

Assessing and comparing carbon credit rating agencies

Report

Ingrid Wawrzynowicz, Matthias Krey, Ximena Samaniego

Funded by Carbon Market Watch

All views expressed are solely those of the authors

FREIBURG, GERMANY, 11.09.2023



Acronyms

| | |
|-------|--|
| ARR | Afforestation, Reforestation and Revegetation |
| APD | Avoided Planned Deforestation |
| AUD | Avoided Unplanned Deforestation |
| CAR | Climate Action Reserve |
| CCQI | Carbon Credit Quality Initiative |
| CDM | Clean Development Mechanism |
| CER | Certified Emission Reduction |
| ESA | European Space Agency |
| ESMA | European Securities and Markets Authority |
| GEDI | Global Ecosystem Dynamics Investigation |
| GFCC | Global Forest Canopy Cover |
| IFM | Improved Forest Management |
| IOSCO | International Organization of Securities Commissions |
| LEDS | Low Emission Development Strategies |
| LiDAR | Light imaging, detection and ranging |
| NbS | Nature-based Solutions |
| NDCs | Nationally Determined Contributions |
| REC | Renewable Energy Certificate |
| REDD | Reducing Emissions from Deforestation and Forest Degradation |
| SD | Sustainable Development |

Contents

| | |
|--|-----------|
| 1. Key findings and recommendations | 1 |
| 2. Overview | 3 |
| 2.1 Background | 3 |
| 2.2 Objective, scope and structure..... | 4 |
| 3. Comparison of carbon credit rating agencies | 7 |
| 3.1 Claims of selected carbon credit rating agencies | 7 |
| 3.2 Key company characteristics and number of rated projects of the selected agencies | 8 |
| 3.3 Comparison of Governance, Monitoring, Framework Transparency | 10 |
| 3.4 Comparison of Frameworks | 12 |
| 3.4.1 COMPARISON OF ENVIRONMENTAL INTEGRITY CRITERIA | 20 |
| 3.5 Overall rating of cross-sectoral approach and recommendations for improvement | 28 |
| 4. Deep Dive – REDD AD | 40 |
| 4.1 Comparison of REDD AD approach | 40 |
| 4.2 Overall rating of REDD AD approach and recommendations for improvement | 50 |
| 4.3 Comparison of overlapping ratings | 55 |
| 5. Conclusions | 58 |
| 6. References | 62 |

Tables

| | |
|--|-----------|
| Table 1: Overview of scoring methodology | 13 |
| Table 2: CCQI Benchmarking – Baseline and Project Scenario analysis..... | 21 |
| Table 3: Comparison of other cross-sectoral baseline and project scenario tests performed by agencies | 22 |
| Table 4: CCQI Benchmarking - Non-Permanence | 22 |
| Table 5: Comparison of other cross-sectoral non-permanence tests performed by agencies. | 23 |
| Table 6: CCQI Benchmarking – Leakage..... | 24 |
| Table 7: Comparison of other cross-sectoral leakage tests performed by agencies | 25 |

| | |
|---|----|
| Table 8: CCQI Benchmarking – Additionality | 26 |
| Table 9: Comparison of other cross-sectoral additionality tests performed by agencies | 27 |
| Table 10: Comparison of cross-sectoral co-benefit and safeguards tests performed by agencies | 28 |
| Table 11: Rating categories and requirements for cross-sectoral approach | 31 |
| Table 12: Overall rating of cross-sectoral approach | 33 |
| Table 13: REDD AD- Baseline and Project scenario approach | 41 |
| Table 14: REDD AD - Non-Permanence approach | 43 |
| Table 15: REDD AD Leakage approach | 46 |
| Table 16: REDD AD Additionality approach | 48 |
| Table 17: Rating categories and requirements for REDD AD approach | 50 |
| Table 18: Overall rating – REDD AD approach..... | 52 |
| Table 19: BeZero's, Calyx's and Sylvera's scales ranked from high to low | 56 |
| Table 20 Overlapping ratings compared on normative scale | 56 |
| Table 21: Overview of rating of the key differences identified for general and REDD AD approach | 59 |

Figures

| | |
|---|----|
| Figure 1: Rating for alignment with CCQI and other tests | 29 |
| Figure 2: Range of limitation of overall rating if additionality is assigned the lowest score ... | 30 |

1. Key findings and recommendations

- The **lack of a universally accepted and in-depth definition of carbon credit quality** is a fundamental issue for the carbon market, as market actors increasingly treat carbon credits with caution.
- The carbon credit rating agencies (BeZero, Calyx, Renoster and Sylvera) assessed in this report aim to address this lack of standardisation by **distinguishing between robust carbon credits and those not delivering on their promises**. The agencies claim to increase transparency, mitigate reputational risk and enable fair pricing.
- Renoster's approach differs the most from other agencies in that the company aims to **limit its qualitative assessment** and instead derive its ratings as much as possible based on an **algorithmic analysis**. Further, Renoster **assesses leakage but excludes it from their overall rating score**.
- Approaches to rate carbon credits differ in many ways across the agencies. For example, most agencies use direct tests for their assessments, while **Renoster's tests are mostly implicit**¹.
- **Other key differences** identified include approaches to limit the overall score of a project based on its additionality score. This is done differently across agencies. Other examples include the assessment of double-issuance and double counting, co-benefits and safeguards, leakage, buffer strength, permanence benchmarks and rating transparency.
- According to the findings, **Calyx has the most stringent cross-sectoral as well as REDD AD approach**.

¹ Implicit in this context means that a test is not directly evaluating a specific criterion by applying tools and methods used under carbon-crediting programmes and the respective methodologies. For example, policy/regulatory additionality for REDD+ projects is not tested by qualitative analysis of the policies and regulations influencing deforestation. Instead, it is implicitly tested for by choosing a reference area in the same municipality as the project area, assuming that the deforestation rate of the reference area must reflect any relevant regulations and policies.

- However, the comparison is conducted at a high-level and needs to be regarded with caution due to strong limitations, such as the exclusion of the assessment and comparison of:
 - Individual models (e.g. statistical models, machine learning), handling of the model or interpretation of its results
 - Quantitative assessments/Calculations
 - Data input (e.g. remote sensing data, sampling data etc.)
 - Assessment reports of multiple overlapping ratings (only compared at a high-level)
 - Undisclosed information

In addition, only a deep dive on REDD AD approaches is provided. Other sector-specific approaches are not discussed.

Further, the agencies state that a thorough understanding of the frameworks usually takes users/customers a significant amount of time to comprehend and fully grasp the scope and depth of the assessment process (multiple days or weeks according to several agencies).

- Improvements for carbon credit rating agencies approaches could be derived at a high-level for key areas such as: governance, transparency, review processes and project-level assessments.
- **On the way forward:** More alignment and oversight are needed for carbon credit rating agencies. Guidelines for **financial credit rating agencies could serve as a blueprint**. Most importantly, leakage must be included in the overall score as well as mitigation activities to avoid leakage and reversals.

Further, more methodological (i.e. framework) transparency is strongly recommended. A complete understanding and comparison of how the selected ratings agencies arrive at their ratings is not possible based on the information publicly available.

2. Overview

2.1 Background

Carbon markets² have been established as a way for developed countries to meet their UN climate targets by being able to buy emission reductions at a lower cost than domestic abatement. The most prominent carbon crediting schemes, the Clean Development Mechanism (CDM) regulated by the UN which started in 2001 and the industry self-regulated carbon-crediting programmes under the Voluntary Carbon Market (VCM), started around a similar time.

Since their emergence carbon crediting schemes have been criticised for being overly complex, non-transparent and lacking equity. This, coupled with the ongoing debate about the efficacy of carbon crediting in reducing emissions, has led to a significant evolution in the past two decades. In particular, the lack of environmental integrity (i.e. additionality, baseline setting, monitoring, permanence, avoidance of leakage and double counting) has been criticised (Betz et al. 2022).

In 2011, carbon markets crashed, highlighting the importance of sufficient and reliable demand, which is partly influenced by the quality and credibility of supply, as well as policy certainty for functioning carbon markets. Particularly, the lack of additionality of registered activities and the lack of sustainable development (SD) co-benefits resulted in bans against the use of certified emission reductions (CERs) from the Clean Development Mechanism (CDM) in the EU Emissions Trading Scheme and consequently, in severe lack of demand (Hoch et al. 2020, Michaelowa et al. 2021).

Carbon credit programmes are challenged with finding a balance between ensuring environmental integrity of mitigation outcomes while keeping transaction costs low. In January 2023, findings of a nine-month investigation by SourceMaterial, The Guardian and Die Zeit regarding Verra's REDD+ projects were published. The overall outcome of the study was that more than 90% of Verra's REDD+ credits (out of 29 of the 87 forest conservation projects certified by Verra) do not represent genuine carbon reductions (The Guardian 2022). This publication was followed by an 85% price decrease for such credits (Carbon Pulse 2023b).

As of today, there is no universally accepted standardised carbon credit quality assessment methodology, making it difficult for buyers to compare carbon credits and enable appropriate pricing. The need for oversight and transparency has been recognized and gave reasoning to several initiatives, such as the Integrity Council for the Voluntary Carbon Market (IC-VCM), to fill this gap by defining minimum quality criteria (Core Carbon Principles) and carbon credit attributes that carbon crediting programmes can choose to comply with and be assessed against (IC-VCM 2023).

² In this report the term carbon markets is used for supply of and demand for credits from carbon crediting programmes that apply a baseline-and-credit approach.

Against this backdrop, carbon credit rating agencies have recently stepped into the market to assess carbon credit quality at project-level. Their ratings aim to distinguish between robust carbon credits and those that do not deliver on their stated benefits. This gives them the possibility to strongly influence the market. Moreover, the agencies claim to increase transparency in the market, while ensuring independence and thereby enabling fair pricing (BeZero 2022b, Clean Technica 2022, Quantum Commodity Intelligence 2022, Sylvera 2022d).

2.2 Objective, scope and structure

Objective

The objective of this report is to qualitatively assess the robustness and reliability of evaluations of **carbon credit quality** (i.e. the likelihood that reported mitigation outcomes indeed reflect the stated volume of emission reductions or removals achieved by the respective project) by selected carbon credit rating agencies, at a high-level (see [Out of Scope](#)). The following chapters aim to address the question of whether the selected rating agencies, within the scope of the analysis, deliver on their claims (see [chapter 3.1](#)). Further, this study aims to identify key differences in the rating assessment approaches of the different agencies and to provide recommendations for improvement of their respective approaches.

Scope and structure

To achieve the objective of this report, an assessment of how agencies determine their rating is necessary and will be based on an evaluation in the following aspects: governance, monitoring, transparency (see [chapter 3.3](#)) and frameworks³ (see [chapter 3.4](#)). The ratings agencies included in this report are: BeZero, Calyx, Renoster and Sylvera. The agencies have been chosen based on their public visibility.

First, the report will provide a **general overview of each rating agency's assessment approach** (see [chapter 3.5](#)). Since approaches by agencies differ depending on the project type, this comparison is only conducted at a high-level basis across all sectors⁴. As an entry point, the general assessment approach of each agency is **benchmarked against the requirements set by the Carbon Credit Quality Initiative** (CCQI 2022). CCQI is chosen as a benchmark due to the initiative's independence as well as the public availability and comprehensiveness of their methodology. Despite the CCQI methodology being a top-down assessment and most criteria being directed at carbon crediting programme and methodology level, rather than project level the CCQI is the best option for a

³ A Framework in this context refers to a set of requirements against which a project is evaluated. All agencies use at least one framework for their analysis.

⁴ Sector in this context refers to the overarching activity type, e.g. Forestry, Energy, Waste etc.

benchmark due to the absence of other methodologies with similar blueprint qualities. The criteria for the assessment conducted for this report (based on the [Objective](#)) are two-fold: First, criteria relevant for evaluating whether the issued volume of carbon credit reflects the actual volume of emission reductions and removals achieved by the respective project which consist of **baseline and project scenario, permanence and leakage are assessed**. Secondly, **Additionality** is assessed for evaluating whether these emission reductions and removals are in fact additional, meaning that the emission reductions or removals would not have occurred in the absence of the carbon market incentive.

As a first layer of the comparison, CCQI criteria are selected for the assessment (i.e. blue tables in [chapter 3.4.1](#)). CCQI tests solely aimed at the carbon crediting programme and methodology level and not transferable to the (cross-sectoral) project level are excluded. This includes for example, the assessment of a crediting programme's governance. Further, aspects not related to baseline and project scenarios, permanence, leakage, and additionality are also excluded. Other tests, such as whether legal requirements have been appropriately considered or whether carbon credits were considered before the decision to proceed with the project, can be equally used for project-level and have, therefore, been included.

The second layer (i.e. orange tables in [chapter 3.4.1](#)) compares **any other cross-sectoral test performed by the agencies**. Since the CCQI approach is directed at carbon-crediting programme and methodology-level, the agencies' frameworks in contrast stipulate additional, more project-level focused, tests, which are compared in this layer. This section also includes a brief comparison of the co-benefits and safeguards ratings. Key similarities and differences of the agencies' assessment approaches will be elaborated and an **overall rating for the cross-sectoral approach** is provided based on the results of the analysis. Co-benefits are not a focus topic for this report and are only included in brackets in the overall score, since they do not impact the mitigation outcome. In addition, they are either handled as a separate score by the rating agencies (i.e. Calyx, Sylvera) or not scored at all but instead information on their assessment is provided (i.e. BeZero, Renoster).

For a more in-depth assessment of a project-type specific approach, a **deep dive on Reducing Emissions from Deforestation and Forest Degradation (REDD) - avoided deforestation (AD) projects is conducted** (see [chapter 4](#)). The tests applied by the agencies for REDD AD projects, according to the available information, are summarised at a high-level and compared in this section. Key similarities and differences are explored. [Chapter 4.2](#) provides an **overall rating of the REDD AD approach of the agencies and recommendations** on how to improve the respective approaches.

In [chapter 4.3](#), a high-level comparison of multiple **overlapping ratings** is conducted. Finally, a **conclusion of the assessment is given in chapter 5**, summarising the key findings of the report and recommendations for improving carbon credit rating agencies frameworks and their approaches to rating REDD AD projects.

The information sources used for the report are publicly available frameworks and whitepapers by the respective agencies as well as information shared by the agencies via interviews (1h per agency) and in written form (i.e. assessment reports, information available on customers platform, questionnaires, e-mail exchanges).

Review by agencies

Each agency was provided with the option to review the draft version of this report. Commented draft reports were received from three of the four agencies selected for this report (i.e. BeZero, Calyx and Sylvera). Renoster provided a few general comments via email but did not share a commented version of the report.

The agencies were given a two-week timeframe to make amendments to their frameworks or other information on the agencies' websites. The cut-off date for implemented changes to be included in the report was the **19th of June**. This option has been utilised by BeZero and Calyx. BeZero published a new document explaining how additionality is a limiting factor in its rating framework (BeZero 2023), while Calyx published multiple new detailed information on their ratings approach (Calyx 2023a,c,d).

A second partial review was offered on the 7th of July. Agencies could provide their comments on the figure and updated tables in chapter [3.5](#), [4](#), [4.1](#) and [4.2](#) until the 12th of July. No comments from Renoster were received.

Out of scope

The report does not cover an assessment and comparison of:

- Individual models (e.g. statistical models, machine learning) used by rating agencies.
It is important to highlight that models and machine learning used for assessment of a carbon credit project do have an impact on the rating, although they are in general not the main drivers. This is the case when models are used to assess measurement errors. However, where they are used for baseline assessment of avoided deforestation (REDD) projects, models can have a significant impact on the rating.
- Quantitative assessments/Calculations.
Methods for calculating carbon stock, leakage, baseline etc. significantly influence each rating and were excluded from the report.
- The datasets used for models and other calculations.
Datasets used by rating agencies to rate projects differ for project types, location, and practice. An overarching assessment of the robustness of datasets across all sectors or on the individual sector level is therefore not possible. However, using inappropriate datasets poses a significant risk to the robustness of a rating, especially for forestry projects (Nomura et al. 2019).
- Assessment reports of overlapping ratings.

While the report provides a high-level comparison of overlapping ratings, it does not include any justification by the agencies for the respective ratings. This is not public information, and not all agencies agreed to make it public. To fully understand how identified differences in assessment approaches influence the final ratings, a thorough analysis of assessments reports of multiple overlapping ratings (ratings of the same projects) is required.

- Undisclosed information

Information not shared due to the protection of intellectual property or other reasons is excluded from the report. Further, information shared upon request and not explicitly mentioned in the respective frameworks cannot be verified.

- Information published after the 19.06.2023.

The 19.06.2023 was set as a cut-off date for new publications to be analysed, assessed, and included in the report (see [Review by agencies](#)).

Further, not every sector-specific approach by the selected agencies is covered on an in-depth level – it focuses instead on a cross-sectoral comparison, combined with one specific deep-dive on REDD AD. The cross-sectoral approach comparison will be conducted as a high-level analysis of aspects assessed for the rating. A deep-dive is provided for the REDD AD approach, since this is the only sector covered by all agencies and more information is available on REDD AD compared to other forestry project-types. It should be noted that the sector coverage between agencies differs strongly (see [Sector coverage](#)) and we did not carry out a detailed assessment of the rating agencies' approach for each individual project type.

It is important to consider the limitations posed by excluding these elements from the comparison of carbon credit rating agencies' approaches. Moreover, the authors have been informed by the agencies that a thorough understanding of the frameworks usually takes a significant amount of time (multiple days or weeks according to several agencies) of explanation and discussion with the agency to grasp the full scope and depth of the assessment process.

3. Comparison of carbon credit rating agencies

3.1 Claims of selected carbon credit rating agencies

Carbon credit quality is a complex term that can either exclude or include sustainable development and safeguards or attach different weights to given criteria. However, all carbon credit rating agencies assessed in this report aim to provide a post-issuance assessment on whether a project's reported mitigation outcomes (emission reductions or removals) indeed reflect real emission reductions or removals. This is tested with two overarching checks by the rating agencies. One check tests if the issued volume of carbon credit reflects the actual volume of mitigation outcome achieved by the respective project. This requires the volume to represent permanent, leakage deducted mitigation outcomes quantified in a robust manner based on sufficient available and accessible data. Another check is focused on whether these mitigation outcomes are in fact additional.

