CLEAN AIR TASK FORCE

February 16, 2024

Chair Rostin Behnam
Commodity Futures Trading Commission
Three Lafayette Centre 1155
21st Street, NW
Washington, DC 20581

Re: Commission Guidance Regarding the Listing of Voluntary Carbon Credit Derivative Contracts, RIN 3038–AF40 (88 Fed. Reg. 89410)

Clean Air Task Force ("CATF") respectfully submits these attachments to our comments on the Commodity Futures Trading Commission's ("CFTC" or "Commission") proposed guidance regarding the listing of voluntary carbon credit derivative contracts ("proposal" or "proposed guidance"), 88 Fed. Reg. 89410 (Dec. 27, 2023).

- William R. L. Anderegg, et al., *Future climate risks from stress, insects and fire across US forests*, 25 Ecology Letters 1510 (2022).
- Ben Filewod & Geoff McCarney, *Avoiding carbon leakage from nature-based offsets by design*, 6 One Earth 790 (2023).
- Annelise Gill-Wiehl, et al., *Pervasive over-crediting from cookstove offset methodologies*, Nature Sustainability (2024).
- Barbara K. Haya, et al., *Comprehensive review of carbon quantification by improved forest management offset protocols*, 6 Frontiers Forests & Global Change, No. 958879 (2023).
- Joan Pinto, *Analysis: Do offset registry revenue models offer perverse incentives to over-credit?*, Carbon Pulse (Mar. 20, 2023).
- Varsha Ramesh Walsh & Michael W. Toffel, *What Every Leader Needs to Know About Carbon Credits*, Harv. Business R. (Dec. 15, 2023).
- Oranuch Wongpiyabovorn, et al., *Challenges to voluntary Ag carbon markets*, 45 Applied Econ. Perspectives & Pol'y 1154 (2022).
- Peter Woods Ellis, et al., *The principles of natural climate solutions*, 15 Nature Comms. No. 547 (2024).

Received: 16 December 2021 | Revised: 28 March 2022 | Accepted: 9 April 2022

DOI: 10.1111/ele.14018

LETTER

ECOLOGY LETTERS cnrs WILEY

Future climate risks from stress, insects and fire across US forests

William R. L. Anderegg[1] | **Oriana S. Chegwidden**[2] | **Grayson Badgley**[3,4] | **Anna T. Trugman**[5] | **Danny Cullenward**[2,6] | **John T. Abatzoglou**[7] | **Jeffrey A. Hicke**[8] | **Jeremy Freeman**[2] | **Joseph J. Hamman**[2,9]

[1]School of Biological Sciences, University of Utah, Salt Lake City, Utah, USA
[2]CarbonPlan, San Francisco, California, USA
[3]Blackrock Forest, Cornwall, New York, USA
[4]Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York, USA
[5]Department of Geography, University of California, Santa Barbara, Santa Barbara, California, USA
[6]Institute for Carbon Removal Law and Policy, American University, Washington, DC, USA
[7]Management of Complex Systems Department, University of California, Merced, Merced, California, USA
[8]Department of Geography, University of Idaho, Moscow, Idaho, USA
[9]National Center for Atmospheric Research, Boulder, Colorado, USA

Correspondence
William R. L. Anderegg, School of Biological Sciences, University of Utah, Salt Lake City, UT 84103 USA.
Email: anderegg@utah.edu

Funding information
Microsoft; David and Lucile Packard Foundation; National Institute of Food and Agriculture, Grant/Award Number: 2018-67012-31496 and 2018-67019-27850; National Science Foundation, Grant/Award Number: 1714972, 1802880, 2003017, 2003205, 2017494 and 2044937

Editor: Joshua Lawler

Abstract

Forests are currently a substantial carbon sink globally. Many climate change mitigation strategies leverage forest preservation and expansion, but rely on forests storing carbon for decades to centuries. Yet climate-driven disturbances pose critical risks to the long-term stability of forest carbon. We quantify the climate drivers that influence wildfire and climate stress-driven tree mortality, including a separate insect-driven tree mortality, for the contiguous United States for current (1984–2018) and project these future disturbance risks over the 21st century. We find that current risks are widespread and projected to increase across different emissions scenarios by a factor of >4 for fire and >1.3 for climate-stress mortality. These forest disturbance risks highlight pervasive climate-sensitive disturbance impacts on US forests and raise questions about the risk management approach taken by forest carbon offset policies. Our results provide US-wide risk maps of key climate-sensitive disturbances for improving carbon cycle modeling, conservation and climate policy.

KEYWORDS
biotic agents, carbon cycle, disturbance, drought, nature-based climate solutions

INTRODUCTION

Earth's forests play a fundamental role in the global carbon (C) cycle and currently are a substantial carbon sink, sequestering up to 25% of human carbon dioxide emissions annually (Bonan, 2008; Pan et al., 2011). Yet the future of forests in a rapidly changing climate is highly uncertain (Brodribb et al., 2020; Friedlingstein et al.,

William R. L. Anderegg and Oriana S. Chegwidden contributed equally.

2014). In particular, increasing climate stresses and disturbance could compromise forest C storage, yielding manifold impacts on biodiversity, ecosystem services and carbon cycle feedbacks and undermining the potential of forests as a climate solution (Holland et al., 2019; Hurteau et al., 2009; Seidl et al., 2017). For example, an unprecedented and climate-fuelled bark beetle outbreak in Canada drove immense swaths of tree mortality and reversed an entire region of boreal forest from a C sink to a C source over a decade with large implications for climate policy (Kurz, Dymond, et al., 2008; Kurz, Stinson, et al., 2008). In addition to insect outbreaks, wildfires and climate stress have been widely documented as prominent risks because they strongly regulate forest C stocks and are likely to increase in future climates (Bentz et al., 2010; Buotte et al., 2019; Hicke et al., 2012; Wang et al., 2021). Thus, it is essential to rigorously quantify and understand drivers of historical risks and use this understanding to project future climate-driven risks for forest ecosystem functions and services, including long-term C storage (Anderegg et al., 2020; Clark et al., 2016).

Due to forests' role as a C sink and important co-benefits for biodiversity and ecosystem services, governments, corporations and non-governmental organisations have shown widespread and growing interest in leveraging forests as 'nature-based climate solutions' to sequester and store C as part of meeting climate policy goals (Cook-Patton et al., 2020; Griscom et al., 2017; Roe et al., 2019). Yet significant scientific gaps remain that greatly limit the effective use of forest-based climate solutions in an evidence-based climate policy framework. Crucially, to be used for climate mitigation, forests must achieve some level of 'permanence' whereby a management or policy action leads to more ecosystem C storage, averaged over time, compared to a rigorous baseline (Hurteau et al., 2009; Ruseva et al., 2017). Although fossil C emissions persist in the atmosphere for hundreds to thousands of years (Archer et al., 2009), many public and private carbon markets only require increased forest C storage to last at most for up to 100 years (Ruseva et al., 2017) — and sometimes only up to a few decades.

Rigorous forest climate risk assessment is crucial for understanding climate impacts on ecosystems and biodiversity, informing conservation and management prioritisation, and guiding climate policies and programmes relying on forest carbon uptake and storage. Continental-scale risk assessment is currently lacking and urgently needed (Anderegg et al., 2020; Buotte et al., 2019; Lecina-Diaz et al., 2021). Spatial quantification of risks can inform forest protocols in climate policy by ensuring that climate risks are adequately considered in programme design — for example, through the construction of 'buffer pools' and other insurance mechanisms — and can inform forest project development and conservation (Hurteau et al., 2013). However, current forest offset protocols tend to include fixed, spatially invariant risks that do not incorporate future climate impacts and likely underestimate the integrated risks to forests over long time scales (e.g. the 100 year horizon used by the Climate Action Reserve (Anderegg et al., 2020)).

In this paper, we combine forest inventory data across United States (US) forests, remote-sensing data of wildfires, high resolution climate data and downscaled climate model projections to assess climate-sensitive risks for forest C stocks in the US. We first quantify how forest structure and climate anomalies mediate major climate-related risks to US forests from wildfire and non-fire, climate stress-mediated tree mortality (defined here as tree mortality that is sensitive to climate, excluding fire-driven mortality). We then model the spatial patterns and magnitudes of these risks over the historical record. Finally, we use downscaled future climate data to project how these risks might evolve in the future due to climate change, revealing where forests are likely to be the most vulnerable in the 21st century.

METHODS

Overview and climate data

We constructed statistical models of climate risks from fire, (non-fire) climate stress-driven tree mortality, and insect-driven tree mortality using high-resolution historical climate data, satellite data for fire burn area and forest inventory plot data for tree mortality. We performed extensive cross-validation and comparisons against independent datasets over the historical period. We then developed a high-resolution downscaled climate dataset from six climate models to project these climate risks across the US for three future climate scenarios. Statistical risk models and validation are described below, and full input dataset details, pre-processing, and climate downscaling are described in the Supplementary Information (SI) Methods.

Risk models

Fire

We developed a statistical model to create gridded (4-km spatial resolution), monthly predictions of burn area as a function of climatic variables. This model built on previous fire risk estimation efforts (Barbero et al., 2014). Many of the methods are similar, although updated with more recent data (through 2018 rather than 2010).

The model was fit to historical fire data from the Monitoring Trends in Burn Severity (MTBS) database (SI Methods), which comprises 30-m annual rasters of burn severity as well as burn area boundary polygons for individual fires (Eidenshink et al., 2007). The dataset covers fires from 1984–2018 and includes fires larger than 202 ha (404 ha in the Western US) for the continental

US. As predictors, we considered both temporally-varying climatic variables as well as time-invariant vegetation variables. Our primary climatic variables were monthly temperature, precipitation and climatic water deficit (CWD) derived from the TerraClimate dataset (SI Methods) for the historical period. For vegetation, we used the National Forest Type Dataset. The forest groups ranged in area between 51k to 67 M ha with the smallest eight groups each representing <~2 M ha. To limit the number of variables and prevent overfitting, spatially sparse forest groups were aggregated into supersets by combining the smallest forest groups with the most spatially similar, larger forest group (we chose an area threshold of 1.76 M ha). This consolidation decreased the number of forest groups from 25 to 17 and had little effect on model behavior. We converted these forest group maps into 17 binary, gridded maps, each of which became a predictor in the model. Every pixel was assigned to one forest group.

We fit a 'hurdle' regression model to predict burn area as a function of climate and vegetation variables. This model jointly predicts the probability of a non-zero value and, if a non-zero value is present, its continuous value (Cragg, 1971). Intuitively, this model can be thought of as combining a classifier ('was there fire?') and a regression ('if there was fire, how large was the burn area?'). For computational reasons, all datasets were aggregated to a 16 km^2 grid for fitting. The model was then applied to create predictions on the 4-km grid. We formally represented the hurdle model using a sequence of two generalised linear models: a Binomial model with logit link function predicting zero versus non-zero values, and a linear Gaussian model with normal link function predicting burn area in the locations where it was non-zero. We implemented the hurdle model in Python using scikit-learn by combining the LogisticRegression and LinearRegression methods (Pedregosa et al., 2012).

In addition to the variables described above, we included two additional predictors to better capture inter-annual trends. To create these predictors we first calculated two timeseries representing a conterminous US-average monthly temperature and precipitation and then calculated a 12-month rolling maximum of each of the two timeseries. Conceptually, these two predictors provide a measure of longer-term drought stress when conterminous US-wide high temperatures and sharp precipitation regimes occur simultaneously. In practice, including these extra predictors improved overall model performance only slightly, but allowed the model to better reproduce both monthly trends, interannual variability, and the observed increase in burn area over the observation period (see Figures S10 and S11).

For the full model, we assessed accuracy using area-under-the-ROC-curve (AUC) from the output of the logistic regression portion of the hurdle model. We consider this AUC the primary metric of interest given the sparse and nearly binary nature of the training data. We report these AUC values obtained using split-halves cross-validation, where the held out set was constructed by sampling years independently (Figure S1). We also assessed performance through the model's ability to reproduce three patterns: annual, seasonal and spatial. For annual and seasonal trends, we computed an R^2 between the value computed directly from the data and the model's prediction allowing for a constant offset difference (Figure S1). For the spatial trends we computed an AUC just as we did for the full model, except for first averaging over time, which we consider the appropriate metric given the sparse nature of the data. While these are not the metrics on which the model is trained, they provide an indication of how well the model captures important and observable patterns in the data. For visualisation purposes, predicted monthly burn areas from the model (fraction/month) were summed across months to estimate predicted burn area for each year (fraction/year).

Climate stress- and insects-driven tree mortality

We constructed 'climate stress' and 'insect' tree mortality models using data from the US Forest Inventory and Analysis (FIA) dataset, which is a nationwide standardised network of >100,000 long-term forest monitoring plots that track growth, mortality and overall health of US forests. We used FIA data from 2000 to 2018. We aggregated FIA forest plot data on live and dead basal area from a tree-level to a 'condition' level, grouping together conditions representing repeated inventories of the same location. To construct climate stress and insect risk models, we screened for plots that had at least 2 or more inventory measurements, which enables the estimation of a true mortality rate. We next screened out plots that had a 'fire' or 'human' disturbance code or a 'cutting' treatment code to remove major confounding disturbances.

We estimated the fraction of mortality based on the concept of a census interval, which we define as a pair of measurements in two measurement years (t_0, t_1). The fraction of mortality is defined as the ratio of new dead basal area in t_1 to the total live basal area in t_0, which was then normalised by the census length to give annual mortality rates. We computed this ratio separately for each condition. Given that many FIA plots only had one repeated measure (only one census interval), we used the first census interval for all conditions. We modelled 'climate stress-driven' mortality as the mortality that occurred during this census interval (with other confounding mortality drivers excluded, see above) and 'insect-driven' mortality using the 'agent code' (AGENTCD) tree-level data, where codes of 10–19 indicate insects as the primary causal agent of death. We note that the climate stress mortality models include mortality from insects, which was a deliberate decision because insects and climate stress such as drought often co-occur and interact to kill trees in many forests across the US and thus cannot be clearly separated (Anderegg et al., 2015), although we performed a sensitivity analysis of the climate stress model when excluding mortality of trees with insect agent codes and observed very similar modelled mortality patterns ($R^2 = 0.60$, $p < 0.0001$). Drought and other climate stress-driven mortality does not have a clear or widely used agent code in the FIA database; instead, climate/drought-driven mortality is often attributed to a wide array of more proximate agents (including insects, disease, weather and other/unknown); see Anderegg et al. (2015) for a detailed discussion. Thus, our attribution is that this mortality is likely driven by "climate stress" broadly defined, as we have aimed to remove other major drivers of mortality, notably fire, human disturbance and account for stand and self-thinning dynamics in model construction (see below). This mortality attribution approach has uncertainties but is generally reasonable and is the standard approach that has been widely used in numerous climate-related FIA studies (Hember et al., 2017; Shaw et al., 2005; Stanke et al., 2021).

We fit a statistical model predicting mortality as a function of climatic and stand variables. We formally represented the hurdle model using a sequence of two generalised linear models: a Binomial model with logit link function predicting zero versus non-zero values, and a linear model with beta-distributed link function, which is used for modeling proportions where values are between zero and one, for predicting mortality in the conditions where it was non-zero. The beta-distributed link function for the linear regression was chosen based on inspecting the behavior of the raw data distributions. We implemented the hurdle model in R using the glm function in the default 'stats' package and the 'betareg' package (Zeileis et al., 2010).

For each condition, we extracted the mean, minimum and maximum over the census interval for six annual climate variables that were selected based on their importance in the drought and insect mortality literature: precipitation, temperature, Palmer Drought Severity Index (PDSI), potential evapotranspiration (PET), climatic water deficit (CWD) and vapor pressure deficit (VPD) (Bentz et al., 2010; Creeden et al., 2014; Williams et al., 2013). We also extracted the stand age for each condition from FIA and the community-weighted mean and range of the functional trait of the water potential at 50% loss of hydraulic conductivity (P50) from maps published in a recent study (Trugman et al., 2020) scaled to 0.25 degree. P50 has been widely linked to drought-driven mortality risk in site-level (Nardini et al., 2013) studies and meta-analyses (Anderegg et al., 2016). We also included in the mortality models two stand variables, age or age-squared, to account for background ecological dynamics such as self-thinning and background mortality, following Hember et al. (2017). All predictor variables were z-scored across the full dataset for that variable to ensure that variable ranges did not drive model outputs.

Climate stress and insect mortality models were fit independently to each FIA 'forest type,' which was chosen as an intermediate compromise of capturing the diversity of responses across US forests but aggregating above a species-level to enable adequate estimation of mortality levels. To ensure that each forest type had 50 or more condition measurements, we aggregated some sparse forest types into more common ones (59 were so aggregated out of the initial set of 171), leading to 112 initial forest types in our dataset. We aggregated condition-level mortality rates, age, climate data and functional traits to a 0.25 degree grid for each forest type. This grid size was chosen through sensitivity analyses to determine the optimal aggregation where the coefficient of variation of mortality rate stabilised but large-scale climate variation was preserved. All climate stress and insect mortality models were fit using this 0.25 degree gridded data for each forest type.

We considered collinearity among predictor variables by examining variance inflation factors. We found that variance inflation factors were too high for comparing mean/min/max of the same variable (e.g. mean vs. min vs. max annual temperature), but were generally within acceptable levels (<5) across the six predictor climate variables, stand age and P50 hydraulic trait. Thus, we conducted a stratified model selection analysis where we fit all possible model combinations with one of each predictor variable (i.e. varying all possible combinations of mean vs min vs. max of each climate variable, age vs. age-squared, P50 mean vs. P50 range) and selected the most parsimonious model via Akaike Information Criterion (Burnham & Anderson, 2004). For all analyses in this paper, we fit the same predictor variables across all forest types to reduce complexity. Thus, individual forest types were not allowed to have separate predictor variables. Model selection analyses were done separately on climate stress and insect mortality dependent variables.

We examined optimal model complexity by comparing the AIC and R^2 of nested sets of models. We compared climate stress and insect mortality models as a function of: (i) a null model of forest type-only (i.e. each forest type would receive only its mean mortality), (ii) a null model of mortality as a function of forest type and age only (i.e. no climate predictors), (iii) mortality as a function of forest type, age and climate predictors, and (iv) mortality as a function of forest type, age, climate and functional traits. We observed that climate variables significantly improved (i.e. deltaic $\ll -3$) model performance beyond both null models for both climate stress and insects and that the range of P50 significantly improved climate stress mortality models, but not insect models.

We assessed model performance with cross-validation and used two primary metrics that reflect the performance of different parts of the hurdle model. We first tested for spatial autocorrelation using Moran's I for each forest type and each of the climate stress and insect mortality models. For forest types and mortality models where significant autocorrelation was detected, we used a comparison of Moran's I by distance bin, using the 'correlog' function in the pgirmess package in R (Giraudoux et al., 2018), to determine the autocorrelation length. We set the spatial autocorrelation length as the midpoint between the last significant bin and the first non-significant bin. We then conducted spatial hold-out cross-validation (Ploton et al., 2020) for each forest type and mortality model, whereby one grid cell was held out from model training and a spatial buffer around that grid cell equal to the autocorrelation length was also removed from model training. The model was then fit on the remaining data and used to predict the hold-out grid cell, and this was repeated 1000-fold for each forest type and mortality model. Similar to the fire model, we examined model performance using cross-validated area under the receiver operating curve (AUC) for the binary component of the hurdle model. We also considered the non-zero-value R^2 for the beta-regression part of the hurdle model (Figure S5), aggregating as above.

Finally, we imposed one further set of criteria on climate stress and insect models to incorporate climate-dependence only where justified based on model performance. For all final model-based analyses (i.e. Figures 1d, f, 2–4), we identified forest types where cross-validated AUC was greater than 0.6 and the forest type had >20 grid cells with mortality observed in the historical record, based on a recent analysis of stability and information criteria in regression models (Jenkins & Quintana-Ascencio, 2020). This led to risks being modelled with climate variables and projected for 30 forest types in the insect models and 61 forest types in the climate stress models out of 112 possible forest types. For all forest types that did not meet these criteria, we projected mortality simply as the mean of historical mortality for that forest type, and thus set all future mortality to that value. We note that this is a very conservative decision and is likely to underestimate future risks.

We further performed two evaluations of our mortality models against independent metrics or datasets. For the climate stress mortality models, we compared our observed mortality rates by species against the recent 'Forest Stability Index' for eight major western US forest species (Stanke et al., 2020) and observed a strong relationship ($R^2 = 0.72$; Figure S2). For the insect mortality models, we compared our observed and modelled spatial patterns to an independent dataset of Aerial Detection Surveys (ADS) done by the US Forest Service to map bark beetle-driven mortality across the US (Williams et al., 2016). Despite large differences in types of dataset (e.g. aerial versus plot; 'bark beetle-driven' mortality versus 'insect-driven' plot agent codes) and spatial scales, we found strong agreement between our models and that independent dataset (AUC = 0.79 and $R^2 = 0.29$ comparing our modelled mortality and ADS observed mortality; Figure S3).

