CARBON MARKET WATCH

RATING THE RATERS

Assessing the quality of carbon credit rating agencies

11/09/2023

EXECUTIVE SUMMARY

A study commissioned by Carbon Market Watch and conducted by Perspectives Climate Group has assessed the performance of four major carbon credit rating agencies: BeZero, Calyx Global, Sylvera and Renoster. Carbon credit rating agencies are organisations that assess and rate the quality of carbon credits in the voluntary carbon market (VCM). These agencies develop and use their own frameworks to comprehensively and impartially evaluate the likelihood that a carbon credit truly represents a tonne of CO2e that has been reduced or removed by a project.

Whilst their actions support a drive for higher integrity in the market, confusion is routinely created due to assessments of the same projects being starkly different. At worst, this risks legitimising the use of low-quality carbon credits. The conclusion is that approaches to rate carbon credit quality are generally well thought out, even though there is still some room for improvement.

The report highlights discrepancies among agencies' assessments by evaluating approaches to additionality (does the project bring additional climate benefits that would not have occurred otherwise?), non-permanence risks (how likely is carbon storage to be temporary or reversed?), leakage risks (how likely are destructive activities to shift elsewhere?), and co-benefits and safeguards.

Buyers should take the time to understand how agencies operate, and what their ratings represent. For example, a "B" rating means something very different from agency to agency (above average for Calyx, but below average for Sylvera and BeZero). It is important that buyers do not disregard the agencies that provide stricter ratings. Rating agencies should be considered as a reliable source of information on credit quality that shine light on some uncomfortable truths, and not as a scapegoat for companies when they might be accused of greenwashing.

Carbon credit rating agencies contribute to improving credit quality and transparency by raising the currently low bar for quality criteria. However, they don't entirely address broader issues within the market, some of which are outside of their control. Therefore, active engagement of buyers with rating agencies is necessary to promote positive change and reliability in the VCM.

CARBON MARKET WATCH

A move away from offsetting claims will offer new opportunities for rating agencies. As buyers move away from absolute "carbon neutrality" assertions, towards more accurate and nuanced claims, they will become more interested in understanding the tradeoffs of what they are buying, for example whether they prefer to support projects that deliver durable benefits but may have a high risk of leakage, or a project that is associated with accurate quantification of impacts but low co-benefits. Rating agencies are well-equipped to provide information on the nuances behind "credit quality" that will become more important in this new contribution focused outlook.

RECOMMENDATIONS FOR RATING AGENCIES	BEZERO	CALYX GLOBAL	SYLVERA	RENOSTER
Reversal risk should be assessed over a 100-year period	👎	☺	☺	👎
Provide different non-permanence ratings based on credit vintages	👎	☺	👎	👎
If you can't measure leakage, don't assume it's 0.	☺	☺	☺	👎
The additionality score should be the maximum score a project can get.	👎	☺	👎	☺
Co-benefits should be assessed for any project, based on the SDGs.	☺	☺	☺	👎
Safeguards ratings should be incorporated into the credit quality score	👎	👎	👎	👎
Frameworks should be publicly available	😏	😏	😏	😏
Some ratings should be publicly available	😏	😏	😏	😏
Install a strong policy that monitors developments and sets a timeframe for updates of ratings.	😏	😏	😏	😏

This table lists some areas where rating agencies can improve. Note that these are suggestions for improvement, and not the quality criteria used to assess agencies.[1] Already meeting many of these recommendations demonstrates a high level of integrity from an agency, but a significant amount of red in the table below does not necessarily mean that the agency is failing its mission.

---

[1] A list of quality criteria applied in the study can be found in the report.

CARBON MARKET WATCH

INTRODUCTION

Despite growing awareness of carbon credit quality, the majority of credits circulating in the voluntary carbon market (VCM) do not live up to their promised impact.[2]

Some buyers and intermediaries attempt their own due diligence by scrutinising project details and methodologies. However, this process has proved very costly and complex. To address these concerns, carbon credit rating agencies emerged as organisations striving to assess and rate the quality of carbon credits in the VCM. These agencies have developed and use their own assessment frameworks to provide a more comprehensive and impartial evaluation of carbon credits.