The service carbon credit rating agencies provide is to assess the claims made by carbon crediting programmes, auditors, and verifiers to establish whether a verified project does in fact achieve the reported mitigation impact and proposes its ratings as a tool to verify this claim. The rating agencies' evaluations are based on the aggregation of individual assessments of criteria such as robust baseline scenario and project emissions, additionality, permanence, and leakage avoidance. The key reason for the demand for rating agencies' services are the observed shortcomings in the market and the lack of quality assurance, which create confusion over carbon credit quality. Moreover, recent controversies focus on the alignment of incentives for various carbon market actors to maximise credit issuance rather than emission reductions (The Guardian 2023). The underlying problem is often summarised by rating agencies as a "tonne is not a tonne", meaning that in the agencies' views not every carbon credit does represent a mitigation outcome equivalent to a tCO_{2e}. This means that rating agencies work under the premises that either the methodology and/or the data and parameters used by the project are not suitable for calculating the real emission reductions or removals resulting from the project. Hence, their rating, if sufficiently robust, can promote improvements for carbon-crediting programmes, methodologies, the carbon crediting process as well as project development, which helps foster integrity on the supply side.

Rating agencies further aim to enable companies that lack the needed expert knowledge and resources to make an informed decision on which credits they should and should not be investing in (e.g. to manage reputational risk). However, in a recent webinar hosted by Ecosystem Marketplace, a survey based on responses from 44 organisations found that some buyers (number was not disclosed) felt that rating agencies were adding to the overall confusion in the carbon market (Carbon Pulse 2023c).

While the overarching claim is broadly the same, the agencies frame their service differently in the respective frameworks and whitepapers. BeZero claims that "a BeZero Carbon Rating is an informed opinion on the quality of the carbon credit, in terms of its likelihood to achieve 1 tonne of CO_{2e} avoidance or removal" (BeZero 2022a). Sylvera uses a very similar wording by explaining that their ratings' "primary function is to assess the likelihood that the claimed Greenhouse Gases (GHGs) have been avoided or removed" (Sylvera n.d.a.). Calyx's GHG ratings reflect the "risk that carbon credits do not meet their claims of reducing or removing the equivalent of one metric tonne of CO₂" (Calyx Global 2023a). In contrast to the other agencies, Renoster defines their score as "the ratio of genuine carbon removals" (Renoster n.d.b).

3.2 Key company characteristics and number of rated projects of the selected agencies

In the following chapter general information on overarching characteristics of the selected agencies are presented.

Business models and rating accessibility

All agencies are for-profit private companies that provide their rating as a subscription service to their customers. Main customers are carbon credit traders, online marketplaces for carbon credits and corporate sustainability departments. In addition, an increasing number of companies retrieve the scores through intermediaries⁵ selling carbon credits (Wall Street Journal 2023).

In terms of public availability of ratings, several differences were observed. BeZero's ratings are all publicly available on their website, including a summary of the justification for each rating. Customers are offered more detailed information on the ratings, for example via BeZero's platform. Marketplaces and exchanges can also host the ratings (Financial IT 2022). Some of Calyx's and Sylvera's ratings are available on Net Zero Marketplace, though the rating scales presented on Net Zero Marketplace are reduced compared to those used by the agencies in their original assessments (see [Table 19](#)). The Net Zero marketplace only lists some of Calyx's and Sylvera's ratings, while the rest are not publicly available. None of the reasoning behind the ratings (i.e. assessment reports) from both of these agencies are public. On Renoster's website, only ratings older than 6-12 months are publicly available. Only customers have access to ratings in this 6-12 month timespan. Published ratings are accompanied by a 30-45 min video, explaining in detail how the rating is composed, paired with an assessment report (Renoster 2022).

Additionally, all agencies aim to reflect on project quality by assessing a project's co-benefits. The co-benefit ratings of Calyx and Sylvera are available for the projects listed on Net Zero Marketplace. Renoster's co-benefit rating is not reflected in a score but rather a qualitative summary, which is included in a project's assessment report and brief rating summary on their website. BeZero follows a similar approach and does not include co-benefits in their score but provides information on their assessments via their customer platform. Hence, all agencies treat co-benefits as a separate assessment, not accumulated in the overall score of a project. Renoster's ratings are focused on biodiversity and community impacts, while Calyx and Sylvera's rating encompasses all Sustainable Development Goals (SDGs), which includes biodiversity and community impacts (Calyx 2023a, Renoster n.d.b., Sylvera n.d.). BeZero states upon request that their assessment is also focused on all SDGs, as opposed to only biodiversity and community impacts.

Staffing

Sylvera employs the largest team of 152 members in total with expertise ranging from earth observation, climate policy, machine learning, nature conservation, geospatial analysis, data science and engineering (LinkedIn 2023a). BeZero employs 113 experts with expertise spread across climate and earth sciences, financial research, data and technology, machine learning, remote sensing, industrial engineering, and public policy (LinkedIn 2023b). The number of Calyx staff members is not available on their website but 21 people are listed as affiliated on LinkedIn (LinkedIn 2023c). Among

⁵ Brokers, retailers, or exchanges of carbon credits.

Calyx's team are GIS and natural resource experts, SDG and GHG Analysts as well as other scientists. Renoster has by far the smallest team, currently consisting of 8 team members, ranging from data scientist and engineers to environmental and remote sensing experts (Renoster n.d.a). It should be noted that staff numbers on LinkedIn might not reflect the actual number of staff employed but give a broad indication.

Sector coverage

BeZero covers the widest variety of sectors, including Energy, Household Devices, Industrial Processes, Nature-based Solutions (NbS) as well as Tech Solutions (i.e. Biochar). Sylvera is focused on Energy as well as NbS. Under Sylvera, NbS is limited to forestry projects (i.e. ARR, IFM and REDD+). Calyx covers NbS (Forest & Land), Household and Community, Manufacturing and Industry, Waste and Renewable Energy sectors (Calyx Global 2023e). The range for Energy project types is much larger at Calyx than at Sylvera. In contrast, Renoster is solely focused on Forestry projects (i.e. ARR, IFM and REDD+) (BeZero 2023c, Calyx Global n.d.b., Renoster n.d.b., Sylvera n.d.)

Number of published ratings

Since their founding year, Calyx has rated over 370 projects (Calyx Global n.d.c) (founded in 2021), BeZero over 315 projects (BeZero 2023c) (founded in 2020) and Sylvera states the company rated over 150 (founded in 2020) (status: 15.06.2023). Renoster is the youngest of the rating agencies and has only rated 10 projects of which 5 are currently publicly available (Renoster 2023), while another 130 projects have been rated at a lower level (i.e. focus on documentation and remote sensing solely), but are not publicly available.

3.3 Comparison of Governance, Monitoring, and Framework Transparency

This chapter presents a comparison of main similarities and key differences, regarding Governance, Monitoring, Transparency and Frameworks used for the rating process of each agency.

Governance - Conflict of interest (Col) policies

There is **no conflict of interest evident for any of the agencies** based on publicly available information. All agencies provide safeguards to mitigate risk for Col. BeZero for example, claims on their website that they are free of any conflict, stating that BeZero is neither selling, nor owning or recommending carbon credits. Further, the company is not developing or advising on the development of carbon credit projects and is not receiving any financial benefit linked to the rating assigned to a project and its credits (BeZero 2022b). Calyx claims to not accept financial incentives from carbon crediting programmes, methodology developers, or project developers. Since the company assesses carbon crediting programmes as well as it rates methodologies and projects, this would otherwise result in a conflict of interest, according to Calyx. In addition, the company has a Col policy that is applied to their GHG and SDG panels. Renoster states in their questionnaire that they do not rely on or consult

with project developers or registry bodies during their review process. They claim that their reviews are entirely independent and solely based on the data that projects are required to share publicly. Further, Renoster states that their assessments are, as much as possible, algorithmic, and repeatable. Reviewers, project developers, verifiers, and clients are not involved and cannot influence the outcome. Besides not selling carbon credits and not accepting payments from verifiers, registries, or developers Sylvera applies a code of conduct (Sylvera 2023). Under point 3 of this code of conduct, the company specifies 14 rules applied to avoid conflict of interest, such as the requirement of rating personnel to disclose any relationship that may create a conflict of interest (Sylvera 2022a).

Review process

BeZero monitors all latest information pertaining to the project, sector, and methodology. Data sources include new research (e.g. alerts from satellite monitoring) and new project documents (e.g. new monitoring reports, changes in regulations, changes in methodology), as well as any other information deemed relevant to the project rating. BeZero explains that monitoring is applied at minimum annually but is also triggered by events (e.g. loss events). The agency further states that they apply a “ratings per vintage” approach.

Calyx follows a similar path by monitoring all new and relevant information (e.g. regulations, incentive schemes, penetration rates, reversal risks etc.) at regular frequencies (not specified in public documents) that are determined by each element. Calyx decides an appropriate frequency for monitoring specific topics, such as regulations that may be reviewed annually. The company further explained that Calyx has moved to “ratings per vintage” for many of their project types due to the variety of changes that may occur (e.g. new policies/regulations affecting additionality).

Renoster stated during the interview that they do not currently have a monitoring system set up, because they started their services in July 2022 and their first ratings are only 6 months old. The company is currently working on a monitoring system that would flag projects that have experienced deforestation or have a new monitoring report. This would then trigger a review.

Sylvera’s rated projects are monitored at least once in any twelve-month period, but the company aims to arrive at a quarterly basis, the company explained during the interview. Monitoring is also triggered after a significant event.

All companies monitor all relevant information, including carbon stock changes, media, and press releases as well as policy updates and new project documentation.

BeZero, Renoster and Sylvera use a fully internal rating and **rating review process**, while Calyx involves external partners for various tasks, such as supporting the drafting of assessment frameworks, providing expert reviews for ratings as well as taking part in their GHG or SDG panels (BeZero 2022a, Calyx Global n.d.e, Renoster n.d.b, Sylvera n.d.). These panels, consisting of experts from Carbon Limits, Climate Focus, Climate Law & Policy, INFRAS and the Stockholm Environment Institute, are

part of Calyx's governance and endorse their rating frameworks (Calyx Global n.d.d). It should be noted that Calyx's rating review is either conducted by a senior staff member with specific expertise or subject matter experts from their partner organisations.

The **framework review process** of BeZero and Renoster are fully internal. In contrast, **Calyx and Sylvera's frameworks are peer-reviewed**, which increases independence. Sylvera's Framework Review Committee comprises representatives of industry bodies and multilateral institutions, carbon crediting programmes bodies, project developers, technical and scientific experts, exchanges and marketplace operators, financial institutions as well as buyers and retirees of carbon credits. Calyx's framework is reviewed by a subject matter expert of the respective external panel (GHG or SDG) (Calyx Global 2023b). Once a relevant change is made to the framework, a review of all past ratings is triggered for all agencies.

Framework Transparency

Transparency is differentiated between information available to the public and that which is only available to customers. All agencies provide further details on their frameworks to customers if requested but provide different levels of transparency to the public.

On the cross-sectoral framework, all agencies provide information on the key elements of their assessment process. Calyx's framework only used to be superficially described on their website and offered the least amount of information of the four agencies (Calyx 2023a). However, the company has recently published more information on their GHG, SDG and REDD approaches (Calyx 2023a,c,d). Regarding project-type specific frameworks, Calyx has only published their REDD framework, while Sylvera and Renoster provide a thorough description of their approach for all project-types that they cover. It should be noted that the co-benefit rating performed by Renoster is not included in the company's current whitepaper, lacking transparency in this regard.

BeZero had not published any project-type specific frameworks by 19.06.2023 (see [Out of scope](#)) but instead provided some examples for project-type specific approaches throughout articles on their website. The company states that they are in the process of rolling out more sector specific frameworks (e.g. ARR), and have published some, which were not assessed as part of this study due to their recent publication (see [Out of scope](#)).

3.4 Comparison of Frameworks

In this chapter, the overarching (i.e. cross-sectoral) frameworks of the selected agencies are compared. First, an overview of the projects' score definition and scale, score aggregation, score components as well as separate (i.e. co-benefits) or excluded scores (i.e. leakage) are summarised in [Table 1](#) below. This is followed by a comparison of the general cross-sectoral approach agencies apply

Table 1: Overview of scoring methodology⁶

| Agency | Score definition | Score scale | Score components ⁷ | Extra score components (as listed in the respective frameworks) |
|---------------|---|---|--|--|
| BeZero | A score of D reflects the lowest, while BBB indicates a moderate and AAA the highest likelihood of achieving 1 tonne of CO ₂ e avoidance or removal. | 8-point scale between D-AAA (low to high): AAA, AA, A, BBB, BB, B, C, D. | <ul style="list-style-type: none"> • Additionality • Over-crediting (incl. Baseline) • Leakage • Non-Permanence • Perverse incentives • Policy & political environment | <ul style="list-style-type: none"> • SDG score (excluded from overall rating) |
| Calyx | A score of E reflects a high risk that carbon credits do not meet their claims of reducing or removing the equivalent of one metric tonne of CO ₂ . The risk then decreases evenly down to a score of A (lowest risk). | 10-point scale between E-A (low to high) (including + for each letter): E, E+, D, D+, C, C+, B, B+, A, A+ | <ul style="list-style-type: none"> • Additionality • Over-crediting (incl. Baseline, Project emissions, Leakage) • Non-Permanence | <ul style="list-style-type: none"> • SDG score (separate) |

⁶ As defined by the agencies in their respective frameworks.

⁷ As defined by the agencies in their respective frameworks. Scores listed for one company but not for another does not indicate that the aspects assessed under this score are not part of other scores of the other agencies. A more detailed explanation can be found below under Score components.

| | | | | |
|-----------------|--|--|---|--|
| | | | <ul style="list-style-type: none"> Overlapping claims | |
| Renoster | A score of 1 reflects a genuine representation of tCO ₂ reduced or removed, while a score higher than 1 means that the project is under-crediting. A score that is lower than 1 indicates that the project is over-crediting. | Numeric score from 0- undefined | <ul style="list-style-type: none"> Additionality Baseline (incl. over-crediting) Permanence Verification | <ul style="list-style-type: none"> Co-benefits score (separate) Leakage (excluded from overall rating) |
| Sylvera | <p>A score of D reflects the lowest, while BBB indicates a moderate and AAA the highest likelihood of achieving 1 tonne of CO₂e avoidance or removal</p> <p>Provisional rating (i.e. rating provided if information is deemed insufficient by Sylvera to give a complete rating):</p> <p>P+ likely reasonable performance against offsetting claims (high likelihood of additionality and permanence potential)</p> <p>P no evidence of major concerns, but some uncertainties/mixed signals to additionality, carbon and/or permanence sub-scores</p> <p>P- at least one major concern across additionality, carbon and/or permanence sub-scores</p> | <p>8-point scale between D-AAA (low to high): AAA, AA, A, BBB, BB, B, C, D.</p> <p>3-point scale for provisional rating: P+, P, P-</p> | <ul style="list-style-type: none"> Additionality (incl. Baseline and over-crediting) Carbon Score⁸ (incl. Leakage, Over-crediting) Non-Permanence | <ul style="list-style-type: none"> Co-benefits score (separate) |

Source: Authors based on BeZero 2022a, Calyx 2023c, Renoster n.d.b, Sylvera n.d.

⁸ The carbon score verifies whether a project is accurately reporting on its activities which directly translates to its overall avoidance (meaning reduction) or removal of CO₂, and other GHGs, measured in CO₂ equivalent (CO₂e) (Sylvera 2022b).

Eligibility criteria

BeZero and Calyx define certain criteria for projects to be eligible for the rating process. To be rated by BeZero, a project must have applied an additionality test or provide sufficient information on how it is deemed additional. A project must also be audited by a recognised third-party auditor to ensure the robustness of the published data and information. Further, sufficient information on the design and ongoing monitoring of the project must be always available in the public domain (BeZero 2022a). Calyx requires carbon crediting programmes to meet a set of minimum criteria in order for a project (registered under the respective carbon crediting programme) to be eligible for a rating. This screening includes the following aspects: governance (e.g. grievance procedures, legal aspects), rules and procedures (i.e. public availability of governance structure and regulatory documentation, and the transparency and robustness of the process for methodology development), Stakeholder engagement (i.e. requirements for public consultation), transparency (i.e. public availability of information on projects, calculation requirements, SDG contributions and auditing), validation and verification (i.e. requirement for independent accredited third party, procedures and guidance and oversight) as well as registry operations (i.e. secure operation, public availability of information on issuance, cancellation and retirement to avoid overlapping claims) (Calyx 2023c). It should be noted that a high performing project can be excluded simply because it is registered under an ineligible carbon-crediting programme, regardless of whether the individual project might be performing better than what the carbon-crediting programme requires.

Renoster and Sylvera both do not stipulate eligibility criteria (Renoster n.d.b, Sylvera n.d.), but Sylvera explains that it gives a provisional rating in case relevant information is missing and a project is deemed not eligible for a complete rating ([Table 1](#) and [17](#)). **However, all agencies check whether the information available for the project is sufficient to assess it and derive a complete rating.** In addition, it should be noted that only few carbon crediting programmes **would be excluded** for assessment due to the eligibility tests applied by Calyx and BeZero. Due to Calyx stipulating more requirements, their screening will result in fewer projects deemed eligible for a rating compared to the other agencies.

Score components

It is important to highlight that **all agencies are broadly assessing the same overarching environmental integrity criteria (e.g. additionality, baseline setting, project emissions, permanence, leakage⁹), though the terminology of the scores differs as well as the factors assessed under each score component.** For example, BeZero's framework has a score for over-crediting, which is a test that checks whether the issued credits have been correctly calculated. This

⁹ Leakage is assessed by Renoster but, in contrast to the other agencies, not included in their final score.

includes baseline carbon stocks, leakage and permanence deductions (BeZero 2022a). Hence, this score component includes other score components. Sylvera has a “carbon score”, that verifies whether the project has delivered on its carbon claims by comparing reported data to third-party and independent data (i.e. detected tree coverage and loss events by Sylvera for forestry, data from grid operators, energy regulators, and off-takers for renewable energy projects) (Sylvera n.d.). This score component is also informed by other permanence and leakage, which are listed as separate score components by other agencies (e.g. Renoster). BeZero has individual scores for policy and perverse incentives (BeZero 2022a). However, Calyx and Sylvera also test for these factors but under different score components (Baseline score under Calyx and Additionality score under Sylvera) (Calyx 2023c, Sylvera n.d.).

However, **Calyx’s framework includes a score for the potential for overlapping claims** ([Table 1](#)). This is also tested by BeZero, though under over-crediting score. Renoster and Sylvera do not assess overlapping claims.