For future CMIP6 projection-driven risk models, we used the same climate variables as chosen by the final model selection analysis and projected climate stress and insect risk (i.e. % basal area killed per year) over each decade from 1950–2100 in different climate models and scenarios. Future decadal climate variables were z-scored against three decadal baselines (1990–1999; 2000–2009; 2010–2019) and the ensemble mean was taken across these baselines for each climate model and decade. All modelled and projection maps (e.g. Figures 1d/f, 3) were made on all conditions in FIA, regardless of number of censuses in the historical record, to cover all US forests, aggregated to 0.25° by forest type, and then aggregated across forest types as described above. For future projections, we used a constant stand age and P50 functional trait based on the 2000–2018 historical values due to uncertainties about future forest dynamics and composition. This is an assumption and uncertainty, but a full exploration of stand age dynamics, species distribution and composition shifts, and demography is beyond the scope of this current analysis.

RESULTS

The fire risk model reliably predicted historical fires (cross-validated AUC: 0.89), capturing interannual variability (cross-validated R^2: 0.64), seasonal patterns (e.g. spring risk in the southeastern US and fall risk in the western US; cross-validated R^2: 0.90), and spatial patterns (cross-validated AUC: 0.78) (Figure 1a-b; Figure S1). The model captured the spatial patterns of more prevalent fire across the western US, in particular in California and the northern Rocky Mountains.

Historical patterns of climate stress-driven tree mortality, which is predominantly drought stress and includes biotic agents/insects, were highest across the western US and intermountain West, which was captured in the mortality model (cross-validated spatial R^2: 0.18; Figure 1c–d). These patterns were consistent with the independent comparison dataset (Figure S2) and other recent studies (31). The inclusion of forest physiological metrics for drought tolerance, specifically community-weighted plant hydraulic traits, substantially improved the predictive accuracy of the climate stress mortality models (ΔAkaike Information Criterion << −10), consistent with drought-physiology studies (Anderegg et al., 2016). Observed historical permanence risks to US forests from insect-driven mortality specifically were highest in the Rocky Mountains and modelled risks captured the key broad spatial patterns in risks (cross-validated spatial R^2: 0.31; Figure 1e–f). These observed and modelled insect risks showed strong spatial agreement with an independent continent-wide insect outbreak dataset (Figure S3).

Under all future shared socioeconomic pathway (SSP) climate scenarios, fire risks are projected to increase

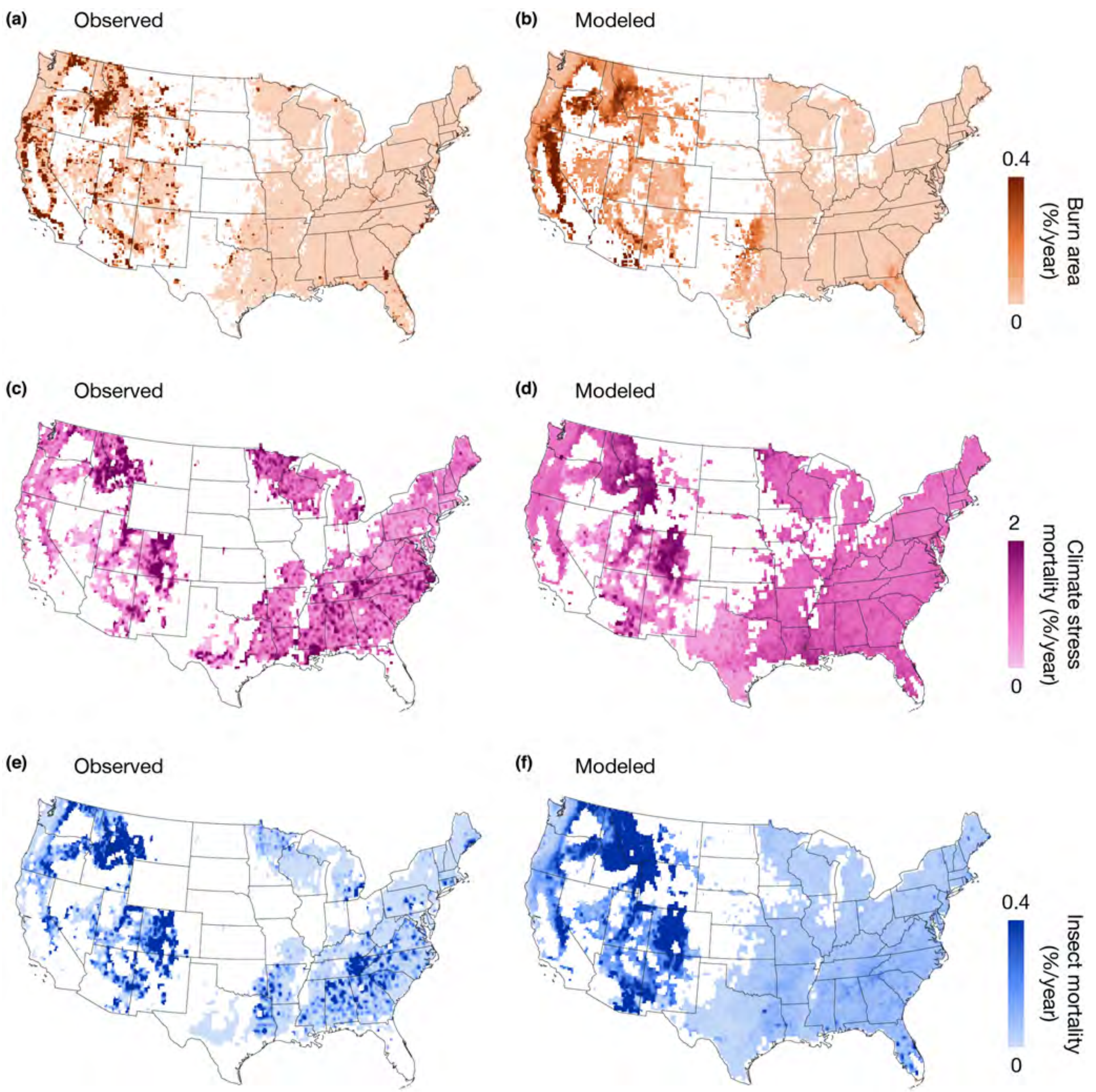


FIGURE 1 Observed (left) and statistically modelled historical (right) risk maps for fire (a&b), non-fire, climate stress-driven tree mortality (c&d), and insect-driven tree mortality (E&F) reveal widespread and spatially varying risks. Fire risk is modelled as burn area by wildfires, that is, fraction of a grid cell burned per year. Climate-stress and insect-driven tree mortality risk are modelled as basal area mortality per year. For each impact risk model (b, d, f), anywhere shaded is considered forested. The forest mask for fire differs slightly from those used for climate stress and insects due to different input data. Data gaps in forest inventory in WY and OK preclude observed risk estimates in climate stress and insect cases (c, e)

substantially throughout the 21st century (Figure 2a). Future risks increase similarly across scenarios through mid-century, but diverge by 2050. By 2080–2099, the multi-model mean projects 4-fold (SSP2-4.5), 9-fold (SSP3-7.0), and 14-fold increases (SSP5-8.5) in US-averaged fire risk compared to historical (average 1990–2019) values. Projected climate stress risks increased substantially and varied by emissions scenario with average mortality increases by a factor of 1.3 in SSP2-4.5, 1.5 in SSP3-7.0, and 1.8 in SSP5-8.5 by 2080–2099 (Figure 2b). Future US-wide insect risk projections indicated increases of 1.2-fold in SSP2-4.5, 1.4-fold in SSP3-7.0, and 1.7-fold in SSP5-8.5 by 2080–2099 (Figure 2c). We note that the climate stress and insect mortality models are not independent and thus should not be considered additive. All three climate-sensitive risks showed large

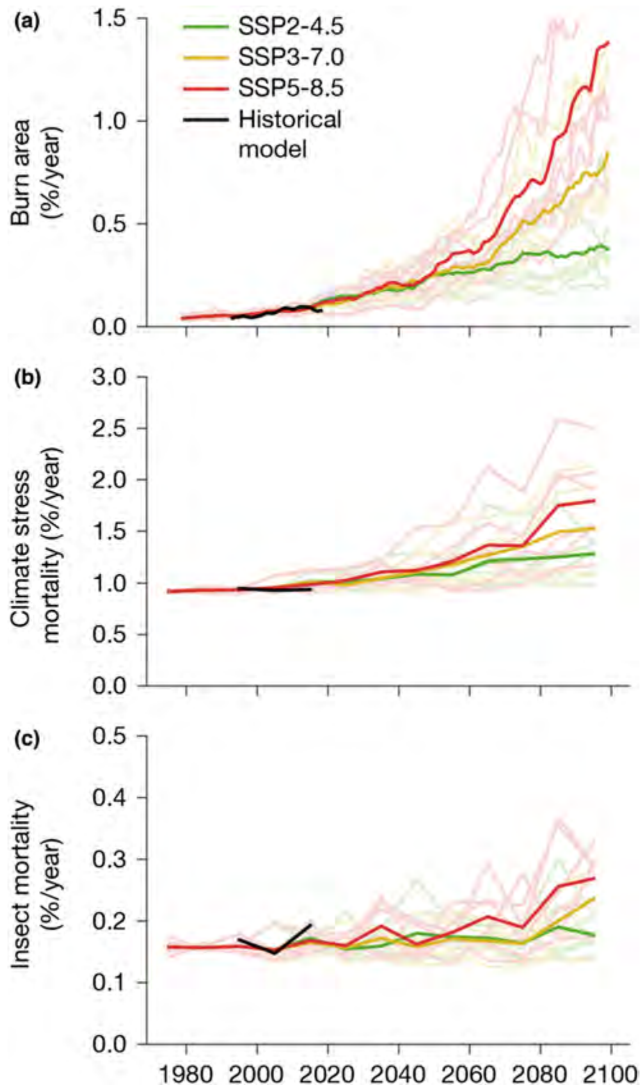


FIGURE 2 Projected 21st century risks for fire (a), non-fire, climate stress-driven tree mortality (b), and insect-driven tree mortality (c) averaged across the US. Statistical risk models forced by simulations from each different climate model shown as transparent lines, coloured according to the three shared socioeconomic pathway (SSP) climate scenarios. The multi-model mean for each SSP is shown opaque. Statistical model projection driven by historical meteorological data (rather than meteorological data derived from a climate model) are shown in black. Fire risks are calculated with a 10-year centered moving average, while climate stress-related and insect-related are presented as decadal averages

differences across climate models, although the relative ranking of risk by SSP was consistent by the end of the century. The substantial differences between high and low emissions SSPs emphasises the critical importance of climate policy to mitigate climate risks to US forests.

We then conducted a risk assessment to quantify which regions and forests are likely to experience the highest risks in the 21st century (Figure 3, average of 2080–2099). By the end of the 21st century, high levels of fire risk, which were historically confined to pockets in California and the intermountain western US, are projected to expand across the entire western US (Figure 4). While these risks are substantially mitigated by emissions reductions (SSP2-4.5, Figure 3a), risks are still projected to increase dramatically in regions like the Great Plains in the central US and southeastern US (Figure 4).

Future climate stress risks increased most across broad swaths of the intermountain and southwestern US, California, and western Texas, although parts of the eastern US and the upper midwestern US also exhibited increased climate stress mortality risk (Figures 3b, 4). Projected insect risk to forest permanence was highest across the Rocky Mountains in the intermountain western US, Sierra Nevada mountains in California, and parts of the northern Midwest (Figures 3c, 4). Climate stress and insect mortality model projections were only made for forest types where models showed skillful cross-validated performance (i.e. AUC > 0.6) and thus lower risk in some regions (e.g. southern pine beetle risk in the southeastern US (Weed et al., 2013)) may reflect data and model limitations rather than inherently lower risks.

DISCUSSION

Our results provide a synthesis of fire-, climate stress- and insect-driven climate risks to forests in an open-source dataset available at continental scales. Climate-sensitive risks to US forests have major impacts on forest C cycling and climate change feedbacks, and thus quantifying forest permanence risks is important for conservation and climate policy efforts. Tree mortality and disturbance are large uncertainties in current land surface and vegetation models (Bugmann et al., 2019; Fisher et al., 2018; Pugh et al., 2019) and better large-scale historical datasets are needed for benchmarking and improving these models. Thus, the disturbance risk and mortality maps and their climate sensitivities derived here can help advance C cycle models. Our results reveal that US forests are very likely to experience increasing risks from climate change that undermine their C sequestration potential, an important factor that should be considered in climate change mitigation policy.

The spatial patterns in our risk models — both historical and future risks — broadly agree with other similar efforts of individual disturbances in the literature, such as the burn area patterns of large fires (e.g. Barbero et al., 2014) or drought risks (Buotte et al., 2019). The spatial patterns of insect model projections are consistent with previous projections for several major insect species (Bentz et al., 2010) and overall magnitude is similar to coarse-level ecoregion projections in parts of the western US (McNellis et al., 2021) (Figure 4). Further, the climate sensitivities of insect mortality for several western US pine species with the highest historical insect-driven mortality were consistent with estimates in the literature (Figure S4) (Bentz et al., 2010; Creeden et al., 2014).

Our statistical modeling approach with static vegetation for estimating future climate risks to forests due to

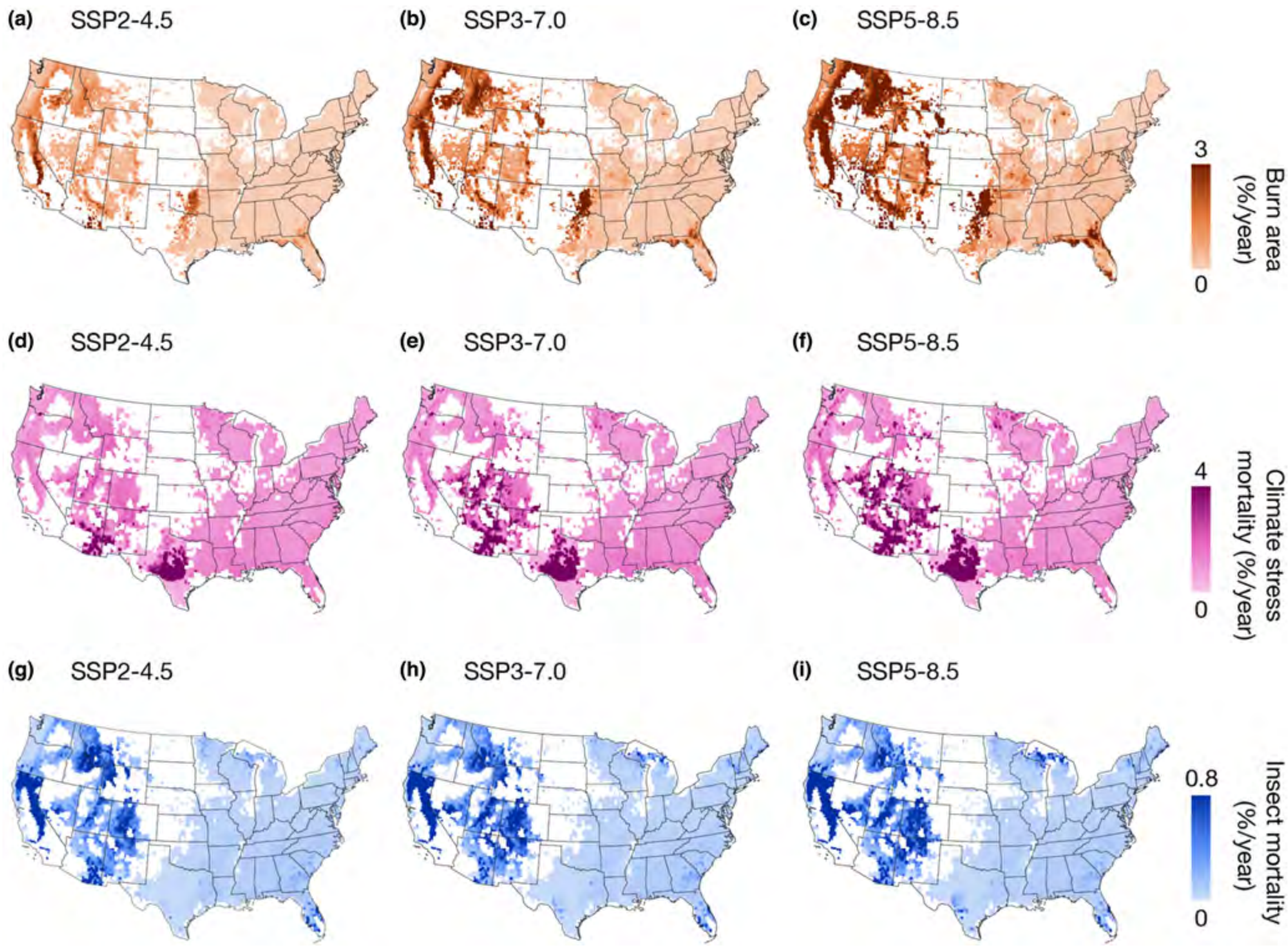


FIGURE 3 Risks for fire, climate stress-driven tree mortality, and insect-driven tree mortality (rows) averaged over the 2080–2099 period, separated by shared socioeconomic pathway (SSP) climate scenario (columns). Note that color-bars are substantially expanded relative to those in Figure 1 in order to visualise future projections that exceed historical risks

fire, climate stress, and insects is an informative 'first-step' analysis and complementary approach to current process-based methods for several reasons. First, accurately capturing mortality due to fire, climate stress and insects is quite challenging and often not included in state-of-the-art land surface models used in CMIP6 projections. Mortality of any type is a major uncertainty in process-based models, and efforts to predict drought mortality from first principles are nascent and still need substantial work (Bugmann et al., 2019). Prognostic insect-driven mortality is completely absent in CMIP6 mechanistic models currently (Fisher et al., 2018). Prognostic fire is present in some CMIP6 models, but none are able to capture the extent of current extreme fire events (Fisher et al., 2018). The statistical approach presented here is rigorously validated against historical observations and likely provides an upper bound of the extent of future disturbance given the lack of vegetation-disturbance feedbacks and dampening factors, further discussed among the several important caveats and limitations below. These models and approaches could be applied in other regions or countries by leveraging global fire data from MODIS and forest loss/disturbance data from Landsat, bringing in ground plot networks where possible and accounting for direct human land-use change.

These climate-sensitive risk maps and projections provide spatial quantification and uncertainty assessment across climate models, climate scenarios, and risk models that can inform risk management and conservation decisions. To support these aims, all data and code underlying these models are publicly available, and can be easily accessed and visualised via a web portal (https://carbonplan.org/research/forest-risks). As with any analysis, these projections are subject to several uncertainties and caveats. In addition to uncertainties in underlying Earth system models and statistical climate downscaling approaches (SI Methods), these projections use empirical models based on static forest composition and structure over the 1984–2018 period. Thus, these projections do not account for shifts in forest composition or distribution, interactions among risks, and carbon dioxide effects on plant drought stress. In particular, large-scale impacts of fires, drought, or insects could substantially reduce

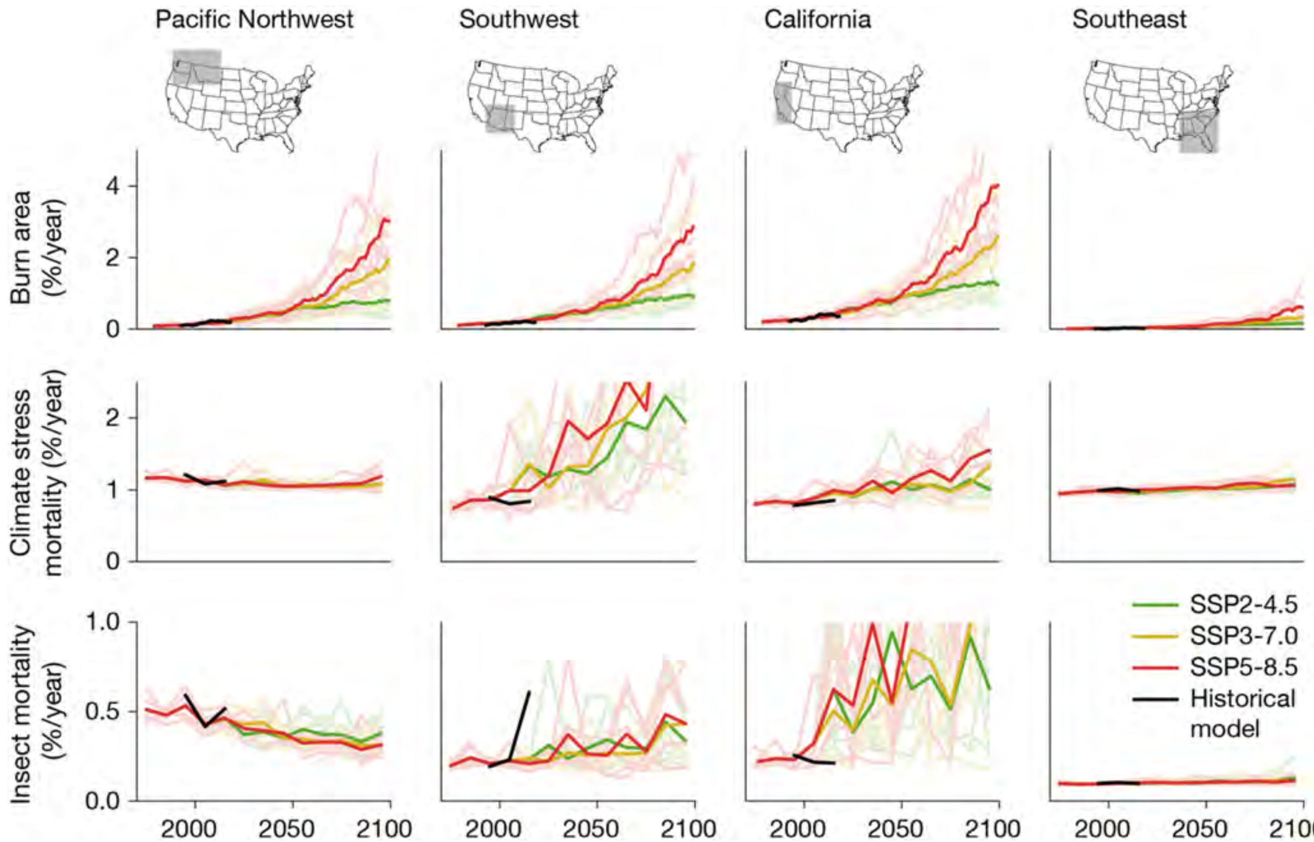


FIGURE 4 Regionally averaged time series of forest risks from three impacts: fire (top), non-fire, climate stress-driven tree mortality (middle), and insect-driven tree mortality (bottom). Regions of interest shown as gray boxes in the maps at the top of each column. As in Figure 2, statistical risk model simulations from each climate model simulation are shown as transparent lines, coloured according to the three shared socioeconomic pathway (SSP) climate scenarios. The multi-model mean for each SSP is shown opaque. Statistical model projections for each region driven by historical meteorological data (rather than meteorological data derived from a climate model) are shown in black. Fire risks are calculated with a 10-year centered moving average, while climate stress-related and insect-related are presented as decadal averages

biomass, and thus risk, although this is not likely to exert a material influence before the 2050s (Abatzoglou et al., 2021; Barbero et al., 2015). The risk projections also do not include impacts of land-use management, considered to be a strong potential lever in fire risk (Smith et al., 2016) and to a lesser degree climate stress and insect risks. We note, however, that there are also many reasons that these risk projections may be conservative or underestimates for insects and climate stress mortality, including projections made only for strong historical models, frequent non-linear impacts of drought and insects that may not be well-characterised in inventory data, and novel pests and pathogens. A comparison of our risk projections with mechanistic land surface models with prognostic fire from CMIP6 results revealed strong and consistent spatial correlations (Figure S7), providing additional confidence in the patterns of future risks and their impact on forest carbon sequestration.

