# Mitigating risk of credit reversal in nature-based climate solutions by optimally anticipating carbon release

E-Ping Rau, James Gross, David Anthony Coomes, Thomas Swinfield, Anil Madhavapeddy, Andrew Balmford & Srinivasan Keshav

Published online: 31 Aug 2024.

Submit your article to this journal ☐

Article views: 81

View related articles ☐

View Crossmark data ☐

Taylor & Francis
Taylor & Francis Group

# Mitigating risk of credit reversal in nature-based climate solutions by optimally anticipating carbon release

E-Ping Rau[a], James Gross[b], David Anthony Coomes[a], Thomas Swinfield[c], Anil Madhavapeddy[d], Andrew Balmford[c] and Srinivasan Keshav[d]

[a]Conservation Research Institute and Department of Plant Sciences, University of Cambridge, Cambridge, UK; [b]School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Stockholm, Sweden; [c]Conservation Research Institute and Department of Zoology, University of Cambridge, Cambridge, UK; [d]Conservation Research Institute and Department of Computer Science and Technology, University of Cambridge, Cambridge, UK

**ABSTRACT**
Nature-based climate solutions supply carbon credits generated from net carbon drawdown in exchange for project funding, but their credibility is challenged by the inherent variability and impermanence of drawdown. By evaluating drawdown benefits from a social cost of carbon perspective, project developers can enhance credibility and estimate impermanence by conservatively anticipating drawdowns to be eventually released following a *release schedule*, issuing additional credits when actual release is less severe than anticipated. We demonstrate how we can use *ex post* observations of drawdowns to construct optimal release schedules that limit the risk of credit reversals (when net drawdown is negative). We simulate both theoretical and real-life projects to examine how this approach balances the trade-off between generating credits evaluated as more permanent and limiting reversal risk. We discuss how this approach incentivizes project performance and provides a pragmatic solution to challenges facing larger-scale implementation of nature-based climate solutions.

## Introduction

The international community has pledged to halt deforestation as a part of a package of measures designed to achieve net zero emissions by 2050 [1–5]. In particular, REDD + projects for tropical forest conservation aim to reduce carbon emissions and concurrently provide biodiversity and social co-benefits [6]: it has been estimated that these nature-based solutions (NBS) could potentially deliver around a third of greenhouse gas emissions reduction needed to meet the 2 °C target set by the Paris Agreement [7]. NBS project developers secure funding by issuing carbon credits, which can be sold on international carbon markets to entities seeking to offset unavoidable emissions from their activities. However, the scale of the voluntary carbon market has thus far lagged behind expectations and current investment in NBS projects is woefully insufficient to contribute meaningfully towards net zero targets [8–10].

Multiple factors limit investment in NBS carbon credits [11], including the uncertainties around the real benefits of NBS projects and reputational risks associated with project failure [9,12]. It is challenging to quantify the net carbon drawdown of a project, also known as **additionality**, relative to a counterfactual scenario representing what would have happened without the project interventions. Developing quasi-experimental approaches to establish reliable counterfactual scenarios and measure drawdown more accurately remains an open area of active research [13–16].

Similarly, concerns over the **permanence** of carbon storage in NBS calls into question their value for climate change mitigation [17,18]. Forest carbon storage is lost when forest is degraded due to disturbance [19,20] or is converted to non-forest uses [21]. Higher carbon loss in a project relative to its counterfactual scenario causes release of net carbon drawdown (additionality) back into the atmosphere, during or especially after the end of the project: this means that NBS projects may only provide temporary carbon storage [12,22]. It is currently common to account for this carbon release by setting aside a

proportion of the credits generated into a buffer pool to compensate for later release events, but this approach neither adequately addresses releases occurring after the end of the project nor robustly couples the risk of carbon loss to the buffer pool size. As a result, the amount of buffer needed can be severely underestimated [23].

In contrast to debates about whether impermanent carbon storage contributes meaningfully to reduction in atmospheric carbon dioxide concentration and warming, we adopt a welfare-centered approach and focus on its social benefits in the context of the Cambridge Permanent Additional Carbon Tonne (PACT) framework. The PACT framework assesses additional carbon drawdown using a pixel-matching counterfactual analysis [24,25] and addresses impermanence by conservatively viewing *all* carbon drawdown in NBS projects as impermanent: the drawdown is projected to be released back into the atmosphere over time, either during or after the end of the project, following a **release schedule**. Both the economic benefit of drawdowns and damage of releases are calculated based on the Social Cost of Carbon (SCC) [26,27], and future damages are discounted into present-day terms [28]. The benefit of drawdowns and damage of releases are used to quantify the **Equivalent Permanence** (EP), a value ranging between 0 and 1: EP = 0 and 1 indicate a drawdown that is immediately released (generating no benefit) and permanent drawdown, respectively (see Appendix A.1 for details). The net amount of carbon credits issued is the amount of drawdown multiplied by its EP. The calculation of EP enables comparison of diverse types of carbon credits on a common scale, notably between NBS and geological storage.

In the PACT framework, credits issued for a given crediting interval are the sum of the actual carbon drawdown computed *ex post* (at the end of the crediting interval) and the amount of release that was *predicted* to occur in that interval. This adjustment is necessary because when a project developer reduces the EP of credits issued in previous periods to account for impermanence, they anticipate the quantity of releases in each future period. If the actual release in a period is smaller than predicted, the amount of unrealized predicted release is accounted for in the form of bonus credits. This means that it is possible to issue credits even in a crediting interval where the project has a negative drawdown.

In an NBS project, where carbon drawdown can vary through time and is potentially reversible, the project developer needs to balance two competing interests: generating credits that are evaluated as being as permanent as possible (i.e. higher EP), and ensuring that there are as few years as possible where no credits can be generated due to release events. This is the crux of the problem that we address with the new work presented in this study.

Since the amount of credits issued is the adjusted drawdown multiplied by its equivalent permanence, the project developer can increase the credits issued by estimating a higher equivalent permanence for the drawdown, deferring anticipated releases farther into the future. However, by doing so, they anticipate fewer releases in the immediate future, so there is a higher chance that observed releases will exceed total predicted releases in those future periods, resulting in a negative adjusted drawdown (i.e. reversal of additionality): we refer to this as **reversal** events. It is not realistic to expect NBS projects to be free from the risk of reversal in every time period. Nevertheless, projects where reversal frequently occurs are likely to be viewed as risky investments with low credibility by credit buyers, threatening the viability of the project. We posit that project developer's optimal behavior is to limit probability of reversal to a project-specific tolerable level, say 5%.

It is evident that both equivalent permanence and probability of reversal are influenced by the release schedule: a front-loaded release schedule, where a developer anticipates more releases to occur in the near future, will lower equivalent permanence, but also reduce reversal risk by increasing the amount of total predicted releases that can compensate for actual releases (see Appendix A.2 for a more formal statement of this dilemma). In this study, we resolve this trade-off between increasing permanence and reducing reversal risk by finding an **optimal release schedule** that constructs the release schedules of carbon drawdown in each year in a risk-averse strategy, so as to prevent the probability of reversal from exceeding a chosen level. We demonstrate with a simulation model how this approach is achieved, and how it ensures balance between credit issuance and reduced reversal risk. Specifically, we ask the following questions:

1. Can the optimal release schedule approach balance project performance, credit permanence and reversal risk, either in a theoretical or a real-life setting?

2. What are the effects of parameter assumptions on model output?
3. Does aggregating multiple projects to be managed as a single "portfolio" improve permanence and reduce reversal risk?
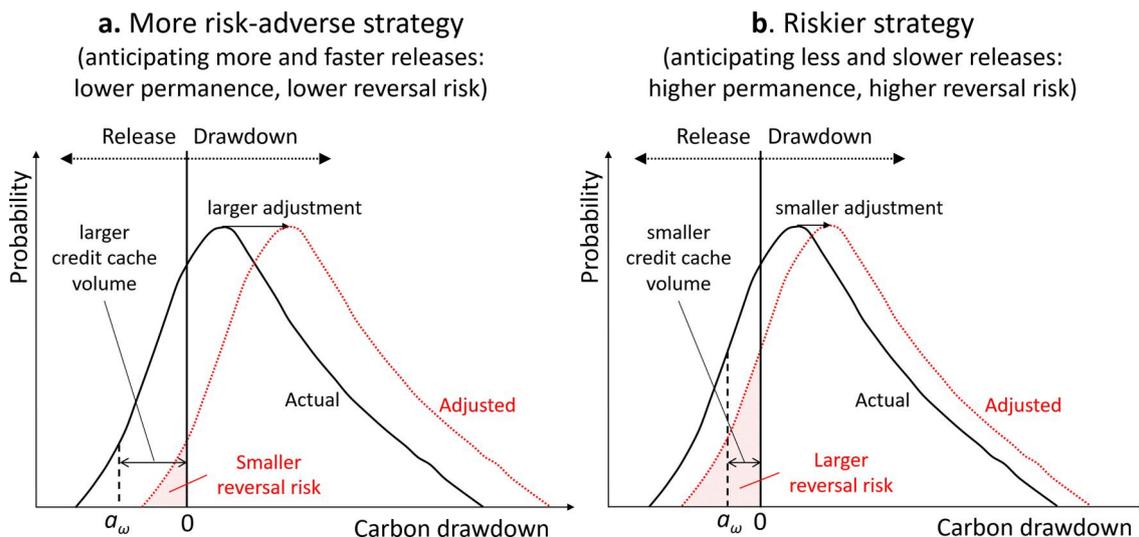
## Materials and methods

### Construction of the optimal release schedule

We model the actual carbon drawdown in year $t$ (Figure 1, black solid curve), computed at the end of year $t$, as a random variable, which may be positive in some years and negative in others, and has a distribution that reflects the uncertainty in project effectiveness. Carbon drawdown in year $t$ is expected to be released in future years $t+j$ according to a release schedule, which can be conceptualized as deposits of drawdowns into future yearly "caches": the drawdown in year $t$ that is anticipated to be released in a future year $t+j$ is viewed as credits taken from year $t$, and deposited into the cache of year $t+j$. The total adjusted drawdown available for credit issuance (Figure 1, red dotted curve) in each year $t$ is calculated as the sum of the year's actual drawdown and the total amount of previously predicted releases in the credit cache (i.e. the sum of deposits taken from drawdown in all past years $t-i$).

A reversal event occurs when the adjusted drawdown in a given year is negative: the probability of reversal is therefore the area under the probability density function of the drawdown to the left of zero. We use the left-tail percentile of the curve, representing the "worst-performing" years, to track this probability. Suppose we aim to limit probability of reversal to 5%, we can find the 5th-percentile of the actual drawdown curve ($a_\omega$, Figure 1, vertical black dashed line), such that the area under the curve to its left is equal to 0.05: this means that the 5% "worst-performing" years will have an actual drawdown lower than the 5th-percentile value. When this 5th-percentile value is negative, the probability of reversal will be larger than 5% unless it is adjusted by prior anticipated releases. By "depositing" anticipating releases, we shift the probability density function of the adjusted drawdown to the right compared to actual drawdown, reducing the probability of reversal. When we shift the curve by the difference between the 5th-percentile and zero, the area under the adjusted drawdown curve to the left of zero (Figure 1, red shaded area) will become 0.05, and the probability of reversal will be successfully limited to 5%. In other words, the absolute value of the 5th-percentile value ($|a_\omega|$, henceforth called the **credit cache volume**) is the amount of total anticipated releases that should be deposited each year to bound the reversal risk to 5%. This can be achieved by depositing issued credits of each year as anticipated releases to fill up the credit cache volumes of subsequent years successively.



**a. More risk-adverse strategy** (anticipating more and faster releases: lower permanence, lower reversal risk)

**b. Riskier strategy** (anticipating less and slower releases: higher permanence, higher reversal risk)

**Figure 1.** At any year $t$ of a project, the probability density function of the actual drawdowns (black solid curve) is derived from empirical data obtained since the start of the project to the year $t$. The probability density function of the adjusted drawdowns (red dotted curve) is derived by adjusting the actual drawdown values with the total amount of anticipated releases. If the total amount of anticipated releases equals the absolute value of the $\omega$-percentile of actual drawdowns ($|a_\omega|$, referred to as the credit cache volume), the probability of reversal (red shaded area) will be limited to $\omega$ (assuming $a_\omega < 0$). Panel (a) represents a more risk-averse strategy, where a developer anticipates a greater chance of future releases, resulting in a larger credit cache volume, lower permanence, and smaller probability of reversal ($\omega$); panel (b) represents a riskier strategy, where a developer anticipates lower releases, resulting in a smaller credit cache volume, higher permanence, and higher probability of reversal.