Further, the weighting of the scores differs (see [Assessment process](#)). For example, additionality is stipulated as a limiting factor by all agencies but with varying degrees of strictness. However, the weighting of other scores is less clearly defined and would require a thorough comparison of overlapping ratings (see [Out of scope](#)).

Assessment process

BeZero’s framework follows a four-step assessment. After passing the eligibility test, the actual assessment of the project starts with step 1, the “Macro factor assessment”. This analysis comprises a country, sector, and methodology-specific risk assessment. Once this first step is complete, initial scores for each of the six risk factors (i.e. score components) are assigned. In step 2, the project-level assessment is performed. Here, the same score components are assessed at project level. All project relevant and publicly available information is reviewed (incl. peer reviewed research papers and geospatial data). After completion of this step, the risk factor scores derived in step 1 may be adjusted up or down to reflect both the macro and project specific risks. Step 3 consists of the risk factor weighting and aggregation into one single risk level estimate. Each of the score components is assigned a weight, based on the following indicative weighting: additionality 50%, over-crediting 20%, non-permanence 10%, leakage 10%, policy & political environment: 5%, perverse incentives 5%. Sufficient available information for each component is required and a threshold for overbearing impact by one factor is applied. Each criterion is assigned one of 5 risk levels from significant to very low risk (adjusted throughout the assessment steps), which is then weighted according to the respective weighting factor.

The overall rating is then obtained by aggregating the weighted risk factors (step 3). A minimum of 80% total score component weighting is required. Hence, sufficient information must be available to obtain the minimum weighting. Like the other agencies, BeZero stipulates that **additionality is a limiting**

factor for the overall rating. However, the strictness of the limitation differs among agencies. For example, if a project is assigned the highest additionality risk score by BeZero, then its overall rating would be limited to BB (i.e. the 4th level of BeZero's 8-point scale) (see [Figure 2](#)).

The last step (step 4) consists of the internal review by the BeZero Carbon Ratings (BCR) Committee Review. The committee consists of BeZero analysts and is chaired by a senior member of the ratings team. The analysts present their analysis and rating recommendations to the committee. The Rating Committee is charged with interrogating the recommendations as well as seeking clarifications until all outstanding issues are deemed resolved. BeZero requires unanimous approval by the Rating Committee for a final BCR to be assigned (BeZero 2022a). BeZero also engages with respective project developers, giving them the opportunity to add missing (publicly available) information (pre-rating, provide feedback, and remedy the rating after its release (BeZero 2023a).

Regarding information sources, BeZero uses industry databases and literature from sources such as the IEA, IRENA, UN-FAO and WHO, as well as peer reviewed papers.

The company also built several bespoke in-house databases that can infer from the existing database of rated projects (e.g. for reported emission factors, penetration of technologies, allowances for national harvests). In addition, machine learning techniques are used for the ratings process (primarily for NbS) regarding verification. For example, machine learning is applied to combine data across different satellite sensor technologies (space- and airborne LiDAR, radar, and optical imagery) to estimate factors such as canopy height and other aspects of vegetation change.

Calyx also follows a step-by-step structure, starting with the eligibility test and moving from a top-down analysis of the "bigger picture" to the bottom-up project-level assessment. First, the Calyx Global GHG integrity assessments are performed. This includes three assessments starting with a carbon crediting programme-level screening (i.e. eligibility test). Carbon crediting programmes are also given the opportunity to comment on this screening and point out errors, which will then be corrected. This is followed by the methodology level rating. Here, Calyx conducts an analysis of the inherent activity type risks and respective provisions of the carbon-crediting programme and methodology. This results in the identification of specific areas of a preliminary risk estimation for each score component Calyx Global assesses. Lastly, the project level assessment is performed, which entails a granular activity type analysis including country specific risks. It is focused on the specific areas of concern identified by the top-down carbon-crediting programme and methodology analysis resulting in a final risk estimation for each GHG integrity risk element score component. **Calyx stipulates additionality as a limiting factor, too. In Calyx's case, however, this means that an overall rating cannot be higher than the assigned score for additionality.** Further, the company states that the assessed risks for the scores are additive, meaning that the risks are accumulated. This ensures that positive assessment in one area cannot "overcome" a risk in another area. A second framework is applied for the separate SDG impact rating. This framework aims to assess claimed project contributions to the UN SDGs (if

these are part of a recognized SDG certification). Aspects covered include the level of change achieved (based on activity, output, outcome, and impact) as well as its attributability (based on predicted, narrated, estimated, and quantified contribution and clear methodology).

Calyx uses, inter alia, peer-reviewed literature, other publications, remote sensing, and machine learning models as information sources. For example, this includes data from the European Space Agency (ESA) or the University of Maryland Global Land Analysis & Discovery's Global Forest Change as well as FAO (Calyx 2023b). Calyx states that the company is currently piloting an approach for engaging with project proponents and that only information in the public domain or deemed commercially sensitive (that has been audited) will be used. A formalised policy is expected to be finalised soon.

Renoster's approach differs in that the tests are only applied at project level. Further the company states during the interview that their general approach focuses as much as possible on quantitative data rather than on qualitative information that is deemed subject to manipulation according to Renoster (e.g. justification for financial additionality). The company states that the overarching goal of their services is to remove individual attestations and circumstances since they are "impossible to verify and easy to generate" according to Renoster. Instead, the agency aims to create a system of algorithmic analysis detached from conflicts of interests (i.e. any manipulation of information leading to increased credit issuance). The score components and their respective tests are either directed at all project types or a specific type (e.g. ARR, IFM, REDD+) in Renoster's framework. Each test either results in a numeric, pass/fail or good/medium/bad score. **If a project fails the additionality test, it will automatically be rated 0.00. All tests, except for leakage and co-benefits tests, inform the final rating score** (see [Leakage approach](#) and [Co-benefits approach](#)).

To arrive at the final score, Renoster uses a formula, which is not discussed since calculations are out of scope. A separate rating is given for co-benefits (see [Co-benefits approach](#)).

Renoster primarily uses remote sensing machine learning models for their assessments, according to the agency. The datasets for these models include Landsat, Sentinel 1, Sentinel 2, ALOS, VIIRS, Global Ecosystem Dynamics Investigation (GEDI), and Airborne LiDAR.

Renoster's approach differs strongly from the "standard" approach (approaches used by carbon crediting programmes, methodologies, and the other agencies) of testing for most criteria in that it uses implicit tests via reference areas. For example, the company uses reference areas in the same municipalities that project area is situated in instead of qualitatively assessing policies and regulations that influence deforestation. The company thereby assumes that the deforestation rate in the reference area must reflect any relevant regulations and policies. This is a similar approach to the one used by the recently approved VM0045 methodology by Verra, which uses matched baselines to determine differences in carbon storage between enrolled and unenrolled forests with the same risks (Verra 2022).

Sylvera's process is also primarily aimed at a bottom-up project level analysis. In contrast to Calyx, Sylvera does not use a carbon-crediting programme and methodology-level assessment framework to identify specific areas for a preliminary risk estimation. The company argues that a project-level assessment captures all relevant aspects of the methodology. Further, the agency explains that this can lead to a biased view. Such a bias needs to be avoided. However, a thorough comparison of overlapping ratings would be required to clarify if this is the case (see [Out of Scope](#)).

Sylvera's framework considers country and sector specific risks. Each framework follows the same score components but applies distinct factors or aspects depending on the specific project type.

The **additionality rating is considered a limiting factor.** If Sylvera assigns the lowest score for additionality, the overall rating is limited to a C or D ([Figure 2](#)). To arrive at the final rating, Sylvera first integrates the additionality of activities and the over-crediting risk to get an overall additionality score. A matrix is used to generate the impact score by combining the carbon and additionality scores. Lastly, the impact and permanence scores are integrated via a matrix to arrive at the overall rating. This same process is followed for the distinct types of projects; however, matrices are adjusted to each project type. A separate rating is given for co-benefits (see [Co-benefits approach](#)).

Sylvera explains that if key data is missing that hinders the issuance of a complete rating, the company provides a **provisional rating instead** based on the carbon score, additionality, and permanence ([Table 1](#) and [17](#)). The matrices used for provisional ratings differ from those used for the fully rated projects. The most common reasons for provisional ratings are unusable shapefiles, limited remote sensing data, incomplete information on baseline or because the project is at pre-issuance stage.

Further, **Sylvera engages with respective project developers**, giving them the opportunity to comment on or challenge the rating and provide feedback as well as additional information. Only if sufficient information or data is given, suggested changes will be implemented. To avoid conflict of interest, the company adopted a Project Proponent Relationship and Grievance Policy (Sylvera 2022e).

Sylvera uses peer-reviewed literature when constructing the frameworks. In case peer-reviewed literature is unavailable, highly corroborated research is used instead. The company also uses datasets from sources such as FAO, World Bank and WHO (Sylvera 2022b). Sylvera utilises different remote sensing or third-party geospatial sources and integrates data limitations (temporal and spatial resolution, metadata) as uncertainties in the framework models. Moreover, Sylvera uses LiDAR as reference data for their forestry project ratings. Besides LiDAR, Sylvera's machine learning models are trained using third-party sources such as Global Forest Canopy Cover (GFCC), GEDI (spaceborne LiDAR) and internal datasets using labelling on high-resolution imagery. Of the four companies, Sylvera is the only that uses terrestrial LiDAR, which can increase accuracy on above-ground biomass estimates.

It is important to note that **AI and LiDAR or similar tools used by carbon credit rating agencies can be helpful but do not automatically improve the assessment process**. They must be fed with the appropriate datasets and complemented with a robust framework. Further, they must be implemented by experts to enable a thorough project evaluation.

Scoring scales

Three of the four agencies use alphabetical scores (BeZero, Calyx and Sylvera), which range from an 8 to a 10-point scale. ([Table 1](#)). Renoster on the other hand uses a numeric score ([Table 1](#)). For example, a rating of 0,50 from Renoster means that, according to the company, only 50% of the carbon credits issued by a project are genuine while other agencies' ratings reflect the risk that the mitigation outcomes are not achieved/likelihood that they are indeed achieved. Renoster's rating can also exceed a score of 1.0 in case a project is subject to under-crediting ([Table 1](#)).

3.4.1 Comparison of environmental integrity criteria

In the following chapter, the agencies' cross-sectoral assessment approaches are summarised. In the first step, benchmarking via the CCQI methodology **at the project level** is applied. These are: baseline and project scenario approach, non-permanence approach, leakage approach, and the additionality approach. The co-benefits and safeguards approach will be discussed, but not used as a benchmark to evaluate the agency's performance.

Please note that respective tests performed by CCQI and the selected agencies are not presented under the score components stipulated in their respective frameworks. Instead, they are sorted into the same set of environmental integrity criteria to accommodate comparability. For example, aspects covered under the over-crediting score by Bezero are listed under leakage and/or permanence.

In the second step, other cross-sectoral tests performed by the agencies to assess the baseline and project scenario, non-permanence, leakage, and additionality are compared. It must be noted that Renoster only rates forestry projects. Therefore, for Renoster, the forestry approach is analysed at a high level, whereas the other agencies undergo a high-level analysis for all their project types.

Baseline and Project Scenario approach

The baseline is defined as the emissions level against which emission reductions or removals of a mitigation activity are determined (CCQI 2022).

Table 2: CCQI Benchmarking – Baseline and Project Scenario analysis

| Agency | Existing policies and legal requirements | Newly adopted government policies and legal requirements | Conservativeness/dealing with uncertainties | Perverse incentives | Consideration of mitigation targets and actions in NDCs or LEDS (where applicable) |
|-----------------|--|--|---|---------------------|--|
| BeZero | X | X | X | X | X |
| Calyx | X | X | X | X | X |
| Renoster | (X) | (X) | (X) | X | (X) |
| Sylvera | X | X | X | X | X |

Source: Authors based on BeZero 2022a, Calyx Global 2023c, Renoster n.d.b, Sylvera n.d, interviews and questionnaires answered by the selected agencies. Legend: X = direct test (X), = implicit test, * = partial test.

While BeZero, Calyx and Sylvera all check explicitly for the elements displayed in [Table 2](#), **Renoster’s approach differs for all tests**. Renoster calculates the dynamic baseline scenario and compares it to the reported scenario. Instead of directly checking whether existing and newly adopted policies as well as mitigation targets and actions in nationally determined contributions (NDCs) or Low Emission Development Strategies (LEDS) have been appropriately considered, Renoster uses an implicit check by choosing reference areas in the same municipality as the project area to inform the baseline scenario calculated by Renoster. It has been pointed out by the agencies that the information needed to test for the consideration of mitigation targets and actions in NDCs or LEDSDs is often not available.

It is also important to note that Sylvera’s project-type specific frameworks do not list a test for perverse incentives. However, on their website and in their whitepaper a check for perverse incentives is mentioned (Sylvera 2022b).

Table 3: Comparison of other cross-sectoral baseline and project scenario tests performed by agencies

| Agency | Data validity |
|----------|---------------|
| BeZero | X |
| Calyx | X |
| Renoster | (X) |
| Sylvera | X |

Source: Authors based on BeZero 2022a, Calyx Global 2023c, Renoster n.d.b, Sylvera n.d, Interviews and Questionnaires answered by the selected agencies. Legend: X = direct test (X), = implicit test.

All agencies assess data validity. This is a test to determine whether the project’s reported information is appropriate and comprehensive and if the amount of information is sufficient for the rating assessment. Sylvera assigns a provisional rating in case the information is found to be insufficient to provide a final rating (Table 1 and 17). Renoster uses an implicit check by calculating the dynamic baseline scenario and comparing it to the reported baseline scenario. However, the company also qualitatively checks whether there is sufficient information provided for the project to perform an assessment.

Non-Permanence approach

Non-permanence occurs when the emission reductions or removals generated by the mitigation activity are later reversed. This can be caused by a natural disaster or project mismanagement and results in a temporary greenhouse gas benefit for the atmosphere (CCQI 2022).

Table 4: CCQI Benchmarking - Non-Permanence

| Agency | Significance of non-permanence risk (for project type) | Measures regarding accounting or compensating for reversals | Measures for avoiding or reducing non-permanence risks |
|----------|--|---|--|
| BeZero | X | X | X |
| Calyx | X | X | X |
| Renoster | — | — | X* |
| Sylvera | X | X | X |

Source: Authors based on BeZero 2022a, Calyx Global 2023c, Renoster n.d.b, Sylvera n.d., Interviews and Questionnaires answered by the selected agencies. Legend: X = direct test (X), = implicit test, * = partial test, - = no test applied.

BeZero, Calyx and Sylvera all check for the tests displayed in [Table 4](#). However, **Renoster’s approach differs**. They do not test the significance of the project-specific non-permanence risk or measures regarding accounting or compensating for reversals, according to their whitepaper. The company explains during the interview that its reasoning behind excluding such measures (i.e. buffers) from the test is that they are not being used for their intended insurance purpose (i.e. the buffer is not sufficiently deducted after a loss event and/or the used buffer credits are not of equivalent quality). The company, therefore, deems most of the current buffers illegitimate. As for the “measures for avoiding or reducing non-permanence risks” test, Renoster does not test for human risks and mitigation measures in their non-permanence assessment but in their co-benefit analysis, which is excluded from the numeric score of the rating (see [REDD AD - Non-Permanence approach](#)).

Table 5: Comparison of other cross-sectoral non-permanence tests performed by agencies.

| Agency | Commitment period ¹⁰ | Data validity |
|----------|---------------------------------|---------------|
| BeZero | X | X |
| Calyx | X | X |
| Renoster | X | (X) |
| Sylvera | X | X |

Source: Authors based on BeZero 2022a, Calyx Global 2023c, Renoster n.d.b, Sylvera n.d, interviews and questionnaires answered by the selected agencies. Legend: X = direct test (X), = implicit test, * = partial test.

All agencies perform mostly the same tests ([Table 5](#)). As for data validity (i.e. testing whether the project has real and comprehensive documentation), Renoster uses an implicit test via the self-derived baseline.

Leakage approach

Leakage refers to the net change of greenhouse gas emissions or removals that are attributable to the mitigation activity, occurring outside the activity’s boundary. These can, inter alia, include indirect emission changes upstream or downstream of the mitigation activity or rebound effects (CCQI 2022).

¹⁰ Project implementation duration that the project has committed to.

Table 6: CCQI Benchmarking – Leakage

| Agency | Selection of leakage sources | Conservativeness/dealing with uncertainties |
|----------|------------------------------|---|
| BeZero | X | X |
| Calyx | X | X |
| Renoster | X* | (X) |
| Sylvera | “X” | X |

Source: Authors based on BeZero 2022a, Calyx Global 2023c, Renoster n.d.b, Sylvera n.d, Interviews and Questionnaires answered by the selected agencies. Legend: X = direct test (X), = implicit test, * = partial test, “X” = test not stringently applied for all project types.

BeZero and Calyx both perform the two tests listed in [Table 6](#). The tests include an assessment of whether a project measures and deducts leakage as well as whether the reported estimates are correct. Both agencies test for geographical as well as market leakage. **Renoster also checks** whether a project measures and deducts leakage for both leakage types. **However, the company only assess whether geographical leakage¹¹** estimates are correct: ([Table 6](#)). Renoster explained during the interview that market leakage¹² is too difficult to measure and is therefore excluded from the calculation.

Although Renoster assesses geographical leakage, it **excludes the leakage rating from the overall score**. The company deems leakage to be a nebulous concept which, despite VCM actors’ efforts, cannot be genuinely measured. Further, Renoster states that regional deforestation, a source of leakage in the forestry sector, is **driven by macro-economics and policies rather than carbon projects**. Moreover, the company deems **leakage to be at odds with additionality**, meaning that in areas with high deforestation rates, projects are simultaneously highly additional and highly prone to leakage. The company is therefore of the opinion that leakage requires a nuanced conversation, which is arguably being met with their assessment and its separation from the overall rating.

Sylvera does not stipulate both leakage tests for all project types in the respective frameworks (see [Table 6](#) and [chapter 4.1](#)). For example, Sylvera lists the leakage as “n/a” in their ARR framework (Sylvera 2022h). Sylvera explains in the questionnaire that the reasoning for this decision is that geographical leakage is considered minimal for ARR. In addition, market leakage is often not considered a risk, according to Sylvera. The company states that many ARR projects introduce

¹¹ Leakage as a result of activity displacement outside of the project's boundaries (IPCC n.d.)