Our results clearly show both spatial heterogeneity and future increases in risk across broad swaths of the continental US. While some current forest offset protocols incorporate risk, for example through the construction of 'buffer pools' or related insurance-like mechanisms, current risk estimates do not incorporate either form of variability (Anderegg et al., 2020). Thus, our findings raise serious questions about the integrity of these programmes. Further work could incorporate observed heterogeneity and future projections to better inform the construction of these climate programmes, such as by translating these risks into specific C loss estimates that could parameterise a better-grounded buffer pool or other insurance programme. Our results provide a critical starting point in quantifying risks over space and time and can inform management, conservation and policy actions. Taken in sum, our results increase the urgency and magnitude of response needed for reducing greenhouse gas emissions to mitigate climate change given the increasing risks of climate change to nature-based climate solutions.

ACKNOWLEDGEMENTS

The authors thank H Stanke and J Shaw for insights and assistance with FIA data analysis and processing. WRLA acknowledges support from the David and Lucille Packard Foundation, US National Science Foundation grants 1714972, 1802880, 2003017, and 2044937, and USDA National Institute of Food and Agriculture, Agricultural and Food Research

Initiative Competitive Programme, Ecosystem Services and Agro-Ecosystem Management, grant no. 2018-67019-27850. CarbonPlan acknowledges support from Microsoft AI for Earth. ATT acknowledges funding from the NSF Grant 2003205 and 2017494, the USDA National Institute of Food and Agriculture, Agricultural and Food Research Initiative Competitive Programme Grant No. 2018-67012-31496 and the University of California Laboratory Fees Research Program Award No. LFR-20-652467. Any errors are the authors' sole responsibility.

AUTHOR CONTRIBUTION

W.R.L.A., O.S.C., J.F., and J.J.H designed the project. W.R.L.A., O.S.C., G.B., A.T.T., J.F., and J.J.H conducted the analyses. W.R.L.A. and O.S.C. drafted the initial paper and all co-authors provided input on writing and analyses.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/ele.14018.

DATA AVAILABILITY STATEMENT

The source code to reproduce our analysis is available in https://doi.org/10.5281/zenodo.4741329. Archival versions of this project's data products are available in https://doi.org/10.5281/zenodo.4741333.

ORCID

William R. L. Anderegg https://orcid.org/0000-0001-6551-3331
Oriana S. Chegwidden https://orcid.org/0000-0003-1376-3835
Grayson Badgley https://orcid.org/0000-0003-1011-4573
Anna T. Trugman https://orcid.org/0000-0002-7903-9711
Danny Cullenward https://orcid.org/0000-0002-6803-9572
John T. Abatzoglou https://orcid.org/0000-0001-7599-9750
Jeffrey A. Hicke https://orcid.org/0000-0003-0494-2866
Jeremy Freeman https://orcid.org/0000-0001-7077-7972
Joseph J. Hamman https://orcid.org/0000-0001-7479-8439

REFERENCES

Abatzoglou, J.T., Battisti, D.S., Williams, A.P., Hansen, W.D., Harvey, B.J. & Kolden, C.A. (2021) Projected increases in western US forest fire despite growing fuel constraints. *Communications Earth & Environment*, 2, 227. https://doi.org/10.1038/s43247-021-00299-0

Anderegg, W.R., Hicke, J.A., Fisher, R.A., Allen, C.D., Aukema, J., Bentz, B. et al. (2015) Tree mortality from drought, insects, and their interactions in a changing climate. *New Phytologist*, 208, 674–683.

Anderegg, W.R.L., Klein, T., Bartlett, M., Sack, L., Pellegrini, A.F.A., Choat, B. et al. (2016) Meta-analysis reveals that hydraulic traits explain cross-species patterns of drought-induced tree mortality across the globe. *Proceedings of the National Academy of Sciences*, 113(18), 5024–5029.

Anderegg, W.R.L., Trugman, A.T., Badgley, G., Anderson, C.M., Bartuska, A., Ciais, P. et al. (2020) Climate-driven risks to the climate mitigation potential of forests. *Science*, 368, eaaz7005.

Archer, D., Eby, M., Brovkin, V., Ridgwell, A., Cao, L., Mikolajewicz, U. et al. (2009) Atmospheric lifetime of fossil fuel carbon dioxide. *Annual Review of Earth and Planetary Sciences*, 37, 117–134.

Barbero, R., Abatzoglou, J.T., Larkin, N.K., Kolden, C.A. & Stocks, B. (2015) Climate change presents increased potential for very large fires in the contiguous United States. *International Journal of Wildland Fire*, 24, 892–899.

Barbero, R., Abatzoglou, J.T., Steel, E.A. & Larkin, K.N. (2014) Modeling very large-fire occurrences over the continental United States from weather and climate forcing. *Environmental Research Letters*, 9, 124009.

Bentz, B.J., Régnière, J., Fettig, C.J., Hansen, E.M., Hayes, J.L., Hicke, J.A. et al. (2010) Climate change and bark beetles of the Western United States and Canada: direct and indirect effects. *BioScience*, 60, 602–613.

Bonan, G.B. (2008) Forests and climate change: forcings, feedbacks, and the climate benefits of forests. *Science*, 320, 1444–1449.

Brodribb, T.J., Powers, J., Cochard, H. & Choat, B. (2020) Hanging by a thread? Forests and drought. *Science*, 368, 261–266.

Bugmann, H., Seidl, R., Hartig, F., Bohn, F., Brůna, J., Cailleret, M. et al. (2019) Tree mortality submodels drive simulated long-term forest dynamics: assessing 15 models from the stand to global scale. *Ecosphere*, 10, e02616.

Buotte, P.C., Levis, S., Law, B.E., Hudiburg, T.W., Rupp, D.E. & Kent, J.J. (2019) Near-future forest vulnerability to drought and fire varies across the western United States. *Global Change Biology*, 25, 290–303.

Burnham, K.P. & Anderson, D.R. (2004) Multimodel inference understanding AIC and BIC in model selection. *Sociological Methods & Research*, 33, 261–304.

Clark, J.S., Iverson, L., Woodall, C.W., Allen, C.D., Bell, D.M., Bragg, D.C. et al. (2016) The impacts of increasing drought on forest dynamics, structure, and biodiversity in the United States. *Global Change Biology*, 22(7), 2329–2352.

Cook-Patton, S.C., Leavitt, S.M., Gibbs, D., Harris, N.L., Lister, K., Anderson-Teixeira, K.J. et al. (2020) Mapping carbon accumulation potential from global natural forest regrowth. *Nature*, 585, 545–550.

Cragg, J.G. (1971) Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica*, 39, 829.

Creeden, E.P., Hicke, J.A. & Buotte, P.C. (2014) Climate, weather, and recent mountain pine beetle outbreaks in the western United States. *Forest Ecology and Management*, 312, 239–251.

Eidenshink, J., Schwind, B., Brewer, K., Zhu, Z.L., Quayle, B. & Howard, S. (2007) A project for monitoring trends in burn severity. *Fire ecology*, 3(1), 3–21.

Fisher, R.A., Koven, C.D., Anderegg, W.R.L., Christoffersen, B.O., Dietze, M.C., Farrior, C.E. et al. (2018) Vegetation demographics in Earth system models: a review of progress and priorities. *Global Change Biology*, 24, 35–54.

Friedlingstein, P., Meinshausen, M., Arora, V.K., Jones, C.D., Anav, A., Liddicoat, S.K. et al. (2014) Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks. *Journal of Climate*, 27, 511–526.

Giraudoux, P., Giraudoux, M.P. & Mass, S. (2018) *Package 'pgirmess'*.

Griscom, B.W., Adams, J., Ellis, P.W., Houghton, R.A., Lomax, G., Miteva, D.A. et al. (2017) Natural climate solutions. *Proceedings of the National Academy of Sciences*, 114, 11645–11650.

Hember, R.A., Kurz, W.A. & Coops, N.C. (2017) Relationships between individual-tree mortality and water-balance variables indicate positive trends in water stress-induced tree mortality across North America. *Global Change Biology*, 23, 1691–1710.

Hicke, J.A., Allen, C.D., Desai, A.R., Dietze, M.C., Hall, R.J., (Ted) Hogg, E.H. et al. (2012) Effects of biotic disturbances on forest carbon cycling in the United States and Canada. *Global Change Biology*, 18, 7–34.

Holland, T.G., Stewart, W. & Potts, M.D. (2019) Source or Sink? A comparison of Landfire-and FIA-based estimates of change in aboveground live tree carbon in California's forests. *Environmental Research Letters*, 14, 074008.

Hurteau, M.D., Hungate, B.A. & Koch, G.W. (2009) Accounting for risk in valuing forest carbon offsets. *Carbon Balance and Managment*, 4, 1.

Hurteau, M.D., Hungate, B.A., Koch, G.W., North, M.P. & Smith, G.R. (2013) Aligning ecology and markets in the forest carbon cycle. *Frontiers in Ecology and the Environment*, 11, 37–42.

Jenkins, D.G. & Quintana-Ascencio, P.F. (2020) A solution to minimum sample size for regressions. *PLoS One*, 15, e0229345.

Kurz, W.A., Dymond, C.C., Stinson, G., Rampley, G.J., Neilson, E.T., Carroll, A.L. et al. (2008) Mountain pine beetle and forest carbon feedback to climate change. *Nature*, 452, 987–990.

Kurz, W.A., Stinson, G., Rampley, G.J., Dymond, C.C. & Neilson, E.T. (2008) Risk of natural disturbances makes future contribution of Canada's forests to the global carbon cycle highly uncertain. *Proceedings of the National Academy of Sciences*, 105, 1551–1555.

Lecina-Diaz, J., Martínez-Vilalta, J., Alvarez, A., Banqué, M., Birkmann, J., Feldmeyer, D. et al. (2021) Characterizing forest vulnerability and risk to climate-change hazards. *Frontiers in Ecology and the Environment*, 19, 126–133.

McNellis, B.E., Smith, A.M., Hudak, A.T. & Strand, E.K. (2021) Tree mortality in western US forests forecasted using forest inventory and Random Forest classification. *Ecosphere*, 12, e03419.

Nardini, A., Battistuzzo, M. & Savi, T. (2013) Shoot desiccation and hydraulic failure in temperate woody angiosperms during an extreme summer drought. *New Phytologist*, 200, 322–329.

Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A. et al. (2011) A large and persistent carbon sink in the World's forests. *Science*, 333, 988–993.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O. et al. (2012) Scikit-learn: machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.

Ploton, P., Mortier, F., Réjou-Méchain, M., Barbier, N., Picard, N. Rossi, V. et al. (2020) Spatial validation reveals poor predictive performance of large-scale ecological mapping models. *Nature Communications*, 11, 4540.

Pugh, T.A., Arneth, A., Kautz, M., Poulter, B. & Smith, B. (2019) Important role of forest disturbances in the global biomass turnover and carbon sinks. *Nature Geoscience*, 12, 730–735.

Roe, S., Streck, C., Obersteiner, M., Frank, S., Griscom, B., Drouet, L. et al. (2019) Contribution of the land sector to a 1.5°C world. *Nature Climate Change*, 9, 1–12.

Ruseva, T., Marland, E., Szymanski, C., Hoyle, J., Marland, G. & Kowalczyk, T. (2017) Additionality and permanence standards in California's Forest Offset Protocol: a review of project and program level implications. *Journal of Environmental Management*, 198, 277–288.

Seidl, R., Thom, D., Kautz, M., Martin-Benito, D., Peltoniemi, M., Vacchiano, G. et al. (2017) Forest disturbances under climate change. *Nature Climate Change*, 7, 395.

Shaw, J.D., Steed, B.E. & DeBlander, L.T. (2005) Forest inventory and analysis (FIA) annual inventory answers the question: what is happening to pinyon-juniper woodlands? *Journal of Forestry - Washington*, 103, 280–285.

Smith, A.M.S., Kolden, C.A., Paveglio, T.B., Cochrane, M.A., Bowman, D.M.J.S., Moritz, M.A. et al. (2016) The science of firescapes: achieving fire-resilient communities. *BioScience*, 66, 130–146.

Stanke, H., Finley, A.O., Domke, G.M., Weed, A.S. & MacFarlane, D.W. (2021) Over half of western United States' most abundant tree species in decline. *Nature Communications*, 12, 1–11.

Stanke, H., Finley, A.O., Weed, A.S., Walters, B.F. & Domke, G.M. (2020) rFIA: an R package for estimation of forest attributes with the US Forest Inventory and Analysis database. *Environmental Modelling and Software*, 127, 104664.

Trugman, A.T., Anderegg, L.D., Shaw, J.D. & Anderegg, W.R. (2020) Trait velocities reveal that mortality has driven widespread coordinated shifts in forest hydraulic trait composition. *Proceedings of the National Academy of Sciences*, 117, 8532–8538.

Wang, J.A., Baccini, A., Farina, M., Randerson, J.T. & Friedl, M.A. (2021) Disturbance suppresses the aboveground carbon sink in North American boreal forests. *Nature Climate Change*, 11(5), 435–441. https://doi.org/10.1038/s41558-021-01027-4

Weed, A.S., Ayres, M.P. & Hicke, J. (2013) Consequences of climate change for biotic disturbances in North American forests. *Ecological Monographs*, 83(4), 441–470.

Williams, A.P., Allen, C.D., Macalady, A.K., Griffin, D., Woodhouse, C.A., Meko, D.M. et al. (2013) Temperature as a potent driver of regional forest drought stress and tree mortality. *Nature Climate Change*, 3, 292–297.

Williams, C.A., Gu, H., MacLean, R., Masek, J.G. & Collatz, G.J. (2016) Disturbance and the carbon balance of US forests: a quantitative review of impacts from harvests, fires, insects, and droughts. *Global and Planetary Change*, 143, 66–80.

Zeileis, A., Cribari-Neto, F., Grün, B. & Kos-midis, I. (2010) Beta regression in R. *Journal of Statistical Software*, 34, 1–24.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Anderegg, W.R.L., Chegwidden, O.S., Badgley, G., Trugman, A.T., Cullenward, D., Abatzoglou, J.T., et al. (2022) Future climate risks from stress, insects and fire across US forests. *Ecology Letters*, 25, 1510–1520. Available from: https://doi.org/10.1111/ele.14018

CellPress
OPEN ACCESS

One Earth

Perspective

Avoiding carbon leakage from nature-based offsets by design

Ben Filewod[1,]* and Geoff McCarney[2]
[1]Grantham Research Institute on Climate Change and the Environment, London School of Economics and Political Science, London, UK
[2]School of International Development & Global Studies and Institute of the Environment, Faculty of Social Sciences, University of Ottawa, Ottawa, ON, Canada
*Correspondence: filewod@gmail.com
https://doi.org/10.1016/j.oneear.2023.05.024

SUMMARY

With nature-based offsets emerging as a core strategy for meeting near-term climate targets, it is essential they deliver real and verifiable mitigation gains. However, the interventions that generate offsets can have unintended effects that cause carbon leakage and ultimately reduce mitigation. Although leakage is "old news" and various anti-leakage measures have been considered, there is little evidence that current practices to address leakage actually work. In this perspective, we present evidence that leakage is vastly underestimated in practice and argue that current efforts to improve accounting methods are unlikely to deliver the accuracy required. We therefore propose and elaborate an alternative approach to address leakage by design, based on a new conceptual framework for understanding leakage in nature-based interventions. We further outline three principles that offset developers, certifiers, and consumers can implement now to improve the credibility of nature-based offsets, without negating further ambition and investment in nature-based solutions.

INTRODUCTION

Amid current enthusiasm for decentralized, market-led climate change solutions, ecosystems are widely seen as near-term linchpins of global mitigation strategies. Land use (notably forests) provides a quarter of planned mitigation under the Paris Agreement,[1] and COP26 (Glasgow) signaled global willingness to allow international transfers of nature-based mitigation. Nature-based offsets already feature in most market emissions pricing schemes[2] and are central to corporate net-zero pledges.[3] Driven by corporate commitments,[4] voluntary offset markets neared US$2 billion in traded value in 2021, with 67% originating in forestry and land-use projects.[5] Support for nature-based mitigation is broad (projects offer both low-cost mitigation and environmental benefits), and estimates of total potential are high (20%–30% of mitigation needed to keep global warming to 1.5°C).[6] Not all nature-based mitigation will (or should[7,8]) substitute for emissions reductions, but the role of nature-based offsets is rapidly expanding.

Most pathways to nature-based mitigation depend on altering the state of coupled ecological-economic systems.[9,10] The effect of such interventions can be tracked as credits in a carbon accounting framework, which become offsets when used to substitute for other mitigation actions. If the underlying carbon accounting is inaccurate, offsetting may prove a dangerous distraction. Recent work[11–13] has highlighted widespread overestimation of the degree to which nature-based interventions have altered the state of the world to generate mitigation (i.e., their additionality), but another source of inaccuracy—carbon leakage—merits equal attention. Leakage occurs when some effects of an intervention fall outside the accounting boundary used to track mitigation effects (e.g., an action causing emissions reductions in one place may also cause increases elsewhere). These beyond-boundary effects are extremely difficult to measure,[14,15] particularly when economic markets are implicated (so-called market leakage).

Today's voluntary and compliance markets routinely transact nature-based offsets on the premise that methods to account for leakage are sufficiently accurate,[16–18] and leakage is often framed as a tractable problem for which "sophisticated and robust tools"[2] and "policy levers to manage risks"[7] are available. Yet the main approach currently in use (adjusting issued credits using a leakage discount factor) has been recognized as insufficiently rigorous "in the long run,"[19] and how best to deal with leakage is controversial.[20,21] Among specialists, the problems are well known: Richards and Andersson[22] argued over 20 years ago that both theoretical and practical challenges prevent the accurate measurement of leakage associated with a specific intervention, concluding that "either the reliability of project analysis will be low or the costs of analysis will be high, and quite possibly both." Concern about leakage followed the first forestry offset projects in the early 1990s,[23–25] and three decades of work have now thoroughly explored the issue: several reviews exist,[19,26,27] and research interest remains high.[9,14,28] However, the solutions proposed are either politically intractable (i.e., scaling up accounting frameworks to include all beyond-boundary effects[21,29–33]) or poorly understood and inconsistently applied.[15,30] In the absence of a viable alternative, nature-based offsets continue to be issued and retired using ad hoc leakage accounting methods of unknown accuracy.