In Figure 2 below, we present a step-by-step illustration of how to evaluate yearly carbon drawdown and available credits and constructing optimal release schedules in order to satisfy the condition that the probability of reversal should not exceed $\omega$, assuming that the carbon drawdown distribution, and consequently the credit cache volume value, is constant over time.

**Year 1**: We measure the actual drawdown ($a_1$, dark blue solid bar). Since there are no previously anticipated releases in year 1, the number of credits available to be issued is $c_1 = a_1$ (blue bracket). We anticipate the releases of this amount of credits by allocating it successively to each future year (light blue solid bars), for each year allocating enough credits to meet the credit cache volume or until all available credits have been accounted for. We see that the anticipated release in year 2 ($\hat{r}_1^2$) has reached the credit cache volume $a_\omega$, while the

anticipated release in year 3 ($\hat{r}_1^3$) is less than this value, and that $c_1 = \hat{r}_1^2 + \hat{r}_1^3$. The equivalent permanence of $c_1$ can be calculated based on this release schedule, and the amount of PACT of $c_1 \times EP_1$ can be issued.

**Year 2**: Since there are previously anticipated releases for this year ($\hat{r}_1^2$), the credits available to be issued is the sum of the actual drawdown ($a_2$, dark red striped bar) and the previously anticipated releases: $c_2 = a_2 + \sum_i \hat{r}_i^2 = a_2 + \hat{r}_1^2$ (red bracket). Once again, we anticipate the release of this amount of credits (light red striped bars) by successively allocating it in future years. We see that in years 3 and 4, the total anticipated releases now amount to $a_\omega$, while there is still a gap between the total anticipated releases in year 5 ($\hat{r}_2^5$) and $a_\omega$.

**Year 3**: The actual drawdown is negative ($a_3 < 0$, carbon loss in the project area exceeds that of the
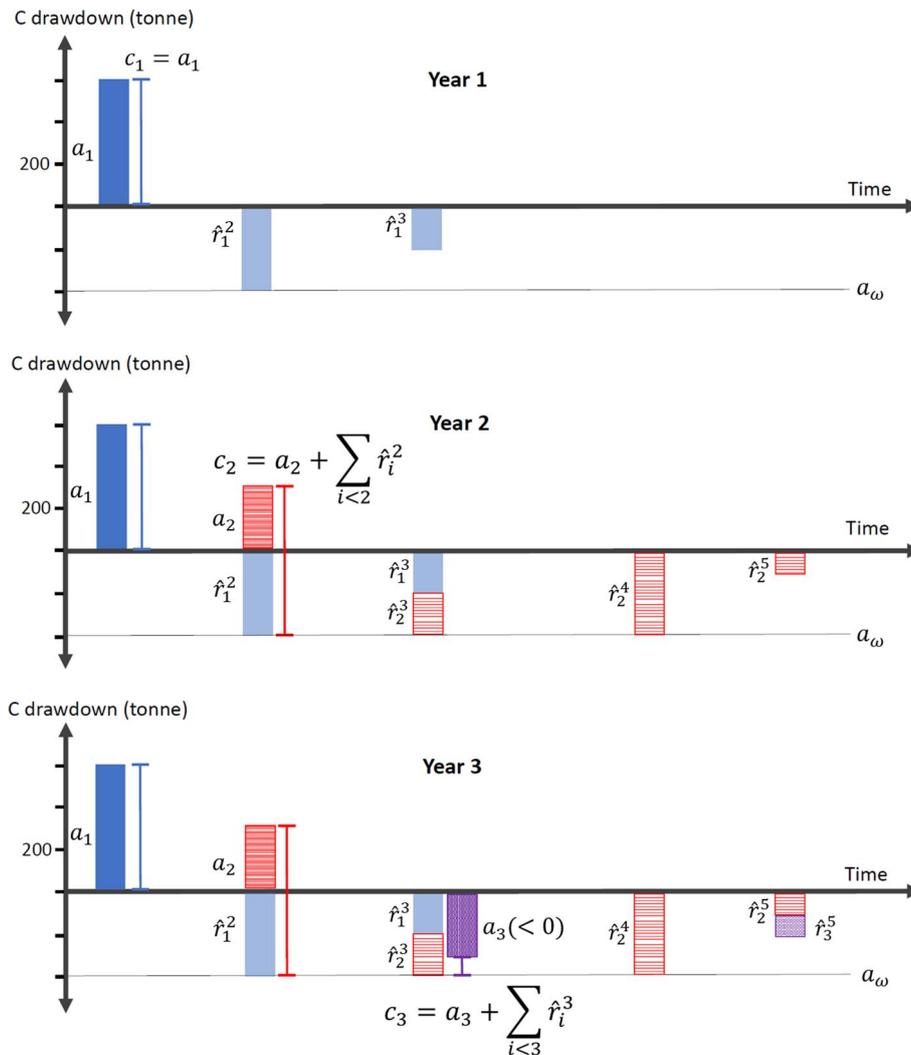


**Figure 2.** Step-by-step illustration of the construction of the release schedule. In each year $t$ (represented by a different color), the actual drawdown ($a_t$, dark-colored bar) is measured, and the credits available to be issued ($c_t$, brackets) is calculated as the sum of the $a_t$ and total previously anticipated releases ($\Sigma_i \hat{r}_i^t$, light-colored bars). The credits are then deposited in the credit caches of future years successively, such that the total anticipated releases in each year amounts to the credit cache volume ($a_\omega = 400$, thin horizontal line), and that each year's credit is fully released over time (for each color, the height of the bracket is equal to the sum of those of the light-colored bars).

counterfactual pixels, resulting in a release event) (dark purple dotted bar). However, because we had already anticipated releases to happen ($\sum_i \hat{r}_i^3$), the anticipated releases can be used to compensate for actual releases, and the credits available to be issued are the adjusted drawdown is $a_3 + \sum_i \hat{r}_i^3 = a_3 + \hat{r}_1^3 + \hat{r}_2^3 > 0$ (purple bracket). We anticipate the release of this amount of credits in the same way as in previous years (light purple dotted bar). Note that because the total anticipated releases in year 4 already amounts to $a_\omega$, the anticipated release of $c_3$ is allocated to year 5 ($\hat{r}_3^5$).

Figure 3 provides a ledger-like summary of the actual and adjusted drawdown, issued credits and credit permanence of each year in the example above. In short, we construct optimal release schedules by allocating the issued credits in each year ($c_t$) successively to future years, in each year allocating the amount necessary for the total anticipated releases in that year to amount to $a_\omega$. To calculate equivalent permanence for the credits issued every year, even near the end of the project, we anticipate releases to continue to occur past the end of the project: this also provides the possibility to issue credits even after the project ends, as long as carbon release occurs at a slower rate than previously anticipated.

If the project developer is willing to tolerate a higher percentage $\omega$ of reversal risk, the $\omega$-percentile value will be less negative, and the credit cache volume will be smaller (Figure 1b). This represents a more risky strategy where credits are deposited in the credit caches of each year in a more gradual manner, which results in higher equivalent permanence overall but also a higher

reversal risk (see Appendix A.2 for mathematical formulation and formal proof).

This approach requires the probability density function of annual net carbon drawdown (Figure 1, black solid curve) to be known, which relies on a reasonably accurate estimation of carbon fluxes in a project, for example through remote sensing technologies. In the simplest cases, parametric distributions can be used to provide an analytical approximation for the value of the credit cache volume. In more complex cases, random sampling can be used to calculate the credit cache volume more accurately as the left-tail $\omega$-percentile of the sampled drawdown distribution. This provides us with an easy-to-implement approach to construct optimal release schedules that limit reversal risk to a percentage of $\omega$ from empirical observations.

As new measurements and estimations of carbon drawdown or release are obtained every year, they can be included in the re-calculation of the carbon drawdown distribution and the credit cache volume. This approach thus allows for positive changes in project performance to be rewarded, creating incentives for project improvement: if positive carbon drawdown is observed over successive years, the drawdown distribution will shift to become more positive, meaning that the credit cache volume will become smaller with the same tolerance threshold of reversal risk. This means that less anticipated release will be deposited in each future year, leading to an increase in equivalent permanence.

### Model simulation

For all simulations in this study, we assumed annual crediting intervals for simplicity. Although

| | | During project | | | After project | | |
|---|---|---|---|---|---|---|---|
| Year | | 1 | 2 | 3 | 4 | 5 | 6 |
| Net actual drawdown ($a$) | | 600 | 300 | -300 | -200 | -200 | -200 |
| Anticipated releases ($r$) of drawdown from | | | | | | | |
| Year 1 | | | 400 | 200 | | | |
| Year 2 | | | | 200 | 400 | 100 | |
| Year 3 | | | | | 0 | 100 | |
| Year 4 | | | | | | 200 | |
| Year 5 | | | | | | | 200 |
| Total anticipated releases ($\Sigma r$) | | 0 | 400 | 400 | 400 | 400 | 200 |
| Adjusted drawdown ($c = a + \Sigma r$) | | 600 | 700 | 100 | 200 | 200 | 0 |
| Equivalent Permanence (EP) | | 0.039 | 0.053 | 0.057 | 0.029 | 0.029 | |
| Credits issued ($c \times$ EP) | | 40 | 47 | 7 | 13 | 13 | 0 |
| Total net carbon drawdown | | 600 | | | | | |
| Annual carbon drawdown rate | | 200 | | | | | |
| Post-project carbon release rate | | 200 | | | | | |

**Figure 3.** A Tabulated summary of the yearly actual drawdown, issued credits and credit permanence of the example given in Figure 2, including the period after the project ends, where the drawdown was assumed to be released at a rate equal to annual drawdown rate during the project, until all drawdowns have been released. Total anticipated releases per year is equal to the credit cache volume ($a_\omega = 400$), except for the last year where all the remaining drawdowns are released.

this differs from current common practice (e.g. five-year verification periods in the Verra methodology), our analysis can be trivially extended to permit longer crediting intervals. A similar assumption has also been used in other modeling studies [29,30]. We have released a detailed methodology for every step of the computation as well as a full implementation of our approach as open-source code [25,31,32]. The simulation code used in this study was written and performed in R 4.3.2, and the *dplyr*, *magrittr*, and *ggplot2* packages were used for data processing and visualization [33–35].

### Simulation of theoretical projects

To illustrate how the optimal release schedule approach balances project performance, credit permanence and reversal risk, we performed Monte Carlo simulations of theoretical projects with different carbon drawdown rates. In each project, carbon loss is modeled with exponential distributions, whose parameter is denoted $\lambda$. Specifically, the project carbon loss distribution is parameterized with $1/\lambda_p = 1$, and the counterfactual carbon loss distribution is parameterized with $1/\lambda_c > 1$. As $1/\lambda$ is the mean of the exponential distribution, a higher $1/\lambda_c$ value indicates higher average annual carbon loss in the counterfactual scenario compared to the annual carbon loss in the project area, which leads to an overall higher net carbon drawdown rate ($1/\lambda_c - 1/\lambda_p = 1/\lambda_c - 1$).

At each year $t$ in the simulation, we randomly drew values of carbon loss in the project ($l_p$) and counterfactual ($l_c$) from the respective exponential distributions. We calculated the drawdown as the difference between the two sampled values ($a_t = l_c - l_p$). We calculated the credit cache volume analytically, assuming it to be constant over the duration of the simulation (see Appendix B.1.1).

To approximate real-life projects where observed data on carbon losses are limited in the initial years, we defined the first five years of the simulation as the **warm-up period**, where observed drawdown and release values are used to inform the estimation of carbon loss distributions in later years. During the warm-up period, releases were ignored, and drawdowns were not issued as credits but stored in a project-level pool as an additional safeguard. In each year $t$ after the warm-up period ($t > 5$), whenever there was a gap between the credit cache volume and total anticipated releases allocated to that year, the drawdown stored in this pool was allocated to fill in the gap as additional anticipated releases, until the pool was depleted.

We calculated total credits ($c_t$) as the sum of the drawdown ($a_t$) and total anticipated releases ($\sigma_t$) (Equation A1). These credits were then allocated as anticipated releases in future years, by filling up successively the credit cache volume in each year, as illustrated in the previous section. We assumed that carbon releases continue to occur past the end of the project, at double the mean counterfactual carbon loss rate ($2/\lambda_c$), and allocated anticipated releases according to the same rule until all credits had been released. We calculated the equivalent permanence (EP) of the credits issued at each year (if any) following Equation A3, assuming $t_0 = 2021$ for the determination of Social Cost of Carbon (SCC) values.