¹² Leakage as an increase in GHG emissions when due to changes in the supply and demand equilibrium, causing other market actors to shift their activities (Mongabay n.d.)

increased supply, thereby depressing timber prices and reducing the incentive to establish new timber plantations. The company added that leakage assumptions included in the design of an ARR project are typically based on the modelling of a hypothetical scenario rather than an observable one because causality between the project and the geographical leakage could not be demonstrated. Sylvera further commented that this hypothetical leakage modelling in the ARR project type results in a standardised leakage discount rate¹³ applied to projects. Leakage deduction, in turn, is included in the Sylvera score. However, Sylvera states during the interview that both leakage tests are conducted if deemed relevant and that the current frameworks might be subject to change in the near future. The company states during the interview that a more balanced approach across all project types will be applied soon. BeZero and Calyx do not follow this approach and apply the same testing for all project types. Furthermore, Calyx looks at the original land use that the project is replacing and assesses whether material leakage from the replacement of this land use is an issue.

Table 7: Comparison of other cross-sectoral leakage tests performed by agencies

| Agency | Leakage extent, discount and mitigation measures | Data validity |
|----------|--|---------------|
| BeZero | X | X |
| Calyx | X | X |
| Renoster | X* | X |
| Sylvera | X | X |

Source: Authors based on BeZero 2022a, Calyx Global 2023d, Renoster n.d.b, Sylvera n.d, interviews and questionnaires answered by the selected agencies. Legend: X = direct test (X), = implicit test, * = partial test.

All agencies test for the leakage extent and discount ([Table 7](#)). However, Renoster excludes leakage from the overall rating. BeZero, Calyx and Sylvera also test for mitigation measures to avoid leakage, such as the development of alternative livelihoods, while Renoster only lists these activities in their co-benefits score but does not perform an assessment on whether these activities have an impact (see [chapter 4](#)) and excludes the co-benefit analysis from the numeric score of the rating.

¹³ An amount deducted from the project's total mitigation outcomes in order to prevent overcrediting.
 Perspectives Climate Group GmbH · www.perspectives.cc · info@perspectives.cc Page 25

Additionality approach

Emission reductions or removals resulting from a mitigation activity are considered additional if the activity would not have occurred without the added incentive generated by the carbon credits (CCQI 2022).

Table 8: CCQI Benchmarking – Additionality

| Agency | Eligibility of mitigation activities that are triggered by legal requirements (regulatory additionality) | Consideration of carbon credits before the decision to proceed with the project | Financial attractiveness (financial additionality) | Barriers |
|-----------------|--|---|--|----------|
| BeZero | X | X | X | X |
| Calyx | X | X | X | X* |
| Renoster | (X) | — | (X) | (X) |
| Sylvera | X | X | X | X |

Source: Authors based on BeZero 2022a, Calyx Global 2023c, Renoster n.d.b, Sylvera n.d, interviews and questionnaires answered by the selected agencies. Legend: X = direct test (X), = implicit test, * = partial test, - = no test applied.

The first and third tests are conducted by all agencies, though Renoster uses an implicit test for regulatory and financial additionality. As described in the baseline scenario comparison, Renoster identifies reference area analogues. The agency will identify reference areas subject to the same regulatory requirements and practices of the project being analysed. The analogues are then used to create a common practice scenario ([Table 9](#)) to inform the financial additionality and regulatory scenario derived by Renoster.

The same principle applies to barrier testing ([Table 8](#)). All agencies, except Renoster, test for the consideration of carbon credits before the decision to proceed with the project.

Calyx only applies a partial barrier test and explains that barriers (i.e. barriers that would prevent the implementation of the proposed project activity, e.g. institutional, technological, financial barriers) are not viewed as decisive by the agency. Calyx only assesses barriers that materially affect financial attractiveness and, therefore, would lower inherent additionality risk. Further, Calyx deems barrier tests unreliable and subjective in nature. The company further explains that project documentation justifying the existence and relevance of barriers is difficult to assess and verify objectively against external

sources. In addition, Calyx deems methodological rules on barrier testing insufficiently robust. Subsequently, these tests would add little to the overall rigour of an additionality determination. Instead, the company relies on the common practice test as well as other circumstantial evidence of barriers such as penetration rates (i.e. a measure of the diffusion of a technology, product, or practice in a market) (Table 8, Table 9). After identifying the claimed barriers, they are evaluated in relation to the potential additional carbon revenues (i.e. to check if the order of magnitude of the revenues seem sufficient to overcome the claimed barriers) as well as other objective proxies for barriers such as regulatory environment, country socioeconomic data and known practice on the ground.

Table 9: Comparison of other cross-sectoral additionality tests performed by agencies

| Agency | Appropriateness of used test | Data validity | Common practice |
|----------|------------------------------|---------------|-----------------|
| BeZero | X | X | X |
| Calyx | (X) | X | X |
| Renoster | (X) | (X) | X |
| Sylvera | (X) | X | X |

Source: Authors based on BeZero 2022a, Calyx Global 2023c, Renoster n.d.b, Sylvera n.d, interviews and questionnaires answered by the selected agencies. Legend: X = direct test (X), = implicit test, * = partial test.

BeZero is the only company clearly reviewing the appropriateness of additionality tests applied to projects. However, when assessing the additionality of a project, the agencies will implicitly test if the additionality test used is suitable. This is done by reviewing the additionality of the respective project with the tests deemed appropriate by the agency. Thereby, comparing their approach to the one used by the project proponent and identifying whether the test used was suitable or not. All agencies conduct a common practice analysis (Table 9).

Co-Benefits and Safeguards

This section is only briefly discussed **without benchmarking**.

Table 10: Comparison of cross-sectoral co-benefit and safeguards tests performed by agencies

| Agency | Community benefits | Biodiversity benefits | SDGs | Safeguards against negative impacts | Data validity |
|-----------------|--------------------|-----------------------|------|-------------------------------------|---------------|
| BeZero | X | X | X | — | X |
| Calyx | X | X | X | — (currently under development) | X |
| Renoster | X | X | — | — | X |
| Sylvera | X | X | X | X | X |

Source: Authors based on BeZero 2022a, Calyx Global 2023c, Renoster n.d.b, Sylvera n.d, interviews and questionnaires answered by the selected agencies. Legend: X = direct test (X), = implicit test, * = partial test, - = no test applied.

All agencies assess co-benefits. Renoster’s ratings address biodiversity and community impacts, while Calyx and Sylvera’s ratings assess all Sustainable Development Goals (SDGs), including biodiversity and community impacts (Calyx 2023a, Renoster n.d.b., Sylvera n.d.). BeZero states upon request that the agency also assesses SDGs and that information on this assessment is provided on their customer platform.

Sylvera is the only agency stipulating a safeguards assessment for their co-benefit score (Table 10). Calyx states that it is in the process of including this assessment as well and is close to finalising this process.

It must be noted that Sylvera provides a separate score for co-benefits and safeguards. Similarly, Calyx provides a co-benefit score and a separate score for safeguards is currently under development. However, BeZero and Renoster do only discuss these features.

3.5 Overall rating of cross-sectoral approach and recommendations for improvement

In this section of the report, the four agencies’ cross-sectoral approaches are assessed. The methodology used assesses the risk or likelihood of a mitigation outcome being achieved or not (see [Objective](#)). It is based on the following requirements:

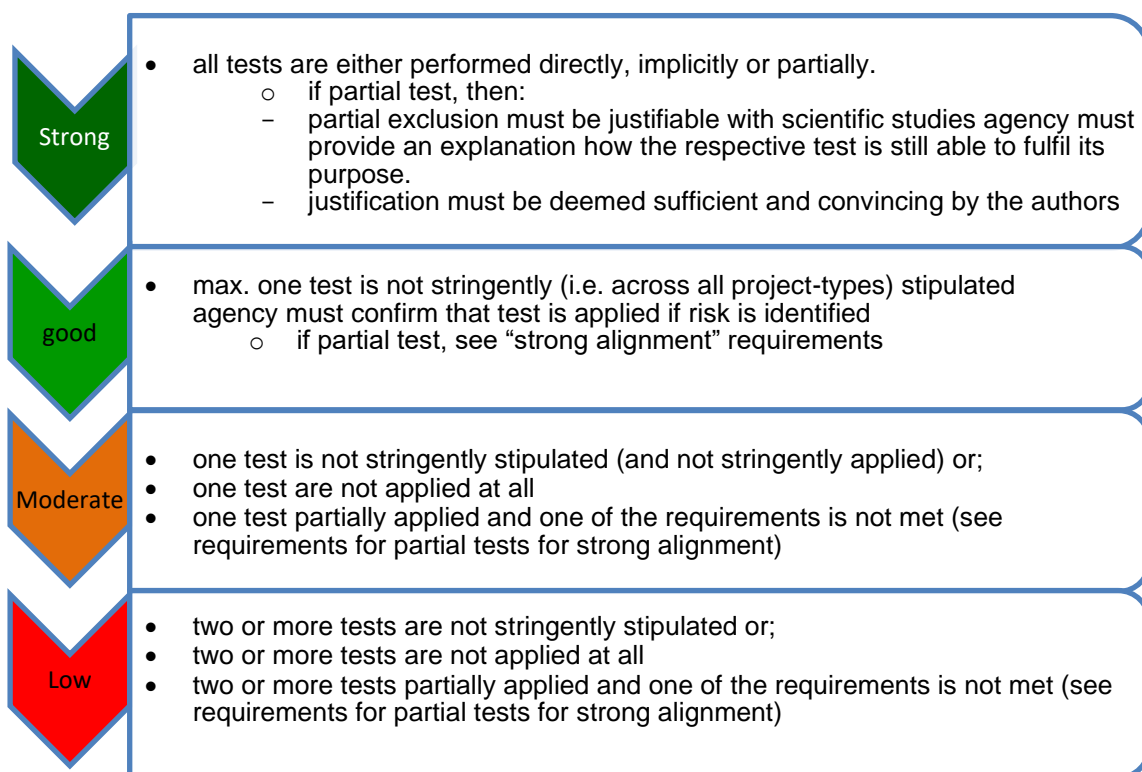
Elements considered in the overall assessment of rating agencies’ approach:

- The alignment with CCQI and other tests ([Figure 1](#))
- The additionality score’s limiting effect on the overall score ([Additionality as a limiting factor](#))
- The application of rigorous tests (i.e. tests going beyond the assessment of the risk/likelihood of a mitigation outcome being achieved and/or beyond what is required by the majority of carbon-crediting programmes) ([Rigorous stringent tests](#))

Bare minimum requirements for the analysis:

- Governance (CoI) ([Table 11](#))
- Review process ([Table 11](#))
- Framework and Ratings Transparency (i.e. the availability of general information on the overarching approach and availability of at least some ratings) ([Table 11](#))

Figure 1: Rating for alignment with CCQI and other tests



Source: Authors.

Additionality as a limiting factor

Robust additionality is a core element for ensuring environmental integrity (Ahonen et al. 2021) and the most important prerequisite to providing an emissions benefit (Cames et al. 2016). Since additionality per definition ensures that only real mitigation outcomes requiring carbon market support are credited

(Ahonen et al. 2021), it is of utmost importance that projects with an elevated risk of non-additionality cannot receive a high rating.

However, the limiting effect of the additionality score on the overall score differs among agencies (Figure 2). A rigorous approach must stipulate additionality as a strictly limiting factor, meaning that if additionality is assigned the lowest score (or fails the test as possible under Renoster’s framework), the overall rating is also limited to the lowest score. An adequate approach must at least limit the overall score to the degree that it is no longer considered a high or good quality project.

Figure 2: Range of limitation of the overall rating if additionality is assigned the lowest score

| | | | | | | | | | | | |
|----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| BeZero | AAA | AA | A | BBB | BB | B | C | D | | | |
| Calyx | A+ | A | B+ | B | C+ | C | D+ | D | E+ | E | |
| Sylvera | AAA | AA | A | BBB | BB | B | C | D | | | |
| Renoster | 1.0 | 0.9 | 0.8 | 0.7 | 0.6 | 0.5 | 0.4 | 0.3 | 0.2 | 0.1 | 0.0 |

Source: Authors based on BeZero 2022a, BeZero 2023d, Calyx Global 2023c, Renoster n.d.b, Sylvera n.d. Please Note: Renoster’s scale goes beyond 1.0 if the project is subject to under-crediting (see Table 1). 1.0 was chosen as a limit to allow for visual comparison with other agencies’ scales.

Rigorous tests

Aspects assessed beyond the assessment of the risk/likelihood of a mitigation outcome being achieved and/or above what most carbon-crediting programme and methodologies stipulate, such as double-issuance¹⁴, double-claiming¹⁵ or a 100-year permanence benchmark for NbS, are rated as rigorous (see Table 11). Co-benefits and safeguards are also rigorous tests but are **only included in brackets in the overall rating** (see Table 12) since they do not impact the mitigation outcomes. Therefore, they are either handled as a separate score by the agencies or not being rated at all and instead only included as additional information.

¹⁴ Double-issuance refers to a scenario where the same mitigation outcome is credited and issued multiple times. This can occur by accident or purposely.

¹⁵ Double-claiming refers to a scenario where a credit issued for a mitigation outcome is claimed by multiple entities.

Bare minimum requirements

Other aspects, such as governance, monitoring, and rating updates as well as transparency are assessed and compared, but only minimum requirements must be met by the respective agency regarding these aspects to be eligible **for all rating categories beyond “insufficient”** (see [Table 11](#)). This is because their impact on the rating is rather indirect compared to the other aspects.

Table 11: Rating categories and requirements for cross-sectoral approach

| Rating | Elements considered | Bare minimum requirements |
|---------------------|---|---|
| Rigorous | All criteria under “Adequate” must be met as well as: <ul style="list-style-type: none"> ● Strong alignment with CCQI Benchmark and other tests ● Additionality stipulated as strictly limiting factor (i.e. no trade-off with other scores possible) ● Application of at least one rigorous test | All criteria must be met to be eligible for a rating above “Insufficient”: <ul style="list-style-type: none"> ● Governance (Col) ● Review process in place (incl. monitoring and update process) ● Framework and Ratings Transparency (i.e. the availability of general information on the overarching approach and availability of at least some ratings) |
| Adequate | All criteria under “Mixed” must be met as well as: <ul style="list-style-type: none"> ● Good alignment with CCQI Benchmark and other tests ● Additionality stipulated as limiting factor (i.e. trade-off with other scores possible, but overall rating is limited to lower range) | |
| Mixed | Good alignment with CCQI Benchmark and other tests | |
| Insufficient | Low alignment with CCQI Benchmark and other tests | |

Source: Authors

Rating summary - Foreword

This analysis of the different rating agencies is based on publicly available information and information shared by the rating agencies (in interviews or other non-public exchanges) specifically for this report. However, this information allows for a high-level comparison of the general approaches of the rating agencies rather than for a thorough in-depth analysis of how the respective approaches, or the models used by the rating agencies work (see [Out of scope](#)). In this regard, rating agencies are less transparent (due to competitiveness) than carbon crediting programmes who provide the public with highly detailed set of documents that explain on which grounds a carbon credit was issued for a specific project in a specific timeframe (including for example the general carbon crediting programme requirements, processes, methodologies and tools, as well as project documentation).

This analysis is focused on carbon credit rating agencies, not carbon crediting programmes. An evaluation of rating agencies does not automatically include a judgement of the performance of carbon crediting programmes and standard setting organisations. Hence, when the authors conclude that the carbon rating agencies are “adequate”, it is not intended to be understood that carbon crediting programmes and standards are considered “inadequate”. Firstly, because such an analysis was not in the scope of the study. Secondly, and more importantly, such an attempt would be comparing apples with oranges. It is the role of carbon crediting programmes and standard setting organisations to operate a complex framework that allows to issue ex-post carbon credits for mitigation outcomes achieved using a baseline-and-credit approach. What is even more important is that in this framework it is the currently accepted norm that at the time of project validation, an independent third-party checks if the project proponent (based on the available information) truthfully and rightfully claims the following:

- that the project must be considered additional, based on the rules to determine additionality set forth in the methodologies.
- the correct counterfactual baseline scenario (“what would happen in the absence of the project”?) based on the rules to determine the baseline set forth in the methodologies.

This means that for the first crediting period (e.g. 10 years) the project is considered additional, and the baseline scenario is assumed to be valid.

This framework has its limitations. This is commonly acknowledged by carbon crediting programmes and standard setting organisations (Hewlett 2023) and is increasingly challenged by carbon rating agencies and other players in the market, but it is still accepted by the mainstream. This is a conflict that results out of the distinct roles, objectives and approaches of carbon crediting programmes and standard setting organisations and carbon credit agencies. In principle, carbon crediting programmes, standard setting organisations, and carbon rating agencies can all be “adequate”. Similar to an issued government bond, a financial market instrument is adequate and the rating agency that rates this bond (also referred to as “second party opinion”) is adequate, irrespective of the rating for the bond. In the best-case scenario the activities of carbon crediting programmes, standard setting organisations and rating agencies result in a race to the top while practices of both actors improve overtime.

Renoster’s approach of self-matched reference areas is not only used to compare deforestation levels between the reference area and the project area, but also for implicitly testing for policies and regulations, which serves as an example for this limitation. It highlights how rating agencies call the current carbon crediting system into question and apply their own approaches to assess carbon credit quality beyond what is used under the current system. Moreover, Calyx and Sylvera apply a 100 year-benchmark for evaluating the commitment and monitoring period for reversals in NbS projects. Calyx stipulates that the lowest score is given for a period of less than 30 years (highest score is given for 100 years). The companies apply the benchmark to all projects and assign successively lower scores for shorter timeframes. It must be noted that only the Climate Action Reserve (CAR) requires such a long monitoring and compensation period for reversals.

Rating summary

[Table 12](#) presents the overall rating of cross-sectoral carbon credit quality assessment approaches (definition as per claim) of the selected agencies, based on [chapter 3.3](#) and [3.4](#). Overall, the agencies have either adequate (i.e. BeZero and Sylvera) rigorous (i.e. Calyx) rating approaches. Renoster was excluded from the comparison due to the strong differences between their and the other agencies’ approaches.