In this perspective, we aim to develop a viable alternative for dealing with carbon leakage from nature-based offsets. We

CellPress
OPEN ACCESS

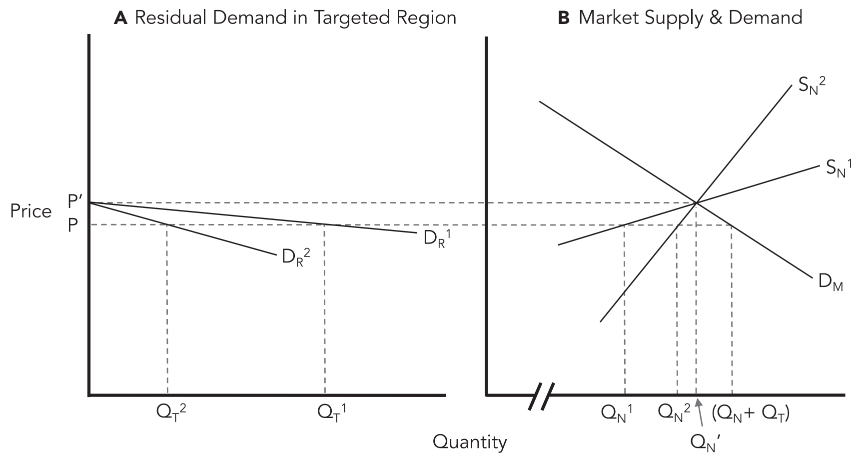


Figure 1. Market leakage occurs because price signals induce agents outside a targeted region to change behavior

Supply and demand diagrams allow a precise illustration of this phenomenon. S_N is supply from all non-targeted producers and is shown for two cases, relatively elastic supply (S_N^1) and relatively inelastic supply (S_N^2). The initial equilibrium quantity is $Q_N + Q_T$, the sum of supply from the non-targeted area (right panel) and targeted area (left panel). An intervention resulting in $Q_T = 0$ will cause the price to increase from P to P′, producing a new equilibrium at Q_N'. For the case of relatively inelastic supply, non-targeted producers had been producing Q_N^2 prior to the intervention, and market leakage is $Q_N' - Q_N^2$. For the case of relatively elastic supply, market leakage is $Q_N' - Q_N^1$. In both cases market leakage results from producers outside the targeted region moving up their supply curves due to the change in price resulting from the supply restriction in the targeted region. Note that $(Q_N' - Q_N^1)/Q_T^1 > (Q_N' - Q_N^2)/Q_T^2$, i.e., market leakage is proportionately greater for more competitive markets, such as those with fewer barriers to entry or lower transaction costs to displacing supply.

focus on market leakage, reviewing its theoretical basis, the empirical literature, and current accounting practice. We present evidence that suggests nature-based offsets in use today are systematically underestimating market leakage effects, and argue that attempts to refine leakage measurement will not deliver the accuracy needed in practice. We then propose a novel approach to designing nature-based interventions and associated offset markets to avoid the leakage problem: a framework for thinking about market leakage that enables identification of how and when it arises from nature-based interventions. Leveraging the available evidence through our framework, we outline three principles for a design-based solution to the market leakage problem that preclude the need for (currently unachievable) precision in leakage accounting. Our proposed alternative approach provides a conservative solution that can be immediately implemented within today's decentralized offset regimes, as well as clarifying ongoing misunderstandings about market leakage in a nature-based context.

THE PROBLEM WITH MARKET LEAKAGE

While a number of classifications of leakage[14,15,19,24,34] now exist (including impacts via connected ecological systems,[15] information, motivation, and institutions,[24] or spatial interactions[35]), we focus on two canonical types involving economic agents: "direct" leakage and "market" leakage. Direct (or "activity") leakage arises when the economic agents targeted by an intervention shift activities outside of the accounting boundary, whereas market leakage arises when non-targeted agents adjust their behavior in response to altered economic incentives. While activity leakage is relatively tractable (targeted agents are known, and their actions are observable), market leakage is not. Teasing out market leakage effects from background economic activity is extremely difficult, since which agents are responding to changes in incentives and how much of their behavior is due to this response depends on unobservable motivations.

The mechanism by which market leakage operates is information transmission through price. In an economic model of a single market equilibrium, reducing output from one producer causes prices to rise, moving the system out of equilibrium and incentivizing producers and consumers to respond as the system adjusts along a new equilibrium path. The net result of such equilibrium adjustments in interconnected markets is complex, but theory provides some general guidance.[36,37] All else being equal, market leakage will be lower if demand is elastic with respect to price or alternative products are not substitutable, and higher when supply is more elastic with respect to price, substitutable goods have a higher net carbon footprint, or supply restrictions are small (Figure 1). These theoretical conditions raise concerns about market leakage from nature-based offsets, which typically operate within globalized "food, fuel, and fiber" commodity markets where products are highly substitutable, demand is relatively inelastic with respect to price,[38–40] and markets are very large relative to the size of interventions.

While there are several approaches to measuring market leakage, each has its limitations. Partial or general equilibrium models are arguably the most suitable because they are developed specifically to capture market interdependencies (equilibrium effects) by simulating the actions of economic agents. Unfortunately, subtle changes in parameters can substantially alter results[39] (as demonstrated under our first principle below), and building such models requires highly specialized personnel and abundant data. Accounting-based approaches (e.g., input-output analysis or material flow analysis) offer an alternative and can provide compelling circumstantial evidence[41,42] but cannot separate out the causal impact of a specific intervention. Quasi-experimental econometric techniques can, but rely on assumptions about the independence of pseudo-controls that are violated by the presence of leakage effects. Many researchers therefore apply simple zone-based methods,[43,44] also widely used to quantify activity-shifting leakage (e.g., from protected areas[45,46]). This requires the assumption that leakage occurs within known areas that are unaffected by background economic incentives, an untenable premise in the well-functioning and large-scale markets where indirect effects are of most concern. Although innovations are ongoing,[47,48] an accessible, replicable, and accurate method of estimating market leakage has not yet been found.

Thus far, assessments of market leakage from nature-based interventions in the literature have focused primarily on "stop-harvest" forest mitigation projects, frequently leveraging established partial equilibrium models. Leakage estimates for developed countries from these models are typically at least 70% of reduced output, measured either as forestry production[37,49–52] or carbon stocks.[53] Lower estimates (50% or less) have been found for the specific case of a global carbon price,[54] or in a developing country context when international leakage is deemed negligible[55] (e.g., Kuik[56] estimates 0.5%–11.3% market leakage from large national supply restrictions in developing countries, but this result depends on methodological choices and is contradicted by other evidence[47]). Afforestation scenarios may possibly produce lower leakage than avoided conversion (e.g., $\leq$43% vs. $\leq$ 92% in one estimate[36]) because of productivity differences or the availability of underutilized land.

Carbon leakage from non-forest interventions is less well studied. Kim et al.[57] find about 15% leakage from crop conversion, while econometric studies of leakage from conservation reserves (also known as "slippage") suggest that leakage is important[58] but possibly low[59]: estimates include 4% activity leakage (measured as forest cover loss)[43] and 20% market leakage (measured as farm area,[60] although criticisms of this estimate[61–63] highlight measurement challenges). Since marginal farmland may be preferentially enrolled in conservation programs,[43] estimated leakage in this context may in fact be low because additionality is weak (an idea we return to in the next section). Further econometric evidence is available for forest-to-agriculture conversion in Brazil,[47,64,65] but only adds to the uncertainty, as leakage estimates range from insignificant to essentially all program gains. A growing literature also considers carbon leakage associated with the unilateral adoption of various climate policies (e.g., carbon taxes), but the potential for these studies to inform leakage estimates for nature-based offset projects is unclear.

In our view, the empirical literature on market leakage supports two inferences. First, leakage from nature-based interventions can be very high. Failing to identify the true level of leakage could lead to credits that grossly overstate actual mitigation impacts, and widespread use of these credits as offsets could put climate policy targets in jeopardy. Second, leakage is context specific.[36] Results from one intervention (or averages of them[66]) do not provide an accurate estimate of leakage from another intervention. The problem with market leakage is thus a problem of measurement: to accurately assess market leakage, the dynamic adjustment of a complex system must be measured every time.

A CONCEPTUAL FRAMEWORK FOR AVOIDING LEAKAGE BY DESIGN

If market leakage cannot be accurately measured in practice, nature-based interventions must be designed to minimize leakage risk while ensuring that offset markets are able to identify and apply high-risk offerings in ways that do not jeopardize critical near-term mitigation goals. Doing so requires a nuanced understanding of when and how leakage arises. We therefore elaborate a conceptual framework for understanding leakage in coupled economic-ecological systems and apply this framework to the complete set of possible nature-based mitigation interventions to identify design choices that lead to leakage risk.

Our approach is derived from the observation that when nature-based interventions are used as offsets, the solution space is bounded by the need to simultaneously satisfy three well-known criteria: permanence, additionality, and (no) leakage.[67] Reframing these criteria as simultaneous constraints is important. Recent debate[68–74] around the biophysical potential of nature-based mitigation has tended to consider each issue separately, obscuring overlaps and trade-offs between them. Reframing them as simultaneous constraints unifies the challenges to real nature-based mitigation, thereby scoping down the set of possible nature-based interventions and enabling a design-based solution to market leakage.

In implementing this new simultaneous constraints framework, our main concern is the existence and implications of a conceptual "duality" between additionality and leakage that emerges from our approach. When a market system is in equilibrium, any intervention that achieves additionality by altering supply or demand in a particular market will transmit information to connected markets through price changes (both for products in that market and for markets for secondary products or production inputs). The resulting adjustments across the entire economic system are known to economists as general equilibrium effects; those that fall outside a carbon accounting boundary are also known as leakage. This is the duality at the heart of our design-based solution: when carbon accounting does not cover the entire economic system, additionality and leakage are two sides of the same coin. Carbon credits that rely on altering supply or demand to claim additional mitigation also inherently create market leakage risk. Conversely, market interventions that do not generate leakage risk are likely not additional.

To apply this insight to design credible nature-based offset markets, we must focus on the additionality claim(s) made to issue carbon credits. Nature-based interventions do not necessarily rely on altering market supply or demand to achieve additionality, and a single offset project may involve multiple additionality claims; for example, forest cover loss may be avoided by banning logging while also improving household fuelwood efficiency. Whether market leakage is important relative to intended mitigation therefore depends on the degree to which a project alters market equilibria to claim additionality. To identify at-risk additionality claims, we trace out the set of possible mitigation interventions in coupled economic-ecological systems (Figure 2) and apply our concept of simultaneous constraints to identify those interventions at risk of market leakage.

Of course, there are other constraints beyond permanence, additionality, and leakage (e.g., maintaining non-carbon values or establishing the certainty of emissions baselines). Including these challenges leads us to identify three most credible categories of nature-based intervention (gray highlights in Figure 2). In economies, interventions that reduce aggregate demand or decarbonize production are at relatively lower risk of market leakage (especially if decarbonization results from non-transferable innovation). Interventions that reduce supply are at high risk. Interventions that transition economically unused ecosystems between stable states can also generate leakage-free mitigation while avoiding the problems associated with reducing disturbance or establishing non-native ecosystem states. Thus, by reframing challenges to nature-based mitigation as simultaneous

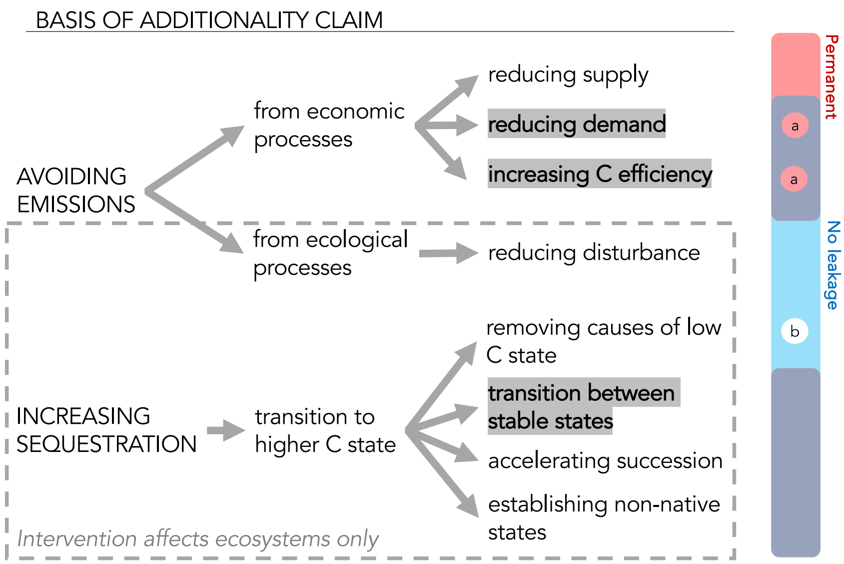


Figure 2. Generalized classification of nature-based interventions (flowchart) showing permanence or leakage constraints (colored bars)
Starting with the basic dichotomy of avoiding emissions or increasing sequestration, we identify eight ways in which nature-based interventions can generate additional mitigation. Three of these (gray highlight) have minimal leakage risk, are permanent, and satisfy other criteria such as preserving biodiversity (see main text). Note that (a) our preferred economic interventions (gray highlight) can still cause market leakage in some circumstances, and (b) removing causes of low-carbon states can cause ecological leakage.

constraints and focusing on the additionality claims made to issue carbon credits, it becomes possible to identify design choices that lead (or not) to market leakage risks.

To justify these claims, it is helpful to examine the different additionality pathways for nature-based mitigation presented in Figure 2. We first consider interventions in ecosystems alone (i.e., no economic implications and no market leakage; dashed box in Figure 2). The primary way such interventions deliver mitigation is by increasing sequestration, which requires transitioning ecosystems to higher carbon states. If the new state is naturally occurring, it may be a later successional stage (e.g., shrubland to forest). Since succession would have occurred anyway, baseline dynamics must be netted out in crediting (i.e., the offset is additional in time only; accelerating succession in this way can still deliver useful mitigation if baseline succession is slow). If the new state is not naturally occurring (e.g., shrubland to non-native plantation), non-carbon values such as preferences for native biodiversity restrict large-scale deployment.

However, unlike the economic concept of a unique equilibrium, ecological systems can exist in multiple naturally occurring stable states[75] (e.g., savanna/closed forest, rock/kelp). Understanding why ecosystems persist in a stable low-carbon state when naturally occurring high-carbon states are possible allows us to differentiate two important subcategories of ecosystem interventions. Both currently active processes (abiotic or biotic disturbance agents, such as fire or grazing megafauna[76]) and the effects of past actions (path dependency) can cause stable low-carbon states. Removing the most important active processes (fire, pests, pathogens) is unlikely to be permanent, and accurately modeling baselines is extremely challenging. Meanwhile, more tractable removals of "ecosystem engineers" (e.g., large grazers) is constrained by ecological leakage if relocated or non-carbon values if eliminated. Conversely, if low-carbon stable states exist because of the history of past events alone (i.e., due to path dependency), interventions can shift ecosystems between stable equilibria to achieve both additionality and permanence without market implications or leakage risk. Restoring degraded but abandoned land is the most prominent example.

The pathways classified as "avoiding emissions from economic processes" in Figure 2 reintroduce economic considerations. Following the "duality" we identify above, any economically additional intervention in this category alters market equilibrium by definition, thus producing price changes that result in altered economic behavior (leakage) throughout the connected economic system. There are only two exceptions: if the carbon intensity of ecologically derived goods can be reduced while maintaining the flow of such goods into economies (e.g., via material or process substitutions), price changes may not occur, and if no substitutes exist for such goods then price changes are irrelevant. When neither condition is met, reducing market leakage below 100% of claimed mitigation requires that alternative output is only available at higher prices, thereby causing quantities demanded to fall. This assumption is a problem for (relatively) small projects without market power, a category that arguably encompasses most nature-based offsets issued to date.

Nature-based interventions that reduce supply to markets are accordingly highly likely to be interventions at high risk of leakage. This is true whether what is being reduced is the supply of goods or of factors of production such as land, and even if this reduction is temporary. We recognize, however, that reducing supply is not the only way market-exposed nature-based offsets can claim additionality. Reducing the carbon footprint of economic activity (i.e., increasing emissions efficiency) can deliver economic additionality without leakage (provided the price and quantity of outputs remains unchanged) and is an important mitigation strategy.[77] This broad category of interventions includes projects that maintain output while reducing inputs (e.g., optimal rotation grazing) and those that substitute low-carbon for high-carbon service delivery (e.g., green infrastructure). The other option is to reduce demand, which can avoid emissions or generate sequestration without causing adverse effects outside the accounting boundary by removing anthropogenic causes of low-carbon ecosystem states.

Of course, general equilibrium effects must still be considered whenever markets beyond the accounting boundary are implicated. For example, increases in efficiency can lead to increased production via price reductions or firm entry (i.e., rebound effects[26]), and reduced demand can suppress prices and incentivize increased consumption elsewhere. Such effects depend on the connection between interventions and external markets. For example, rebound effects from culturally specific changes in resource management are less likely, while rebound effects from transferable technological innovation are more likely. The primary concern, however, are interventions that alter markets by increasing prices or reducing the supply of outputs, which are more directly tied to leakage risk. Finally, note that efficiency gains must be demonstrated within a credited project to credibly avoid the leakage problem. It is not sufficient to assume strong spatial and temporal coordination between multiple projects when issuing credits, although this is a common assumption in large-scale assessments of nature-based mitigation potential.[10,78,79]

THREE PRINCIPLES TO AVOID LEAKAGE BY DESIGN

The conceptual framework outlined above allows leakage risk to be located within specific nature-based additionality claims. Doing so provides a means for project proponents to design nature-based interventions that avoid market leakage risks and for buyers (or evaluators) to assess the likely leakage risk of an intervention. To facilitate application, the remainder of this perspective outlines three key principles for moving offset markets toward the design-based solutions our framework implies.

Our first principle captures a necessary design feature to control market leakage at the level of an individual project issuing carbon credits, while our second principle provides a critical demand-side safeguard for compliance carbon markets. Our third principle proposes a general rule for designing offset markets to deal with leakage (and other sources of uncertainty) and recognizes that at-risk interventions can still advance mitigation goals if the use of associated credits is appropriately targeted. Implementing these principles does not require coordination between market participants, but we stress their complementary nature. Correctly applying principle 1 will be essential to the proper application of principle 3, while principle 2 provides an important restriction and safeguard on the overall implementation of our framework.

Principle 1: When the design of a nature-based intervention implies market leakage risks, upper-bound estimates of potential leakage should be used

There is widespread agreement that accounting methods for market leakage should be conservative (i.e., biased toward overestimating leakage effects).[15] We present evidence which strongly suggests that the opposite is true in practice. In the absence of reliable, low-cost methods for market leakage accounting, third-party certification standards have been forced to rely on ad hoc approaches with mixed (often low) evidential standards. Since research-quality estimates are costly and are highly context specific, ensuring the use of upper-bound estimates is a conservative design-based alternative. Some steps have been taken in this direction (e.g., the current verified carbon standard [VCS] "Agriculture, Forestry, and Other Land Use" requirements apply a 100% discount factor to calculate leakage in some cases), but the use of arbitrarily low estimates of possible leakage appears widespread.

To assess whether current crediting practices are conservative, we reviewed a small random sample of credited projects from the most important nature-based offset methodologies (by issued credit volumes). All-time issuances for the selected methodologies are about 480 Mt, or roughly 2 years' worth of the annual reductions Canada needs to hit its 2030 emissions target. All the methodologies we reviewed are for forest-based interventions, and all adjust issuances for market leakage by applying a discount factor at the project level. The range of possible discount factors included in a methodology defines minimum and maximum potential leakage rates for credits issued under that methodology; these ranges were from 10% or 20% up to a maximum of 70% in our review (Table 1, column 8). In our random sample, the discount factors selected and applied by projects (Table 1, column 9) were almost always the minimum possible value. For Verra-registered projects, additional data showed a similar phenomenon in projections of total (market + activity) leakage for sampled projects (median values: VM0007 11%, VM0009 10%, AR-ACM0003 0%, VM0015 6%, VM0006 4%; not reported in Table 1). A recent study[28] of avoided forest conversion or degradation projects corroborates our results, with 26 out of 68 projects claiming no leakage and 28 deducting expected leakage at a median rate of just 6%.