We simulated projects with net carbon drawdown rate varying from 0.1 to 5 ($1/\lambda_c$ from 1 to 6). For each drawdown rate value, we performed 100 repetitions of simulations lasting 50 years. For each repetition, we calculated the following three statistics: (1) mean annual credits issued across all years (excluding the warm-up period), (2) maximum equivalent permanence (EP) across all years, and (3) reversal risk (proportion of years with negative credits) across all years. For illustrative purposes, we also selected three of the simulation settings where drawdown rate is 0.1, 1, 4, respectively, and plotted the yearly time series of (1) credits issued, (2) equivalent permanence (EP), and (3) reversal risk (proportion of years with negative credits over all 100 repetitions).

### Sensitivity analyses

We examined the sensitivity of simulation outputs to three model parameters: (1) length of the warm-up period, (2) post-project carbon release rate, and (3) project duration.

Firstly, based on the observation that reversal risk increases when drawdown rate is low, we posited that a longer warm-up period may help mitigate this phenomenon, and performed sensitivity analysis of reversal risk to the length of the warm-up period to test this hypothesis. We performed simulations of theoretical projects with exponential carbon loss distributions, with settings where drawdown rate is 0.1, 0.3, 0.5, and 0.7, respectively ($1/\lambda_c = 1.1, 1.3, 1.5, 1.7$), and where the length of warm-up period ranges from 1 to 20 years with steps of one year, resulting in $4 \times 20 = 80$ simulation settings in total. For each simulation setting, we performed 100 repetitions of simulations

following the same procedure as described in the first section, apart from the length of warm-up period. In all simulations, the project carbon loss distribution is parameterized with $1/\lambda_p = 1$, and the project duration was set at 50 years. We calculated the reversal risk (proportion of years with negative credits across all years) of each simulation, and plotted the reversal risk against the length of the warm-up period for each group of simulations with the same drawdown rate.

Secondly, we performed sensitivity analysis of credit permanence to post-project carbon release rate. We performed simulations of theoretical projects with exponential distributions of annual carbon loss in the project area ($1/\lambda_p = 1$) and in the counterfactual scenario ($1/\lambda_c = 1.3$). We constructed settings where the carbon release rate after the project ends was $n \times \lambda_c$, with $n$ ranging from within the interval [1,5] with steps of 0.1, resulting in 50 simulation settings in total and the post-project release rate ranging from 1.3 to 6.5. For each simulation setting, we performed 100 repetitions of simulations following the same procedure as described in the previous section, apart from the post-project carbon release rate. In all simulations, the warm-up period was set at 5 years, and the project duration was set at 50 years. We calculated the maximum equivalent permanence (EP) across all years for each simulation and plotted the maximum EP against $n$.

Lastly, we performed sensitivity analysis of credit permanence to project duration. We performed simulations of theoretical projects with exponential carbon loss distributions, with settings where the project duration ranges from within the interval [10, 100] with steps of 5, resulting in 19 simulation settings in total. For each simulation setting, we performed 100 repetitions of simulations following the same procedure as described in the previous section, apart from the project duration. In all simulations, the project carbon loss distribution is parameterized with $1/\lambda_p = 1$, the counterfactual carbon loss distribution is parameterized with $1/\lambda_c = 1.3$, the warm-up period was set at 5 years, and the post-project carbon release rate was set at $2/\lambda_c$. We calculated the maximum equivalent permanence (EP) across all years for each simulation and plotted the maximum EP against the project duration.

## Simulation of real-life projects

To exemplify how the optimal release schedule approach could be applied to real-life NBS projects, We performed Monte Carlo simulations of four ongoing REDD + projects: Rio Pepe y ACABA (RPA), Gola, Alto Mayo, and Mai Ndombe. We used satellite-based datasets from the year of the start of each real-life project ($t_0$) to 2021 to track forest cover and aboveground biomass through time, and applied a pixel-matching approach to quantify observed annual carbon losses in the project and in the counterfactual scenario. (For details, see Appendix C and the PACT Tropical Moist Forest Accreditation Methodology document [25]).

We derived net carbon drawdown distributions from the project and counterfactual carbon loss distributions, and calculated the credit cache volume with a sampling approach: in each year $t$ before 2021, we fitted statistical distributions to the carbon loss values in the project and counterfactual from the start of the project to year $t$; in each year $t$ after 2021, we fitted statistical distributions to the carbon loss values in the project and counterfactual from the start of the project to 2021 (for details of distribution fitting, see Appendix C.4). We then randomly sampled 1000 carbon loss values in the project and in the counterfactual, calculated their differences as the sampled drawdown distribution, and calculated the credit cache volume as the absolute value of its 5% percentile. As we expect the carbon loss distributions in the project area and in the counterfactual scenario to be variable through time, we also expect the credit cache volume not to be constant over time, and therefore in each year include new observed carbon loss values to update the credit cache volume. We defined the first five years of the simulation as the **warm-up period**, where observed drawdown and release values are used to inform the estimation of carbon loss distributions in later years. During the warm-up period, releases were ignored, and drawdowns were not issued as credits but stored in a project-level pool as an additional safeguard. In each year $t$ after the warm-up period ($t > 5$), whenever there is a gap between the credit cache volume and total anticipated releases allocated to that year, the drawdown stored in this pool was allocated to fill in the gap as additional anticipated releases, until the pool is depleted.

We calculated total credits ($c_t$) as the sum of the drawdown ($a_t$) and total anticipated releases ($\sigma_t$) (Equation A1). These credits were then allocated as anticipated releases in future years, by filling up successively the credit cache volume in each year, as described above. We assumed that carbon

releases continue to occur past the end of the project occurs, at double the mean observed annual carbon loss rate in the counterfactual scenario, and allocated anticipated releases according to the same rule until all credits had been released. We calculated the equivalent permanence (EP) of the credits issued at each year (if any) following Equation A3.

For each project, we performed 100 repetitions of simulations lasting 50 years, starting from the year of the start of each real-life project ($t_0$). For each repetition, we calculated the following three statistics: (1) mean annual credits across all years (excluding the warm-up period), (2) maximum equivalent permanence (EP) across all years, and (3) reversal risk (proportion of years with negative credits) across all years.

Below is a step-by-step summary of the simulation procedure of real-life projects. At each yearly time step $t$:

1. Obtain project carbon loss ($l_p$) and counterfactual carbon loss ($l_c$) from $t_0$ to $t$ (if $t \le 2021$) or from $t_0$ to 2021 (if $t > 2021$).
2. Fit both carbon loss distributions, draw 1000 random values from each, calculate sampled drawdown as the difference between the two: $a_t = l_c - l_p$.
3. Calculate credit cache volume ($|a_\omega|$) by finding the left-tail 5% percentile of the sampled drawdown distribution, with a upper bound of zero.
4. Obtain drawdown values
   a. If $t \le 2021$: use observed values at year $t$
   b. If $t > 2021$: draw random values from carbon loss distribution and calculate drawdown as the difference between the two
5. Calculate credits
   a. If $t \le 5$: zero credits in the warm-up period, but positive drawdown is placed in a project-level pool.
   b. If $t > 5$: calculate credits ($c_t$) as the sum of drawdown and total anticipated releases, which should be equal to the credit cache volume (extract drawdown from project-level pool to fill up gap if total anticipated releases is smaller than credit cache volume and if the pool hasn't been depleted).
6. If $c_t > 0$, calculate anticipated releases (release schedule) based on the optimal release allocation rule: calculate the available space in the credit cache of each following year $j$ (credit

cache volume – anticipated releases already allocated), and fill them up successively from the smallest $j$ to the largest, until all credits have been allocated to a future year.
   a. For years after the end of the project ($j > 50$), the available space for release is set to be double the mean counterfactual carbon loss rate (i.e. $a_\omega = 2/\lambda_c$).
7. If $c_t > 0$, calculate the equivalent permanence (EP) of the credits following Equation A3.

### Simulation of portfolios of aggregated projects

To test whether aggregating different projects together and managing them as a single project increases permanence and reduces reversal risk, we performed Monte Carlo simulations of aggregated theoretical or real-life projects. We adopted a simple way of aggregating projects, where the annual carbon drawdown and credits of the aggregated project is simply the sum of the carbon drawdown and credits of the individual projects: it can be imagined as being essentially a single large project with many geographically distinct project areas with different deforestation drivers and carbon loss patterns. We calculated the credit cache volume with a sampling approach: in each year, we generated random samples of carbon loss values in the project and in the counterfactual for each project, and calculated the drawdown distribution of the aggregated project ($A$) as the difference between the sum of carbon loss values over all projects and the sum of carbon loss values over all counterfactuals. We then calculated the credit cache volume as the percentile value of $A$ such that $Pr(a_t < a_\omega) = \omega$, with a upper bound of zero.

In each year $t$ before 2021, we fitted statistical distributions to the carbon loss values in the project and counterfactual from the start of the project to year $t$; in each year $t$ after 2021, we fitted statistical distributions to the carbon loss values of each project area and their respective counterfactual scenario, from the start of the project to 2021 (for details, see Appendix C.4. We then randomly sampled 1000 carbon loss values in the project and in the counterfactual, calculated their differences as the sampled drawdown distribution, and calculated the credit cache volume as the absolute value of its 5% percentile.

We first simulated the following aggregated projects consisting of four **theoretical** projects with exponentially distributed carbon loss distributions: (1) drawdown rate $= 0.1, 0.1, 0.1, 4$, respectively (one high-drawdown and three low-drawdown); (2)
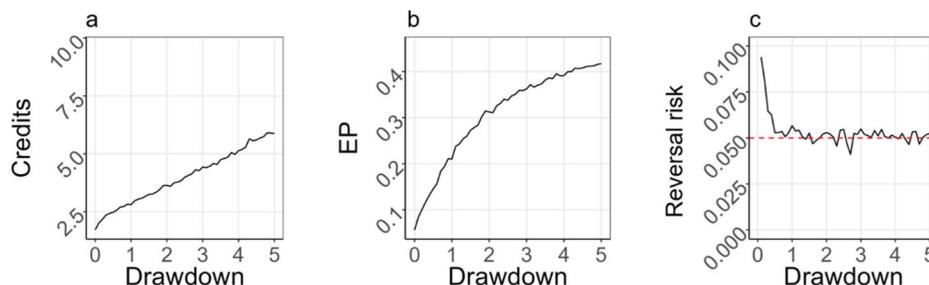
drawdown rate = 0.1, 0.1, 4, 4, respectively (two high-drawdown and two low-drawdown); (3) drawdown rate = 0.1, 4, 4, 4, respectively (three high-drawdown and one low-drawdown).

We then simulated the following aggregated projects based on the selected real-life NBS projects: (1) all four projects (Gola, Alto Mayo, RPA, Mai Ndombe); (2) three projects out of the four with the highest observed carbon drawdown rates (Gola, Alto Mayo, RPA). At each yearly time step $t$ in the simulation, we obtained yearly carbon loss values in the project and counterfactual, respectively, summed over all projects in the portfolio, and spanning the period from $t_0$ to $t$ (or to the latest year with available observations for year $t$ in the future). We performed distribution fittings for each project, drew 1000 random values from each of the fitted carbon loss distributions, respectively ($l_p^i$ and $l_c^i$), and calculated the sampled drawdown as $a_t = \sum_{i=1} l_c^i - \sum_{i=1} l_p^i$. We then calculated $a_\omega$ as the percentile value of $A$ such that $Pr(a_t < a_\omega) = \omega$, with a upper bound of zero.

## Results

### Simulation of theoretical projects

Assuming a threshold of 5% reversal risk, projects with higher carbon drawdown obtained more credits (Figure 4a, Figure 5, top row), as expected. Their credits also had higher equivalent permanence (Figure 4b, Figure 5, middle row). This can be explained by the fact that with an overall higher drawdown, the estimated credit cache volume is smaller (fewer credits need to be deposited as anticipated releases in each year to cover the risk), leading to more gradual release schedules that increase the equivalent permanence. By following the optimal release schedule, we can achieve the desired 5% reversal risk except when carbon drawdown is very close to zero (Figure 4c, Figure 5, bottom row).

### Sensitivity analyses

Increasing the length of the warm-up period consistently resulted in decreases in reversal risk until around 5%, and the lower the drawdown rate, the longer the warm-up period needed to be for reversal risk to descend to 5%. When drawdown rate is 0.1, even with a warm-up period of 20 years, the mean reversal risk stayed above 5% (Figure 6). Nevertheless, even with the shortest warm-up period (one year) and the lowest drawdown rate examined (0.1), mean reversal risk was never higher than 0.15. Increasing post-project carbon release rate slightly decreased maximum EP, but only when at its lower value range: increasing the post-project release rate above 3 did not result in further decrease in maximum EP, which reaches its minimum at 0.1 (Figure 7). Increasing the project duration led to a continuous increase in maximum EP (Figure 8).