Table 12: Overall rating of cross-sectoral approach

| Agency | Rating and reasoning |
|--------|---|
| BeZero | <p>Elements considered:</p> <ul style="list-style-type: none"> + Strong alignment with CCQI Benchmark, as well as wide coverage of other tests + Additionality is stipulated as a limiting factor <p>Rigorous tests:</p> <ul style="list-style-type: none"> + Double issuance and double-claiming test + [Co-benefits] (not rated but measured and considered in assessment report (i.e. BeZero’s platform, accessible for customers)) <p>Bare minimum requirements:</p> <p>Governance:</p> <ul style="list-style-type: none"> + Safeguards to mitigate risk for CoI are provided <p>Review process:</p> <ul style="list-style-type: none"> - Fully internal rating review and framework review process + Monitoring and rating update process in place <p>Framework Transparency:</p> <ul style="list-style-type: none"> + Good Transparency on cross-sectoral framework (but lack of concise and centralised information) - Low transparency on project-type specific framework <p>Please Note: This represents the state of information until 19.06.2023. More information will be published after this date according to BeZero and some of it has been published since.</p> <p>Rating Transparency:</p> <ul style="list-style-type: none"> + Strong public rating accessibility |

| | |
|------------------------|--|
| <p>Calyx</p> | <p>Elements considered:</p> <ul style="list-style-type: none"> + Strong alignment with CCQI Benchmark, as well as wide coverage of other tests + Additionality is stipulated as a strictly limiting factor (i.e. lowest additionality rating leads to lowest overall rating) <p>Rigorous tests:</p> <ul style="list-style-type: none"> + Double issuance and double-claiming test (i.e. overlapping claims score) + [Co-benefits] (separate score) <p>Please Note: This represents the state of information until 19.06.2023. Calyx is currently developing a test for safeguards, which will result in a separate score.</p> <p>Bare minimum requirements:</p> <p>Governance:</p> <ul style="list-style-type: none"> + Safeguards to mitigate risk for Col are provided <p>Review process:</p> <ul style="list-style-type: none"> + Internal/external rating review and external framework review process + Monitoring and rating update process in place <p>Framework Transparency:</p> <ul style="list-style-type: none"> + Strong transparency on cross-sectoral framework • Moderate to low transparency on project-type specific frameworks (only REDD framework publicly available) <p>Please Note: This represents the state of information until 19.06.2023. Calyx's renewable energy framework will soon be published according to the company.</p> <p>Rating Transparency:</p> <ul style="list-style-type: none"> • Medium public accessibility on ratings (some ratings available on Net Zero Marketplace at a reduced, without reasoning/justification of the rating) |
| <p>Renoster</p> | <p>Incomparable due to data-focused approach</p> <p>Elements considered:</p> <ul style="list-style-type: none"> - Low alignment with CCQI Benchmark, as well as wide coverage of other tests. Not all scores are included in the overall score. - Leakage tested but not included in overall score - Land tenure and carbon rights tested not included in the non-permanence score. Instead under co-benefits (not included in publicly available framework) - Additionality as a strictly limiting factor <p>Rigorous tests:</p> <ul style="list-style-type: none"> + [Co-benefits] (not rated but measured and considered in assessment report, accessible for customers as well as the public) <p>Bare minimum requirements:</p> <p>Governance:</p> <ul style="list-style-type: none"> + Safeguards to mitigate risk for Col are provided <p>Review process:</p> <ul style="list-style-type: none"> - Fully internal rating review and framework review process - No monitoring and rating update process yet, but under development <p>Framework Transparency:</p> <ul style="list-style-type: none"> - Strong transparency on framework and reasoning behind ratings <p>Rating Transparency:</p> <ul style="list-style-type: none"> • Moderate public rating accessibility (only older ratings available on website but detailed assessment report is shared) |
| <p>Sylvera</p> | <p>Elements considered:</p> <ul style="list-style-type: none"> + Good alignment with CCQI Benchmark, as well as wide coverage of other tests + Additionality is stipulated as a limiting factor <p>Rigorous tests:</p> <ul style="list-style-type: none"> + [Co-benefits] (separate score) + [Safeguards] (included in separate co-benefit score) |

| | |
|--|--|
| | <p>Bare minimum requirements:</p> <p>Governance:</p> <ul style="list-style-type: none"> + Safeguards to mitigate risk for Col are provided <p>Review process:</p> <ul style="list-style-type: none"> + Internal rating review and external framework review process + Monitoring and rating update process in place <p>Framework Transparency:</p> <ul style="list-style-type: none"> + Strong framework transparency <p>Rating Transparency:</p> <ul style="list-style-type: none"> • Medium public accessibility on ratings (some ratings available on Net Zero Marketplace at a reduced, without reasoning/justification of the rating) |
|--|--|

Source: Authors based on assessment in chapters [3.3](#) and [3.4](#). Legend: ■ = rigorous, ■ = adequate, ■ = mixed (not applicable), ■ = insufficient (not applicable), ■ = incomparable, [] = test is either applied under separate score or not scored.

Rating per agency

BeZero's approach strongly aligns with CCQI and other tests ([Table 12](#)). Additionality is stipulated as a limiting factor by the company (BeZero 2023). Projects assigned the lowest additionality score are limited to a BB (moderately low) rating ([Figure 2](#)). Most projects receiving the lowest additionality score are assigned ratings of 'B' or 'C' in the overall assessment (BeZero 2023). The company applies a rigorous test for double-issuance and double-claiming. The former is assessed by auditing the project boundaries for the specific project being rated and for any project operating or under development within a radius of 50 km. As an example of how the company assesses double-claiming, BeZero states that the agency checks for overlaps with national schemes or other crediting mechanisms such as the renewable energy certificate (REC) market. Further, they state that they assess co-benefits but do not rate them. Instead, they are measured and considered in the assessment reports (i.e. on BeZero's platform, accessible for customers).

Regarding governance, BeZero applies safeguards to mitigate the risk for Col. The company has a process for monitoring and rating updates. Both, the rating review and framework review process are fully internal. The agency's overall framework is publicly available but lacks conciseness. Relevant information is spread across multiple articles on the website. Within the two-week timeframe given to implement changes considered for this report (see [Review by agencies](#)), BeZero included links to respective articles on their website to ease access to relevant information. Project-type specific frameworks were not available until 19.06.2023 (see [Out of Scope](#)) but the agency states that more information on sub-sector/project-type specific frameworks will be rolled out soon. Some, such as a piece on ARR, have already been published after 19.06.2023. It should be noted that BeZero is the only agency that makes all their ratings publicly available. The ratings are also published using the same scale that is presented to customers. Only a detailed reasoning behind the rating is not publicly available. Instead, a summary is provided, which enables the public to understand the key drivers for the rating.

Due to the strong alignment with CCQI and wider coverage of other tests and additionality being a limiting factor, BeZero's cross-sectoral approach is rated "adequate".

Calyx's approach also strongly aligns with CCQI and other tests ([Table 12](#)). However, Calyx is only partially assessing barriers ([Table 8](#)). The reasoning is justifiable with scientific studies (Karthi et al. 2005, Michaelowa and Purohit 2007, Schneider 2011) as well as concerns raised by CCQI (CCQI 2022, p. 28). Further, the justification (see summary under [Table 8](#)) is deemed sufficient and convincing by the authors. In addition, the overall purpose of the test is still fulfilled according to the authors.

Additionality is stipulated as a strictly limiting factor by the agency ([Figure 2](#)). Calyx further applies a rigorous test for double-issuance and double-claiming. Double-issuance is assessed by checking for overlaps with accounting boundaries (e.g. REDD projects) and for double-claiming, the agency compares national mitigation targets with the renewable energy certificate (REC) market (e.g. renewable energy projects). In addition, the agency assesses co-benefits (in a separate score) and is currently developing a test for safeguards, which will also result in a separate score.

To mitigate the risk for CoI, Calyx applies respective safeguards. Further, a monitoring and rating update process are in place. The framework reviews are approved by an external panel and the ratings are reviewed either by internal staff or by experts from their panels, which increases independence. With regard to framework transparency, Calyx's detailed GHG framework as well as their SDG and REDD framework were published within the two-week timeframe given to agencies to implement changes considered for this report (see [Review by agencies](#)). Yet, other project-type specific frameworks were not publicly available as of 19.06.2023. The agency states that the renewable energy framework will be published soon. It should be noted only a selection of the ratings is publicly available (on Net Zero Marketplace), and the scale is reduced compared to the scale visible to customers (see [Figure 2](#)).

Due to the strong alignment with CCQI and wider coverage of other tests, additionality being a strictly limiting factor and rigorous tests such as overlapping claims, Calyx is thereby rated "rigorous".

Renoster's approach shows low alignment with CCQI and other tests. Further, leakage as well as land tenure (included in the co-benefit test) are (though assessed by the company) excluded from the overall rating (see summary below [Table 6](#)). It must be noted that the rating agency shares their leakage and co-benefit assessment with customers and the public (for the publicly available ratings).

Conceptually, the other agencies agreed with Renoster's argument that leakage is at odds with additionality for forestry projects. However, the two are not mutually exclusive since the risk for leakage is also strongly tied to forest accessibility. For example, forests in remote areas with no direct access to roads minimise the risk for leakage (Schwarze et al. 2002). In addition, implementing mitigation measures can minimise leakage risk. Yet, such measures are not included in Renoster's overall score but in the assessment report only. Excluding leakage and mitigation measures from the overall rating

results in a less robust and less comparable rating. Furthermore, it makes Renoster's rating definition of reflecting on the ratio of genuine carbon removals achieved, problematic. A carbon project has, by definition, to increase the net amount of carbon reduced or removed. The risk for leakage strongly influences this effect. If the number of trees cut down remains the same and only the location has moved from the project to outside of this area, the effect for the climate remains zero. As a response to this argument, Renoster explains that they aim to reform how leakage is viewed in the market. The company assumes that every tree will be eventually cut down, if no regulation to protect the trees is implemented. Therefore, Renoster argues that the market needs to move towards thinking about leakage in terms of time lost. This would enable a framework where leakage is expressed as, for example, trees that may have been due to be cut down in 2030 but were cut down in 2023 because of the project. Leakage would be quantified in terms of damage done to the environment over a specific timeframe. A concept similar to tonne-year accounting¹⁶.

Further, the agency does not assess market leakage, arguing that it is too difficult to genuinely measure. In the scientific realm, this topic has been highly debated for over two decades, leading to the exclusion of tropical deforestation from the Kyoto Protocol (Bellassen et al. 2008) as well as the EU Emissions Trading System (European Commission 2008). Streck (2021) published a paper in 2021 stating, inter alia, that market leakage (for REDD+ projects) is particularly complex and the attribution of leakage across markets and supply chains can be extremely difficult. Especially when compared to geographical leakage, the paper claims that market leakage is difficult to manage. Furthermore, it claims that market leakage can be modelled and accounted for but only to a certain extent (Streck 2021). However, leakage being hard to estimate is not a reason for excluding it from accounting – rather a conservative estimation should be used and deducted from calculated emission reductions.

With regard to additionality, Renoster stipulates this score as a strictly limiting factor ([Figure 2](#)). Due to Renoster's rating score reflecting the ratio of genuine carbon removals, projects are assigned a 0.00 rating if they do not pass the additionality test conducted by the agency. Further, the agency assesses co-benefits. Yet, they are not rated but measured and reported on in the respective assessment report.

The company applies safeguards to mitigate the risk for Col. Regarding the rating and framework review, Renoster applies a fully internal process. No monitoring and rating update process is in place yet, but under development. Renoster's framework is detailed and publicly available. With regard to

¹⁶ Tonne/year accounting refers to the idea that a larger quantity of CO₂ stored for a shorter period of time equals a smaller quantity of CO₂ stored for a longer period of time. A tonne-year is a time-specified unit of carbon dioxide equivalence. One Tonne-year is defined as a metric tonne (MT) of CO₂e stored for one year. This approach allows project activities with a shorter duration to credit climate benefits annually (Chay et al. 2022).

ratings transparency, only Renoster's older ratings are publicly available on their website, but a detailed assessment report is provided alongside a video explaining the reasoning behind the rating.

Renoster's approach **varies strongly** from the other agencies in that it is **very data-focused and solely focuses on forestry**. The quality of Renoster's data-focused approach strongly depends on the quality and appropriateness of the underlying calculations and data since the qualitative assessment is much more reduced compared to the approaches of the other three agencies. Most tests directly conducted by agencies are implicitly tested by Renoster. Hence, the company's approach cannot be thoroughly assessed due to lack of in-depth comparison of multiple overlapping ratings assessment reports to analyse how the approach impacts the rating result. Further, it is important to note that the very aspects Renoster claims to focus on (e.g. calculations, models etc.) are out of scope for this report. Therefore, this approach was not comparable to the approaches of the other three selected agencies and is **excluded from the overall rating** ([Table 12](#)).

Sylvera's approach shows good alignment with CCQI and other tests. This is because leakage is only partially and not stringently (i.e. across all project-type frameworks) stipulated. Upon request, the agency clarified that leakage is considered for all project-types if deemed relevant. Additionality is stipulated as a limiting factor ([Figure 2](#)). The company states that in approximately 90% of all scoring combinations, the highest score would be limited to D (i.e. lowest score). The alternate case of C (i.e. second lowest score), is only achievable through highly exceptional and unlikely combinations of scores according to Sylvera. Moreover, C is still defined as "very high risk". The agency further states that projects rated C are judged to be either fundamentally flawed with regards to their additionality or permanence potential or have significantly failed to achieve their mitigation/emission reduction or removal targets. Further it should be noted that the range for additionality as a limiting factor is smaller for Sylvera's ratings compared to BeZero' ratings. Hence Sylvera is stricter regarding this element ([Figure 2](#)). Further, Sylvera applies two rigorous tests (i.e. co-benefits and safeguards), though they are rated in a separate score.

Sylvera, applies safeguards in order to mitigate the risk for Col. In addition, a monitoring and rating update process is in place. Although the rating review process is fully internal, Sylvera's frameworks are peer-reviewed. The company's overall framework as well as project-type specific frameworks are publicly available and detailed. Similar to Calyx, only a selection of the ratings is publicly available (on Net Zero Marketplace) and only at a reduced scale compared to the scale provided to customers (see [Figure 2](#)).

Due to the good alignment with CCQI and other tests and additionality being stipulated as a limiting factor, Sylvera's cross-sectoral approach is rated "adequate".

Recommendations for improvement

Multiple areas of improvement options have been identified. Most importantly, leakage, as well as land tenure (as a permanence risk factor) should be included in the overall score. Moreover, frameworks should clearly stipulate all tests that are applied. If a test is only applied when deemed relevant, a thorough explanation of when this is the case should be provided.

Additionality should ideally be stipulated as a strictly limiting factor. In addition, including rigorous tests on double-issuance, double-claiming, co-benefits, and safeguards increases the robustness of the rating and is therefore recommended. Co-benefits and safeguards are especially important in the context of NbS. The agencies stated during the interviews and in the questionnaires that measuring these factors is challenging. However, causing negative effects on the environment and/or communities is contradictory to the concept of carbon credit quality. Moreover, a score (or separate scores) for co-benefits and safeguards would contribute to market transparency, integrity as well as fair pricing.

Regarding transparency, providing sufficient information on the project-type specific framework and accessibility of the ratings is vital since this demonstrates how the rating approaches work in practice. A lack of both hinders constructive criticism as well as comparability and is counterproductive to the overarching goal set by all rating agencies to increase transparency in the market. Moreover, the frameworks should ideally be comprehensive, clearly structured, and utilise comparable terminology. Information relevant to the assessment process should be included in each framework or at least clearly reference another document (e.g. main framework). Furthermore, it needs to be presented in a concise and centralised manner to increase accessibility and comprehensibility. Relevant factors influencing the robustness of a rating, such as models and datasets, should at least be briefly explained. In contrast to models, which are often part of the intellectual property of the agencies, datasets used are often obtained from third parties. In this case, providing a list of the datasets used for certain project types and project locations would significantly improve transparency of ratings.

As for the ratings and framework review process, an external team of experts provides increased independence, which should ideally be implemented by all agencies. This can take different shapes or forms (e.g. advisory body, expert panel).

When **compared to financial rating agencies**, carbon credit rating agencies are lacking regulation and oversight. Financial rating agencies follow guidelines such as the IOSCO (International Organization of Securities Commissions) Code of Conduct Fundamentals for Credit Rating Agencies (OICV-IOSCO 2015) or the ESMA (European Securities and Markets Authority) Guidelines on the validation and review of Credit Rating Agencies' methodologies (ESMA 2017). These include requirements such as the use of normative scales to compare the ratings among the individual scales. Further, requirements for transparency and methodologies are stipulated. However, it must be noted that the carbon credit rating industry is nascent and considerably smaller in market size compared to the market for financial ratings. With this in mind, it seems natural that the industry is going through a stage of unregulated growth. Self-regulation and governmental regulation can realistically only be

expected if the market continues to grow and reaches a certain size at which market participants (particularly clients) and governments might request more regulation. It must further be noted that BeZero recently aligned their rating scale with Sylvera's scale (BeZero 2023b) exemplifying how there is (to a certain extent) a trend by rating agencies to harmonise approaches to a certain extent.

4. Deep Dive – REDD AD

4.1 Comparison of REDD AD approach

This chapter provides a high-level analysis of the agencies' assessment approaches for REDD AD projects. This is an addition to [chapter 3.4](#) and therefore should be considered in unison with the tables 2-12.

Disclaimer: This chapter compares the shared/available and comparable information. Other factors assessed by the agencies, but not stipulated in the available sources might be missing. Hence, the tables below are not exhaustive. The factors listed here should, therefore, only be viewed and compared with caution.

Table 13: REDD AD- Baseline and Project scenario approach

| Agency | Factors assessed for baseline | Factors assessed for project emissions/removals |
|-----------------|---|--|
| BeZero | <ul style="list-style-type: none"> ● Deforestation rate (via dynamic baseline) ● RA ● Drivers for deforestation ● Governance ● Projection approach ● Historical reference data ● Baseline reassessment period ● Spatial allocation of deforestation | <ul style="list-style-type: none"> ● Emission sources ● Sampling approaches ● Approach used for estimations (e.g. forest loss) ● Monitoring frequency ● Conservativeness, (i.e. reporting and deducting for uncertainties) ● Biomass estimates (i.e. check if estimates are within the range of publicly available data) |
| Calyx | <ul style="list-style-type: none"> ● Deforestation rate (via dynamic baseline) ● RA ● Drivers of deforestation ● Governance ● Projection approach ● Historical reference data ● Baseline reassessment period ● Spatial allocation of deforestation | <ul style="list-style-type: none"> ● Emission sources ● Sampling approaches ● Approach used for estimations (e.g. forest loss) ● Monitoring frequency ● Conservativeness, (i.e. reporting and deducting for uncertainties) ● Biomass estimates (i.e. check if estimates are within the range of publicly available data) |
| Renoster | <ul style="list-style-type: none"> ● Deforestation rate (via dynamic baseline) ● RA ● Drivers for deforestation ● Governance ● Projection approach ● Historical reference data | <ul style="list-style-type: none"> ● Emission sources ● Sampling approaches ● Approach used for estimations (e.g. forest loss) ● Conservativeness, (i.e. reporting and deducting for uncertainties) ● Biomass estimates (i.e. check if estimates are within the range of publicly available data) |

| | | |
|----------------|---|--|
| Sylvera | <ul style="list-style-type: none"> ● Deforestation rate (via dynamic baseline) ● RA ● Drivers of deforestation ● Governance ● Projection approach ● Historical reference data | <ul style="list-style-type: none"> ● Emission sources ● Sampling approaches ● Approach used for estimations (e.g. forest loss) ● Monitoring frequency ● Conservativeness, (i.e. reporting and deducting for uncertainties) ● Biomass estimates (i.e. check if estimates are within the range of publicly available data) |
|----------------|---|--|

Source: Authors based on BeZero 2022a, Calyx Global 2023b, Renoster n.d.b, Sylvera n.d, interviews and questionnaires answered by the selected agencies. Legend: RA = Reference Area.