In principle, rating agencies assess the likelihood that a carbon credit truly represents a tonne of carbon dioxide equivalent (CO2e) that has been reduced or removed by a project. Any classification less than the maximum score indicates that a project is likely not delivering the equivalent of one tonne of CO2e emissions reduction or removal per credit issued. Figure 1 below shows that the quality of carbon credits on the market is heterogeneous but tends towards the lower end of the spectrum.

While suggesting that there is an assortment of credit quality in the market, the graph should not be used to draw conclusions about an agency's strictness, or interpreted to show that one agency is stricter than another. For example, Sylvera has a higher share of low-rated projects than other agencies, but it has also rated fewer projects overall. It could be that all of the projects rated by Sylvera actually score poorly under the other agencies' ratings, thus making Sylvera appear less strict by comparison.

---

[2] See for example CCQI (2023): "CCQI presentation on new scores released on 31 January 2023" or The Guardian (2023): "Revealed: more than 90% of forest carbon offsets by biggest certifier are worthless, analysis shows"

Figure 1: Distribution of ratings per rating agency
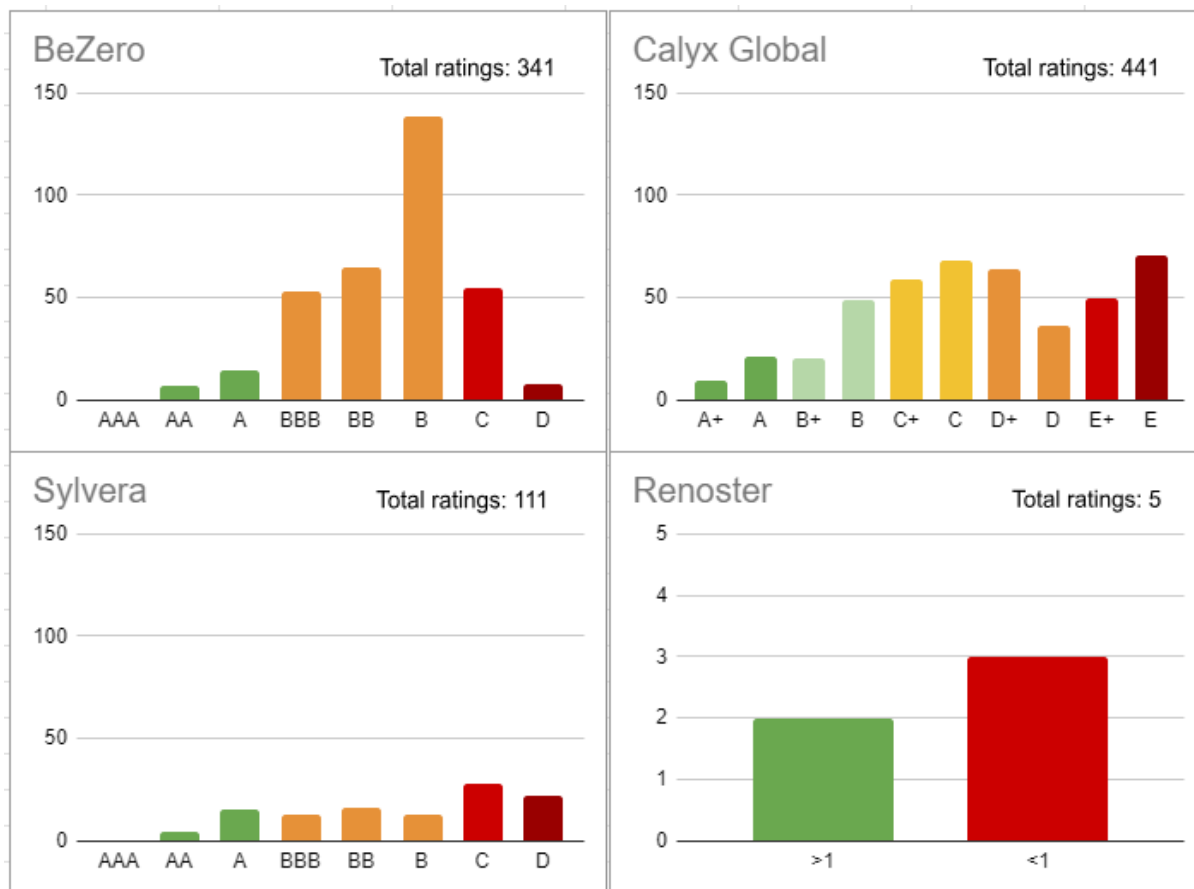
Data retrieved from agencies' websites on 9.8.2023, and from Sylvera via mail on 14.08. Note: This is a simplification of Renosters continuous numerical scoring, into a binary over/under-crediting categorisation. A score below 1 suggests overcrediting, while scores above 1 suggest undercrediting.