### Simulation of real-life projects

In three of the four projects simulated (RPA, Gola, and Alto Mayo), by following the optimal release schedule, reversal risk could be limited to around the 5% threshold (Figure 9, bottom row). Lower-risk projects (RPA, Gola) consistently generated credits to be issued in all years in virtually all repetitions, whereas higher-risk projects (Alto Mayo, Mai Ndombe) had a non-negligible chance of not having credits to issue in certain years (Figure 9, top row). Lower-risk projects generated credits with high equivalent permanence in the beginning that decreases gradually over time as the project approaches its end: this can be explained by our assumption of higher carbon release rate after the project ends [22]. Higher-risk projects had credits with low equivalent permanence throughout the simulation (Figure 9, middle row).
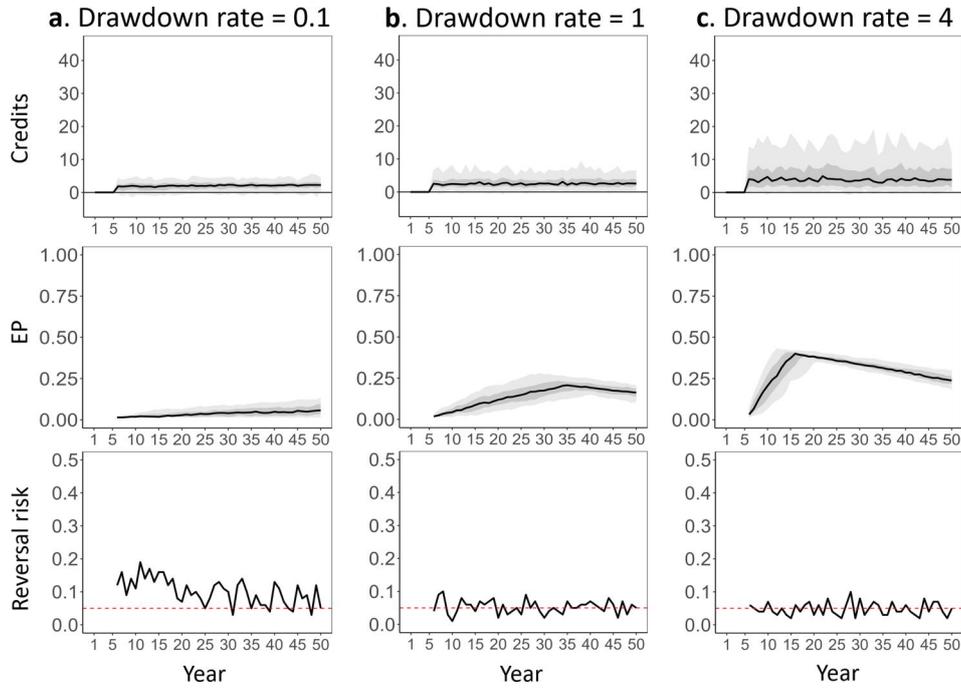


**Figure 4.** The relationship between net carbon drawdown rate (unitless) and (a) credits issued (unitless, median across all years); (b) equivalent permanence (EP) (maximum across all years); (c) reversal risk (proportion of years with negative credits across all years). The curves indicate mean values out of 100 repetitions of Monte Carlo simulation over 50 years. The horizontal red dashed line in the panel on the right indicates the 5% reversal risk threshold.

**a.** Drawdown rate = 0.1     **b.** Drawdown rate = 1     **c.** Drawdown rate = 4



**Figure 5.** Simulated yearly time series of credits (unitless), equivalent permanence (EP), and reversal risk for theoretical projects with net carbon drawdown rates (unitless) of (a) 0.1, (b) 1, and (c) 4. Light gray and dark gray shaded areas show intervals between [5%–95%] and [25%–75%] percentiles, respectively, and black curves represent median values. Horizontal black solid lines in the top row indicate zero credits, and horizontal red dashed lines in the bottom row indicate the 5% reversal risk threshold ($\omega$).
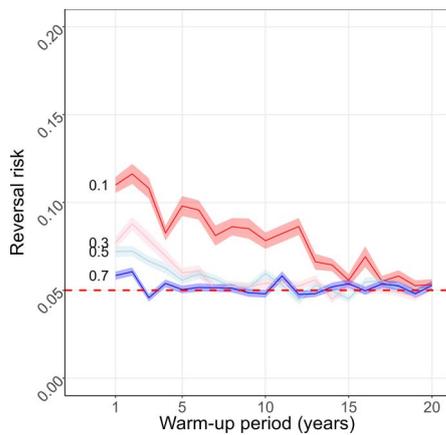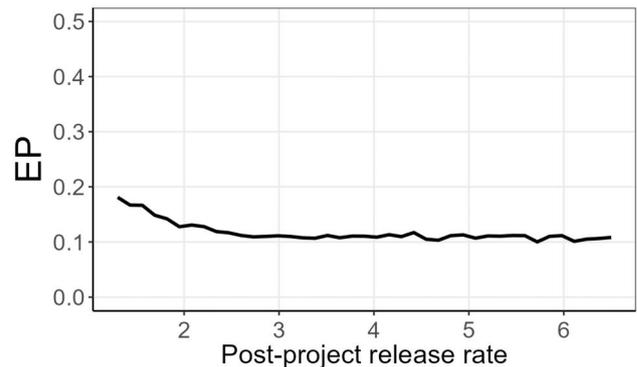


**Figure 6.** The relationship between length of warm-up period and reversal risk (proportion of years with negative credits), with drawdown rates labelled next to each curve. Curves represent mean over 1000 repetitions, and shaded areas represent the 95% confidence interval.



**Figure 7.** Effect of post-project release rate on the maximum equivalent permanence (EP) across all years of the simulation period (project duration held constant at 50 years). In all simulations, $1/\lambda_p = 1$, $1/\lambda_c = 1.3$, and the warm-up period length is five years. Black curves indicate the mean of 100 repetitions.

### Simulation of portfolios of aggregated projects

For theoretical projects, as more high-drawdown projects exist in the aggregated project, overall amount of credits issued increased, and the equivalent permanence was able to reach a higher level. The presence of even one high-drawdown project was sufficient to keep the overall reversal risk at around the 5% threshold level. (Figure 10).

For real-life projects, the inclusion of the low-drawdown Mai Ndombe project resulted in a considerably lower equivalent permanence (EP) and higher reversal risk: nevertheless, the aggregated project performs much better than the Mai Ndombe project alone, and the overall reversal risk rarely exceeds 0.1 (Figure 11a). The aggregated project containing only the two high-drawdown projects and one intermediate project performed well, with consistently high EP and low reversal risk (Figure 11b).

### Discussion

Our approach computes an optimal release schedule that a NBS project developer should adopt to limit the risks of credit reversal arising from annual variability in net carbon drawdown (additionality),

provided that they also adopt the PACT framework of evaluating the social benefits of impermanent carbon credits. We demonstrate that this approach succeeds in limiting reversal risk to a desirable level, except in projects with near-zero effectiveness at carbon drawdown. This incidentally allows us to identify ineffective projects.

Our approach dynamically adjusts the optimal release schedule based on observed carbon drawdown distributions updated with new observed data, allowing for credit permanence to be evaluated from an early stage, and for additional credits to be generated whenever the project performs better than predicted. This creates incentives for
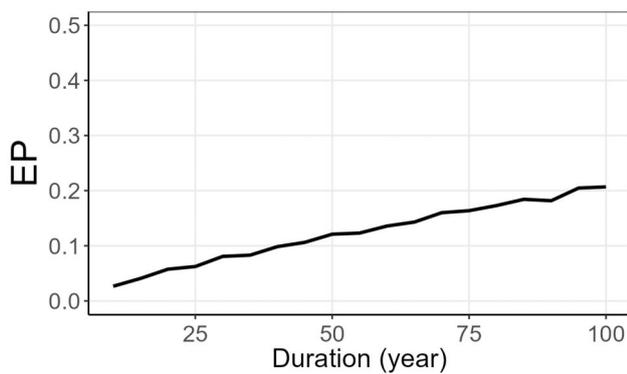
project developers and local communities to maintain long-term project custodianship, which has been identified as one of the key components needed for projects to be successful and effective [36–38]. It may also provide a form of inter-generational equity, as future custodians of a project can receive credits from safeguarding the drawdown achieved in the past, as opposed to simply being expected to look after it without any reward.

We believe this approach to be a large improvement over the current buffer pool approach, which does not base the size of the buffer pool on an empirical assessment of the expected reversal risk. In addition, in the buffer pool approach, the credits stored in the buffer pool are typically canceled at the end of the project, which neither creates incentives for project improvement during the project nor after the project ends (although mechanisms such as a long-term monitoring system for reversals have been proposed to address this issue [39]). In contrast, by estimating anticipated releases past the end of the project, our approach provides incentives even *after* the project ends.

Sensitivity analysis revealed that longer initial warm-up period could help maintain reversal risk below the desired threshold. This warm-up period was designed in our simulations as a provisional method to be able to obtain sufficient observed data on carbon losses in early years, and to build an additional safeguard against reversal events.
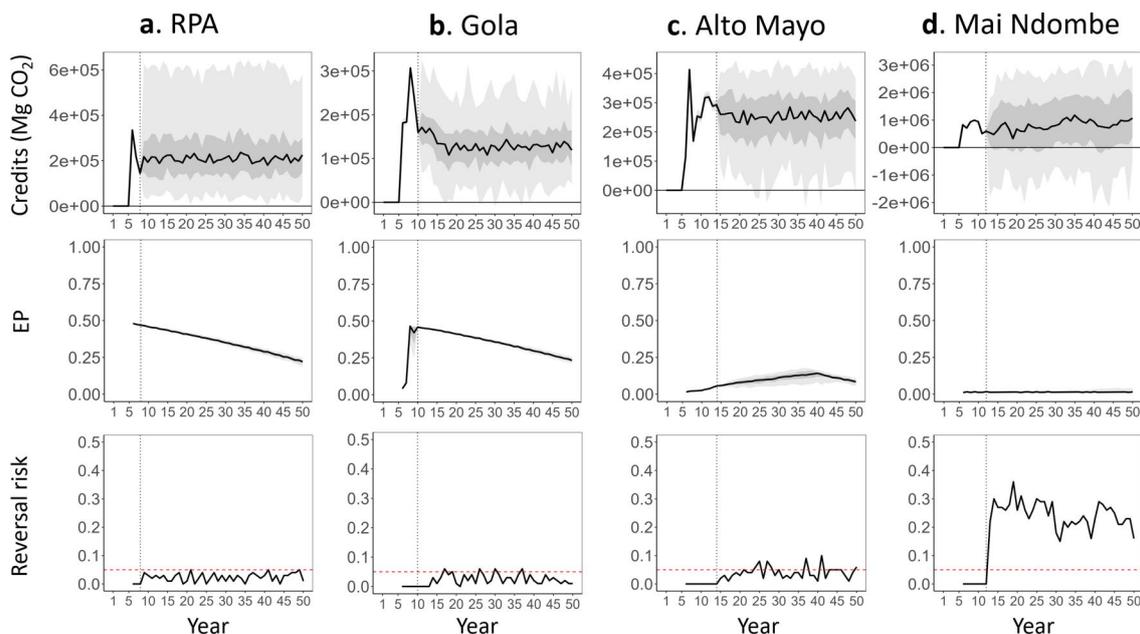


**Figure 8.** Effect of project duration on the maximum equivalent permanence (EP) across all years of the simulation period (post-project release held constant at $2/\lambda_c$).
In all simulations, $1/\lambda_p = 1$, $1/\lambda_c = 1.3$, and the warm-up period length is five years. Black curves indicate the mean of 100 repetitions.