All agencies conduct the key tests (see [Table 13](#)) for baseline setting and project emissions/removals and use a dynamic baseline to estimate the deforestation rate. Renoster also applies dynamic baselines, which implicitly tests for governance via choosing a reference area in the same municipality. A key difference is that BeZero and Calyx both assess risks associated with the requirements stipulated in the methodology (i.e., inter alia, baseline reassessment period and spatial allocation of deforestation), while Sylvera and Renoster do not follow this approach. Sylvera argues that all relevant risks associated with a project can be identified without a methodology risk assessment and that such an assessment can lead to bias (see [Assessment process, Sylvera](#)). How these differences in the approach affect the rating can only be thoroughly examined via an in-depth comparison of overlapping ratings, which is out of scope for this report. However, it should be noted that one advantage of assessing risk related to the carbon-crediting programme- and methodology-level requirements could be identified for the permanence risk assessment (e.g. buffer strength) (see [Table 14](#)).

It must be noted that the reference areas chosen by the agencies for their calculation of the deforestation rate, besides other factors, can strongly influence the baseline setting score. However, during the interviews and in the questionnaires, all agencies reported the same criteria, such as: topography, climate, accessibility, and biota. The actual differences in choosing the reference areas cannot be observed at a high-level but only through in-depth comparison of multiple overlapping ratings and therefore, not included in the table and analysis. Further, it should be noted that models used for calculating the deforestation rate strongly influence the rating but are out of scope for this report (see [Out of scope](#)).

REDD AD - Non-Permanence approach

Table 14: REDD AD - Non-Permanence approach

| Agency | Risks assessed | Mitigation measures assessed |
|----------------------|---|---|
| <p>BeZero</p> | <ul style="list-style-type: none"> ● Natural, <u>inter alia</u>: <ul style="list-style-type: none"> ○ Fires ○ Extreme weather (e.g. Storm/Wind) ○ Pests and diseases ○ Droughts ○ Sea-level rise ○ Floods ● Human, <u>inter alia</u>: <ul style="list-style-type: none"> ○ Political risks ○ Negative trends (e.g. forest loss) ○ Land tenure & carbon rights ○ Overlap with reserves and protected areas ○ Country context (governance) ○ Project proponent’s risk analysis (incl., inter alia, finance and track record of compensation for reversals) ○ FPIC ○ Commodity price risk on secondary products where applicable | <ul style="list-style-type: none"> ● Mitigation measures for natural & human risk (incl. stakeholder engagement) ● Appropriateness of buffer pool deduction ● Commitment period and monitoring period ● Buffer pool requirements (incl. monitoring of project level contributions to the buffer pool, the extent and direction of contribution changes, cancellation of credits, release of credits back out of the buffer pool, and justification for buffer pool releases) ● Strength of buffer pool (incl. sufficiency, track record of compensation, credit quality) |

| | | |
|------------------------|--|--|
| <p>Calyx</p> | <ul style="list-style-type: none"> ● Natural, <u>inter alia</u>: <ul style="list-style-type: none"> ○ Fires ○ Extreme weather (e.g. Storm/Wind) ○ Pests and diseases ● Human, <u>inter alia</u>: <ul style="list-style-type: none"> ○ Political risks ○ Negative trends (e.g. forest loss) ○ Land tenure & carbon rights ○ Country context (governance) ○ Overlap with reserves and protected areas ○ Project proponent's risk analysis (incl., inter alia, finance and track record of compensation for reversals) | <ul style="list-style-type: none"> ● Mitigation measures for natural & human risk (incl. stakeholder engagement and employment) ● Commitment and monitoring period (100-year benchmark) ● Treatment of avoidable reversals ● Buffer pool requirements (incl. monitoring of project level contributions to the buffer pool, the extent and direction of contribution changes, cancellation of credits, release of credits back out of the buffer pool, and justification for buffer pool releases) ● Strength of buffer pool (incl. sufficiency, track record of compensation, credit quality) |
| <p>Renoster</p> | <ul style="list-style-type: none"> ● Natural, <u>inter alia</u>: <ul style="list-style-type: none"> ○ Fires ○ Extreme weather (e.g. Storm/Wind) ○ Pests and diseases ○ Floods ● Human, <u>inter alia</u>:^x <ul style="list-style-type: none"> ○ Land tenure | <ul style="list-style-type: none"> ● Mitigation measures for natural & human risk (incl. stakeholder engagement)^x ● Commitment period ● Appropriateness of buffer pool deduction |

| | | |
|-----------------------|---|--|
| <p>Sylvera</p> | <ul style="list-style-type: none"> ● Natural, <u>inter alia</u>: <ul style="list-style-type: none"> ○ Fires ○ Pests and diseases ○ Droughts ○ Storm/Wind ○ Floods ● Human, <u>inter alia</u>: <ul style="list-style-type: none"> ○ Political risks ○ Negative trends (e.g. forest loss) ○ Land tenure & carbon rights ○ Country context (governance) ○ Overlap with reserves and protected areas ○ Project proponent’s risk analysis (incl., inter alia, finance, track record of compensation for reversals) ○ FPIC ○ Commodity price risk on secondary products where applicable | <ul style="list-style-type: none"> ● Mitigation measures for natural & human risk (incl. stakeholder engagement) ● Appropriateness of buffer pool deduction ● Buffer pool requirements (incl. monitoring of project level contributions to the buffer pool, the extent and direction of contribution changes, cancellation of credits, release of credits back out of the buffer pool, and justification for buffer pool releases) ● Commitment and monitoring period (100-year benchmark) |
|-----------------------|---|--|

Source: Authors based on BeZero 2022a, Calyx Global 2023b, Renoster n.d.b, Sylvera n.d, interviews and questionnaires answered by the selected agencies. Legend: ^x = not included in score.

All agencies assess human and natural risks as well as mitigation measures. Please note that the list of risks assessed is not exhaustive. Hence, the risks are listed including the term “inter alia”. Due to different levels of information disclosure the comparability of the respective risks is hampered (status 19.06.2023, see [Out of Scope](#)). With regard to mitigation measures, Calyx and Sylvera use a 100-year benchmark for the commitment and monitoring

period REDD AD projects (Table 14). Calyx states that the lowest score is assigned for less than 30 years. Calyx applies a “rating per vintage”, where for example, vintage year 1 with a longer remaining monitoring time frame is rated more favourably than a later vintage year with a shorter time frame of monitoring and compensation. Renoster only accounts for natural risks in their score. Other factors, such as land tenure are assessed by Renoster under their co-benefit rating. However, the assessment process for this score is not included in the publicly available framework and not included in the overall score.

BeZero and Renoster, similarly, also assess the commitment period and rank a project according to the remaining time frame. However, a set timeframe for the monitoring beyond the compensation period (as under Calyx’s and Sylvera’s framework) is not stipulated by the two agencies. Further, none of the agencies stipulate a test in their frameworks for evaluating whether the percentage of the buffer pool contribution by the project is appropriate regarding the risks associated. Calyx and BeZero, however, assess the strength of the buffer pool. This test aims to assess the quality of the permanence buffer as an insurance mechanism. It includes, inter alia, an assessment of the carbon credit quality in the buffer. For example, Calyx explains that buffer pools consisting of REDD credits are deemed less robust since such credits are more prone to permanence and leakage risks. Renoster, Calyx and Sylvera include tonne/year accounting and adjust the risks according to the timeframe of the credit. Renoster is the only agency not assessing buffer pool requirements.

REDD AD - Leakage approach

Table 15: REDD AD Leakage approach

| Agency | Risks assessed | Factors assessed for conservativeness and mitigation activities |
|--------|----------------|---|
|--------|----------------|---|

| | | |
|-----------------|---|--|
| BeZero | <ul style="list-style-type: none"> Market leakage & geographical leakage | <ul style="list-style-type: none"> Leakage belts and leakage area baseline Dedication of leakage allocations Mitigation activities Discount factor |
| Calyx | <ul style="list-style-type: none"> Market leakage & geographical leakage | <ul style="list-style-type: none"> Leakage belts and leakage area baseline Dedication of leakage allocations Mitigation activities Discount factor |
| Renoster | <ul style="list-style-type: none"> Market leakage*^x Geographical leakage^x | <ul style="list-style-type: none"> Leakage belts and leakage area baseline Dedication of leakage allocations Discount factor |
| Sylvera | <ul style="list-style-type: none"> Market leakage* Geographical leakage | <ul style="list-style-type: none"> Leakage belts and leakage area baseline Dedication of leakage allocations Mitigation activities Discount factor |

Source: Authors based on BeZero 2022a, Calyx Global 2023b, Renoster n.d.b, Sylvera n.d, Interviews and Questionnaires answered by the selected agencies. Legend: * = partial test, ^x = not included in score.

BeZero and Calyx both assess the same risks and mitigation measures. Sylvera does not stipulate a market leakage test for REDD+ Avoided Unplanned Deforestation (AUD) in their framework, arguing that it is not relevant since most REDD+ projects protect against unplanned deforestation from

subsistence-based practices. The company further explains that in cases where market leakage is relevant, for example in an Avoided Planned Deforestation REDD+ project, market leakage is included within the Sylvera score.

Renoster tests for whether both market and geographical leakage have been measured and discounted by the project. Renoster also tests for reverse/positive leakage, where trees may be planted outside of projects due to the project's activities (i.e. in some ARR projects). Yet, the company only measures geographical leakage (see [chapter 3.5](#)). All agencies analyse leakage belts, leakage area baseline, the dedication of the leakage allocations as well as discount factors. Renoster does not include mitigation activities to avoid leakage in their score. Although technically tested by Renoster, the mitigation activities are part of their co-benefit assessment which is not included in their publicly available framework and not included in the overall score ([Table 15](#)).

REDD AD - Additionality approach

Table 16: REDD AD Additionality approach

| Agency | Factors assessed |
|--------|--|
| BeZero | <ul style="list-style-type: none"> ● Prior consideration of carbon credits before implementation of project ● Regulatory surplus ● Common practice ● Financial attractiveness ● Barriers |
| Calyx | <ul style="list-style-type: none"> ● Prior consideration of carbon credits before implementation of project ● Regulatory surplus ● Common practice ● Financial attractiveness ● Barriers* |

| | |
|-----------------|---|
| Renoster | <ul style="list-style-type: none"> • Regulatory surplus • Common practice • Financial attractiveness • Barriers • Area boundaries analysis* |
| Sylvera | <ul style="list-style-type: none"> • Prior consideration of carbon credits before implementation of project • Regulatory surplus • Common practice • Financial attractiveness • Barriers |

Source: Authors based on BeZero 2022a, Calyx Global 2023b, Renoster n.d.b, Sylvera n.d, interviews and questionnaires answered by the selected agencies. Legend: * = Test also conducted by other agencies but under a different score or test (e.g. eligibility test, baseline score).

All agencies apply the same key tests for REDD AD as already reflected in the CCQI benchmarks ([Table 2](#) and [3](#)). Renoster stipulates an area boundaries analysis in their framework. This is conducted to determine whether the project's boundaries have been manipulated to favour carbon credit issuance. The test is also applied by the other agencies but in other scores or eligibility criteria.

4.2 Overall rating of REDD AD approach and recommendations for improvement

The overall rating of the REDD AD approach is informed by the chapters [3.4](#) and [4.1](#) does not present separate but coherent ratings.

This subchapter focuses on aspects related to assessing the risk or likelihood of a mitigation outcome being achieved or not (see [Objective](#)). It is based on the following requirements:

Elements considered in the overall assessment of rating agencies' REDD AD approach:

- The additionality score's limiting effect on the overall score ([Additionality as a limiting factor](#))
- The application of rigorous tests ([Rigorous REDD AD stringent tests](#))
- The application of key REDD AD tests ([Table 17](#))

Bare minimum requirements for the analysis:

- Framework Transparency (i.e. the availability of general information on the REDD AD approach) ([Table 17](#))

Further:

- All score components must be included in overall score
- All tests required under cross-sectoral approach must be applied for REDD AD

Rigorous REDD AD tests

Aspects assessed beyond the objective and/or above what most carbon-crediting programmes and methodologies stipulate, such as 100-year permanence benchmark for compensation and monitoring of reversal, buffer strength (i.e., quality of buffer pool as insurance mechanism) or dynamic baselines, are rated as rigorous (see [Table 18](#)). Co-benefits and safeguards are also rigorous tests but are **only included in brackets in the overall rating** (see [Table 12](#)) since they do not impact the mitigation outcomes. Therefore, they are either handled as a separate score by the agencies or not being rated at all and instead only included as additional information.

Table 17: Rating categories and requirements for REDD AD approach

| Rating | Elements considered | Bare minimum requirements |
|-----------------|---|---|
| Rigorous | <p>All criteria under "Adequate" must be met as well as:</p> <ul style="list-style-type: none"> • Additionality stipulated as a strictly limiting factor (i.e. no trade-off with other scores possible) • Application of at least one rigorous REDD AD test (i.e. a test going beyond the | <p>Criterion has to be met to be eligible for a rating above "Insufficient":</p> <ul style="list-style-type: none"> ○ Framework Transparency (i.e. |

| | | |
|---------------------|---|--|
| | assessment of the risk/likelihood of a mitigation outcome being achieved and/or beyond what is required by most REDD AD methodologies). | general information on REDD AD approach) |
| Adequate | <p>All criteria under “Mixed” must be met as well as:</p> <ul style="list-style-type: none"> ● Application of key REDD AD tests for: <ul style="list-style-type: none"> ○ Baseline setting (i.e. deforestation rate, reference areas, projection approach, drivers for deforestation, governance, historical reference data) ○ Project emissions (i.e. sampling approach, emission sources, approach for estimations, monitoring frequency) ○ Additionality (i.e. prior consideration of carbon credits, regulatory surplus, common practice, financial attractiveness, barriers) ○ Leakage (i.e. market and geographical leakage, mitigation activities) ○ Non-Permanence (i.e. natural and human risks, mitigation activities, commitment period, buffer pool deduction and requirements) ○ Other tests (either under additionality, baseline setting or eligibility): <ul style="list-style-type: none"> ▪ Area boundaries analysis ▪ No clearing of native forest 10 years prior to project implementation | |
| Mixed | <ul style="list-style-type: none"> ● All score components must be included in overall score ● All tests required under cross-sectoral approach must be applied for REDD AD | |
| Insufficient | At least one of the “Mixed” criteria is not met | No general information on REDD AD approach available |

Source: Authors

Table 18: Overall rating – REDD AD approach

| Agency | Rating and reasoning |
|------------------------|---|
| <p>BeZero</p> | <p>Elements considered:</p> <ul style="list-style-type: none"> + All key REDD AD tests are applied + Additionality stipulated as a limiting factor <p>Rigorous REDD AD test:</p> <ul style="list-style-type: none"> + Dynamic baseline + Strength of buffer test (incl. credit quality, sufficiency, track record of compensation) + [Co-benefits] (not rated but measured and considered in assessment report (i.e. BeZero’s platform, accessible for customers)) <p>Bare minimum requirements:</p> <p>REDD AD Framework Transparency:</p> <ul style="list-style-type: none"> – Low Transparency (i.e. no REDD framework publicly available, but certain aspects of approach referenced across multiple articles on website). <p>Please Note: This represents the state of information until 19.06.2023. More information will be published after this date according to BeZero and some of it has been published since.</p> |
| <p>Calyx</p> | <p>Elements considered:</p> <ul style="list-style-type: none"> + All key REDD AD tests are applied + Additionality stipulated as a strictly limiting factor <p>Rigorous REDD AD tests:</p> <ul style="list-style-type: none"> + Dynamic baseline + Strength of buffer test (incl. credit quality, sufficiency, track record of compensation) + 100-year permanence benchmark + [Co-benefits] (separate score) <p>Please Note: Calyx is currently developing a test for safeguards, which will result in a separate score.</p> <p>Bare minimum requirements:</p> <p>REDD AD Framework Transparency</p> <ul style="list-style-type: none"> + Strong Transparency |
| <p>Renoster</p> | <p>Incomparable due to data-focused approach</p> <p>Elements considered:</p> <ul style="list-style-type: none"> – Exclusion of certain score components from overall score <ul style="list-style-type: none"> ○ Leakage tested but not included in overall score |

| | |
|-----------------------|--|
| | <ul style="list-style-type: none"> ○ Land tenure and carbon rights tested not included in permanence score. Instead under co-benefits (not included in publicly available framework) + Additionality stipulated as a limiting factor <p>Rigorous REDD AD test:</p> <ul style="list-style-type: none"> + Dynamic baseline + [Co-benefits] (not rated but measured and considered in assessment report, accessible for customers as well as the public) <p>Bare minimum requirements:</p> <p>REDD AD Framework Transparency:</p> <ul style="list-style-type: none"> + Strong Transparency |
| <p>Sylvera</p> | <p>Elements considered:</p> <ul style="list-style-type: none"> + All key REDD AD tests are applied + Additionality stipulated as a limiting factor <p>Rigorous REDD AD tests:</p> <ul style="list-style-type: none"> + Dynamic baseline + 100-year permanence benchmark + [Co-benefits] (separate score) + [Safeguards] (included in separate co-benefit score) <p>Bare minimum requirements:</p> <p>REDD AD Framework Transparency:</p> <ul style="list-style-type: none"> + Strong Transparency |

Source: Authors based on assessment in chapters [3.3](#) and [3.4](#). Legend: ■ = rigorous, ■ = adequate, ■ = mixed (not applicable), ■ = insufficient (not applicable), ■ = incomparable, [] = test is either applied under separate score or not scored.