These results contrast sharply with research estimates of market leakage from forest-based interventions, which are typically above 70% (and can reach >100%; see "the problem with market leakage"). Since the methodologies we reviewed include a range of interventions and market contexts, why are the leakage rates allowed by standards and applied by projects uniformly so much lower than those suggested by the research literature? One explanation relates to technical complexity in market leakage accounting. Because accuracy is difficult and therefore costly, standards must negotiate a compromise between scientific rigor and financial viability (see work by Cashore and colleagues[80,81] for related political economy concerns). However, the compromises that have been made in practice can seriously distort carbon accounting systems. Consider the general exclusion of difficult-to-measure leakage beyond country borders (unwarranted in light of research results,[42,49,50] but in alignment with international norms in carbon accounting). One randomly sampled project (#1175 on the Verra Registry) applied this principle to rule out leakage effects from the 87% of foregone output destined for export. An alternative explanation for low leakage rates is expediency. Leakage deductions can make or break the financial case for an offset project and are a key concern for project developers.[82] Once rules are in place, project proponents are financially incentivized to apply the lowest possible discount factor, and there are minimal controls on strategic behavior. The evidential standard for selecting a discount factor is weak (typically subjective assessment of likely leakage location, expert opinion, and/or selective appeal to research literature), and the effectiveness of auditing is limited by a lack of external sources of information and potential conflicts of interest.[20]

Refinements to leakage accounting are unlikely to solve the problem of achieving accuracy in practice. Consider efforts to develop tractable leakage estimation formulas,[56,57,83] which aim to approximate the adjustment of an economic-ecological system toward a new equilibrium using a limited set of parameters. The formula of Murray et al.[36] (Box 1) is the best known and is widely positioned as the most rigorous option for project-level

Table 1. Market leakage in third-party forest carbon standards

Registry	Methodology	Project type[a]	Volume (Mt)[b]	Trigger	International leakage	Approach	Possible range	Median value[c]
Verra (verified carbon standard)	VM0007 v1.6 (framework) (VMD0011 v1.1)	multiple (REDD+)	145.7	reduction of wood products supply (to markets >50 km from project area)	no	discount factor (*wood products*) or VMD0037	20%–70% of foregone supply (timber)	0% (0%–40%)
							40% of foregone supply (fuelwood/charcoal)	
	VM0009 v3.0 (VMD0037 v1.0)	AC (forest, grassland)	102.78	reduction in commodity supply	no	discount factor (*wood products*) or VMD0037	10%–70% of foregone supply (discount factor)	0%
							≪30% of foregone supply (VMD0037 approach)	
	AR-ACM0003 v2.0	A/R	14.86	market leakage is not monitored				
	VM0015 v1.1	A(U)C (forest)	73.2	market leakage is not monitored				
	VM0006 v2.2 (VCS AFOLU Requirements v4.1)	A(U)C (forest), A(U)D (forest)	10.9	reduction in wood products supply	no	discount factor (*per pool*)	20%–70%	20% (0%–20%)
Gold Standard	Afforestation/ Reforestation v1	A/R	0.46	market leakage is not monitored				
American Carbon Registry	IFM, US Non-Federal v1.3	IFM	6.66	reduction in wood products supply (>5%)	no	discount factor (*total credits*)	10%–40%	40%
	A/R Degraded Land v1.2 (AR-TOOL15 v2.0)	A/R	3.69	market leakage is not monitored				
	US Forest Projects v1 (compliance protocol)	A/R, IFM, AC	121.84	reduction in wood products supply	no	discount factor (*wood products*)	20%	20%
In development	ART-TREES v2.0	REDD+	NA	subnational scale	no	discount factor (*total credits*)	0%–20%	NA
	BC Forest Carbon v2.0 (compliance)	A/R, IFM, AC	NA	reduction in wood products supply	yes	discount factor (*total credits*)	47.37%–71.89% (default)	NA

Italic text in "Approach" indicates which carbon pool is discounted to adjust crediting for leakage. Median values fall below possible ranges when projects report no market leakage. Median values are based on *ex ante* projections.

[a]A(U)C, avoided (unplanned) conversion; A/R, afforestation/reforestation; IFM, improved forest management; REDD+, multiple forest pathways.

[b]Issuances on public registries, all-time.

[c]Mean value (range), based on a random sample of five or ten registered projects. Ranges are not reported where all values were identical.

Box 1. A leakage calculation formula

Murray et al.[61] provide a widely used formula for estimating market leakage from foregone forest harvest, which approximates the adjustment of an economic-ecological system toward a new equilibrium:

$$\text{Leakage}\ (\%) = \frac{100 * e * \gamma * C_N}{[e - E(1+\gamma * \varphi)]C_R}.$$

The physical subsystem is represented by C_R and C_N, the carbon "footprints" of harvest in a reserved and non-reserved forest area, respectively. The size of the supply restriction is represented by $\varphi \in [0, 1]$ (i.e., the fraction of total supply restricted by the offset). The adjustment of the economic subsystem is captured by the substitutability of timber from the reserved and non-reserved area $\gamma \in [0, 1]$, the price elasticity of supply e (the percent change in supply caused by a percent change in price), and the price elasticity of demand E (the percent change in demand caused by a percent change in price).

This simple approach clearly demonstrates the core mechanics of market leakage for a good experiencing a supply restriction. Market leakage will be higher when production is displaced to a location with a higher carbon "footprint" ($C_N > C_R$), when suppliers are more responsive to changes in price ($|e|$ large), or when demanders are less responsive ($|E|$ small). It will be smaller when foregone output is less substitutable ($\gamma < 1$) and proportionately larger when the supply restriction φ is small, because price increases (and hence reductions in demand) will be less.

leakage accounting in both voluntary and compliance methodologies. For results from applications of this formula to be accurate, the parameter estimates used must correctly describe the measured system. This is not a trivial problem,[67] not least because key economic parameters (e.g., the price elasticities of demand and supply, which describe consumer and producer behavior) are not stable over space or time[39,84] and are difficult to estimate.

To recognize the uncertainty in this approach (and thus the need for upper-bound estimates), consider the following example. A forest conservation project (Verra Registry #607) issues carbon credits based on reducing lumber output in southern British Columbia (BC), Canada. Applying the Murray et al. formula yields a leakage estimate of about 69% (in contrast, project documents indicate a discount factor of 20% and an actual deduction of about 11% on recent issuances, with 2.9 Mt retired so far). This estimate employs default regional parameters provided by BC's draft forest carbon protocol (row 12 in Table 1). Varying the elasticity parameters (e and E in Box 1) by 25% to approximate reasonable confidence intervals yields estimates ranging from 58% to about 78%. This sensitivity is a problem: regionally specific estimates of these parameters reported in the literature span two orders of magnitude[85] and vary markedly over time.[84] In less data-rich contexts (for example, many developing countries) the market data necessary to estimate these parameters are unlikely to be available, forcing proponents to apply estimates out of context.

Given the potentially systematic underestimation of market leakage suggested by our evidence, the lack of low-cost and accurate methods for leakage estimation, and high inherent uncertainty in leakage estimates, we argue that mandating the use of upper-bound possibilities for projects including market leakage in their design is essential to build conservativeness into the accounting system. In practice, this may often mean the application of leakage rates approaching 100% for high-risk interventions. Since both theory and evidence suggest that the risk of leakage from market-exposed nature-based offsets is generally high, the burden of proof ought to be on project proponents to credibly demonstrate low rates. This requires a reasonably complete model of the economic system (including international markets, if implicated) and context-specific parameterization. In general, it will be preferable to design interventions that avoid market leakage from the start.

Principle 2: Nature-based credits which include market leakage risk in their design should not substitute for avoided emissions in compliance settings

The integrity of nature-based offset schemes as a mitigation strategy depends on the accuracy of offset accounting methods. The relationship between additionality and leakage outlined in our conceptual framework implies that some degree of market leakage is inevitable when additionality results from altering economic behavior and market linkages extend beyond project accounting boundaries. Accounting for market leakage is most risky in compliance settings (e.g., in cap-and-trade markets), where nature-based offsets substitute on a one-to-one basis for avoided emissions in meeting policy objectives. Substituting uncertain offsets for certain emissions reductions risks decoupling measured progress toward policy targets from physical changes in stocks of atmospheric greenhouse gases, but a design-based approach can circumvent the problem by avoiding additionality claims that rest on leakage-generating market interventions.

To substantiate our concerns about leakage in compliance settings—and illustrate the implicit relationship between leakage and additionality articulated in our framework—we present a global assessment of forest cover loss and protected areas in Earth's tropical forest biomes (Figure 3). Our calculations show that, over the last two decades, steady increases in protected forest areas have not been associated with falling forest cover loss (Figure 3A). This pattern of non-declining forest cover loss despite ongoing protection raises the possibility of widespread leakage—or, following the duality we highlight, the possibility of widespread non-additionality. Credits could be issued if it could be shown that forest cover loss would have been higher in the absence of protection, but doing so would require a credible counterfactual baseline (and a robust leakage estimate).

Figure 3B illustrates several challenges in teasing out market leakage from "background" economic activity. Zonal statistics

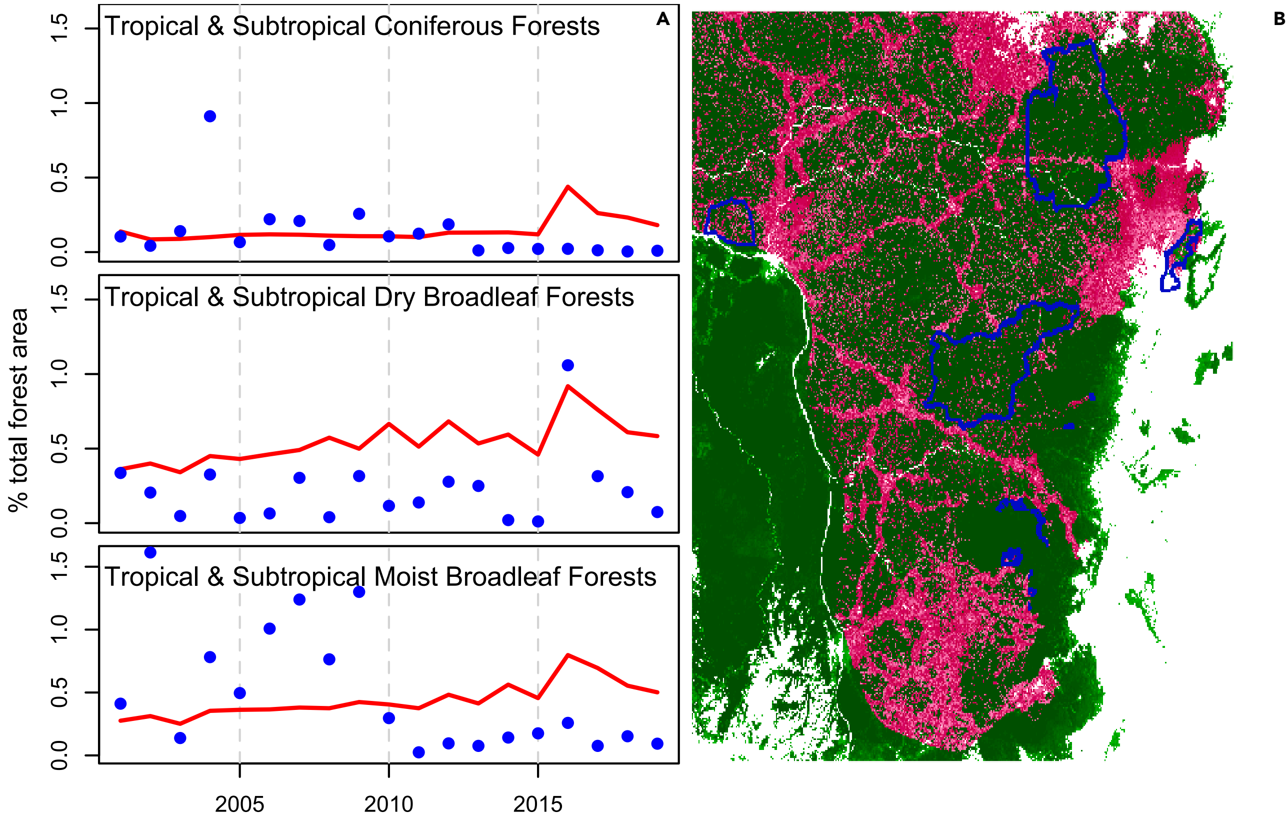


Figure 3. Forest cover loss and area protected in Earth's tropical forest biomes
(A) Summaries by biome of annual area protected (blue dots) and forest cover loss (red lines).
(B) Colorblind palette shows part of one ecoregion-level calculation (n = 279) used to generate the biome-level summaries, which illustrates challenges to leakage accounting (see main text). For each ecoregion, we tabulated annual changes in area protected as well as forest cover loss (pink shading; darker is more recent) outside protected areas (blue outlines) in forests with similar canopy closure (green shading; darker is higher closure).
Aggregating results at the biome level (A) reveals non-declining cover loss despite steady annual increases in forest protection, and a simple panel regression (not reported) of cover loss on area protected at the ecoregion level confirmed the lack of a significant relationship. Zooming in on loss in one randomly selected ecoregion (pink shading in B) shows the difficulty of teasing out leakage effects around protected areas from the economic "background" of landscape-level forest cover loss.

(e.g., using a fixed-width buffer around protected areas) require making assumptions about where leakage occurs, and leakage estimates will clearly vary with these assumptions. Quasi-experimental methods require pseudo-controls, and since all ecologically similar forests exhibit cover loss which may be due to leakage, the independence of these is doubtful. Measuring leakage by modeling the behavior of economic agents requires rich microdata, which are unlikely to be available in our example (the Northeast Congolian lowland forest ecoregion, located mostly within the Democratic Republic of the Congo). The insight provided by our framework is that because protection reduces the supply of forested land, anywhere that protection is additional is also at high leakage risk by design—thereby raising the stakes for accurate leakage estimation. Because measuring this leakage risk accurately is extremely difficult, credits from such projects are highly uncertain.

Principle 2 offers a conservative design-based solution to irreducible uncertainty about the true impact of an intervention (such as forest protection), by prohibiting high-risk carbon credits from being used as offsets in compliance settings. In effect, principle 2 can be viewed as an extension of principle 1, which requires application of a 100% leakage rate in circumstances where the costs of incorrect leakage estimates are high. Given the high rates of market leakage estimated in the literature for at-risk projects and that any attempt to justify lower leakage rates will necessarily rely on uncertain estimation practices, the cost of allowing real emissions for uncertain offsets in compliance settings is simply too great. In practice, this means prohibiting prominent nature-based interventions, likely including many current-generation REDD+ projects, from being used as offsets to meet legally required climate targets. However, our underlying conceptual framework reveals that—contrary to prior concerns[86]—avoiding high-risk projects does not cut all nature-based interventions off from compliance financing. Non-economic interventions can still qualify, as can interventions that reduce aggregate demand, and market-exposed interventions (such as REDD+) can be designed to minimize or avoid leakage risk.

Current efforts to avoid introducing uncertain credits into carbon accounting simply rule out broad project categories (both the Gold Standard and the European Emissions Trading Scheme exclude REDD+ credits, in part because of uncontrolled leakage risks[21]). In contrast, our framework implies that specific additionality claims within these project categories can be allowed when projects are designed to credibly avoid the leakage problem. REDD+ projects, for example, could be designed to maintain the supply of goods and services resulting from deforestation or degradation by improving landscape-level production efficiencies (thereby shifting the additionality claim from ‘‘reducing supply’’ to ‘‘increasing C efficiency’’ in Figure 2). However, such efforts must carefully consider the mechanics of market leakage and should be cautious about exceptions (such as relying on claims of future production increases to disregard leakage risks from current supply restrictions).

Implementing our principle 2 is likely to increase the average cost of nature-based offsets in compliance settings. Applying our framework to a well-known global estimate[10] (see https://doi.org/10.5281/zenodo.7924179) suggests that 59%–70% of low-cost mitigation potential from nature should be classified as high leakage risk because the associated additionality claims are based at least partly on reducing supply to markets. However, accurate prices are essential to enable efficient market-based mitigation solutions, maintain the integrity of these systems, and provide incentives for low-carbon innovation. The importance of these outcomes argues strongly for avoiding uncertain offsets. If allowed to substitute for avoided emissions in compliance settings, highly risky offsets reduce the cost of compliance, thereby substituting for more reliable (but potentially more costly) offsets and other emissions reduction solutions, while also reducing incentives to invest in R&D to drive innovation in new low-carbon solutions.

Principle 3: In non-compliance settings, the level of uncertainty that is acceptable in a (nature-based) offset should be set by the action for which it substitutes

When used as offsets, carbon credits substitute for alternative mitigation actions by definition. This substitution is premised on fungibility: in a carbon accounting system, a credit used as an offset is deemed equivalent to a unit of foregone emissions. We have argued that uncertainty about market leakage breaks this equivalence for specific categories of nature-based additionality claims, and on this basis have proposed prohibiting the substitution of highly uncertain offsets for relatively certain regulated emissions (principle 2). Our final principle focuses on the underlying substitution dynamic to guide market design in non-compliance settings. Where credits are used for a variety of purposes with uncertain mitigation effects, accounting integrity can be preserved by matching credits to actions on the basis of comparable certainty. Put simply, an uncertain credit should not substitute for a certain emission—but it can substitute for an uncertain one (for example, as part of corporate branding initiatives in voluntary markets or standard systems).

Operationalizing matching on uncertainty requires understanding why firms purchase non-compliance credits, as well as significant advances in how purchases are claimed and monitored—issues which are currently the focus of multiple governance initiatives in voluntary carbon markets. We suggest two possible approaches. First, if different categories of substitution can be identified, markets can be stratified such that carbon credits substitute only for comparably uncertain mitigation actions. A simple version of this approach is already in use by The Science-Based Targets initiative, which allows highly uncertain ‘‘avoidance’’ credits to be applied only to offset beyond-value-chain emissions. Further stratification could potentially unlock other sources of finance for uncertain nature-based interventions, for example by allowing their use in sustainability claims that demonstrate commitment but do not assert progress toward net-zero targets. Growing interest in designing biodiversity offset markets argues strongly for exploring the possibilities of market stratification now, since the true fungibility of such non-carbon interventions is also likely to be very uncertain.

One important potential application of stratification is de-risking investments in innovation. Firms or governments already invest in portfolios of low-carbon innovations, the ultimate impact of which is uncertain. Matching these investments with purchases of uncertain nature-based offsets provides a hedge against failure, increasing the chances of achieving mitigation goals. Following our principle, such purchases would substitute for the uncertainty in innovation outcomes, with market stratification providing for a pool of lower-cost, but less certain, emission offsets. Where a low-carbon transition strategy involves some probability of failure, access to such a lower-cost pool of uncertain offsets could be beneficial.

A second approach is possible when uncertainty can be quantified (as risk) or resolved over time. Risky offsets can be combined into portfolios, reducing volatility[87] and creating mitigation assets whose expected value is more certain. If the true effect of interventions can be tracked over time (and carbon accounting updated accordingly[88]), a portfolio approach could unlock significant financing for nature-based interventions without jeopardizing mitigation incentives. Making portfolios work for leaky offsets requires pinning leakage down to known ranges, which would require a considerable amount of further research into approaches to estimate the potential leakage associated with novel nature-based interventions and geographic settings. We therefore suggest our third principle as a general guide to future market development, and emphasize the urgency of implementing our first and second principles immediately to control leakage in rapidly growing offset markets.

CONCLUSION AND OUTLOOK

Nature-based offsets can play a vital role in enabling deeper and cheaper net emissions reductions, but only if credited offsets are real. Scaling up nature-based solutions is challenged by the continued lack of an accurate and cost-effective method for measuring market leakage. Current approaches appear to significantly underestimate the likely magnitude of market leakage effects, introducing a risk of silent failure into nature-based offset regimes. To correct this course, we present a conceptual framework for avoiding market leakage by design and identify three principles that can be put into practice now. Our first principle can be implemented by project developers alone, while our second and third principles depend on the use to which offsets are put and should be applied by the buyers of nature-based offsets and the designers of offset schemes.

Prior work[15,19,30,55,89,90] has suggested similar ‘‘design-based’’ options to reduce or mitigate leakage, for example by avoiding leaky interventions, reducing demand, substituting foregone livelihoods or output, or constraining leakage agents. These suggestions have been inconsistently applied and lack an underlying conceptual framework, significantly reducing their potential for broader implementation to control market leakage from nature-based offsets. In this article, we have aimed to establish a more consistent and robust basis for understanding market leakage that helps to resolve the problem. As we have shown, decades of economic research have not produced a reliable and low-cost approach to estimating the leakage associated with a particular offsetting intervention, leading most third-party standards to instead apply discount factors to account for potential market leakage by rule of thumb. Table 1 provides evidence suggesting that this system is not working, as actual leakage estimates applied in practice appear to diverge sharply from peer-reviewed estimates of market leakage in nature-based offsets.

Early proponents of nature-based offsets have tended to see inaccuracy as acceptable given the need to pioneer new financing models or achieve urgent conservation objectives (e.g., reduced tropical deforestation[21]). Our criticisms rest on the observation that more than 30 years after the first nature-based offset projects[23] (and 28 years since the concept of leakage from them was introduced[24]), a robust and low-cost method for market leakage accounting has not yet been found. As nature-based offsets take an increasingly central role in critical near-term mitigation efforts, it is time for a new approach.

We acknowledge that our proposals would prohibit important categories of (uncertain, highly leaky) nature-based offsets from substituting for reduced emissions. Some may see this as throwing the nature-based offsets "baby" out with the bathwater, but this need not be the case. High uncertainty[91] and a lack of credible leakage accounting[18,20] are major barriers to scaling up nature-based mitigation. In the words of the CEO of the International Emissions Trading Association,[92] "a market without trust will never be successful." We have argued that controlling market leakage via carbon accounting cannot deliver credible leakage estimates, primarily because of the difficulty of obtaining accuracy in practice. Abandoning inaccurate accounting in favor of a conservative design-based approach is a necessary step to building trust and, therefore, to boosting demand for credible nature-based offsets. We are trying to help the "baby" grow.