Figure 1 also shows that a credit's quality is not always clear from looking at a given rating, due to the scaling systems adopted by rating agencies. For example, a score of "B" might sound relatively positive, but for BeZero or Sylvera, it would actually be the 6th lowest possible score on an 8-point scale.

Figure 2 demonstrates the potential for confusion on the market as the different rating approaches lead to a range of values for a given project. (see Figure 2). For example, the Ecomapua Amazon REDD+ project received a high ranking from Sylvera, whereas Calyx and BeZero both attribute a low rating to the project.

Figure 2: Ratings for a set of projects

ECOMAPUA AMAZON REDD+ PROJECT (VCS 1094)

Scale								
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)

Scale								
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	

GUANARÉ FOREST RESTORATION PROJECT (VCS 959)

Scale											
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	0.0

*Adapted from Perspectives Climate Group (2023): "Table 20: Overlapping ratings compared on normative scale". *NZM = Net Zero Marketplace, an exchange platform on which ratings by Sylvera and Calyx are displayed in toned down tiers.*

The differences in these ratings is one of the main reasons why it is necessary to assess the work of rating agencies, and to understand what value they add to the market and, more importantly, the climate.

Drawing on the results of a study conducted on behalf of Carbon Market Watch by Perspectives Climate Group (PCG), this briefing provides recommendations to rating agencies as well as buyers.

CARBON MARKET WATCH

THE STUDY: ASSESSING AND COMPARING CARBON CREDIT RATING AGENCIES

The PCG study, assessing the rating approaches of Calyx Global, Sylvera, BeZero and Renoster, is structured around three key parts:

1. An analysis that is applicable to all project types (i.e. "cross sectoral"), and which focuses on key quality elements such as baseline setting, non-permanence, leakage, additionality, co-benefits and safeguards. This section compares the rating agencies' approaches to that of the Carbon Credit Quality Initiative (CCQI), a methodology assessment tool produced by the US chapter of the World Wide Fund for Nature (WWF-US), Environmental Defense Fund (EDF) and Öko-Institut.

2. An analysis of the rating frameworks that are specific to REDD+ avoided deforestation projects. This section offers substantial insight into how the agencies' overall framework is applied to specific project types.

3. A comparison of overlapping projects, displaying how the different frameworks result in contrasting rating scores.

CARBON MARKET WATCH

WHAT CREDIT QUALITY ELEMENTS
ARE BEING ASSESSED IN THE PCG STUDY?

Baseline setting is crucial to measuring the impact of a given project by assessing if it satisfactorily answers "what would have happened without the project?"

Non-permanence is an element specific to projects which involve the storage of carbon in sinks, and the risk that this stored carbon is released. This is important because storing carbon in a sink temporarily is proven not to generate an equivalent benefit to the climate as emissions reduction.

Leakage is what occurs when the project activity leads to (an increase in) emissions outside of the project area. Instead of truly reducing emissions, the project is shifting these to somewhere else, which would therefore not be considered a real reduction.

Additionality analyses whether or not emission reductions or removals financed through the carbon market contribute an extra climatic benefit. It establishes if a project is making an impact, or if it would have happened anyway without intervention. Additionality is therefore a crucial aspect to consider when assessing credit quality.

Co-benefits and safeguards are a set of rules and impacts related to two objectives: 1) Co-benefits - if a project delivers benefits beyond reducing emissions, such as through improved health or conservation of biodiversity. 2) Safeguards - ensuring that projects do not harm local communities, indigenous peoples, or the environment.

CARBON MARKET WATCH

FINDINGS

The report finds that while all agencies test similar criteria, several discrepancies exist. Subsequently, this sometimes leads to vastly divergent ratings for the same projects. The following section summarises various key lessons learned from the study.

- **Baseline setting is largely aligned.** The rating agencies all test for the aspects featured in the CCQI methodology. However, while other agencies set their baselines solely by evaluating the assumptions and other elements detailed by the project, Renoster first applies its own calculation of baseline emissions level before comparing this to the information in the project documentation.