**Figure 9.** Simulated yearly time series of credits (Mg $CO_2$), equivalent permanence (EP), and reversal risk (proportion of years with negative credits) for the four NBS projects, simulated over 50 years with 100 repetitions, starting from $t_0$ (year of start of the project). Light and dark gray shaded areas show intervals between [5%–95%] and [25%–75%] percentiles, respectively, and red curves represent median values. Horizontal black solid lines in the top row indicate zero credits, and horizontal red dashed lines in the bottom row indicate the 5% reversal risk threshold. Gray vertical dotted lines indicate year 2021 (the latest year with available remote sensing observations).
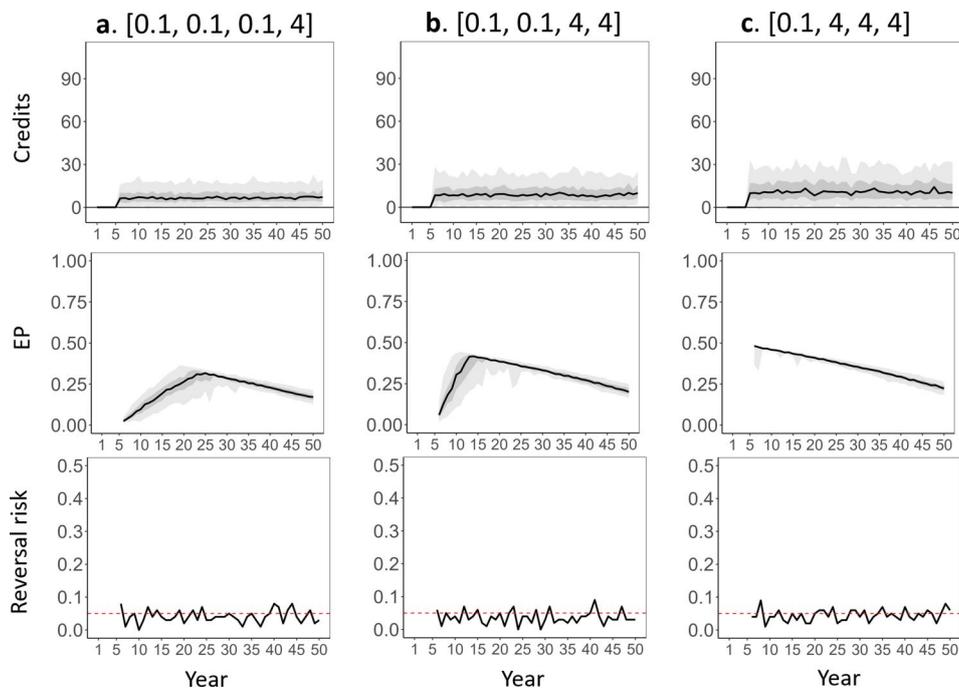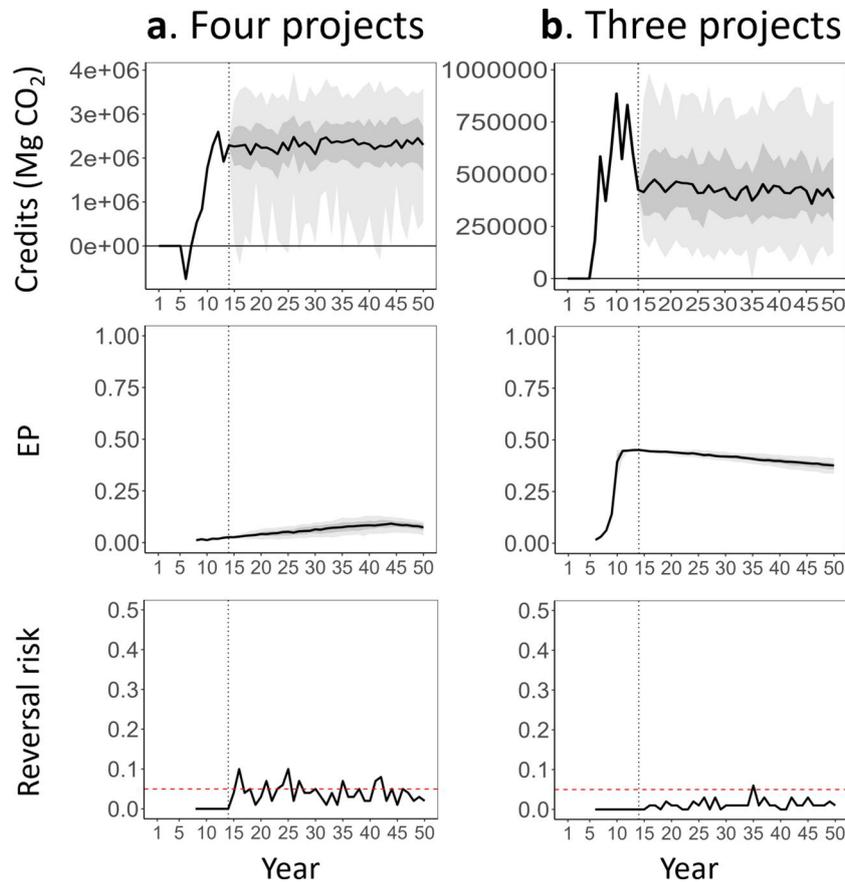
**Figure 10.** Simulated yearly time series of credits (unitless), equivalent permanence (EP), and reversal risk for aggregated projects consisting of four theoretical projects with (a) drawdown rate = 0.1, 0.1, 0.1, 4, respectively (one high-drawdown, three low-drawdown); (b) drawdown rate = 0.1, 0.1, 4, 4, respectively (two high-drawdown and two low-drawdown); (c) drawdown rate = 0.1, 4, 4, 4, respectively (three high-drawdown and one low-drawdown). Light gray and dark gray shaded areas show intervals between [5%–95%] and [25%–75%] percentiles, respectively, and red curves represent median values. Horizontal black solid lines indicate zero credits, and horizontal red dashed lines in the bottom row indicate the 5% reversal risk threshold $(1 - \omega)$.

Newer implementation of the pixel-based matching approach for drawdown estimation, which generates a large number of project and counterfactual sub-samples, may eliminate the need to define a warm-up period because the sub-samples could be used to estimate carbon drawdown distributions even in early years. Nevertheless, in cases where remote sensing approach is not adequate (due to lower quality of remote sensing data, for example), alternative methods such as community-based monitoring may be employed, and time may be need for data availability and quality to improve [40]. Under these circumstances, implementing a warm-up period may be a reasonable approach. Its length represents an additional trade-off: longer warm-up period means lower reversal risks, but at a cost of delaying the start of credit issuance, which is a challenge that many projects currently face [8]. Strengthening infrastructural and financial support in forest monitoring and project planning and implementation in the initial stages, such as *via* readiness-related programs and funds [41], will be crucial in overcoming this challenge.

We also found a clear positive effect of longer project duration on credit permanence (SI **section F**). Given that project duration can be limited by uncertainty of the stakeholders about the durability of funding flows [37], we propose that rather than making overly optimistic claims about long project duration, continual and dynamic monitoring should be leveraged to demonstrate project improvement and update credit permanence, so that the more steady income flow itself can become an incentive for risk-averse local communities to maintain participation [42].

This approach is straightforward to implement as long as forest monitoring provides reasonably accurate estimates of carbon drawdown, which can be facilitated by advances in remote sensing technology and accumulation of monitoring data, in particular those who provide more accurate estimates of forest biomass [43]. Using publicly available remote sensing can also help alleviate the cost of forest monitoring, especially when it is needed over a large spatial scale [44,45]. Moreover, a combination of spatially explicit data on drivers of forest carbon loss [46] and models that predict forest dynamics (e.g. growth and mortality), disturbance risks, and carbon loss patterns in response to climate change [20], could help evaluate credit permanence by improving project-counterfactual matching in disturbance regime, detecting temporal non-stationarity of carbon drawdown distributions, and infer the impact of rare but severe disturbance events.

**Figure 11.** Simulated yearly time series of credits (Mg $CO_2$), equivalent permanence (EP), and reversal risk for aggregated projects consisting of (a) four projects and (b) three projects with the highest observed net carbon drawdown rates out of the four. The time series starts from the earliest $t_0$ (year of start of the project) out of all aggregated projects. Light gray and dark gray shaded areas show intervals between [5%–95%] and [25%–75%] percentiles, respectively, and red curves represent median values. Horizontal black solid lines in the top row indicate zero credits, and horizontal red dashed lines in the bottom row indicate the 5% reversal risk threshold.

Our approach also easily allows for multiple projects to be aggregated and evaluated as a single project "portfolio," which leads to higher permanence and lower reversal risk than could be achieved in each project individually. Aggregating projects (especially those that are diverse in geographic location and methodology) and assessing their ongoing risks with observed data allows credit buyers to manage risk sensibly in the volatile carbon market. It safeguards each individual project against rare but catastrophic disturbances (as the chance that multiple projects will be affected by the same kind of sporadic catastrophe at the same time is exceedingly low), especially for projects considered as "high-risk" either due to being at an early stage of implementation or having high deforestation pressure. Aggregating projects also increases the amount of observed data that can be utilized to inform drawdown distributions, especially where long-term data are still scarce [24,47].

In this study, we optimized release schedules given the constraint of a maximum 5% reversal risk, a simple threshold that represents no more than two reversal events during a 40-year period.

Future research directions include exploring the implications of jointly optimizing the release schedule and the maximum reversal risk threshold, for example by maximizing the expected revenue from credit sales, assuming that reversal events cause the credit price to fall.

Although our study focuses on avoided deforestation projects, the central principle of issuing impermanent credits with anticipated future releases can be applied to other types of NBS projects, provided that we can evaluate their benefits with counterfactual analysis and anticipate costs of future reversal of those benefits. For an example, for an afforestation project, we can model carbon releases as following an extreme value distribution, estimate the amount of releases expected to occur (return level) within a given time period [48], and allocate anticipated releases such that the long-term cumulative anticipated releases in a given period match the expected return level.

While there are significant differences between our proposed approach and the common practice in current methodologies used by certifying standards (e.g. Verra VM0048), we believe that active

discussions between academics, certifying organizations, and project developers will be instrumental in helping bridge this gap and increasing appreciation for evaluating impermanence through the lens of social cost of carbon. The development of an easy-to-use toolkit that automates the modelling process will further facilitate the adoption of our approach by project developers.

In conclusion, our study proposes a novel yet simple way to anticipate risks of carbon release and evaluate permanence of carbon credits generated by nature-based solutions, achieving a balance between sustainable income flows and project credibility and incentivizing project improvement. Assuming that the social benefits of impermanent carbon credits are increasingly being valued, our findings offer a pragmatic solution to improve market confidence in NBS projects, paving the way for scaling up the implementation of this crucial climate mitigation strategy.

## Acknowledgments

## Authors' contributions

S. K. and D. A. C. supervised the project. E-P. R., S. K., J. G. and D. A. C. conceived the research plan. S. K. and J. G. constructed the mathematical and theoretical models. T. S. assembled input data. E-P. R. wrote code, ran model simulations, and analyzed simulation outputs. E-P. R., S. K. and D. A. C. wrote the manuscript. All co-authors revised the manuscript.

## Disclosure statement

A. B. is a trustee of the World Land Trust, a non-governmental organization that supports forest-based carbon projects. The Cambridge Centre for Carbon Credits (4C) has no commercial interest in carbon credits.

## Funding

## Data availability statement

The data that support the findings of this study are openly available in Apollo, the University of Cambridge repository, at https://doi.org/10.17863/CAM.110933.

## References

1. The Paris Agreement. Accessed: 2023-12-12; Available from: https://unfccc.int/process-and-meetings/the-paris-agreement.
2. Dinerstein E, Vynne C, Sala E, et al. A global deal for nature: guiding principles, milestones, and targets. Sci Adv. 2019;5(4):eaaw2869. doi: 10.1126/sciadv.aaw2869.
3. Crowe O, Beresford AE, Buchanan GM, et al. A global assessment of forest integrity within key biodiversity areas. Biol Conserv. 2023;286:110293. doi: 10.1016/j.biocon.2023.110293.
4. Mitchard ETA. The tropical forest carbon cycle and climate change. Nature. 2018;559(7715):527–534. doi: 10.1038/s41586-018-0300-2.
5. Canadell JG, Raupach MR. Managing forests for climate change mitigation. Science. 2008;320(5882):1456–1457. doi: 10.1126/science.1155458.
6. Angelsen A, Martius C, Sy V, et al. Introduction: REDD+ enters its second decade. In: Transforming REDD+: Lessons and new directions. Indonesia: Center for International Forestry Research; 2018. p. 1–13. doi: 10.17528/cifor/007045.
7. Griscom BW, Adams J, Ellis PW, et al. Natural climate solutions. Proc Natl Acad Sci U S A. 2017;114(44):11645–11650. doi: 10.1073/pnas.1710465114.
8. Atmadja SS, Duchelle AE, Sy VD, et al. How do REDD+ projects contribute to the goals of the Paris Agreement? Environ Res Lett. 2022;17(4):044038. doi: 10.1088/1748-9326/ac5669.
9. Laing T, Taschini L, Palmer C. Understanding the demand for REDD+ credits. Envir Conserv. 2016;43(4):389–396. doi: 10.1017/S0376892916000187.
10. Seddon N, Chausson A, Berry P, et al. Understanding the value and limits of nature-based solutions to climate change and other global challenges. Philos Trans R Soc Lond B Biol Sci. 2020;375(1794):20190120. doi: 10.1098/rstb.2019.0120.
11. CDP. Harnessing the potential of the private sector to deliver REDD+; 2018. Policy briefing https://www.greenindustryplatform.org/sites/default/files/downloads/resource/REDD__Policy_Briefing_EN.pdf.
12. Joppa L, Luers A, Willmott E, et al. Microsoft's million-tonne $CO_2$-removal purchase—lessons for net zero. Nature. 2021;597(7878):629–632. doi: 10.1038/d41586-021-02606-3.
13. West TAP, Caviglia-Harris JL, Martins FSRV, et al. Potential conservation gains from improved protected area management in the Brazilian amazon. Biol Conserv. 2022;269:109526. doi: 10.1016/j.biocon.2022.109526.
14. West TAP, Wunder S, Sills EO, et al. Action needed to make carbon offsets from forest conservation work for climate change mitigation. Science. 2023;381(6660):873–877. doi: 10.1126/science.ade3535.