One objection to our proposals is that (correctly) applying high discount rates may make projects uneconomic. This misunderstands the premise of market-based mitigation schemes, which require accurate information to deliver economically efficient outcomes. Allowing bad offsets depresses prices and crowds out good projects. Such price dilution appears to be widespread today (in the past, fears of it have cut off nature-based solutions from offset-based finance[21,93]). Prices for forestry and land-use offsets in voluntary markets continue to hover around US$5 per ton[5] and roughly scale[4] inversely with leakage risk. True carbon prices are much higher: Paris-consistent prices were estimated at US$40–80 per ton in 2020,[94] and the median internal carbon price employed by corporations was US$25.[95] Estimates of the social cost of carbon (used in national policy-making) range higher still.[96] Building trust in the credibility of nature-based offsets can unlock these higher prices, potentially making more nature-based mitigation available and unleashing innovation to identify lower-cost mitigation solutions.

A second objection is a lack of alternatives. For example, Streck[21, p.849] argues that "concerns about leakage cannot be an excuse for inaction [on tropical forest loss]," and nature-based offsets are often presented as most suitable for difficult-to-abate industrial emissions. We agree with these views but contend that bad accounting is not the solution. The choice is not between current practice and nothing; it is between credible and non-credible interventions. Taking a conservative approach to avoiding market leakage will direct finance toward projects that actually deliver claimed mitigation while appropriately pricing offsets, which in turn can help to drive innovation in emissions-intensive sectors and leaky project categories. Conservativeness is particularly urgent because problems stack: the additionality of offsets is extremely difficult to demonstrate,[97] and recent work has highlighted high-profile cases of non-additional issuances.[12,13,98,99] By contrast, a design-based approach can credibly avoid the market leakage problem.

Finally, we stress that our concern with market leakage is most acute in the current context of decentralized implementation of many (relatively) small interventions. Coordinated actions and large-scale implementation can provide market substitutes or mobilize the resources necessary for accurate accounting. However, timing matters: believing that complementary actions will occur in the future is not sufficient for ignoring market leakage now (nor can a national program ignore international effects if consistent accounting approaches do not yet exist). We hope that our conceptual framework helps resolve such misunderstandings about how and where market leakage matters, but the outline we have provided is necessarily incomplete. Wealth effects, the rebound effects of intensification, and long- vs. short-run equilibrium dynamics deserve more consideration within our framework. A deeper exploration of the problems we note with quasi-experimental statistical methods is also warranted, given rapidly growing applications in offset monitoring and verification. Nevertheless, our framework and principles for a design-based approach would contribute to improving the credibility of nature-based offset markets, helping this important set of mitigation strategies to realize their potential.

EXPERIMENTAL PROCEDURES

Resource availability

Lead contact

Inquiries should be directed to the lead contact, Ben Filewod (b.filewod@lse.ac.uk).

Materials availability

This study did not generate new unique materials.

Data and code availability

Figure 3 was generated using publicly available datasets pre-loaded on the freely available Google Earth Engine GIS. The Earth Engine script used to process these datasets is available at https://doi.org/10.5281/zenodo.7924179. Data on leakage rates presented in column 9 of Table 1 are drawn from publicly available offset registries as explained below. The random sample we report is available from the lead contact upon request.

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Global assessment of tropical forest cover loss

We used high-resolution data on global forest cover (Global Forest Change[100]) and a database of protected area boundaries (World Database on Protected Areas[101]; polygons only) to analyze forest cover loss in protected forests and ecologically similar areas, as illustrated in Figure 3. We used the Google Earth Engine GIS for analysis, structuring processing by ecoregion (RESOLVE Ecoregions[102]) to facilitate parallelization. We preserved original data resolutions (raster data) and did not allow error margins in vector analysis; for one raster operation (percentile calculations) we allowed Earth Engine to rescale resolution on-the-fly to avoid resource limits.

We report aggregate results for n = 279 tropical forest ecoregions (i.e., those located within tropical and subtropical forest biomes in the RESOLVE database). For each ecoregion, we obtained and merged the spatial boundaries of "Designated," "Established," and "Inscribed" protected areas in management categories prohibiting resource extraction ("Ia," "Ib," "II," "III," "IV," and "Not Reported"), and calculated the 10th and 90th percentiles of the pixel-level distribution of (year 2000) forest canopy closure for the resulting area. We applied these percentiles to all forest cover calculations to increase comparability between protected and non-protected forest. We then calculated start-of-period (year 2000) forest area and protected area per ecoregion, and forest cover loss and total area protected for each year from 2001 to 2019 (inclusive). We applied a medium-resolution fire mask (MODIS CCI Burned Area, v5.1[103]) within each annual calculation to reduce the inclusion of non-anthropogenic forest cover loss in our analysis. We differenced annual totals to obtain year-by-year changes and generated Figure 3 using R.

The resulting data provide an approximate view of forest area protected and forest cover loss in ecologically similar forests for Earth's tropical forest biomes. This is a demonstrative analysis, with important limitations affecting accuracy: Global Forest Change data do not detect small-scale disturbances (e.g., selective logging), comparison of changes over time is complicated by differences in Landsat sensor technology and data processing, not all non-anthropogenic disturbance is due to fires (and pixel size artifacts prevent full fire masking in our approach), the choice of a 10th–90th percentile constraint is arbitrary, and incomplete fields in the World Database on Protected Areas may cause true area protected to be overstated due to filtering (conversely, unknown management effectiveness implies that effective protected area may be overstated).

Analysis of leakage in issued nature-based carbon offsets

We downloaded public registry data on credit issuances from Verra (https://registry.verra.org/) Gold Standard (https://registry.goldstandard.org/), and the American Carbon Reserve (https://americancarbonregistry.org) in April/May 2022, and selected the nature-based offset methodologies with the most issuances (per registry) for analysis, as reported in Table 1. We include two methodologies currently in development (no issuances) for comparison. We used the most up-to-date version of each methodology, noting that the issued volumes we report include credits issued according to earlier versions. We analyzed methodologies and reported the conditions under which market leakage must be assessed (Table 1, column "Trigger"), whether international leakage is considered (column "International leakage"), the approach used to account for leakage (column "Approach"), and the range of market leakage values possible under the methodology (column "Possible range").

To assess average market leakage values in practice (column "Median value"), we took a pseudorandom sample of five unique project identifiers for each methodology in R using sample_n {dplyr}. We took ten samples for VM0007. For each project, we obtained or calculated market leakage values using best available information from public documents linked on the relevant registry. We used *ex ante* data (i.e., projected mitigation and leakage from project design documents). For total leakage from VCS (main text), we report *ex ante* estimates of cumulative total leakage (typically given over a 30-year horizon) divided by the claimed emission reductions (baseline emissions minus project emissions). We note that issued credits are based on *ex poste* values, which may differ from the *ex ante* data we report if methodologies require ongoing monitoring (e.g., of a designated leakage zone) to calculate discount factors. However, *ex ante* estimates are typically conservative (in the sense of reflecting the upper bound of project proponent's views on the market leakage deductions they may incur); in several cases, project documents asserted proponents' views that *ex poste* leakage values would be lower.

ACKNOWLEDGMENTS

We thank, without implicating, Vic Adamowicz, John Caspersen, Rob Macquarie, Brian Murray, Kenneth Richards, Nicholas Rivers, and Sean C. Thomas for their comments at various stages of this paper's development. During the preparation of portions of this work, B.F. was funded by a Social Sciences and Humanities Research Council of Canada (SSHRC) Vanier scholarship and by the ForestLAB project at the Grantham Research Institute on Climate Change and the Environment at the London School of Economics and Political Science. G.M. was supported in part by funding from SSHRC as part of the Greening Growth Partnership project at the University of Ottawa (SSHRC award no. 895-2017-1018).

AUTHOR CONTRIBUTIONS

Conceptualization, B.F. and G.M.; investigation and methodology, B.F.; writing – original draft and writing – review & editing, B.F. and G.M.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Grassi, G., House, J., Dentener, F., Federici, S., den Elzen, M., and Penman, J. (2017). The key role of forests in meeting climate targets requires science for credible mitigation. Nat. Clim. Chang. *7*, 220–226. https://doi.org/10.1038/nclimate3227.
2. Webb, C., and Zakir, Z. (2019). What the Paris Agreement Means for Carbon Pricing and Natural Climate Solutions: A Business Guide.
3. Kreibich, N., and Hermwille, L. (2021). Caught in between: credibility and feasibility of the voluntary carbon market post-2020. Clim. Pol. *21*, 939–957. https://doi.org/10.1080/14693062.2021.1948384.
4. The World Bank (2021). State and Trends of Carbon Pricing 2021 (World Bank).
5. Donogrio, S., Maguire, P., Myers, K., Daley, C., and Lin, K. (2021). State of the Voluntary Carbon Markets 2021: Installment 1 (Ecosystem Marketplace).
6. Roe, S., Streck, C., Beach, R., Busch, J., Chapman, M., Daioglou, V., Deppermann, A., Doelman, J., Emmet-Booth, J., Engelmann, J., et al. (2021). Land-based measures to mitigate climate change: potential and feasibility by country. Glob. Chang. Biol. *27*, 6025–6058. https://doi.org/10.1111/gcb.15873.
7. Anderson, C.M., DeFries, R.S., Litterman, R., Matson, P.A., Nepstad, D.C., Pacala, S., Schlesinger, W.H., Shaw, M.R., Smith, P., Weber, C., and Field, C.B. (2019). Natural climate solutions are not enough. Science *363*, 933–934. https://doi.org/10.1126/science.aaw2741.
8. Griscom, B.W., Lomax, G., Kroeger, T., Fargione, J.E., Adams, J., Almond, L., Bossio, D., Cook-Patton, S.C., Ellis, P.W., Kennedy, C.M., and Kiesecker, J. (2019). We need both natural and energy solutions to stabilize our climate. Glob. Chang. Biol. *25*, 1889–1890. https://doi.org/10.1111/gcb.14612.
9. Seddon, N. (2022). Harnessing the potential of nature-based solutions for mitigating and adapting to climate change. Science *376*, 1410–1416.
10. Griscom, B.W., Adams, J., Ellis, P.W., Houghton, R.A., Lomax, G., Miteva, D.A., Schlesinger, W.H., Shoch, D., Siikamäki, J.V., Smith, P., et al. (2017). Natural climate solutions. Proc. Natl. Acad. Sci. USA *114*, 11645–11650. https://doi.org/10.1073/pnas.1710465114.
11. West, T.A.P., Börner, J., Sills, E.O., and Kontoleon, A. (2020). Overstated carbon emission reductions from voluntary REDD+ projects in the Brazilian Amazon. Proc. Natl. Acad. Sci. USA *117*, 24188–24194. https://doi.org/10.1073/pnas.2004334117.
12. Badgley, G., Freeman, J., Hamman, J.J., Haya, B., Trugman, A.T., Anderegg, W.R.L., and Cullenward, D. (2022). Systematic over-crediting in California's forest carbon offsets program. Glob. Chang. Biol. *28*, 1433–1445. https://doi.org/10.1111/gcb.15943.
13. Guizar-Coutiño, A., Jones, J.P.G., Balmford, A., Carmenta, R., and Coomes, D.A. (2022). A global evaluation of the effectiveness of voluntary REDD+ projects at reducing deforestation and degradation in the moist tropics. Conserv. Biol. *36*, e13970. https://doi.org/10.1111/cobi.13970.
14. Meyfroidt, P., Börner, J., Garrett, R., Gardner, T., Godar, J., Kis-Katos, K., Soares-Filho, B.S., and Wunder, S. (2020). Focus on leakage and spillovers: informing land-use governance in a tele-coupled world. Environ. Res. Lett. *15*, 090202. https://doi.org/10.1088/1748-9326/ab7397.
15. Schwarze, R., Niles, J.O., and Olander, J. (2002). Understanding and managing leakage in forest–based greenhouse–gas–mitigation projects. Philos. Trans. A Math. Phys. Eng. Sci. *360*, 1685–1703. https://doi.org/10.1098/rsta.2002.1040.
16. Chagas, T., Galt, H., Lee, D., Neeff, T., and Streck, C. (2020). A Close Look at the Quality of REDD+ Carbon Credits.
17. Espejo, A.B., Becerra-Leal, M.C., and Aguilar-Amuchastegui, N. (2020). Comparing the environmental integrity of emission reductions from REDD programs with renewable energy projects. Forests *11*, 1360. https://doi.org/10.3390/f11121360.
18. van der Gaast, W., Sikkema, R., and Vohrer, M. (2018). The contribution of forest carbon credit projects to addressing the climate change challenge. Clim. Pol. *18*, 42–48. https://doi.org/10.1080/14693062.2016.1242056.
19. Atmadja, S., and Verchot, L. (2012). A review of the state of research, policies and strategies in addressing leakage from reducing emissions from deforestation and forest degradation (REDD+). Mitig. Adapt. Strateg. Glob. Chang. *17*, 311–336. https://doi.org/10.1007/s11027-011-9328-4.

20. Peterson St-Laurent, G., Hagerman, S., and Hoberg, G. (2017). Barriers to the development of forest carbon offsetting: insights from British Columbia, Canada. J. Environ. Manage. *203*, 208–217. https://doi.org/10.1016/j.jenvman.2017.07.051.

21. Streck, C. (2021). REDD+ and leakage: debunking myths and promoting integrated solutions. Clim. Pol. *21*, 843–852. https://doi.org/10.1080/14693062.2021.1920363.

22. Richards, K., and Andersson, K. (2001). The leaky sink: persistent obstacles to a forest carbon sequestration program based on individual projects. Clim. Pol. *1*, 41–54.

23. Putz, F.E., and Pinard, M.A. (1993). Reduced-impact logging as a carbon-offset method. Conserv. Biol. *7*, 755–757. https://doi.org/10.1046/j.1523-1739.1993.7407551.x.

24. Bastos Lima, M.G., Persson, U.M., and Meyfroidt, P. (2019). Leakage and boosting effects in environmental governance: a framework for analysis. Environ. Res. Lett. *14*, 105006. https://doi.org/10.1088/1748-9326/ab4551.

25. Brown, K., and Neil Adger, W. (1994). Economic and political feasibility of international carbon offsets. For. Ecol. Manag. *68*, 217–229. https://doi.org/10.1016/0378-1127(94)90047-7.

26. Hertel, T.W. (2018). Economic perspectives on land use change and leakage. Environ. Res. Lett. *13*, 075012. https://doi.org/10.1088/1748-9326/aad2a4.

27. Henders, S., and Ostwald, M. (2012). Forest carbon leakage quantification methods and their suitability for assessing leakage in REDD. Forests *3*, 33–58. https://doi.org/10.3390/f3010033.

28. Atmadja, S.S., Duchelle, A.E., De Sy, V., Selviana, V., Komalasari, M., Sills, E.O., and Angelsen, A. (2022). How do REDD+ projects contribute to the goals of the Paris Agreement? Environ. Res. Lett. *17*, 044038. https://doi.org/10.1088/1748-9326/ac5669.

29. Andersson, K., and Richards, K.R. (2001). Implementing an international carbon sequestration program: can the leaky sink be fixed? Clim. Pol. *16*.

30. Wunder, S. (2008). How do we deal with leakage? In Moving ahead with REDD: Issues, Options, and Implications (Center for International Forestry Research), pp. 65–76.

31. Boucher, D., and Elias, P. (2013). From REDD to deforestation-free supply chains: the persistent problem of leakage and scale. Carbon Manag. *4*, 473–475. https://doi.org/10.4155/cmt.13.47.

32. Schwartzman, S., Lubowski, R.N., Pacala, S.W., Keohane, N.O., Kerr, S., Oppenheimer, M., and Hamburg, S.P. (2021). Environmental integrity of emissions reductions depends on scale and systemic changes, not sector of origin. Environ. Res. Lett. *16*, 091001. https://doi.org/10.1088/1748-9326/ac18e8.

33. Keith, H., Vardon, M., Obst, C., Young, V., Houghton, R.A., and Mackey, B. (2021). Evaluating nature-based solutions for climate mitigation and conservation requires comprehensive carbon accounting. Sci. Total Environ. *769*, 144341. https://doi.org/10.1016/j.scitotenv.2020.144341.

34. Meyfroidt, P., Roy Chowdhury, R., de Bremond, A., Ellis, E.C., Erb, K.-H., Filatova, T., Garrett, R.D., Grove, J.M., Heinimann, A., Kuemmerle, T., et al. (2018). Middle-range theories of land system change. Global Environ. Change *53*, 52–67. https://doi.org/10.1016/j.gloenvcha.2018.08.006.

35. Delacote, P., Robinson, E.J., and Roussel, S. (2016). Deforestation, leakage and avoided deforestation policies: a spatial analysis. Resour. Energy Econ. *45*, 192–210. https://doi.org/10.1016/j.reseneeco.2016.06.006.

36. Murray, B.C., McCarl, B.A., and Lee, H.-C. (2004). Estimating leakage from forest carbon sequestration programs. Land Econ. *80*, 109–124. https://doi.org/10.2307/3147147.

37. Gan, J., and McCarl, B.A. (2007). Measuring transnational leakage of forest conservation. Ecol. Econ. *64*, 423–432. https://doi.org/10.1016/j.ecolecon.2007.02.032.

38. Turner, J.A., and Buongiorno, J. (2004). Estimating price and income elasticities of demand for imports of forest products from panel data. Scand. J. For. Res. *19*, 358–373. https://doi.org/10.1080/02827580410030127.

39. Buongiorno, J. (2015). Income and time dependence of forest product demand elasticities and implications for forecasting. Silva Fenn. *49*, 17. https://doi.org/10.14214/sf.1395.

40. Roberts, M.J., and Schlenker, W. (2013). Identifying supply and demand elasticities of agricultural commodities: implications for the US ethanol mandate. Am. Econ. Rev. *103*, 2265–2295. https://doi.org/10.1257/aer.103.6.2265.

41. Meyfroidt, P., and Lambin, E.F. (2009). Forest transition in Vietnam and displacement of deforestation abroad. Proc. Natl. Acad. Sci. USA *106*, 16139–16144. https://doi.org/10.1073/pnas.0904942106.

42. Meyfroidt, P., Rudel, T.K., and Lambin, E.F. (2010). Forest transitions, trade, and the global displacement of land use. Proc. Natl. Acad. Sci. USA *107*, 20917–20922. https://doi.org/10.1073/pnas.1014773107.

43. Alix-Garcia, J.M., Shapiro, E.N., and Sims, K.R.E. (2012). Forest conservation and slippage: evidence from Mexico's national payments for ecosystem services program. Land Econ. *88*, 613–638. https://doi.org/10.3368/le.88.4.613.

44. Roopsind, A., Sohngen, B., and Brandt, J. (2019). Evidence that a national REDD+ program reduces tree cover loss and carbon emissions in a high forest cover, low deforestation country. Proc. Natl. Acad. Sci. USA *116*, 24492–24499. https://doi.org/10.1073/pnas.1904027116.

45. Ford, S.A., Jepsen, M.R., Kingston, N., Lewis, E., Brooks, T.M., MacSharry, B., and Mertz, O. (2020). Deforestation leakage undermines conservation value of tropical and subtropical forest protected areas. Global Ecol. Biogeogr. *29*, 2014–2024. https://doi.org/10.1111/geb.13172.

46. Robalino, J., Pfaff, A., and Villalobos, L. (2017). Heterogeneous local spillovers from protected areas in Costa Rica. J. Assoc. Environ. Resour. Econ. *4*, 795–820. https://doi.org/10.1086/692089.

47. Moffette, F., and Gibbs, H.K. (2021). Agricultural displacement and deforestation leakage in the Brazilian legal Amazon. Land Econ. *97*, 155–179.

48. Groom, B., Palmer, C., and Sileci, L. (2022). Carbon emissions reductions from Indonesia's moratorium on forest concessions are cost-effective yet contribute little to Paris pledges. Proc. Natl. Acad. Sci. USA *119*, e2102613119. https://doi.org/10.1073/pnas.2102613119.

49. I Kallio, A.M., and Solberg, B. (2018). Leakage of forest harvest changes in a small open economy: case Norway. Scand. J. Forest Res. *33*, 502–510. https://doi.org/10.1080/02827581.2018.1427787.

50. N Wear, D., Murray, B.C., and Murray, B.C. (2004). Federal timber restrictions, interregional spillovers, and the impact on US softwood markets. J. Environ. Econ. Manag. *47*, 307–330. https://doi.org/10.1016/S0095-0696(03)00081-0.

51. Hu, X., Shi, G., and Hodges, D. (2014). International market leakage from China's forestry policies. Forests *5*, 2613–2625. https://doi.org/10.3390/f5112613.

52. Kallio, A.M.I., Solberg, B., Käär, L., and Päivinen, R. (2018). Economic impacts of setting reference levels for the forest carbon sinks in the EU on the European forest sector. For. Pol. Econ. *92*, 193–201. https://doi.org/10.1016/j.forpol.2018.04.010.

53. Nepal, P., Ince, P.J., Skog, K.E., and Chang, S.J. (2013). Forest carbon benefits, costs and leakage effects of carbon reserve scenarios in the United States. J. For. Econ. *19*, 286–306. https://doi.org/10.1016/j.jfe.2013.06.001.