The rating for the REDD AD approach by all agencies reflect similar results of the overall rating in [3.6](#). Based on the analysis, **BeZero’s and Sylvera’s approaches are rated “adequate”** and **Calyx’s approach is rated “rigorous”**.

BeZero’s approach is rated “adequate”. The agency applies all key REDD AD tests and stipulates additionality as a **limiting factor** ([Figure 2](#)). The company further uses a dynamic baseline to estimate the deforestation rate and analyse the buffer strength. The company also analyses co-benefits. However, these are not rated but the assessment is provided on their customer platform. As of 19.06.2023 (see [Out of Scope](#)) no project-type specific framework has been published. However, BeZero states that it is in the process of rolling out more information (e.g. on ARR), and some of it has been published since.

Calyx’s approach is rated “rigorous”. The company also applies all key REDD AD tests and stipulates additionality as a **strictly limiting factor** ([Figure 2](#)). The agency also applied a 100-year benchmark for the monitoring and compensation period for reversals, uses a dynamic baseline to

estimate the deforestation rate and analyses the buffer strength. In addition, Calyx assesses co-benefits, though under a separate score. Calyx states that the agency is currently developing an assessment process for safeguards and that it is close to finalisation and will result in a separate score. With regard to transparency on their REDD approach, Calyx published their detailed REDD framework within the two-week timeframe given to the agencies to implement changes considered for this report (see [Review by agencies](#)).

Sylvera's approach is rated "adequate". Sylvera stipulates additionality as a **limiting factor** ([Figure 2](#)) and applies all key REDD AD tests. However, market leakage is not stipulated in their REDD framework (see Text below [Table 15](#)) but is applied if deemed relevant. Sylvera's REDD framework is detailed and publicly available. As for Sylvera's approach to leakage, information on when market leakage is considered, should be added to the framework. It is important to note that the leakage sections of the frameworks are still under development and will be more balanced in the future, according to Sylvera.

Sylvera uses a dynamic baseline to estimate the deforestation rate and assesses co-benefits, though under a separate score. The company also applies a 100-year benchmark for compensating and monitoring for reversals. Moreover, Sylvera is the only agency assessing safeguards against negative impacts (included in their co-benefit score) (status 19.06.2023), a rigorous test that is especially important for NbS project types and adds further nuance to Sylvera's co-benefit score.

Renoster's approach is not rated (see [Renoster, cross-sectoral approach](#)). Overall, the **data-focused approach chosen by Renoster** for their assessment process is an interesting concept (since it aims to reduce human error), but it requires an in-depth assessment of multiple overlapping ratings to understand the appropriateness of this approach on the resulting rating compared to the approaches taken by the other agencies. Therefore, the approach has been excluded from the overall rating.

Renoster does apply all key REDD AD tests but is the only agency that excludes a key score component. Leakage as well as mitigation activities for leakage and permanence (e.g. land tenure) are excluded from the overall score (see [chapter 3.5](#)). The agency uses a dynamic baseline to estimate the deforestation rate and assess co-benefits. Yet, the co-benefit analysis does not influence the overall rating but is included in the assessment report. Renoster's framework is publicly available. The videos explaining the reasoning behind their publicly available ratings provide additional transparency.

Recommendations for improvement

The most important recommendation regarding the REDD AD (as well as the cross-sectoral) approach is that leakage must be included in the overall rating. Although leakage calculation is challenging, it cannot be assumed to be zero. Excluding leakage from the overall score is, therefore, an insufficient approach. Further, leakage should be clearly stipulated for all project-types. Even if the risk for leakage is deemed low, this does not justify the exclusion of the test from a project-type specific framework and

poses a risk to robustness and transparency. If leakage is not stipulated and only analysed when deemed relevant, this should be stated in the framework accordingly.

To increase robustness, additionality should be stipulated as a strictly limiting factor. In addition, dynamic baselines to estimate deforestation rates, as used by all agencies, are an especially useful tool (Haya et al. 2023). This goes beyond what is required by most carbon crediting programmes and methodologies and increases conservativeness. Moreover, a set permanence benchmark of 100 years for monitoring and compensation, is another rigorous approach and is applied by Calyx and Sylvera. Since forestry projects are prone to reversals, a longer time period benchmark for monitoring and compensation ensures a conservative permanence assessment. This is also in alignment with CCQI requirements (CCQI 2022). In the scientific realm even much longer time periods are recommended to reach equivalence with the release of fossil carbon in the atmosphere (Scott et al. 2015). Another important aspect to include in the assessment is the strength of the buffer. Analysing the quality of the buffer pool increases the robustness of the non-permanence score and thus the overall score.

Beyond the aspects relating to the risk or likelihood of a mitigation outcome being achieved or not, co-benefits and safeguards are important aspects for carbon credit quality, especially for NbS projects. Although both are either assessed under a separate score (e.g. Calyx and Sylvera) or as an additional and non-rated information (e.g. BeZero and Renoster), co-benefits provide valuable information on the respective project. Thus, this enables carbon credit buyers and the public to make an informed decision on the broader quality of the credits. Including safeguards against negative impacts in this score, as done by Sylvera, adds important nuance and is therefore strongly recommended. Calyx states that an assessment of safeguards (in a separate score) is currently under development.

Another important recommendation is to increase transparency. The available frameworks do not enable the public to compare all relevant aspects (for all project types) considered by the agencies in their assessment process. Full disclosure of all details of the assessment approach would be ideal, but the coverage of the following factors and aspects should be listed in publicly available frameworks as a starting point to enable the public to make an informed decision on whether a rating is robust or not:

- Reference area criteria considered for baseline assessment
- Aspects considered for assessment of conservativeness
- Risks and mitigation measures considered for non-permanence
- Risks and mitigation activities considered for leakage
- Type of leakage sources considered (and reasoning for exclusion, if applicable)

4.3 Comparison of overlapping ratings

This chapter comprises a high-level comparison of overlapping ratings (i.e. ratings of the same projects) among the selected agencies for each assigned rating. Most ratings (5 out of 6) overlap between BeZero, Calyx and Sylvera. Only one project (i.e. Guanaré) has been rated by all agencies. It is

important to note that Calyx's and Sylvera's scale differ on Net Zero Marketplace, compared to the ratings visible to customers, i.e., the scale points are reduced, so that the 10-point scale of Calyx is reduced to 5 points and the 8-point scale of Sylvera is reduced to 3 points (tiers).

Another key difference is that Renoster's scale represents **the ratio** of genuine carbon removals in contrast to the other agencies' scores that reflect on **the likelihood or risk** of a tCO_{2e} mitigation outcome being achieved or not (see [Table 1](#)). Hence, Renoster's scale is numeric while the other agencies use letter scales.

Table 19: BeZero's, Calyx's and Sylvera's scales ranked from high to low

| BeZero | | | | | | | | | | | |
|-------------|--|-----|-----|---|-----|-----|-----|--|-----|-----|-----|
| Scale | AAA | AA | A | BBB | BB | B | C | D | | | |
| Calyx | | | | | | | | | | | |
| Scale | A+ | A | B+ | B | C+ | C | D+ | D | E+ | E | |
| NZM scale | A | | B | | C | | D | | E | | |
| Sylvera | | | | | | | | | | | |
| Scale | AAA | AA | A | BBB | BB | B | C | D | | | |
| Provisional | P+ | | | P | | | | P- | | | |
| NZM scale | Tier 1 (includes AAA and AA ratings on watch) | | | Tier 2 (includes A ratings on watch) | | | | Tier 3 (includes BB and B ratings on watch) | | | |
| Renoster | | | | | | | | | | | |
| Scale | 1.0 | 0.9 | 0.8 | 0.7 | 0.6 | 0.5 | 0.4 | 0.3 | 0.2 | 0.1 | 0.0 |

Source: Authors based on BeZero 2022a, Calyx Global 2023c, Net Zero Marketplace n.d., Renoster n.d.b, Sylvera n.d. Please Note: Renoster's scale goes beyond 1.0 if the project is subject to under-crediting (see [Table 1](#)). 1.0 was chosen as a limit to allow for visual comparison with other agencies' scales.

Due to these differences, comparability is hampered and only a broad comparison can be provided. Limitations to this comparison are listed in chapter 2.2, [Out of scope](#). Further, no detailed reasoning or justification for the ratings was publicly provided by the agencies, which would be required in order to identify reasons for the observable differences between overlapping ratings.

Table 20: Overlapping ratings compared on normative scale

| | | | | | | | | | |
|--|-------------------------|----------------------|----|---|----------------------|----|---|--------|---|
| Ecomapua Amazon REDD+ Project (VCS 1094) | BeZero | AAA | AA | A | BBB | BB | B | C | D |
| | Calyx NZM scale | A | | B | | C | | D | E |
| | Sylvera NZM scale | Tier 1 | | | Tier 2 | | | Tier 3 | |
| Keo Seima Wildlife Sanctuary (VCS 1650) | BeZero | AAA | AA | A | BBB | BB | B | C | D |
| | Calyx NZM scale | A | | B | | C | | D | E |
| | Sylvera NZM scale | Tier 1 (on watch) | | | Tier 2 | | | Tier 3 | |
| Mai Ndombe REDD+ (VCS 394) | BeZero | AAA | AA | A | BBB | BB | B | C | D |
| | Calyx NZM scale | A | | B | | C | | D | E |
| | Sylvera NZM scale | Tier 1 | | | Tier 2 (on watch) | | | Tier 3 | |
| Envira Amazonia Project (VCS 1382) | BeZero | AAA | AA | A | BBB | BB | B | C | D |
| | Calyx NZM scale | A | | B | | C | | D | E |
| | Sylvera NZM scale | Tier 1 | | | Tier 2 | | | Tier 3 | |
| Luangwa Community Forests | BeZero | AAA | AA | A | BBB | BB | B | C | D |
| | Calyx NZM scale | A | | B | | C | | D | E |

| | | | | | | | | | | | |
|--|-------------------|--------|-----|-----|-------------------|-----|-----|-----|--------|-----|-----|
| Project (VCS 1382) | Sylvera NZM scale | Tier 1 | | | Tier 2 (on watch) | | | | Tier 3 | | |
| Guanaré Forest Restoration Project (VCS 959) | BeZero | AAA | AA | A | BBB | BB | B | C | D | | |
| | Calyx NZM scale | A | | B | | C | | D | | E | |
| | Sylvera NZM scale | Tier 1 | | | Tier 2 | | | | Tier 3 | | |
| | Renoster | 1.0 | 0.9 | 0.8 | 0.7 | 0.6 | 0.5 | 0.4 | 0.3 | 0.2 | 0.1 |

Source: BeZero 2023c, Net Zero Marketplace n.d., Renoster 2022. Legend: ■ = assigned rating. All ratings have been retrieved from the respective website on the 2nd of June 2023. Please Note: Renoster’s scale goes beyond 1.0 if the project is subject to under-crediting (see [Table 1](#)). 1.0 was chosen as a limit to allow for visual comparison with other agencies’ scales.

The Ecomapua Amazon REDD+ Project has been rated quite differently by all agencies. Sylvera assigned a high score (i.e. Tier 1), while BeZero and Calyx assigned a low score (i.e. C and D). Keo Seima has been rated higher by BeZero and Sylvera (A and Tier 1 (on watch)), compared to Calyx who assigned their lowest rating (i.e. E). For Mai Ndombe the ratings differ less, with BeZero assigning a B and Sylvera giving a Tier 2 rating. Calyx, however, again assigns their lowest rating. Envira is rated very similar to Mai Ndombe, as well as Luangwa and Guanaré that are rated low by all agencies.

Aside from Ecomapua, Calyx assigns the lowest rating for all other projects ([Table 19](#)). Further, it has to be noted that Sylvera has put Luangwa, Mai Ndombe as well as Keo Seima “on watch” (i.e. under re-evaluation).

The observable differences across overlapping ratings can have multiple reasons. Based on the findings in prior chapters, these could, inter alia, include differences in handling additionality as a limiting factor, the assessment of double-issuance and double-claiming, as well as buffer strength.

5. Conclusions

Carbon credit rating agencies are a new type of player in the market and have taken on the ambitious task to assess carbon credit quality, increase transparency, mitigate corporate reputational risk, and enable fair pricing. To this end, each agency uses their own individual framework(s). Several similarities as well as differences between the cross-sectoral and REDD AD approaches have been identified.

Due to the limitations of the report (see [Out of Scope](#)), many aspects could not be assessed and compared, such as whether a methodology-level assessment to identify specific areas for a preliminary risk estimation creates bias, how exactly other key differences in rating approaches (e.g. leakage) affect the resulting rating or what impact differences between datasets, models and calculations may have. A thorough and in-depth assessment of overlapping ratings would be required to fully grasp the differences between agencies' approaches and the resulting ratings. Hence, the following conclusions are limited to the scope of the assignment.

According to the analysis conducted in chapters 3.4-4.3, all agencies are providing either an **“adequate”** or **“rigorous”** approach to derive their ratings (Renoster’s approach has been excluded). [Table 21](#) summarises the most relevant differences identified. Calyx has been rated rigorous for their cross-sectoral as well as their REDD AD approach, while BeZero and Sylvera are both rated “adequate” for both approaches.

Table 21: Overview of rating of the key differences identified for general and REDD AD approach

| Agency | General approach | | | | REDD AD approach | |
|----------|----------------------------|------------------------------------|---------|-----------------------------|------------------|-------------------------------|
| | Public Rating Transparency | Additionality as a limiting factor | Leakage | Double-issuance & -claiming | Buffer strength | 100-year Permanence benchmark |
| BeZero | ++ | + | + | ++ | ++ | - |
| Calyx | - | ++ | + | ++ | ++ | ++ |
| Renoster | + | ++ | - | - | - | - |
| Sylvera | - | + | + | - | - | ++ |

Source: Authors based on assessment in chapters [3.4](#) and [4.1](#). Legend: ++ = rigorous, + =adequate, - = to be improved, = assessed but not included in the final score.

Key differences between agencies' have been identified for the transparency of their ratings. While BeZero’s ratings are all publicly available, only a fraction of Calyx’s and Sylvera’s ratings is accessible by the public. Moreover, the scales on publicly available ratings are “bucketed”, leading to a lack of nuance and less comparability with other ratings with higher point scale. In addition, for Sylvera’s

ratings on Net Zero Marketplace, it is also unclear whether they are on watch or not. BeZero also offers a brief summary of the reasoning behind the rating, which is not offered by Calyx and Sylvera. Renoster provides only ratings older than 6-12 months on their website but published ratings are accompanied by a 30-45 min video (providing an explanation of the rating) and an assessment report. Transparency is vital to enable the public as well as customers to make an informed decision when buying carbon credits. Hence, it needs to be increased not only for the ratings but also for the frameworks.

Another key difference has been identified for handling additionality as a limiting factor. While Calyx and Renoster strictly limit their overall rating based on the additionality score, BeZero and Sylvera allow for a rating higher than the additionality score. If additionality is assigned the lowest possible score, Sylvera limits the overall rating to their lowest and in some cases (max. 10%) their second lowest score. BeZero limits the overall rating for projects receiving the lowest additionality score to a BB. The majority of such projects are assigned a 'B' or 'C' in the overall assessment (BeZero 2023). However, the possible range for the overall rating is larger compared to Sylvera ([Figure 2](#)). Renoster assigns a 0.00 rating, if a project fails its additionality test. In this regard, the company is stricter than Calyx. However, this is rooted in Renoster's rating definition, which reflects the ratio of genuine carbon removals. Hence, the rating strictly follows the logic that credits stemming from projects found to not be additional are "hot air".

Multiple rigorous tests are stipulated in all frameworks across agencies. This includes double-issuance and double-claiming, as well as buffer strength as tested by Calyx and BeZero. Further, a 100-year permanence benchmark for REDD AD projects, as applied by Calyx and Sylvera, increases the robustness of their ratings.

Lack of robustness has been identified for Renoster's leakage and permanence approach. The former as well as natural risks in the latter are not included in Renoster's overall rating. In general, the approach applied by Renoster differs from other agencies. The data-focused approach aims to detach the assessment from the majority of qualitative assessments and the associated human error, according to Renoster. However, this approach is only as good as the underlying calculations, datasets, models, and interpretation of results, which cannot be compared due to lack of in-depth comparison of overlapping ratings. Therefore, this approach was not comparable to the approaches of the other three selected agencies and is therefore excluded from the overview table and overall rating. It has to be noted that the reasoning for these differences given by Renoster are challenging the use of current requirements for environmental integrity criteria (see chapters [3.4](#) and [3.5](#)). Further, leakage as well as permanence (e.g. buffer pool and land tenure) tests are not only relevant for present but for future issuances. This is due to the tests being partially forward looking and not only evaluating how a project has accounted for leakage and reversals so far but also the future risks and respective mitigation activities associated with a project. In contrast, baseline setting and additionality are fixed by the project developer at the beginning of a project (but need to be adjusted in the future if relevant changes occur).

Besides the key factors listed in [Table 21](#), co-benefits are also assessed by all agencies. However, they are not listed in [Table 21](#), since they do not impact the mitigation outcomes. Further, they are not counted towards the overall score, but either rated in a separate score (Calyx and Sylvera) or not rated at all (e.g. BeZero and Renoster). It should be noted that safeguards against negative effects are only assessed by Sylvera under this score. Calyx states that the agency is currently finalising a framework for safeguards, which will be a third separate score next to their GHG and SDG scores. Both aspects are vital for broader carbon credit quality. Their assessment is thus strongly recommended.

Finally, **this assessment would have greatly profited from more readily available information**, especially on overlapping ratings. Since increasing transparency is one of the overarching goals of carbon credit rating agencies, more transparency is expected. Further, the existence of tests that have only been confirmed to be conducted by agencies upon request (e.g. interview) but not explicitly mentioned in the respective framework, cannot be verified. Hence, lack of transparency hampers accountability and should therefore be avoided as much as possible. **Any test not stipulated in the respective frameworks should be considered with caution.** In addition, the comparison of overlapping ratings shows that the same project can be rated very differently by each agency. This can lead to confusion for consumers. Considering the lack of sufficient transparency on respective frameworks used for the assessment, it becomes impossible for the public to understand the differences between ratings and judge the appropriateness as well as quality of the rating. More transparency will help **reduce opaqueness of carbon credit rating agencies frameworks.**

Overall, many of the issues raised in this report could be mitigated by aligning oversight and guidelines for carbon credit rating agencies with financial credit rating agencies (see [chapter 3.5](#)). As reflected in [chapter 2.1](#), there is no universally agreed and in-depth definition of carbon credit quality. Carbon markets are in constant dynamic development. To date, there is no silver bullet for resolving certain issues such as leakage being at odds with additionality and the accounting difficulties for leakage and permanence. The complexity of carbon credit quality assessment is immense and without a solid and universally agreed benchmark, variety in quality and approaches to assess it will remain. In contrast, some issues identified in the respective frameworks, such as leakage being assessed but not included in the overall rating, are clearly posing a risk to the robustness of a rating.