- **Non-permanence requires stronger rules.** Calyx Global and Sylvera test the strength of the buffer pool used to compensate for reversals, and use a 100-year time frame to assess the non-permanence risk. Neither Renoster nor BeZero do this. Calyx is the only agency that assigns different non-permanence risk ratings to credits based on their vintage. This is an important element to consider as following the *start of a project,* most monitoring obligations end after a certain number of years. For example, a project that monitors reversals for 30 years will only be monitoring carbon stored in year 28 for two years. This creates a difference in quality, because there is only a quality assurance in place for two years for the carbon stored in year 28, whereas carbon stored in year 1 has enjoyed assurance for 29 years. In its non-permanence tests, Renoster takes a more lenient approach as it is the only agency that limits its assessment to natural risk factors, at the exclusion of those that are human caused.

- **Leakage is approached differently by all agencies.** Unlike other agencies that include leakage in their final score rating of credit quality, Renoster excludes their assessment of leakage risk from their main credit quality evaluation. Also, Sylvera does not systematically assess market leakage for all project types and tests for it only if deemed relevant. This means that market leakage is not assessed for REDD+ unplanned deforestation, for example, on the basis that unplanned deforestation is driven by subsistence farming rather than market driven demand. This may lead to underestimating the leakage risk.

- **Additionality is not assessed uniformly by all agencies.** Not only do agencies test the component aspects of additionality (e.g., financial, regulatory, and common practice additionality) in contrasting methods, but the overall

CARBON MARKET WATCH

additionality of a project is also weighed differently in final scores. Even though all agencies claim that additionality is the most important score component, this is not always reflected in the final scores. Under Renoster and Calyx, projects with a high non-additionality score can only receive the lowest grade (i.e. 0 or E, respectively). It is possible that high-risk projects receive a score of 2 out of 8 with Sylvera, and for BeZero a score of 4 out of 8 (where 8 is the best and 1 is the worst), meaning that credits with very low climate integrity could still receive a medium rating.

- **Co-benefit approaches are unaligned.** When assessing co-benefits, each agency applies their own tools and definitions. Calyx and Sylvera have a separate scoring framework for co-benefits, whereas Renoster and BeZero do not go as far, only providing the qualitative information. Furthermore, Calyx, Sylvera and BeZero use the Sustainable Development Goals (SDGs) as a matrix to assess co-benefits, whilst Renoster takes a narrower approach by focusing only on biodiversity and community gains.

- **Safeguards that would prevent the potential for harm are not incorporated into credit quality**. Not one agency considers safeguards as a component that influences the overall rating. In fact, Sylvera is the only agency which considers safeguards at all, but only as a part of the co-benefits score that is separate from the mitigation outcome score. If a project infringes upon human rights, this should be reflected in the main criteria reflecting credit quality and not only as a part of an optional separate rating.

- **Transparency is uneven and lacking.** Publicly available information on assessment frameworks and ratings varies. Whilst over time more and more information is being published, and agencies show willingness to provide information upon request, it remains the case that more information should be publicly available. Moreover, project ratings are not always made fully available to the public. On their website BeZero publishes both ratings and high level summaries of all the projects they have assessed, whilst Calyx and Sylvera publish toned down versions of some ratings on the third party carbon credit (re)seller "Net Zero Marketplace". Renoster does publish all of its ratings, but with a delay of a few months during which time the ratings are only available to customers.

TECHNICAL RECOMMENDATIONS FOR RATING AGENCIES

While rating agencies develop their frameworks based on their own considerations and assumptions, our examination demonstrates how standards could be elevated further. In highlighting the specific aspects below we deliver essential recommendations for rating agencies to consider:

- ### Reversal risk should be assessed over a 100-year period.
  Already does: Calyx, Sylvera | Should still adopt: BeZero, Renoster
  Rating agencies should use a conservative estimate over a 100-year period to check the reversal risk calculated by projects. This would still not be sufficient to guarantee equivalence between carbon storage and emission reductions, but is a step in the right direction towards better accounting of the temporary benefits associated with carbon credits.