54. Sun, B., and Sohngen, B. (2009). Set-asides for carbon sequestration: implications for permanence and leakage. Climatic Change *96*, 409–419. https://doi.org/10.1007/s10584-009-9628-9.

55. Sohngen, B., and Brown, S. (2004). Measuring leakage from carbon projects in open economies: a stop timber harvesting project in Bolivia as a case study. Can. J. For. Res. *34*, 829–839. https://doi.org/10.1139/x03-249.

56. Kuik, O. (2014). REDD+ and international leakage via food and timber markets: a CGE analysis. Mitig. Adapt. Strateg. Glob. Chang. *19*, 641–655. https://doi.org/10.1007/s11027-013-9527-2.

57. Kim, M.-K., Peralta, D., and McCarl, B.A. (2014). Land-based greenhouse gas emission offset and leakage discounting. Ecol. Econ. *105*, 265–273. https://doi.org/10.1016/j.ecolecon.2014.06.009.

58. Lichtenberg, E., and Smith-Ramírez, R. (2011). Slippage in conservation cost sharing. Am. J. Agric. Econ. *93*, 113–129. https://doi.org/10.1093/ajae/aaq124.

59. Fortmann, L., Sohngen, B., and Southgate, D. (2017). Assessing the role of group heterogeneity in community forest concessions in Guatemala's Maya biosphere reserve. Land Econ. *93*, 503–526. https://doi.org/10.3368/le.93.3.503.

60. Wu, J. (2000). Slippage effects of the conservation reserve program. Am. J. Agric. Econ. *82*, 979–992.

61. Roberts, M.J., and Bucholtz, S. (2005). Slippage in the conservation reserve program or spurious correlation? A comment. Am. J. Agric. Econ. *87*, 244–250. https://doi.org/10.1111/j.0002-9092.2005.00715.x.

62. Roberts, M.J., and Bucholz, S. (2006). Slippage in the conservation reserve program or spurious correlation? A rejoinder. Am. J. Agric. Econ. *88*, 512–514. https://doi.org/10.1111/j.1467-8276.2006.00875.x.

63. Wu, J. (2005). Slippage effects of the conservation reserve program: reply. Am. J. Agric. Econ. *87*, 251–254. https://doi.org/10.1111/j.0002-9092.2005.00716.x.

64. Alix-Garcia, J., and Gibbs, H.K. (2017). Forest conservation effects of Brazil's zero deforestation cattle agreements undermined by leakage.

Global Environ. Change *47*, 201–217. https://doi.org/10.1016/j.gloenvcha.2017.08.009.
65. Cisneros, E., Zhou, S.L., and Börner, J. (2015). Naming and shaming for conservation: evidence from the Brazilian Amazon. PLoS One *10*, e0136402. https://doi.org/10.1371/journal.pone.0136402.
66. Pan, W., Kim, M.-K., Ning, Z., and Yang, H. (2020). Carbon leakage in energy/forest sectors and climate policy implications using meta-analysis. For. Pol. Econ. *115*, 102161. https://doi.org/10.1016/j.forpol.2020.102161.
67. Murray, B.C., Sohngen, B., and Ross, M.T. (2007). Economic consequences of consideration of permanence, leakage and additionality for soil carbon sequestration projects. Climatic Change *80*, 127–143. https://doi.org/10.1007/s10584-006-9169-4.
68. Bastin, J.-F., Finegold, Y., Garcia, C., Mollicone, D., Rezende, M., Routh, D., Zohner, C.M., and Crowther, T.W. (2019). The global tree restoration potential. Science *365*, 76–79. https://doi.org/10.1126/science.aax0848.
69. Bastin, J.-F., Finegold, Y., Garcia, C., Gellie, N., Lowe, A., Mollicone, D., Rezende, M., Routh, D., Sacande, M., Sparrow, B., et al. (2019). Response to Comments on "The global tree restoration potential. Science *366*, eaay8108. https://doi.org/10.1126/science.aay8108.
70. Lewis, S.L., Mitchard, E.T.A., Prentice, C., Maslin, M., and Poulter, B. (2019). Comment on "The global tree restoration potential. Science *366*, eaaz0388. https://doi.org/10.1126/science.aaz0388.
71. Veldman, J.W., Aleman, J.C., Alvarado, S.T., Anderson, T.M., Archibald, S., Bond, W.J., Boutton, T.W., Buchmann, N., Buisson, E., Canadell, J.G., et al. (2019). Comment on "The global tree restoration potential. Science *366*, eaay7976. https://doi.org/10.1126/science.aay7976.
72. Skidmore, A.K., Wang, T., de Bie, K., and Pilesjö, P. (2019). Comment on "The global tree restoration potential. Science *366*, eaaz0111. https://doi.org/10.1126/science.aaz0111.
73. Friedlingstein, P., Allen, M., Canadell, J.G., Peters, G.P., and Seneviratne, S.I. (2019). Comment on "The global tree restoration potential. Science *366*, eaay8060. https://doi.org/10.1126/science.aay8060.
74. Baldocchi, D., and Penuelas, J. (2019). The physics and ecology of mining carbon dioxide from the atmosphere by ecosystems. Glob. Chang. Biol. *25*, 1191–1197. https://doi.org/10.1111/gcb.14559.
75. Beisner, B., Haydon, D., and Cuddington, K. (2003). Alternative stable states in ecology. Front. Ecol. Environ. *1*, 376–382. https://doi.org/10.1890/1540-9295(2003)001[0376:ASSIE]2.0.CO;2.
76. Pausas, J.G., and Bond, W.J. (2019). Humboldt and the reinvention of nature. J. Ecol. *107*, 1031–1037. https://doi.org/10.1111/1365-2745.13109.
77. Sohngen, B., Golub, A., and Hertel, T.W. (2009). The Role of Forestry in Carbon Sequestration in General Equilibrium Models (Routledge), p. 25.
78. Drever, C.R., Cook-Patton, S.C., Akhter, F., Badiou, P.H., Chmura, G.L., Davidson, S.J., Desjardins, R.L., Dyk, A., Fargione, J.E., Fellows, M., et al. (2021). Natural climate solutions for Canada. Sci. Adv. *7*, eabd6034. https://doi.org/10.1126/sciadv.abd6034.
79. Fargione, J.E., Bassett, S., Boucher, T., Bridgham, S.D., Conant, R.T., Cook-Patton, S.C., Ellis, P.W., Falcucci, A., Fourqurean, J.W., Gopalakrishna, T., et al. (2018). Natural climate solutions for the United States. Sci. Adv. *4*, eaat1869. https://doi.org/10.1126/sciadv.aat1869.
80. Judge-Lord, D., McDermott, C.L., and Cashore, B. (2020). Do private regulations ratchet up? How to distinguish types of regulatory stringency and patterns of change. Organ. Environ. *33*, 96–125. https://doi.org/10.1177/1086026619858874.
81. Cashore, B., Auld, G., and Newsom, D. (2004). Governing through Markets (Yale University Press).
82. British Columbia Ministry of Environment and Climate Change Strategy (2021). Forest Carbon Offset Protocol: What We Heard Report (Government of British Columbia).
83. Magnani, F., Dewar, R.C., and Borghetti, M. (2009). Leakage and spillover effects of forest management on carbon storage: theoretical insights from a simple model. Tellus B Chem. Phys. Meteorol. *61*, 385–393. https://doi.org/10.1111/j.1600-0889.2008.00386.x.
84. Latta, G.S., and Adams, D.M. (2000). An econometric analysis of output supply and input demand in the Canadian softwood lumber industry. Can. J. For. Res. *30*, 1419–1428.
85. Song, N., Chang, S.J., and Aguilar, F.X. (2011). U.S. softwood lumber demand and supply estimation using cointegration in dynamic equations. J. For. Econ. *17*, 19–33. https://doi.org/10.1016/j.jfe.2010.07.002.
86. Richards, K.R. (2004). A brief overview of carbon sequestration economics and policy. Environ. Manage. *33*, 545–558. https://doi.org/10.1007/s00267-003-9111-8.
87. Biffis, E., Brandi, G., Lee, L., and Snavely, A. (2022). Forestry-backed Assets Design (Imperial College London).
88. Balmford, A., Keshav, S., Venmans, F., Coomes, D., Groom, B., Madhavapeddy, A., and Swinfield, T. The Value of Impermanent Carbon Credits.
89. Bode, M., Tulloch, A.I.T., Mills, M., Venter, O., and Ando, A.W. (2015). A conservation planning approach to mitigate the impacts of leakage from protected area networks: conservation Planning with Leakage. Conserv. Biol. *29*, 765–774. https://doi.org/10.1111/cobi.12434.
90. Chomitz, K. (1999). Evaluating carbon offsets from forestry and energy projects: how do they compare? (The World Bank). https://doi.org/10.1596/1813-9450-2357.
91. Seddon, N., Chausson, A., Berry, P., Girardin, C.A.J., Smith, A., and Turner, B. (2020). Understanding the value and limits of nature-based solutions to climate change and other global challenges. Philos. Trans. R. Soc. Lond. B Biol. Sci. *375*, 20190120. https://doi.org/10.1098/rstb.2019.0120.
92. International Emissions Trading Association The Anatomy of the Carbon Market.
93. Bosello, F., Eboli, F., Parrado, R., and Rosa, R. (2010). REDD in the Carbon Market: A General Equilibrium Analysis (Fondazione Eni Enrico Mattei).
94. High-Level Commission on Carbon Prices (2017). Report of the High-Level Commission on Carbon Prices (The World Bank).
95. Bartlett, N., Coleman, T., and Schmidt, S. (2021). Putting a Price on Carbon (CDP Worldwide).
96. Wagner, G., Anthoff, D., Cropper, M., Dietz, S., Gillingham, K.T., Groom, B., Kelleher, J.P., Moore, F.C., and Stock, J.H. (2021). Eight priorities for calculating the social cost of carbon. Nature *590*, 548–550. https://doi.org/10.1038/d41586-021-00441-0.
97. Carter, P., Van de Sijpe, N., and Calel, R. (2021). The elusive quest for additionality. World Dev. *141*, 105393. https://doi.org/10.1016/j.worlddev.2021.105393.
98. Calel, R., Colmer, J., Dechezleprêtre, A., and Glachant, M. (2021). Do carbon offsets offset carbon? SSRN J. https://doi.org/10.2139/ssrn.3950103.
99. West, T.A.P., Wunder, S., Sills, E.O., Börner, J., Rifai, S.W., Neidermeier, A.N., and Kontoleon, A. (2023). Action needed to make carbon offsets from tropical forest conservation work for climate change mitigation. Preprint at arXiv. https://doi.org/10.48550/arXiv.2301.03354.
100. Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., et al. (2013). High-resolution global maps of 21st-century forest cover change. Science *342*, 850–853. https://doi.org/10.1126/science.1244693.
101. UNEP-WCMC; IUCN (2020). The World Database on Protected Areas.
102. Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N.D., Wikramanayake, E., Hahn, N., Palminteri, S., Hedao, P., Noss, R., et al. (2017). An ecoregion-based approach to protecting half the terrestrial realm. Bioscience *67*, 534–545. https://doi.org/10.1093/biosci/bix014.
103. Chuvieco, E., Pettinari, M.L., Lizundia-Loiola, J., Storm, T., and Padilla Parellada, M. (2018). ESA fire climate change initiative (Fire_cci): MODIS Fire_cci burned area pixel product, version 5.1. https://doi.org/10.5285/58F00D8814064B79A0C49662AD3AF537.

nature sustainability

Analysis

https://doi.org/10.1038/s41893-023-01259-6

Pervasive over-crediting from cookstove offset methodologies

Received: 2 July 2023

Accepted: 12 December 2023

Published online: 23 January 2024

Check for updates

Annelise Gill-Wiehl[1], Daniel M. Kammen[1,2,3,4] & Barbara K. Haya[2,5]

Cookstove carbon offset projects can progress multiple Sustainable Development Goals (SDGs), including climate, energy, health, gender, poverty and deforestation. However, project emission reductions must be accurately or conservatively estimated to avoid undermining climate action and long-term SDG financing. Here we conduct a comprehensive, quantitative, quality assessment of offsets by comparing five cookstove methodologies with published literature and our own analysis. We find misalignment, in order of importance, with fraction of non-renewable biomass, firewood–charcoal conversion, stove adoption, stove usage, fuel consumption, stacking (using multiple stoves), rebound and emission factors. Additionality, leakage, permanence and overlapping claims require more research. We estimate that our project sample is over-credited 9.2 times. Gold Standard's metered methodology, which directly monitors fuel use, is most aligned with our estimates (1.5 times over-credited) and has the largest potential for emission abatement and health benefit. We provide recommendations to align methodologies with current science and SDG progress.

Roughly 2.4 billion people globally cook with smoky solid fuels or kerosene, contributing to 2–3 million premature deaths annually[1] and roughly 2% of global greenhouse gas (GHG) emissions[2]. Efficient cookstoves can support multiple Sustainable Development Goals (SDGs) including climate, energy, health, gender, poverty and deforestation. Monetizing the GHG emission reductions from efficient cookstove projects through the voluntary carbon market (VCM) has the potential to provide substantial financing for these projects.

Efficient stoves can reduce emissions by (1) using less fuel or switching to a less GHG-intensive fuel and/or (2) reducing the release of methane and other pollutants through more complete fuel combustion. While improved stoves are often touted for their health benefits, only solar, electric, gas, ethanol and, currently, two forced-draft pellet stoves reduce smoke enough to meaningfully reduce disease risk and meet the World Health Organization (WHO) definition of 'clean'[3] (Supplementary Information).

Cookstove projects with credits on the VCM are registered under the Gold Standard (GS) and Verified Carbon Standard offset registries, and estimate carbon emission reductions using methodologies primarily developed by GS and the Clean Development Mechanism (CDM). Cookstoves, one of the fastest growing project types on the VCM, represented 1,213 out of the 7,933 project activities (individually registered or included in a programme of activities) on the VCM[4] and generated ~78.9 million total issued credits (as of 10 May 2023). Most VCM cookstove project activities replace three stone fires or inefficient biomass stoves with improved firewood stoves, while 43 project activities distribute only WHO-defined clean stoves/fuels (Fig. 1).

Studies of offset project quality have documented substantial excess crediting (as much as 13 times from single factors) from improved forest management[5,6], avoided deforestation[7,8] and the United Nations system[9,10]. Over-crediting is harmful to effective climate action, the buyer and the cookstove sector. Poor-quality credits can undermine climate action by justifying ongoing emissions and replacing direct emission reduction and other more effective climate mitigation activities, even if some reduction is achieved. Excess crediting obscures the overall effectiveness of climate efforts and progress

[1]Energy and Resources Group, University of California, Berkeley, CA, USA. [2]Goldman School of Public Policy, University of California, Berkeley, CA, USA. [3]Department of Nuclear Engineering, University of California, Berkeley, CA, USA. [4]United States Agency for International Development, Washington, DC, USA. [5]California Institute for Energy and Environment, University of California, Berkeley, CA, USA. e-mail: agillwiehl@berkeley.edu

https://doi.org/10.1038/s41893-023-01259-6

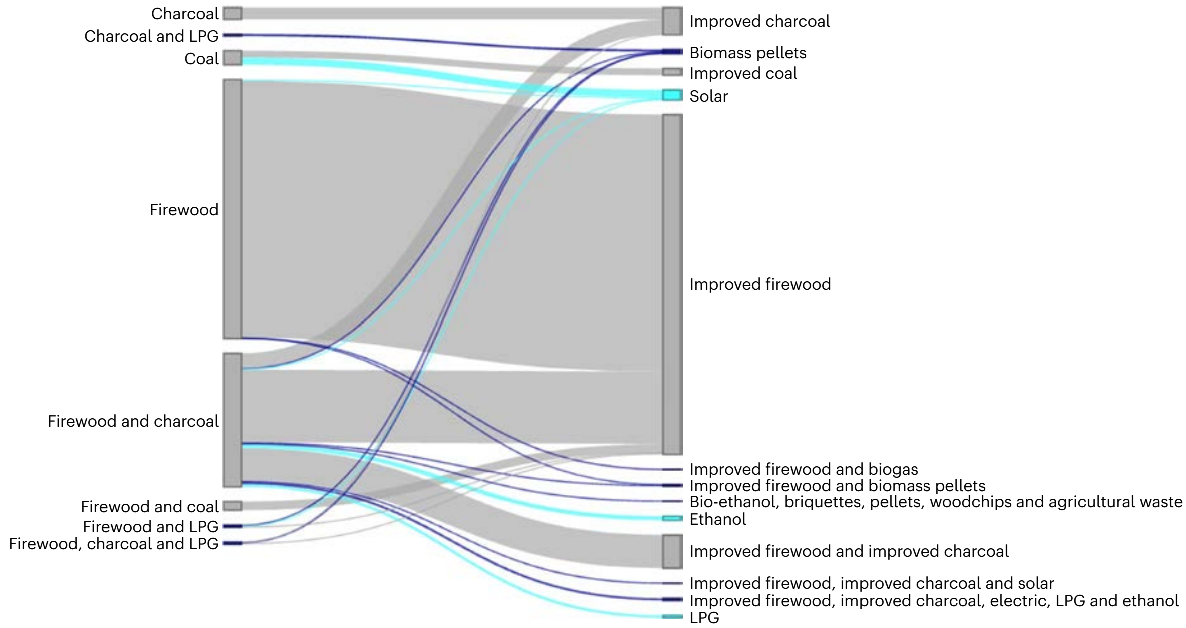


Fig. 1 | Transitions from baseline to project fuels by cookstove carbon offset projects. The left side of the diagram indicates the majority baseline fuels before intervention, and the right side represents the project fuel/stove that the VCM-funded project implemented. The width of the link indicates the relative number of projects. Grey indicates WHO polluting or transitional fuels or stoves (tiers 0–3). Dark blue indicates a mix of WHO clean, transitional or polluting fuels and stoves, while cyan indicates only WHO clean fuels or stoves. We exclude six projects that do not change the stove, but only replaced firewood with agricultural waste. As of 9 November 2022, 4% of cookstove project activities (43 out of 992 projects) registered on the VCM distribute only cooking fuels or stoves that meet the WHO's definition of clean, that is, they meet tier 5 for carbon monoxide and tier 4 for particulate matter.

towards ambitious climate targets. Over-crediting also creates confusion and reputational/legal risk for buyers. Lack of trust that a credit actually represents one metric ton less carbon dioxide equivalent weakens the market and its ability to support efficient cookstoves and all of their SDG benefits.

Studies of cookstove offset projects, covering single or a few factors, found over-crediting from the choice of fraction of non-renewable biomass (fNRB)[11] and methods for track adoption/usage rates[12], and under-crediting from emission factors (EFs)[13,14]. Qualitative studies have discussed quantification challenges and uncertainty[15,16]. This study fills multiple research gaps by performing a comprehensive quantitative assessment of offset credit quality, taking into account interactions in over/under-crediting across all methodology factors for all major cookstoves methodologies, and demonstrating how such a quality assessment can be performed on an offset methodology.

In this Analysis, we (1) discuss the accuracy of all estimation factors used (or not addressed) by the four most prominent cookstoves offset methodologies (GS-technologies and practices to displace decentralized thermal energy consumption (TPDDTEC)[17], GS-simplified methodology for clean and efficient cookstoves (simplified)[18], CDM-energy efficiency measures in thermal applications of non-renewable biomass (AMS-II-G)[19] and CDM-switch from non-renewable biomass for thermal applications by the user (AMS-I-E)[20]) and the recent GS-methodology for metered and measured energy cooking devices (metered)[21] methodology (Table 1; past and current versions), drawing from published literature and our own analysis (Methods). (2) We then recalculate the carbon emission reductions of a purposive sample of 51 cookstove projects, addressing ranges of uncertainty using the Monte Carlo method (MCM), and compare those results with actual credit issuance across eight methodology/stove type categories. (3) We suggest a specific set of methodological reforms to generate high-quality credits. In doing so, (4) we develop and demonstrate an over/under-crediting analysis that can be used to systematically assess quality and inform methodology improvements across all offset project types.

Data and sampling

We identified the 15 countries with the most credits from cookstove projects on the market, and for each country selected the largest projects for each methodology. In addition, we randomly included small- and medium-size projects globally, and covered all types of fuel transition, except electric (Methods). This approach resulted in a sample of 51 projects spanning 25 countries, and accounts for 40% of all issued credits from these cookstove methodologies on the VCM (as of 10 May 2023; Fig. 2).

Results

Here, we first summarize the major factors affecting offset quality assessed in our quantitative analysis and the accuracy of their treatment by each of the methodologies compared with the published literature. For a more detailed discussion of each factor, as well as discussion of additionality, leakage, permanence and overlapping claim, see Supplementary Information.

Adoption, usage and stacking rates

Efficient cookstove projects reduce emissions to the extent that users (1) 'adopt' a more efficient project stove defined as the percentage of distributed stoves actually in use; (2) use the project stove, where 'usage' is defined as the percentage of meals cooked using the project stove; and (3) stop or reduce 'stacking', defined as the percentage of meals cooked using the baseline stove(s) in concert with the project stove. These rates are used to determine the change between pre- and post-project fuel use.