Carbon credits are based on assumptions, probabilities, and estimates. With the improvement of practices, the accuracy of these aspects will increase as well (Hewlett 2023). This automatically leads to methodologies requiring updates and carbon credit rating agencies' frameworks as well as ratings needing reassessment. The way forward for carbon credit rating agencies is strongly tied to the **efforts of the whole market to keep the discussion on what defines a high-quality carbon credit transparent, open for public participation and scientifically sound.** Alignment and oversight of carbon credit rating agencies as well as increased transparency of their approaches could support this endeavour significantly.

6. References

Ahonen, Hanna-Mari; Michaelowa, Axel; Espelage, Aglaja; Kessler, Juliana; Christensen, Johanna; Dalfiume, Sandra; Danford, Erin (2021): Safeguarding integrity of market-based cooperation under Article 6: additionality determination and baseline setting - Background Paper, Freiburg, Perspectives Climate Research

Bellassen, Valentin; Crassous, Renaud; Dietzsch, Laura; Schwartzman Stephan (2008): Reducing emissions from deforestation and degradation: what contribution from carbon markets? PhD diss., auto-saisine

Betz, Regina; Michaelowa, Axel; Castro, Paula; Kotsch, Raphaela; Mehling, Michael; Michaelowa, Katharina; Baranzini, Andrea (2022): The Carbon Market Challenge: Preventing Abuse Through Effective Governance (Elements in Earth System Governance), Cambridge University Press, Cambridge

BeZero (2022a): BeZero Carbon Rating - 22nd November 2022, <https://bezerocarbon.com/pdf/afe76d0800/bezero-carbon-rating-methodology-22nd-november-2022.pdf> (accessed January 9, 2023)

BeZero (2022b): BeZero Carbon Rating - FAQs, <https://a.storyblok.com/f/133849/x/160f9af473/bezero-carbon-faqs-1st-june-2022.pdf> (accessed January 13, 2023)

BeZero (2022c): Additionality, <https://bezerocarbon.com/insights/bezero-carbon-risk-factor-series-additionality/> (accessed February 24, 2023)

BeZero (2022d): Leakage, <https://bezerocarbon.com/insights/bezero-carbon-risk-factor-series-leakage/> (accessed February 24, 2023)

BeZero (2022e): Over-crediting, <https://bezerocarbon.com/insights/bezero-carbon-risk-factor-series-over-crediting/> (accessed February 24, 2023)

BeZero (2023a): An introduction to our Project Developer Engagement Effort <https://bezerocarbon.com/insights/an-introduction-to-the-bezero-carbon-project-developer-engagement-effort/> (accessed March 06, 2023)

BeZero (2023b): Our new rating scale explained, https://bezerocarbon.com/insights/our-new-rating-scale-explained/?utm_medium=email&_hsmi=250040117&_hsenc=p2ANqtz--5CTJQP998fNvnYqcUQgRrgbbuTmmbtBwIT465gcvFR7J3t3YfKGOOnYszFNnfNoJlliQJaoABMdS99wOt-ij9OP55rK9_M725Afj54KJvkMc5wRbQ&utm_content=250040117&utm_source=hs_email (accessed March 27, 2023)

BeZero (2023c): Listings, <https://bezerocarbon.com/listings/> (accessed March 6, 2023)

BeZero (2023d): How additionality limits the BeZero Carbon Rating, https://bezerocarbon.com/insights/how-additionality-limits-the-bezero-carbon-rating/?utm_medium=email&_hsmi=263025645&_hsenc=p2ANqtz-8c3KBQC1ZV01aPq0Oxs7bfFOqVmCym5t7dIVdv

[G5GOGKRAMQWrdYLFQVj0hWFniz3ZvFtyDzHepHkrJnk3QldTHNa8qySFPPr3QnwPleV47-myr2rU&utm_content=263025645&utm_source=hs_email](https://www.perspectivesclimategroup.com/en/press-releases/2023/06/20/g5gogkramqwrdylfqvj0hWFniz3ZvFtyDzHepHkrJnk3QldTHNa8qySFPPr3QnwPleV47-myr2rU&utm_content=263025645&utm_source=hs_email) (accessed June 20, 2023)

Calyx Global (n.d.a): Calyx Global - Our Team, <https://calyxglobal.com/team-bios/> (accessed January 13, 2023)

Calyx Global (n.d.b): Choose Better Carbon Credits, <https://calyxglobal.com/> (accessed February 24, 2023)

Calyx Global (n.d.c): Current Calyx Risk Ratings Distribution, <https://calyxglobal.com/blog-post-dist/> (accessed May 24, 2023)

Calyx Global (n.d.d): Choose Better Carbon Credits, <https://calyxglobal.com/> (accessed June 22, 2023)

Calyx Global (n.d.e): GHG Panel Members, <https://calyxglobal.com/panel-bios/> (accessed June 22, 2023)

Calyx Global (2023a): Calyx Ratings Explained, <https://calyxglobal.com/reportviewer.php?q=TlpEdzk2aWJpRIJ2ZTVTaVhmQVZoZz09&t=r> (accessed April 11, 2023)

Calyx Global (2023b): Turning REDD into Green, <https://calyxglobal.com/reportviewer.php?q=WHp5LzUxcnp1aWJMSnJzTTJBNIJqZz09&t=r> (accessed April 11, 2023)

Calyx Global (2023c): Calyx GHG Ratings Explained, <https://calyxglobal.com/reportviewer.php?q=bDc4YWIEOHM4NnBkYnBLREJkNlp6UT09&t=r> (accessed June 22, 2023)

Calyx Global (2023d): Calyx SDG Ratings Explained, <https://calyxglobal.com/reportviewer.php?q=R1MvWkxRODZvZmZxTFZRSW1mKzRyQT09&t=r> (accessed June 22, 2023)

Calyx Global (2023e): Current Calyx Risk Ratings Distribution, <https://calyxglobal.com/blog-post-dist/> (accessed June 22, 2023)

Cames, Martin, Harthan, Ralf; Füssler, Jürg, Lazarus, Michael; Lee, Carrie; Erickson, Pete; Spalding-Fecher, Randall (2016): How additional is the clean development mechanism - Analysis of the application of current tools and proposed alternatives, Öko-Institut, Berlin

Carbon Credit Qualitative Initiative (2022): Methodology for assessing the quality of carbon credits – Version 3.0, <https://carboncreditquality.org/download/Methodology/CCQI%20Methodology%20-%20Version%203.0.pdf> (accessed January 11, 2023)

Carbon Market Watch (CMW) (2020): Up in smoke – California fires once again highlight dangers of forest offsets, <https://carbonmarketwatch.org/2020/10/22/up-in-smoke-california-fires-once-again-highlight-dangers-of-forest-offsets/> (accessed March 6, 2023)

Carbon Pulse (2023a): COMMENT: Tropical Forest conservation is tricky to measure, but we're running out of time, <https://carbon-pulse.com/189033/> (accessed January 26, 2023)

Carbon Pulse (2023b): VCM Report: REDD+ prices claw background, but large gap remains between standard and OTC credits, <https://carbon-pulse.com/191565/> (accessed February 14, 2023)

Carbon Pulse (2023c): Voluntary carbon market size grows in 2022, but buyers await new standards, <https://carbon-pulse.com/193284/> (accessed February 27, 2023)

Changeblock (2023): Changeblock and Martello Risk Form Partnership to Provide Physical On-Site Verification to Prevent "Phantom Credits" for Environmental Offsets, <https://changeblock.com/articles/changeblock-and-martello-risk-partnership> (accessed February 14, 2023)

CCQI (2022): Methodology for assessing the quality of carbon credits – Version 3.0, <https://www.carboncreditquality.org/download/Methodology/CCQI%20Methodology%20-%20Version%203.0.pdf> (accessed January 13, 2023)

Chay, Freya; Badgley, Grayson; Martin, Kata; Freeman, Jeremy; Hamman, Joe; Cullenward, Danny (2022): Unpacking ton-year accounting, CarbonPlan, <https://carbonplan.org/research/ton-year-explainer> (accessed March 6, 2023)

Clean Technica (2022): Interview with Calyx Carbon Ratings Founder Donna Lee, <https://cleantechnica.com/2022/10/25/interview-with-calyx-carbon-ratings-founder-donna-lee/> (accessed January 13, 2023)

European Commission (2008): Commission staff Working document accompanying the communication from the commission to the European parliament, the council, the European Economic and social Committee, and the Committee of the regions addressing the challenges of deforestation and forest degradation to tackle climate change and biodiversity loss -IMPACT assessment- (SEC (2008) 2619/2)

Financial IT (2022): BeZero Carbon Raises \$50 Million to Accelerate the Net Zero Transition through Carbon Ratings, <https://financialit.net/news/fundraising-news/bezero-carbon-raises-50-million-accelerate-net-zero-transition-through-carbon> (accessed February 14, 2023)

Haya, Barbara; Evans, Samuel; Brown, Letty; Bukoski, Jacob; Butsic, Van; Cabyo, Bodie; Jacobson, Rory; Kerr, Amber; Potts, Matthew; Sanchez, Daniel (2023): Comprehensive review of carbon quantification by improved forest management offset protocols, in: Frontiers in Forests and Global Change, 6

Hewlett, Owen (2023): Hindsight is a wonderful thing - certainty and the carbon markets, https://www.linkedin.com/pulse/hindsight-wonderful-thing-owen-hewlett?trk=public_post_reshare_feed-article-content (accessed March 8, 2023)

Hoch, Stephan; Greiner, Sandra; Mbaye Diagne, El-Hadji; Michaelowa, Axel; Krämer, Nicole; Espelage, Aglaja; Kassaye, Ruth (2020): Closing the deal on 'CDM Transition' – How COP25 defined new guardrails for compromise and what they mean for Africa, <https://www.climatefinanceinnovators.com/publication/closing-the-deal-on-cdm-transition/> (accessed February 14, 2023)

IC-VCM (2023): Core Carbon Principles, Assessment Framework and Assessment Procedure, <https://icvcm.org/wp-content/uploads/2023/03/CCP-Book-FINAL-28Mar23-3.pdf> (accessed March 31, 2023)

IPCC (n.d.): Land Use, Land-Use Change and Forestry – Leakage, https://archive.ipcc.ch/ipccreports/sres/land_use/index.php?idp=71 (accessed March 9, 2023)

LinkedIn (2023a): Sylvera carbon, <https://uk.linkedin.com/company/sylveracarbon> (accessed February 14, 2023)

LinkedIn (2023b): BeZero Carbon, <https://uk.linkedin.com/company/bezero-carbon> (accessed February 14, 2023)

LinkedIn (2023c): Calyx Global, <https://in.linkedin.com/company/calyx-global> (accessed July 7, 2023)

Kartha, Sivan; Lazarus, Michael; LeFranc, Maurice (2005): Market penetration metrics: tools for additionality assessment?, in: Climate Policy, 5.2, p. 147-165

Melanidis, Marina; Hagerman, Shannon (2022): Competing narratives of nature-based solutions: Leveraging the power of nature or dangerous distraction?, in: Environmental Science & Policy 132, pp. 273-281

Michaelowa, Axel; Purohit, Pallav (2007): Additionality determination of Indian CDM projects - Can Indian CDM project developers outwit the CDM Executive Board, in: Climate Strategies,

Michaelowa, Axel; Censkowsky, Philipp; Espelage, Aglaja; Singh, Aayushi; Betz, Regina; Kotsch, Raphaela; Dzukowski, Tim (2021): Volumes and types of unused Certified Emission Reductions (CERs). Lessons learned from CDM transactions under the Kyoto Protocol, transparency gaps and implications for post-2020 international carbon markets, Freiburg, [https://www.perspectives.cc/public/publications/publication/?no_cache=1&tx_pccprojects_plugin\[action\]=show&tx_pccprojects_plugin\[controller\]=Publication&tx_pccprojects_plugin\[uid\]=121](https://www.perspectives.cc/public/publications/publication/?no_cache=1&tx_pccprojects_plugin[action]=show&tx_pccprojects_plugin[controller]=Publication&tx_pccprojects_plugin[uid]=121) (accessed February 14, 2023)

Mongabay (n.d.): Market Leakage, <https://rainforests.mongabay.com/carbon-lexicon/Market-Leakage.html#:~:text=Definition%3A%20Market%20Leakage%20is%20an,actors%20to%20shift%20their%20activities> (accessed March 2, 2023)

Net Zero Marketplace (n.d): Forestry and Land Use projects, <https://netzero.salesforce.com/s/category/all-project-types/forestry-and-land-use/0ZGEZ00000008R4AQ> (accessed March 2, 2023)

Nomura, K; Mitchard, Edward; Bowers, Samuel; Patenaude, Genevieve (2019): Missed carbon emissions from forests: comparing countries' estimates submitted to UNFCCC to biophysical estimates, in: Environmental Research Letters, 14, pp.1-14

OICV-IOSCO (2015): Code of Conduct Fundamentals for Credit Rating Agencies, <https://www.iosco.org/library/pubdocs/pdf/IOSCOPD482.pdf> (accessed March 6, 2023)

Quantum Commodity Intelligence (2022): Scientist sets up 'fully transparent' carbon ratings agency, <https://www.qcintel.com/carbon/article/scientist-sets-up-fully-transparent-carbon-ratings-agency-7111.html> (accessed February 14, 2023)

Renoster (2022): Mercury Rubric - Projects, <https://www.renoster.co/projects>, (accessed February 14, 2023)

Renoster (n.d.a): About, <https://www.renoster.co/about> (accessed February 14, 2023)

Renoster (n.d.b): Whitepaper on methodology and scoring, <https://docsend.com/view/m9di2tbgdvxs7xe9> (accessed January 5, 2023)

Schneider (2011): Assessing the additionality of CDM projects: practical experiences and lessons learned, in: Climate Policy, 9.3, p. 242-254

Schwarze, Reimund; Niles, John; Olander, Jacob (2002): Understanding and managing leakage in forest-based greenhouse-gas-mitigation projects, in: Philosophical Transactions of the Royal Society of London, Series A: Mathematical, Physical and Engineering Sciences 360, pp. 1685-1703.

Scott, Vivian; Haszeldine, Stuart; Tett, Simon; Oschlies, Andreas (2015). Fossil fuels in a trillion tonne world. Nature Climate Change, 5(5), 419-423.

Streck, Charlotte (2021): REDD+ and leakage: debunking myths and promoting integrated solutions, in: Climate Policy, 21.6, pp. 843-852

Sylvera (2022a): Code of Conduct, <https://7608351.fs1.hubspotusercontent-na1.net/hubfs/7608351/1.1%20Code%20of%20Conduct%20.pdf> (accessed February 24, 2023)

Sylvera (2022b): How do we rate carbon credits? The Sylvera Carbon Credit Rating process, <https://www.sylvera.com/blog/how-to-rate-carbon-credits> (accessed February 24, 2023)

Sylvera (2022c): How we're building the world's most accurate dataset on the carbon stored in forests, <https://www.sylvera.com/blog/most-accurate-forest-carbon-dataset> (accessed January 13, 2023)

Sylvera (2022d): Introducing carbon credit pricing, <https://www.sylvera.com/blog/introducing-carbon-credit-pricing> (accessed January 13, 2023)

Sylvera (2022e): Project Proponent Relationship and Grievance Policy, <https://info.sylvera.com/hubfs/1.6%20Project%20Proponent%20Relationship%20and%20Grievance%20Policy.pdf> (accessed January 13, 2023)

Sylvera (2022f): Our Carbon Credit Ratings Framework For REDD+ Projects, <https://www.sylvera.com/resources/redd-ratings-framework> (accessed January 13, 2023)

Sylvera (2022g): Our Carbon Credit Ratings Framework For IFM Projects, <https://www.sylvera.com/resources/improved-forest-management-ratings-framework> (accessed January 13, 2023)

Sylvera (2022h): Our Carbon Credit Ratings Framework For ARR Projects, <https://www.sylvera.com/resources/arr-ratings-framework> (accessed January 13, 2023)

Sylvera (2022i): Our Carbon Credit Ratings Framework For Renewables Projects, <https://www.sylvera.com/resources/renewables-ratings-framework> (accessed January 13, 2023)

Sylvera (2023): Carbon ratings & analytics you can trust, <https://www.sylvera.com/> (accessed June 08, 2023)

Sylvera (n.d.): Sylvera Carbon Credit Ratings - Frameworks & Processes White Paper, https://7608351.fs1.hubspotusercontent-na1.net/hubfs/7608351/Sylvera_Frameworks%26Processes_Whitepaper.pdf (accessed January 6, 2023)

The Guardian (2022): Revealed: more than 90% of rainforest carbon offsets by biggest certifier are worthless, analysis shows, <https://www.theguardian.com/environment/2023/jan/18/revealed-forest-carbon-offsets-biggest-provider-worthless-verra-aoe> (accessed February 14, 2023)

Verra (2022): Methodology For Improved Forest Management Using Dynamic Matched Baselines From National Forest Inventories, <https://verra.org/wp-content/uploads/2022/12/VM0045-IFM-Dynamic-Matched-Baselines-v1.0.pdf> (accessed March 6, 2023)

Wall Street Journal (2023): Carbon-Credit Raters Aim to Assess a Notoriously Opaque Market, www.wsj.com/articles/carbon-credit-raters-aim-to-assess-a-notoriously-opaque-market-11675717717 (accessed March 7, 2023)

Wilkes, Phil; Disney, Matthias; Vicari, Matheus; Calders, Kim; Burt, Andrew (2018): Estimating urban above ground biomass with multi-scale LiDAR, in: Carbon balance and management, 13, pp. 1-2



CONTACTS

Perspectives Climate Group GmbH
Hugstetter Str. 7

79106 Freiburg, Germany
info@perspectives.cc
www.perspectives.cc