- ### Provide different non-permanence ratings based on credit vintage.
  Already does: Calyx | Should still adopt: BeZero, Renoster, Sylvera
  This is to reflect the fact that monitoring carbon storage over the long term is vital for many reasons, most importantly to prevent any potential incentive whereby project owners log a forest they had been protecting as soon as their crediting period ends. If monitoring of reversals ends in year 80 of a project, then a credit issued for storage achieved in year 1 should not be given the same non-permanence risk rating as a credit issued for storage achieved in year 79. The former will be monitored for 79 years, while the latter is only monitored for 1 year.

- ### If you can't measure leakage, don't assume it's 0.
  Already does: BeZero, Calyx, Sylvera | Should still adopt: Renoster
  If accurate measurement is not feasible, a conservative estimation of how the risk of leakage could have impacted the mitigation outcome should influence the final rating. In rejecting leakage deduction from the final score, over-crediting is falsely legitimised.

CARBON MARKET WATCH

- **The additionality score should be set at the maximum score that a project can get.**

  Already does: Calyx, Renoster | Should still adopt: BeZero, Sylvera

  Additionality should not be considered as a binary concept (i.e. it is not either 0 or 100), but projects that rate poorly for additionality should not be able to receive a medium or even a medium-low rating.

- **For any project, co-benefits should be assessed based on the SDGs.**

  Already does: BeZero, Calyx, Sylvera | Should still adopt: Renoster

  The SDGs are a holistic set of goals, split into targets, and even indicators, that are translatable to the project context. We notice a trend for crediting programmes to require co-benefit monitoring and reporting based on the SDGs and welcome an alignment of rating agencies in this respect. The co-benefits rating could be considered as a separate rating (as Calyx does), or be incorporated into the overall rating, so long as it conforms with the SDGs. Co-benefits are the backbone of credit quality and must always be considered.

- **Safeguards ratings should be incorporated into the credit quality score.**

  Already does: none | Should still adopt: BeZero, Calyx, Renoster, Sylvera

  As a bare minimum safeguards should be assessed for credit quality and included in all rating frameworks as part of the overall credit quality score, not only the co-benefits score. Safeguards to avoid potential harm to people or the environment should be considered as fundamental as any other criterion in assessing credit quality.

- **Frameworks should be publicly available.**

  Already does: all to some extent but, requiring continuous improvement

  It is crucial for buyers of carbon credits to understand how ratings have been established, i.e. which factors does a rating agency use to define carbon credit quality. To avoid adding more confusion to the current market, frameworks for the assessment of credits should therefore be made public.

- **Some ratings should be publicly available.**

  Already does: all to some extent but, requiring continuous improvement

  Agencies should make high-level data on ratings publicly available on their own websites. Rating agencies need not necessarily provide free and full disclosure of their ratings, but toned down ratings should be explained in as transparent and granular a way as possible.

CARBON MARKET WATCH

- **Install a strong policy that monitors developments and sets a timeframe for ratings to be updated.**
  Already does: BeZero, Calyx, Sylvera | Should still adopt: Renoster
  Three of the four agencies review their ratings within an annual time frame, or sooner, if rating-relevant changes occur. We urge all rating agencies to install a structured policy that defines review frequency, a monitoring plan for global events that may influence projects, and for the assessment of project documentation. The plans should include a guide that defines which data and media should be tracked.

RECOMMENDATIONS FOR BUYERS

More often than not, projects receive "medium" to "low" scores from rating agencies, demonstrating the scrupulous quality of their evaluations relative to the liberal approval of the carbon crediting standards. Nevertheless, there are many improvements for rating agencies to take on board, and imperfect scores show that carbon credits do not provide a silver bullet solution for solving the climate crisis. It is with this context that we have formulated a series of recommendations for carbon credit buyers.

- **Don't shoot the messenger.** Strict rating frameworks by design assess risks more thoroughly. Don't focus blame on rating agencies for low scores, but rather on low-scoring projects for their shortcomings. Of the rating agencies assessed, Calyx stands out as being more conservative, with a more reliable way of assessing carbon credit quality. This is reflected in the overall lower quality scores assigned by Calyx.

- **Engage with ratings.** Try to understand what the rating truly means.

  - **Look at the scaling system.** Is A the highest score? Or AAA? To be well informed try to understand the whole dimension of the rating scale.

  - **Be careful with provisional ratings.** For example, Sylvera provides simplified provisional scores in case information is missing or incorrect. The latter poses a severe risk to rating legitimacy.