15. Schleicher J, Eklund J, D Barnes M, et al. Statistical matching for conservation science. Conserv Biol. 2020;34(3):538–549. doi: 10.1111/cobi.13448.

16. Guizar-Coutiño A, Jones JPG, Balmford A, et al. A global evaluation of the effectiveness of voluntary REDD + projects at reducing deforestation and degradation in the moist tropics. Conserv Biol. 2022; 36(6):e13970. doi: 10.1111/cobi.13970.

17. Matthews HD, Zickfeld K, Dickau M, et al. Temporary nature-based carbon removal can lower peak warming in a well-below 2 °C scenario. Commun Earth Environ. 2022;3(1):1–8. doi: 10.1038/s43247-022-00391-z.

18. Matthews HD, Zickfeld K, Koch A, et al. Accounting for the climate benefit of temporary carbon storage in nature. Nat Commun. 2023;14(1):5485. doi: 10.1038/s41467-023-41242-5.

19. Anderegg WRL, Trugman AT, Badgley G, et al. Climate-driven risks to the climate mitigation potential of forests. Science. 2020;368(6497):11. doi: 10.1126/science.aaz7005.

20. Anderegg WRL, Wu C, Acil N, et al. A climate risk analysis of earth's forests in the 21st century. Science. 2022; 377(6610):1099–1103. doi: 10.1126/science.abp9723.

21. Busch J, Ferretti-Gallon K. What drives and stops deforestation, reforestation, and forest degradation? an updated meta-analysis. Rev Environ Econ Policy. 2023;17(2):217–250. doi: 10.1086/725051.

22. Carrilho CD, Demarchi G, Duchelle AE, et al. Permanence of avoided deforestation in a transamazon REDD + project (par'a, brazil). Ecol Econ. 2022; 201:107568. doi: 10.1016/j.ecolecon.2022.107568.

23. Badgley G, Chay F, Chegwidden OS, et al. California's forest carbon offsets buffer pool is severely undercapitalized. Front For Glob Change. 2022;5:930426. doi: 10.3389/ffgc.2022.930426.

24. Swinfield T, Balmford A. Cambridge carbon impact: evaluating carbon credit claims and co-benefits; 2023. Early or alternative research output on Cambridge Open Engage at https://www.cambridge.org/engage/coe/article-details/6409c345cc600523a3e778ae.

25. Balmford A, Coomes D, Hartup J, et al. PACT tropical moist forest accreditation methodology. 2023; Early or alternative research output on Cambridge Open Engage at https://www.cambridge.org/engage/coe/article-details/657c8b819138d23161bb055f.

26. Nordhaus W. Estimates of the social cost of carbon: concepts and results from the dice2013r model and alternative approaches. J Assoc Environ Resour Econ. 2014;1(1/2):273–312. doi: 10.1086/676035.

27. Groom B, Venmans F. The social value of offsets. Nature. 2023;619(7971):768–773. doi: 10.1038/s41586-023-06153-x.

28. Balmford A, Keshav S, Venmans F, et al. Realising the social value of impermanent carbon credits. Nat Clim Chang. 2023;13(11):1172–1178. doi: 10.1038/s41558-023-01815-0.

29. Cacho OJ, Hean RL, Wise RM. Carbon-accounting methods and reforestation incentives. Aus J Agri & Res Econ. 2003;47(2):153–179. doi: 10.1111/1467-8489.00208.

30. Chan KK, Golub A, Lubowski R. Performance insurance for jurisdictional REDD+: Unlocking finance and increasing ambition in large-scale carbon crediting systems. Front For Glob Change. 2023;6:1062551. doi: 10.3389/ffgc.2023.1062551.

31. Rau EP, 2024. GitHub repository at https://github.com/epingchris/assess_permanence. commit:a4d52f77f8043

32. Rau EP, Gross J, Coomes DA, et al. Research data supporting "Mitigating risk of credit reversal in nature-based climate solutions by optimally anticipating carbon release. Apollo - University of Cambridge Repository. 2024; doi: 10.17863/CAM.110933.

33. Wickham H, François R, Henry L, et al. dplyr: a grammar of data manipulation; 2023; R package version 1.1.3; Available from: https://CRAN.R-project.org/package=dplyr.

34. Bache SM, Wickham H. magrittr: a forward-pipe operator for R; 2022; R package version 2.0.3; Available from: https://CRAN.R-project.org/package=magrittr.

35. Wickham H. ggplot2: elegant graphics for data analysis. New York: Springer-Verlag; 2016. Available from: https://ggplot2.tidyverse.org.

36. Fortmann L, Salas PC, Sohngen B, et al.; 2014 Incentive contracts for environmental services and their potential in REDD. Policy Research Working Papers. (2419785).

37. Wunder S, Duchelle AE, Sassi C, et al. RREDD + in theory and practice: how lessons from local projects can inform jurisdictional approaches. Front For Glob Change. 2020;3:499592. doi: 10.3389/ffgc.2020.00011.

38. Buchholz T, Gunn J, Springsteen B, et al. Probability-based accounting for carbon in forests to consider wildfire and other stochastic events: synchronizing science, policy, and carbon offsets. Mitig Adapt Strateg Glob Change. 2021;27(1):4. doi: 10.1007/s11027-021-09983-0.

39. Verra. Summary of public comments: proposal to create a long-term reversal monitoring system. Accessed: 5-3-2024; Available from: https://verra.org/wpcontent/uploads/LTRMS-summary-comments.pdf.

40. Brofeldt S, Theilade I, Burgess ND, et al. Community monitoring of carbon stocks for REDD+: does accuracy and cost change over time? Forests. 2014;5(8): 1834–1854. doi: 10.3390/f5081834.

41. Andoh J, Oduro KA, Park J, et al. Towards REDD + implementation: Deforestation and forest degradation drivers, REDD + financing, and readiness activities in participant countries. Front For Glob Change. 2022;5:957550. doi: 10.3389/ffgc.2022.957550.

42. Chesney M, Gheyssens J, Troja B. Market uncertainty and risk transfer in REDD projects. J Sustainable For. 2017; 36(5):535–553. doi: 10.1080/10549811.2017.1326940.

43. Calders K, Jonckheere I, Nightingale J, et al. Remote sensing technology applications in forestry and REDD+. Forests. 2020;11(2):188. doi: 10.3390/f11020188.

44. De Sy V, Herold M, Achard F, et al. Synergies of multiple remote sensing data sources for REDD+ monitoring. Curr Opin Environ Sustainability. 2012;4(6):696–706. doi: 10.1016/j.cosust.2012.09.013.

45. Demarchi G, Subervie J, Catry T, et al. Using publicly available remote sensing products to evaluate REDD+ projects in Brazil. Global Environ Change. 2023; 80:102653. doi: 10.1016/j.gloenvcha.2023.102653.

46. Lapola DM, Pinho P, Barlow J, et al. The drivers and impacts of Amazon forest degradation. Science. 2023; 379(6630):eabp8622. doi: 10.1126/science.abp8622.

47. Herold M, Skutsch M. Monitoring, reporting and verification for national REDD+ programmes: two proposals. Environ Res Lett. 2011;6(1):014002. doi: 10.1088/1748-9326/6/1/014002.

48. Katz RW, Brush GS, Parlange MB. Statistics of extremes: modeling ecological disturbances. Ecology. 2005;86(5):1124–1134. doi: 10.1890/04-0606.

49. Drupp MA, Freeman MC, Groom B, et al. Discounting disentangled. Am Econ J. 2018;10(4):109–134. doi: 10.1257/pol.20160240.

50. Valuation of greenhouse gas emissions: for policy appraisal and evaluation. Accessed: 2023-12-11; Available from: https://www.gov.uk/government/publications/valuinggreenhouse-gas-emissions-in-policy-appraisal/valuation-of-greenhouse-gas-emissions-for-policy-appraisal-and-evaluation.

51. Klar B. A note on gamma difference distributions. J Stat Comput Simul. 2015;85(18):3708–3715. doi: 10.1080/00949655.2014.996566.

52. Verra VCS. Registry. Available from: https://registry.verra.org/app/search/VCS.

53. Verra VCS. Registry, project 944. Available from: https://registry.verra.org/app/projectDetail/VCS/944.

54. Verra VCS. Registry, project 1201. Available from: https://registry.verra.org/app/projectDetail/VCS/1201.

55. Verra VCS. Registry,. project1396. Available from: https://registry.verra.org/app/projectDetail/VCS/1396.

56. Verra VCS. Registry, project 934. Available from: https://registry.verra.org/app/projectDetail/VCS/934.

57. Vancutsem C, Achard F, Pekel JF, et al. Long-term (1990–2019) monitoring of forest cover changes in the humid tropics. Sci Adv. 2021;7(10):eabe1603. doi: 10.1126/sciadv.abe1603.

58. Dubayah R, Blair JB, Goetz S, et al. The global ecosystem dynamics investigation: highresolution laser ranging of the earth's forests and topography. Sci Remote Sens. 2020;1:100002. (doi: 10.1016/j.srs.2020.100002.

59. GEDI L4A footprint level aboveground biomass density, version 1. Accessed: 2023-12-11; Available from: https://daac.ornl.gov/GEDI/guides/GEDIL4AAGBDensity.html.

60. Duncanson L, Kellner JR, Armston J, et al. Aboveground biomass density models for NASA's global ecosystem dynamics investigation (GEDI) lidar mission. Remote Sens Environ. 2022;270:112845. doi: 10.1016/j.rse.2021.112845.

61. Cairns MA, Brown S, Helmer EH, et al. Root biomass allocation in the world's upland forests. Oecologia. 1997;111(1):1–11. doi: 10.1007/s004420050201.

62. Penman J, Gytarsky M, Hiraishi T, et al. 2003; Good practice guidance for land use, land-use change and forestry. Institute for Global Environmental Strategies.

63. Martin AR, Thomas SC. A reassessment of carbon content in tropical trees. PLoS One. 2011;6(8):e23533. doi: 10.1371/journal.pone.0023533.

64. Large scale international boundaries (LSIB). Accessed: 2023-12-11; Available from: https://data.humdata.org/dataset/large-scale-international-boundaries-lsib?

65. Dinerstein E, Olson D, Josh A, et al. An ecoregion-based approach to protecting half the terrestrial realm. Bioscience. 2017;67(6):534–545. doi: 10.1093/biosci/bix014.

66. Mahalanobis PC. On the generalized distance in statistics. Sankhyā: The Indian Journal of Statistics, Series A (2008-). 2018;80: S1–S7.

67. Jarvis A, Reuter HI, Nelson A, et al. 2008; Hole-filled SRTM for the globe version 4, Available from the CGIAR-CSI SRTM 90m database. https://research.utwente.nl/en/publications/hole-filled-srtm-for-the-globe-version-4data-grid.

68. Weiss DJ, Nelson A, Gibson HS, et al. A global map of travel time to cities to assess inequalities in accessibility in 2015. Nature. 2018;553(7688):333–336. doi: 10.1038/nature25181.

69. Scrucca L, Fop M, Murphy TB, et al. 2016. mclust 5: clustering, classification and density estimation using Gaussian finite mixture models;.

