification Schemes	1	0	1	1	0
STAR ProBio SCT (Sustainability Certification Tools)	1	1	0	0	1
Certification Assessment Tool (CAT)	2	1	1	1	1
ADVANCEFUEL	1	2	0	1	1
ITC Standards Map	2	1	2	1	2
SSCI Benchmark and Recognition	2	0	1	1	2
GSSI Global Benchmark Tool	1	2	1	1	2
GIZ Sustainability Standards Comparison Tool (SSCT) and Siegelklarheit	2	1	2	1	2
IDH SIFAV basket of standards	1	1	1	1	2
OECD MNE Guidelines benchmarking	0	0	0	0	2
Seafood Stewardship Index	1	1	1	0	2

SAFA Tool	1	1	2	1	1
FEFAC Responsible Soy Benchmarking Tool	1	1	2	1	2
Blue Angel Comparison of Certification Systems	1	2	1	1	1
SME Compass	2	1	2	0	2
Labelinfo.ch	1	1	1	1	1
SOJA (Soja Netzwerk Schweiz)	1	1	1	1	1
FSI Basket of Standards	2	1	2	0	2
Textile Exchange Corporate Fiber & Materials Benchmark (CFMB) and MCI	0	1	1	0	2

Star4bbs

"Sustainable bio-based systems via effective certification & labelling"

Consortium:

Funded by the European Union

www.star4bbs.eu
info@star4bbs.eu

@STAR4BBS