- **A tonne is not a tonne.** Don't use carbon credits to claim carbon neutrality. Even though the bar is raised by rating agencies, carbon credits are not a miraculous solution that will solve the climate crisis. Carbon credits circulating in the market by and large do not represent the genuine emissions reductions or removals that they claim, and using them for tonne-for-tonne compensation and accounting of emissions is as inaccurate as it is misguided. The purchase of carbon credits should never be communicated by buyers as directly compensating for their own climate impacts.

CARBON MARKET WATCH

PRESENT AND FUTURE ROLE OF RATING AGENCIES IN THE VCM

Contrary to crediting programmes, the bottom line of a rating agency is unaffected when they deliver a verdict on the quality of a project. This independence allows them to raise the bar, set higher standards and even be the catalyst for positive competition that stimulates project developers and crediting programmes into adopting new benchmarks or addressing project- and programme-inherent risks.

Recent developments in the VCM show that rating agencies are not the only organisations taking an interest in determining carbon credit quality. The IC-VCM, in setting a minimum bar for carbon credit quality, has effectively established a sort of rating as well, albeit a binary one.

As buyers move away from viewing carbon credits as an offsetting tool, rating agencies will be influential in assessing the weak and strong suits of projects in this new paradigm of credits used for contributions claims. Under the offsetting model, rating agency checks are akin to assessing a project on whether or not it has managed to achieve the impossible: guarantee the exact delivery and quantification of a tonne of CO_2e reduction or removal. As the ratings distribution of the agencies showed, this guarantee is extremely rare. All grades lower than the highest score underline the conviction that carbon credits should not be used for offsetting purposes and to reach climate targets. As of today, this represents more than 98% of all ratings made by the four agencies assessed in this report.

However, under a contribution model, understanding the quality of a carbon credit becomes much more meaningful. Once it is acknowledged that credits cannot be fully equivalent to a reduction of $1tCO_2e$, it becomes very important for buyers to understand what are the strengths and weaknesses of a given project, and its associated credits. This is simply because buyers can have preferences for certain project attributes. For example, a buyer might want to finance projects that have a shorter durability but higher accuracy in measurement, or prefer projects that have a long durability even though they might be prone to higher risks of leakage. Because there is no longer an expectation for absolute quantitative perfection, rating agencies can shed light on the up- and downsides of what is being bought. The rating assists in better understanding the project, and enables clear communication.

CARBON MARKET WATCH

Still, rating agencies do not address the fundamental problems of the VCM. On the one hand, rating agencies perpetuate typical VCM issues such as unaligned terminology, the high level of complexity, the lack of universal transparency guidelines, and the lack of a consolidated definition of credit quality. On the other hand, there are fundamental issues that lay outside of their influence field.

While they can play a critical role in ensuring quality, in order to demonstrate their effectiveness rating agencies must be transparent and adhere to best practices. Failure to do so undermines trust in the evaluation process and the legitimacy of carbon credits rated through their frameworks. Moreover, if during the evaluation process flawed frameworks or biassed criteria are applied, there is a risk of inadvertently legitimising unworthy carbon credits.

Ultimately, rating agencies may not solve all integrity problems within the VCM, but their impact could be significant depending on how seriously buyers take their assessments. If buyers were to actively consider the recommendations of rating agencies rather than seek out agencies that merely validate their pre-established choices, it would encourage greater credibility and responsibility in the market. In this collaborative approach, rating agencies and informed buyers could work together to drive positive change and foster a more reliable and impactful VCM.

Encouragingly, with regards to credit quality and contribution claims, a positive trend is emerging where some groups, including rating agencies, are setting stricter standards than before. We hope that a race to the top will continue to navigate this domain.

CARBON MARKET WATCH

CARBON MARKET WATCH

Authors

Benja Faecks, Policy Expert - Global Carbon Markets
Gilles Dufrasne, Policy Lead - Global Carbon Markets

Editor

Gavin Mair, Communications Officer

Cover design and layout

Noemí Rodrigo Sabio, Communications Officer

Photo credit

Tero Vesalainen

CONTACT

Benja Faecks
Policy Expert - Global Carbon Markets
benja.faecks@carbonmarketwatch.org

Gilles Dufrasne
Policy Lead - Global Carbon Markets
gilles.dufrasne@carbonmarketwatch.org