70. Millard SP. EnvStats: an R package for environmental statistics. New York: Springer; 2013.

## Appendices

Appendix A provides formal definition of how credits and their equivalent permanence are calculated under the PACT Framework [28] and mathematical representations of the derivation of the optimal release schedules. Appendix B provides analytical solutions of the **credit cache volume**, the key quantity for the construction of optimal release schedule, for special cases when the carbon loss distributions in the project area and in the counterfactual scenario of a project can be expressed as simple parametric distributions. Appendix C provides detailed information on the method used to quantifying annual carbon fluxes in the four real-life REDD+ projects examined. A complete and detailed description of the methods used can also be found in the PACT Tropical Moist Forest Accreditation Methodology document [25].

## Appendix A. Mathematical representation of release schedule and equivalent permanence

### A.1. Release schedule and equivalent permanence

It is widely assumed that carbon sequestration in NBS projects is not permanent, but will be released back into the atmosphere over time: we can account for this by 'depositing' the credits gained in the time period $t$ to 'credit caches' in the future years, and anticipating the amount to be released in future time periods $t+j$ to be $\hat{r}_t^{t+j}, j > 0$. This set of variables is called the **release schedule**, and it is a set of free decision variables to be chosen by the project developer. The choice of a release schedule plays a critical role in three computations.

First, the actual number of credits issued in time period $t$, denoted $c_t$, is equal to the drawdown in time period $t$, adjusted by the amount of total releases anticipated to happen in time period $t$, and thus depends on past release schedules through:

$$c_t = a_t + \sum_{t \geq j > 0} \hat{r}^t_{t-j} = a_t + \sigma_t \qquad \text{(A1)}$$

If past releases were more conservative, then more credits are available to compensate future negative drawdown (release) events.

Second, the probability of reversal of the $i^{th}$ project at time step $t$ depends on whether the observed release exceeds the sum of prior anticipated releases:

$$Pr(\sigma_t{}^i + a^i_t > 0) = Pr(-a^i_t < \sigma_t{}^i) \qquad \text{(A2)}$$

If the release schedule is not risk-averse enough, then a credit purchaser may not purchase the credits from the project due to their concerns about the high occurrences of reversal events.

Finally, the price paid to the project developer depends critically on the **equivalent permanence** (EP). Equivalent permanence is calculated based on the release schedule and the Social Cost of Carbon (SCC) associated with carbon drawdown and release, and the discounting of future damage [28]. The equivalent permanence is the ratio between the value of a given amount of impermanent credits and that of the same amount of permanent credits, such as those from geological storage, and is bound between [0, 1]. under the PACT framework, carbon credits produced in year $t$ ($c_t$) needs to be adjusted downwards to become PACT by multiplying the amount of the credit by its Equivalent Permanence (EP).

Given a discount factor $\delta$, the current value of damage from future release and the Equivalent Permanence are given by:

$$d_t = \sum_{j > 0} \frac{SCC(\hat{r}^{t+j}_t)}{(1 + \delta)^j}$$
$$EP = \frac{SCC(c_t) - d_t}{SCC(c_t)} \qquad \text{(A3)}$$

With a more rapid release schedule, the damage from releases would be discounted less, and the numerator will be smaller: this means that the equivalent permanence would be smaller, and the project developer would get a lower payment from the purchase of the same amount of credits.

In this study, we assumed $\delta$ (discount factor) to be 0.03 [49], and obtained SCC (Social Cost of Carbon) values from the central series of the carbon cost values modeled and provided by the UK government [50]. This dataset contains carbon cost values from 2010 to 2100, and we fitted the values to an exponential model using the *lm* function in R, and used it to extrapolate to values from 2000 to 2500.

## A.2. Optimal release schedule

Due to the stochastic occurrences of forest disturbance and deforestation events, it is not realistic to expect project developers to produce credits that are completely free from reversal events ($\omega = 0$). Instead, we consider the optimal release schedule from the perspective of a risk-averse project developer, who determines future release schedules so that they limit reversal risk to $0 \leq \omega \leq 1$. Lower values of $\omega$ refer to a more risk-averse project developer, i.e.

with lower reversal risk but also lower equivalent permanence and hence lower credit pay outs.

The problem of computing the optimal release schedule can be formulated as a stochastic optimization problem. We view a project as an asset that at time $t$ generates drawdown, modeled as a random variable $a_t$ that can be either positive or negative. This drawdown, when adjusted by prior anticipated releases at this time, $\sigma_t$, gives us the net credits available at time step $t$, $c_t$. If $c_t$ is positive, a project developer can allocate some of these credits, in the form of a release schedule $\hat{r}^t_{t-j}$, $j > 0$ with $\sum_j \hat{r}^{t+j}_t = c_t$, to hedge against future times when it might be negative, so that the probability of reversal is bounded by $\omega$:

$$Pr(c_t < 0) < \omega$$
$$Pr(-a_t > \sigma_t) < \omega \qquad \text{(A4)}$$

Thus, if $\sigma_t$ exceeds a certain quantity, $a_\omega$, for every year $t$ in the project lifetime, then Equation A4 can be satisfied, meeting the constraint. We now state and prove the optimal rule to achieve this goal.

We define the probability of having an drawdown of $x$ tons in project $i$ in year $t$ to be $p^i_t(x^i_t)$ i.e. $p^i_t$ is the probability mass function of $x^i_t$. For simplicity, we assume for now that this probability mass function is time-independent and can be estimated from available observed data, so that:

$$p^i_t(x^i_t) = p^i(x^i_t) \ \forall \ t \qquad \text{(A5)}$$

Let the $\omega$-percentile value of the probability mass function be $a^i_\omega$, so that:

$$p^i(a_t < a^i_\omega) = \omega \qquad \text{(A6)}$$

Let the corresponding cumulative mass function (i.e. probability that the drawdown is less than $x$ in project $i$) be $P^i(x)^i_t$, so that:

$$P^i(-a^i_\omega) = \omega \qquad \text{(A7)}$$

If $\sigma_t{}^i$ of anticipated release have been allocated to this project in time period $t$, the probability that the drawdown is less than the number of previously anticipated releases is $P^i(-\sigma_t{}^i)$, which is also the probability of having to report negative net drawdown (reversal event).

In essence, the **allocation rule** for the optimal release schedule, which maximizes equivalent permanence (EP) at each year while satisfying the boundary condition for maximum reversal risk as much as is feasible, is to allocate all the net adjusted carbon drawdown (if any) to future years such that (a) the sum of allocated drawdown for each year in the project is at most $a_\omega$ and (b) this allocation is done sequentially for each successive year starting from the immediately following one.

Note that a special case arises when $a^i > 0$: in this case, the boundary condition for maximum reversal risk is always satisfied, and we can view the project as having 'low risk', meaning that we do not anticipate any releases to occur during the project, but only past the end of the project. In this case, the credit cache volume is defined to be zero.

As a proof, let the release schedule defined by the allocation rule (call it $S$) not be optimal. Let the optimal schedule, $O$, differ from $S$ for the first time in year $y$. Then, $O$ must allocate either less or more drawdown than $S$ that year. If it allocates more, then it must allocate less to some future year than $S$, since both allocate the same total

amount of drawdown. This results in a lower EP in year $y$, and less credits issued, because EP increases when releases are anticipated to happen further in the future. This means $S$ will have a higher overall credit payment than $O$, which means $O$ is not optimal. On the other hand, if $O$ allocates less drawdown than $S$ in year $y$, then the allocation for that year is less than $a_\omega^i$. But $a_\omega^i$ is the minimum amount that needs to be allocated to meet the boundary condition of the reversal risk. Hence, $O$ will have a higher probability of reversal than $S$. But $S$ is the schedule that has the lowest reversal risk, by construction, so $O$ must have a reversal risk that is higher than $S$, which means it is not optimal. This shows that any schedule other than $S$ is not optimal, which means $S$ is indeed optimal, as claimed.

This analysis is easily generalized to account for a set of projects in a project developer's portfolio. If the project developer has a set of projects indexed by $i$, then we can compute $c_t^i$ for each project for each year $t$. We can then compute the optimal allocation of $C_t = \sum_i c_t^i$ credits as releases across projects and project years, where the project developer wants to limit the upper bound of the probability of reversal to $\omega$ each year. Note that the definition of reversal event for a set of projects, compared to a single project, is when the net total drawdown across all projects is negative.

We assume that the releases allocated to future years are fully interchangeable, so net negative credits in any project can be compensated by anticipated releases in excess of actual releases in any other project in the set. For example, if projects 1 and 2 both have 100 credits available in year 1, and project 1 had a release schedule of [0, 50, 0, 50] and project 2 had a release schedule of [50, 0, 50, 0] in years 2-5, then as long as $a_t^1 + a_t^2 > -50$ for $t \in 2,3,4,5$, neither project would need to report negative credits in those years.

In this case, the total drawdown in year $t$ is $\sum_i a_t^i$. Since the probability mass function of each $a_t^i$ is known, it is possible to compute the probability mass function of their sum. Let $A_\omega$ be the $\omega$-percentile value of this sum. Then, it is obvious that the optimal allocation rule from above continues to hold, with $A_\omega$ taking the role of $a_\omega$, since we can view the aggregation of projects as a single project, albeit with a more complex probability mass function.

# Appendix B.  Finding credit cache volume analytically

A $\omega$-averse strategy in the release schedule computation can only be realized if corresponding values for $a_\omega$ can be found, such that $Pr(a_t + a_\omega \leq 0) = \omega$, where $a_t$ represents the random variable of net carbon drawdown that can be either positive or negative. This, in turn, depends on knowing the distributions of the yearly carbon loss of the project and the counterfactual.

In this section, we consider two cases for these distributions, one special case where both distributions are assumed to be exponential, and one arbitrary case. For the exponential distributed case, we derive in a straightforward fashion closed-form representations for determining $a_\omega < 0$. In the case of the arbitrary distribution, we can at least determine bounds through the application of Chernoff inequality.

## B.1.  Exponentially distributed yearly carbon losses

### B.1.1.  Single project

Let us start with the case of an individual project and its counterfactual scenario, and assume that the random variables corresponding to the amount of carbon loss in project $i$ and in its counterfactual denoted $l_p^i$ and $l_c^i$ are independent random variables with exponential distributions of parameters $\lambda_p^i$ and $\lambda_c^i$ respectively. The drawdown results from the difference of these two random variables as $a_t^i = l_p^i - l_c^i$ with the probability density function given as:

$$p(a) = \frac{\lambda_p^i \lambda_c^i}{\lambda_p^i + \lambda_c^i} e^{a \cdot \lambda_p^i} \text{ if } a < 0$$

$$p(a) = \frac{\lambda_p^i \lambda_c^i}{\lambda_p^i + \lambda_c^i} e^{-a \cdot \lambda_c^i} \text{ if } a > 0 \qquad (B1)$$

For the case where $a_\omega < 0$, we can furthermore obtain the distribution function as:

$$F_A(x) = P(A \leq x) = \int_{-\infty}^{x} \frac{\lambda_p^i \lambda_c^i}{\lambda_p^i + \lambda_c^i} \cdot e^{a \lambda_p^i} \cdot da = \frac{\lambda_c^i}{\lambda_p^i + \lambda_c^i} \cdot e^{a \lambda_p^i}$$

$$(B2)$$

Given the definition of $a_\omega$, namely that $P(A \leq a_\omega) = \omega$, we obtain in a straightforward way a closed-form expression for $a_\omega$:

$$a_\omega = \frac{1}{\lambda_p^i} \cdot \ln\left( \omega \cdot \frac{\lambda_p^i + \lambda_c^i}{\lambda_c^i} \right) \qquad (B3)$$

### B.1.2.  Aggregated projects

The general case with $n$ projects and corresponding counterfactuals is more involved, even under the assumption of the special case with exponential distributions for all carbon loss processes. It is well known that for the sum of $n$ exponentially distributed random variables, the resulting random variable is Gamma distributed if each individual exponentially random variable is i.i.d. Assuming the i.i.d. exponential parameters to be equally $\lambda$, the resulting Gamma distribution follows with shape parameter $n$ and scale parameter $\frac{1}{\lambda}$.

However, the distribution of the resulting difference between two Gamma-distributed random variables is involved. Given two independent Gamma-distributed random variables $X,Y$ with $X \sim \Gamma(\alpha_1, \beta_1)$ and $Y \sim \Gamma(\alpha_2, \beta_2)$, the random variable $Z = X - Y$ follows a Gamma difference distribution [51]. For the case of interest $Z < 0$, the corresponding density results to:

$$f(z) = \frac{\tilde{c}}{\Gamma(\alpha_2)} \cdot (-z)^{\frac{\alpha_1 + \alpha_2}{2} - 1} e^{\frac{\beta_1 + \beta_2}{2} \cdot (-z)} W_{\frac{\alpha_2 - \alpha_1}{2}, \frac{1 - \alpha_1 - \alpha_2}{2}}$$
$$\left( (\beta_1 + \beta_2) \cdot (-z) \right) \qquad (B4)$$

where $W()$ is the Whittaker's W function. Note that this density holds for different scale parameters $\beta_1$, $\beta_2$. This implies the possibility to investigate the drawdown risk characterization also for cases where all projects have the

same exponential rates $\lambda_p$ while all counterfactuals have a different (but homogeneous) exponential rate $\lambda_c$.

## B.2. Arbitrarily-distributed yearly carbon losses

We now generalize the above model to the case where the random variables corresponding to the amount of carbon loss in project $i$ and in its counterfactual denoted $l_p^i$ and $l_c^i$ are independent random variables with arbitrary distributions $F_p$ and $F_c$ (i.e. we start again with a single project and counterfactual carbon loss process, and then expand to the general case with $n$ projects and counterfactuals in the portfolio).

### B.2.1. Single project

As above, the drawdown results from the difference of these two random variables as $a_t^i = l_p^i - l_c^i$ and we are interested in the value $a_\omega$ such that $P(A <= a_\omega) = \omega$. The (left-tail) Chernoff bound of a random variable $X$ is given by

$$\mathbb{P}[X \leq a] \leq \inf_{t<0} M_X(t) \cdot e^{-ta} \qquad (B5)$$

with $M_x(t)$ being the moment-generating function of $X$:

$$M_X(t) = \mathbb{E}[e^{tX}] \qquad (B6)$$

Likewise, for the difference of two random variables $X_1, X_2$ with moment-generating functions $M_{X1}(t), M_{X2}(t)$ we obtain:

$$\mathbb{P}[X_1 - X_2 \leq a] \leq \inf_{t<0} \frac{M_{X_1}(t)}{M_{X_2}(t)} \cdot e^{-ta} \qquad (B7)$$

From this Chernoff bound identity, we can derive a generic formula to determine an upper bound on $a_\omega$. We have:

$$\mathbb{P}[A \leq a] = \mathbb{P}[X_p - X_c \leq a] \leq \inf_{t<0} \frac{M_{X_p}(t)}{M_{X_c}(t)} \cdot e^{-ta_\omega} = \omega \qquad (B8)$$

From this, with some manipulations we obtain the implicit equation:

$$\inf_{t<0} \left( \ln \frac{M_{X_p}(t)}{M_{X_c}(t)} - ta_\omega \right) = \ln \omega \qquad (B9)$$

where $M_{X_p}(t)$ and $M_{X_c}(t)$ represent the moment-generating functions of the carbon loss of the project and counterfactual.

### B.2.2. Aggregated projects

In extending to the case of the project portfolio with $n$ projects and counterfactual carbon loss processes, the above identity is easily adapted. Consider the MGF of the sum of independent random variables $Y = X_1, \ldots X_n$, which is simply given by the product of the individual MGFs of the random variables, i.e.:

$$M_Y(t) = M_{X_1}(t) \cdot \ldots \cdot M_{X_n}(t) \qquad (B10)$$

We can thus obtain the implicit identity:

$$\inf_{t<0} (\ln \frac{\prod_i M_{X_{p_i}}(t)}{\prod_i M_{X_{c_i}}(t)} - ta_\omega) = \ln \omega \qquad (B11)$$

where $M_{Xpi}(t)$ and $M_{Xci}(t)$ represent the moment-generating functions of the $i$-th carbon loss of the project and counterfactual.

## Appendix C. Quantifying carbon loss for real-life projects

### C.1. Study sites

We selected four ongoing REDD+ projects located in South America and West Africa, two of the three worldwide tropical regions (Table C1). All four projects are over 20,000 hectare in size, scheduled to run for 20-30 years, and are registered in the Verra VCS Registry [52]. The projects differ in their land ownership status, historical land use, and main deforestation drivers. The Rio Pepe y ACABA (RPA) project (VCS 1396) is organized by local communities which hold property rights that are protected by law, and agriculture and illegal logging are its main deforestation drivers [53]. The Gola project (VCS 1201) is situated within a national park, and small-scale agriculture is its main deforestation driver [54]. The Alto Mayo project (VCS 944) is a nationally recognized protected area, and coffee production and subsistence farming are its main deforestation drivers [55]. The Mai Ndombe project (VCS 934) area consists of previous commercial logging concessions, and logging remains its main deforestation driver [56]. The projects were selected based on the following criteria:

1. Projects categorized as "reducing deforestation and degradation", with an available polygon of project boundary
2. Projects whose boundary falls entirely within the extent of the tropical moist forest (TMF) biome [57]
3. Ongoing projects, with a VCS Registry status that is not "withdrawn", "on hold", or "inactive"
4. Projects that have started between 1990 and 2018: this is to ensure that the project duration falls within the coverage of the land use class time series, and that the project has been ongoing for at least five years.

We acknowledge that some of these projects, such as in Gola, include areas with historical forest degradation and deforestation that occurred before the start of the project, and in some cases before the start of the availability of the land use class time series. These projects usually also include areas with forest undergoing regrowth after historical degradation and deforestation. We did not use the presence or extent of these areas as exclusion criteria.

### C.2. Tracking forest cover change and carbon flux

To quantify how forest cover changes over time in NBS projects, we used the annual change collection in the JRC-TMF dataset [57], which provides the spatial extent and the annual change of the tropical moist forest (TMF) biome at the 0.09-hectare (30 m × 30 m pixels) resolution from 1990 to 2022, derived from the L1T archive imagery (orthorectified top of atmosphere reflectance). The six following land cover classes were mapped: (1) undisturbed forest, (2) degraded forest, (3) deforested land, (4) forest regrowth, (5) permanent and seasonal water, and (6) other land cover.

To use information on forest cover to quantify how carbon stock changes over time in NBS projects, we assume

**Table C1. Basic description of the NBS projects examined in this study.**

| VCS ID | Full name | Country | Total area (ha) | Start year ($t_0$) | Scheduled end year |
|---|---|---|---|---|---|
| 1396 | Rio Pepe y ACABA REDD + Project | Colombia | 48,177 | 2014 | 2044 |
| 1201 | Gola REDD project | Sierra Leone | 69,714 | 2012 | 2042 |
| 944 | Alto Mayo Conservation Initiative | Peru | ~182,000 | 2008 | 2028 |
| 934 | The Mai Ndombe REDD + Project | DRC | 299,640 | 2011 | 2041 |

that for each project, we can calculate a reference carbon density value for each land cover class that is stable over time. For this, we used the GEDI Level 4 A dataset, which contains footprint-level aboveground biomass density (AGBD) estimates (Mg ha$^{-1}$) for each 25-m GEDI shot [58]. The AGBD estimates are generated from models linking GEDI waveform-derived canopy height metrics with field AGBD estimates for multiple regions and plant functional types [59,60].

We selected GEDI shots occurring from 1st January 2020 to 1st January 2021 and falling within the project area plus a 30-km buffer around it. The inclusion of a 30-km buffer around the project area is to ensure that enough GEDI shots can be found for each land cover class. For each land cover class, we selected the subset of GEDI shots associated with it as shots that overlap with a JRC-TMF pixel (1) that belongs to the land cover class in question and (2) whose eight neighboring pixels also belong to the land cover class in question. The second condition was included to account for the potential geolocation error up to 10 m of GEDI shots [58]. For each land cover class, we calculated the median AGBD value of all the GEDI shots associated with it. We then estimated belowground biomass and deadwood biomass to be 20% and 11% of AGB, respectively, calculated the total biomass as the sum of aboveground, belowground and deadwood biomass, and converted total biomass to total carbon density by multiplying it by the average carbon density of biomass, taken to be 0.47 for this study [61–63].

It is to be noted that we did not consider leakage (displacement of carbon release to areas outside of the project area) in this study.

## C.3. Pixel matching and calculating carbon drawdown

The carbon drawdown of an NBS project of the avoided emission type is the difference between the carbon loss in the project area and the carbon loss in the counterfactual scenario. In order to find the counterfactual scenario, we adopted a pixel-based matching approach [16]: this has the advantage of being transparent, non-arbitrary and replicable. Specifically, we followed the PACT Tropical Moist Forest Accreditation Methodology document, where a complete and detailed description of the methods used can be found [25].

Specifically, we sampled pixels in the project area at a density of 0.25 points/ha for smaller projects ($\leq$ 250k ha) and 0.05 points/ha for large projects (> 250k ha). We then sampled candidate matching pixels from the match destination to the amount of ten times the number of sampled project pixels. The match destination is defined as the area of a 2000-km buffer around the project that falls within the project's country boundary (from the LSIB dataset [64]) and the RESOLVE ecoregion boundaries for all the ecoregions that lie within the project [65], excluding all other NBS

**Table C2. Kolmogorov-Smirnov Goodness-of-fit test statistics and *p*-value of empirical distributions of annual carbon loss fitted using gaussian mixture model (GMM). Bold texts indicate empirical distributions that are significantly different from the GMM fitted distributions.**

| | RPA | Gola | Alto Mayo | Mai Ndombe |
|---|---|---|---|---|
| Project | 0.029 (1) | 0.068 (0.32) | 0.043 (0.7) | **0.094 (0.033)** |
| Counterfactual | **0.14 (0.003)** | 0.094 (0.063) | **0.45 (< 0.001)** | **0.1 (0.012)** |

project areas and a 5-km leakage buffer around each of the NBS projects (including the project being matched).

For each project pixel in a 10% sample of the sampled project pixel set, we matched it to one candidate matching pixel which has the exact same value for the following categorical variables:

1. Land cover class at $t_{-10}$, $t_{-5}$, and $t_0$ (where $t_0$ is the project start year)
2. Country
3. Ecoregion

, and which has the minimum Mahalanobis distance [66] across the following continuous variables:

1. Elevation (from the SRTM data [67])
2. SRTM-derived slope
3. Accessibility [68]
4. Coarsened proportional cover of undisturbed forest and deforested land, at 1200 m × 1200 m resolution, within a 1-km radius buffer around the pixels at $t_{-10}$, $t_{-5}$, and $t_0$.

We deemed the matching results valid if all the standardized mean differences (SMD) of each continuous matching variable between the sampled project pixels and the matched pixels is smaller than 0.2 (unless if a continuous matching variable with an SMD > 0.2 is distributed in the range [0, 1], and the value in one of the pixel sets is close to 0 or 1: this is because near those values SMD becomes misleading). We performed 20 repetitions of the matching process, each time using an independent sample of the project pixels as input and producing a set of matched pixels as output. The matched pixel sets are the" counterfactual scenarios" of the project area, and can be considered to be representative of the trajectory of forest cover change and carbon flux in the project area if the project had not existed.

We then evaluated the carbon losses of both the pixels in the project area and the pixels in the counterfactual scenarios, at a yearly interval. For each year within the JRC-TMF time series (1990–2021), for both the project area and the counterfactual scenarios, we calculated the proportion of pixels in each JRC-TMF land cover class, and used the GEDI L4A-derived estimates of total carbon density for each land cover class, described in the previous section, to calculate the total carbon stock (Mg $CO_2$), and calculated the mean total carbon stock value of all 100 counterfactual repetitions. We then

calculated carbon losses ($l_t$) of each year $t$ in both the project area ($p$) and the counterfactual ($c$) as the difference between the carbon stock of that year ($b_t$) and that of the previous year ($b_{t-1}$): $l_t = b_{t-1} - b_t$. Finally, we calculated annual carbon drawdown ($a_t$) as the difference between the project carbon loss ($l_t^p$) and counterfactual carbon loss ($l_t^c$): $a_t = l_t^c - l_t^p$.

### C.4. Fitting observed carbon loss distributions

As visual examinations showed that the carbon loss distributions often exhibited a multi-modal shape in many projects, especially for the counterfactual scenarios, we performed model-based clustering based on parameterized Gaussian mixture models [69], fitting the empirical distributions as the proportional mixture of two Gaussian components, using the *Mclust* function in the R package *mclust* [69]. To evaluate the goodness of fit, we generated random samples of equal size to the original empirical distribution from each fitted distribution, and performed chi-squared tests using the *gofTest* function in the R package *EnvStats* [70], and reported the test statistics and *p*-value of each fitting result (Table C2). A lower bound of zero was used to remove negative values from the sampled values before the goodness-of-fit test.