Methods for monitoring adoption, usage and stacking fall into three categories: the AMS-I-E, GS-simplified and AMS-II-G, which track them through short cross-sectional surveys. GS-TPDDTEC requires in-field multi-day kitchen performance tests (KPTs) for a sample of households, capturing both usage and stacking rates by directly measuring daily fuel usage. The results are then applied to the full set of project households through surveyed adoption rates. GS-metered

https://doi.org/10.1038/s41893-023-01259-6

Table 1 | Outlining the quantification equations, approaches and data sources of five cookstove methodologies on the VCM

Parameter (typical units)								
Methodology	**Adoption × (stove-days)**	**[Baseline fuel use × (tons per stove per day)**	**Baseline EFs[a] (CO_2–equivalent per ton of fuel)**	**– Project fuel use × (tons per stove per day)**	**Project EFs (CO_2–equivalent per ton of fuel)]**	**× Adjustment for stacking (%)**	**×/– Leakage**	**Scope/applicability**
GS-TPDDTEC	Surveys	Default value, historical data, literature or KPT (rarely chosen)	Methodology defaults, IPCC values, or literature non-CO_2 gases and upstream emissions are optional	KPT	Same EF as the Baseline EF	N/A captured in KPT	Choice between discounting emission reductions by 5%, evaluating leakage through surveys or ignore entirely if deemed very low. This methodology can include stacking in the leakage figure	Credits for less GHG-intensive stoves (for example, improved biomass, heat retention, solar, LPG, electric). This is the most versatile methodology of the five. As of 2020, clean stoves must be registered under GS-metered.
CDM-AMS-I-E	Surveys	Ethanol projects determine the equivalent amount of biomass equal to the amount of energy from ethanol used in the project scenario from a KPT, surveys or literature.	Uses EF of a projected fossil fuel constructed per region from a weighted average of fossil fuel types	KPT or surveys	Differs by fuel type Ethanol projects typically include ethanol production, electricity consumption and transport emissions here	Surveys	Choice between discounting emission reductions by 5%, evaluating leakage through surveys or ignore entirely if deemed very low	Replaces non-renewable biomass with renewable energy (for example, renewable biomass, biogas, bioethanol, electric stoves)
GS-simplified	Surveys	Default value, historical data or sample surveys	Methodology defaults, IPCC values, or literature non-CO_2 gases and upstream emissions are optional	Multiply baseline consumption by the ratio baseline and project stove efficiencies derived from a water boil test or other efficiency assessment	N/A emission reductions are calculated by stove efficiency improvement	Surveys	Choice between discounting emission reductions by 5%, evaluating leakage through surveys or ignore entirely if deemed very low	Replaces traditional cookstoves with improved wood or charcoal stoves. Designed for smaller projects, capped at 10,000 CO_2e per year
CDM-AMS-II-G	Surveys	Default value, historical or survey data, or country specific default	Uses EF of a projected fossil fuel constructed per region from a weighted average of fossil fuel types	Multiply baseline consumption by the difference between baseline and project stove efficiencies derived from a water boil test or other efficiency assessment	N/A emission reductions are calculated by stove efficiency improvement	Surveys	Discount emission reductions by 5% This methodology can include upstream emissions for charcoal or processed biomass production in leakage.	Replaces traditional cookstoves with improved biomass (wood, charcoal, pellet and so on) cookstoves, ovens or dryers Designed for smaller projects, capped at 60 GWh
	Project energy (TJ) ×		**Baseline EFs (–equivalent per TJ)**		**– Project emissions (–equivalent per TJ)**		**×/– Leakage**	
GS-metered[b]	Meters or sales data		Constructed from a weighted average of baseline fuel types displaced in the project activity from historical, literature or survey data		Differs for fuel type (electric or fossil fuel); however, for pellets, which is the GS-metered stove type our sample covers, project emissions are included as either additional leakage (that is, 5–10%) or a standard tons CO_2 per year value for pellet production.		Choice between discounting emission reductions by 5%, evaluating leakage through surveys or ignore entirely if deemed very low. This methodology can include upstream emissions and stacking in leakage figure	Designed for cookstoves with metered or other direct fuel monitoring (for example, purchase records) such as electric, LPG, biogas, bioethanol or biomass pellet

All methodologies allow projects to use fNRB default values or calculate their own with a CDM tool. Nominal caloric value values are from IPCC estimates or project specific testing. We include the exact equations in Supplementary Information, but here outline the approaches using the language defined in the main text (for example, our definition of adoption rate). In the column heads, the '×' and '–' are multiplication and subtraction, respectively, to indicate the following formula: adoption × (baseline fuel use × baseline EF – project fuel use × project EF) ×/– leakage. N/A, not applicable. [a]EFs are calculated as fNRB (%)×(CO_2 emissions (ton of CO_2 per terajoule (TJ) of fuel)+non-CO_2 emissions (ton of CO_2–equivalent per TJ))×nominal caloric value of fuel (TJ per ton). [b]GS-metered also has an option based on specific energy consumption, where thermal energy efficiency is complicated by other factors such as pressure.

https://doi.org/10.1038/s41893-023-01259-6

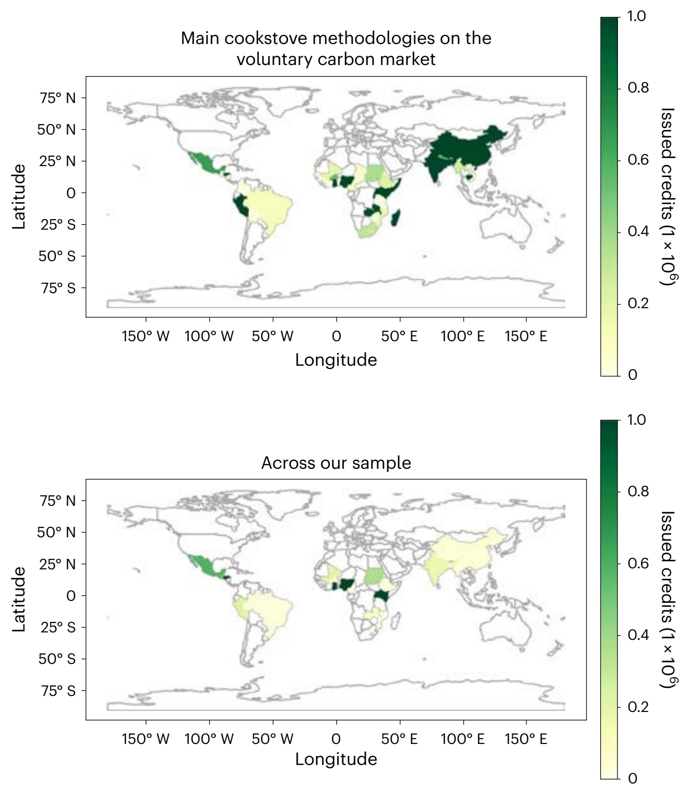


Fig. 2 | Issued credits across the VCM and our sample. Credits issued so far on the VCM across the five methodologies covered as of 9 November 2022 (top panel) and from the 51 cookstove project activities in our sample (bottom panel). We cover the GS-TPDDTEC, GS-simplified, CDM-AMS-II-G, CDM-AMS-I-E and GS-metered.

uses the most robust approach, directly tracking project stove and fuel use in all participating households through meters or fuel sales data.

The methodologies' default surveys range in quality, but all are infrequent and vulnerable to social desirability[22,23] and recall[23,24] biases. For example, AMS-II-G's default survey simply asks households if they used the improved stove in the last week or month. Credits are generated for all households that reply 'yes' as if they used the stove 100% of the time for the entire 1–2 year crediting period, with a discount if they also reported using the baseline stove in the last week or month. In 2017, GS updated their methodologies to provide projects different monitoring options, varying in rigour and capping the survey-derived adoption rate according to the rigour of the option; however, none of the surveys is designed to avoid social desirability bias, which has been well documented in survey methods across disciplines[22] as well as specifically and systematically in cookstove projects[23]. Social desirability bias occurs when participants provide responses (for example, inflating adoption/usage up to two times[24]), which they believe the surveyors (hired by the cookstove project developer) want to hear. Survey-based methods are further complicated as households may suffer from recall bias in remembering stove use over the past year[23].

KPTs, if done well, are reasonably robust, yet still have weaknesses. As a form of social desirability bias, called the Hawthorne effect, households may change their behaviour in the presence of project staff who can observe their stove choices while weighing the fuel[25]. Due to cost, KPTs are only required biennially on a sample of households; however, stove usage, stacking, and fuel quality and availability can be seasonal and highly variable[26]. Thus, KPTs might not accurately represent stove use across the participant pool over the 2 year crediting period.

Our sampled projects use surveys, and report adoption and usage rates much higher than rates documented in the literature (86% adoption rate and 98% usage rate compared with 58% and 52% from our literature reviews[27–37]), and stacking rates that are much lower (2% stacking rate compared with 68% in the literature[26–32,38,39]). These empirical studies, performed on cookstove projects very similar to those participating in the studied offset programmes (Supplementary Information), are designed to avoid bias with frequent, comprehensive, longitudinal surveys, triangulated with photos, field tests and/or stove monitors, and conducted by trained enumerators, unaffiliated with the project.

Since the offset project surveys have known biases, to estimate project carbon emission reductions, we replace all survey-derived adoption and usage rates with literature values as the best data available (Supplementary Information). We use empirical ranges in the MCM using a triangle distribution: adoption 58% (40%, 92%)[27–35], usage 52% (16%, 85%)[36] and stacking 68% (19.3%, 100%)[26–32,38,39]. We discount KPT-derived (that is, GS-TPDDTEC) usage and stacking rates with the MCM using a uniform distribution with the maximum based on an empirical study estimating the Hawthorne effect (−53% in usage and 29% in stacking)[25]. We do not correct GS-metered.

Fuel consumption

Methodologies use three approaches to estimate the difference between baseline and project fuel consumption. AMS-II-G and GS-simplified start by estimating baseline fuel use, and then use differences in the baseline and project stove efficiencies to estimate fuel use savings on the basis of surveyed adoption and usage rates. GS-TPDDTEC determines baseline and project fuel consumption separately and calculates emission reductions as the difference between the two. GS-metered/AMS-I-E start with measured/surveyed project fuel use and back-calculate baseline fuel consumption, assuming the equivalent energy would have been used in the baseline by the less-efficient baseline stove.

Methodologies give projects several options to determine most inputs. AMS-II-G, GS-simplified and GS-TPDDTEC allow projects to determine total baseline fuel use using a default value (0.4–0.5 tons of firewood per capita per year[40]), literature, national/project survey data or a KPT (rarely chosen)[16] (Table 1). AMS-II-G and GS-simplified use default values for the baseline stove efficiency and determine the project stove efficiency with a laboratory test. GS-metered and AMS-I-E determine baseline fuel consumption with default values, literature or surveys. GS-TPDDTEC and GS-metered require KPTs and metered or sales data, respectively.

CDM's previous default baseline stove efficiencies are lower than those found in the literature[41], while laboratory-derived project stove efficiencies are higher than actual performance in the field[42]. For projects that use default efficiencies, we update them to the CDM Methodology Panel's 2022 recommendations, which reflect current literature[40] (for example, from 10% and 20% to 15% and 25%, respectively, for firewood and charcoal).

Baselines constructed with project-led and national[43,44] fuel consumption surveys are vulnerable to social desirability[22,23] and recall[23,24] biases as households may want to present affluence and struggle to estimate kilograms of fuel used[23]. These biases can result in abnormally high baseline and/or low consumption values, especially when used together. Without a way to ground truth fuel consumption, we simply confine fuel consumption values to a reasonable literature-derived range of 2–4 MJ per capita per day[45,46] energy delivered to the pot (Supplementary Information).

fNRB

Projects that reduce biomass use should only be credited for the proportion of CO_2 emissions reduced from non-renewable sources. Previously, all methodologies relied on inaccurate CDM fNRB default values. As these defaults have now expired, projects may calculate fNRB values from a CDM tool[47] or assume a 30% default (rarely chosen). Both the earlier defaults and the tool overstate forest degradation compared

Table 2 | Outlining the factors and adjustments to each methodology based on published literature and our own analysis (Methods and Supplementary Information) and then the amount of over- or under-crediting from each individual factor across the issued credits from our sample of projects

Total amount of over-crediting across issued credits of studied projects from the average in our Monte Carlo Method (95% confidence interval)								
All factors	**Adoption rates[a]**	**Usage rates[a]**	**Stacking rates[a,b]**	**Fuel consumption**	**fNRB**	**EFs**	**Firewood–charcoal conversion**	**Rebound**
Definition	Percentage of distributed stoves actually in use	Percentage of meals cooked using the project stove	Percentage of meals cooked using the baseline stove in concert with the project stove	Amount of cooking fuel used by project households before and after obtaining the project stove	fNRB	The carbon dioxide equivalent emissions of fuel used, including upstream and non-carbon dioxide gases	Amount of firewood (on a wet basis) needed to produce the equivalent weight of charcoal (on a dry basis)	Increase in a household's overall cooking energy consumption with access to an improved stove
9.2 (7.0, 11.5)	1.4 (1.0, 1.7)	1.4 (1.1, 1.8)	1.1 (0.8, 1.4)	1.4 (1.1, 1.7)	1.7 (1.3, 2.1)	0.6 (0.5, 0.8)	1.5 (1.1, 2.0)	1.0 (0.8, 1.3)
Adjusted with	MCM using a triangle distribution: 58% (40%, 92%)	MCM using a triangle distribution: 52% (16%, 85%)	MCM using a triangle distribution: 68% (19.3%, 100%)	CDM's updated default baseline stove efficiencies if used and contained values within 2–4 MJ per capita per day delivered energy.	MCM using a triangle distribution from 'Scenario B–low yield' of Bailis et al. [2]	EFs for each cooking fuel from Floess et al. [49]	Charcoal upstream and point-of-use emissions factors from Floess et al. [49]	Literature-derived rebound effect: 22%
GS-TPDDTEC	✓	Discounted with an MCM using a uniform distribution with a maximum of a 53% decrease in usage	Discounted with MCM using a uniform distribution with a maxiimum of a 29% increase in stacking	✓	✓	✓	✓	
CDM-AMS-I-E (specific to ethanol projects)	✓	✓		✓	✓	✓	✓	✓
GS-simplified	✓	✓	✓	✓	✓	✓	✓	✓
CDM-AMS-II-G	✓	✓	✓	✓	✓	✓	✓	✓
GS-metered				✓	✓	✓	✓	✓

A check mark means that the approach outlined in the 'Adjusted with' row was applied; a blank cell means no adjustment was made and the text describes our approach. [a]One GS-TPDDTEC requires the removal of the baseline stove, and one AMS-II-G builds the improved stove in the exact spot of the baseline stove. We use slightly different Monte Carlo method distributions for these projects (see Supplementary Information). [b]Projects typically report a percentage of baseline stove use, which is then incorporated into the fuel consumption calculation. Using the project's documentation, we separate these two parameters.

with published literature[2]. WISDOM model of Bailis et al. [2] estimates fNRB, accounting for biomass regrowth and geographical, ecological and land use heterogeneity at the subnational level[2]. The most robust fNRB approach so far is a dynamic landscape model, Modelling Fuelwood Sustainability Scenarios[48]. When our study was conducted, few national values were available. Using the MCM, we replace project fNRB values with the 'Scenario B–low yield' 'minimum value' of Bailis et al. as the low boundary, 'expected value' as the mode and 10% over the expected value as the high boundary. On average, the projects chosen fNRBs are 3.0 (minimum 1.1, maximum 16.4) times the values of Bailis et al. [2] (Table 2 and Supplementary Information).

EFs

To translate fuel use into GHG emissions, GS uses 2006 Intergovernmental Panel on Climate Change (IPCC) default EFs and allows, but does not require, the inclusion of upstream emissions. Counterintuitively, to work around an early agreement prohibiting the crediting of reduced deforestation, CDM cookstoves methodologies apply a baseline EF assuming future fossil fuel use rather than biomass. This is a source of under-crediting[9]. We replace each approach with cooking fuel-specific EFs, including upstream emissions, from Floess et al. [49], the most comprehensive, up-to-date cooking fuel EF database. We also update all global warming potentials to the most recent IPCC values, accounting for distinctions for renewable/non-renewable biomass[50]. Due to high uncertainty around the climate impacts of black carbon emissions from cookstove projects[51], we, like the current methodologies, exclude black carbon.

Firewood–charcoal conversion

All methodologies allow projects replacing charcoal to use a firewood–charcoal conversion factor to estimate the amount of firewood (on a wet basis) needed to produce the equivalent weight of charcoal (on a dry basis). All used a default of six, which a CDM methodology panel updated to four in 2022 after our sample selection, based on literature[40]. Alternatively, methodologies allow projects to use literature to establish this conversion factor. All projects using

https://doi.org/10.1038/s41893-023-01259-6

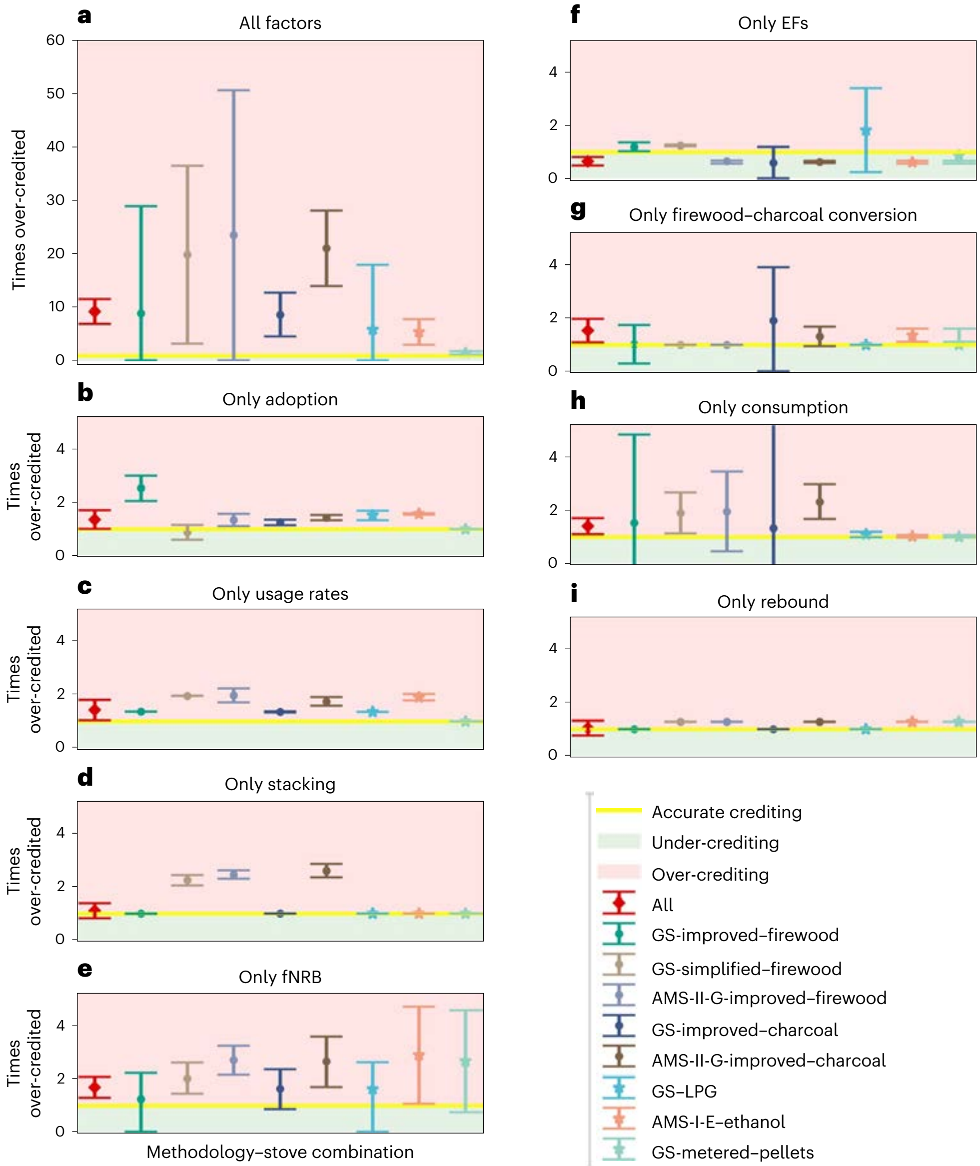


Fig. 3 | Over/under-crediting across factors. a–i, The mean amount of total over/under-crediting after quantifying all factors (n = 51 projects) (**a**) and individual factors by methodology–stove combinations for adoption (**b**), usage rates (**c**), stacking (**d**), fNRB (**e**), EFs (**f**), firewood–charcoal conversion (**g**), consumption (**h**) and rebound (**i**) methodologies only. GS–firewood (n = 9 projects), GS-simplified–firewood (n = 9 projects), GS–charcoal (n = 7 projects), GS–LPG (n = 4 projects), CDM-AMS-II-G–firewood (n = 13 projects), CDM-AMS-II-G–charcoal (n = 3 projects), CDM-AMS-II-E–ethanol (n = 4 projects) and GS-metered–pellet (n = 3 projects). The points indicate the total over- or under-crediting, while the error bars refer to the 95% CI for the total over-crediting across our sample of projects and the categories we delineate. We limit the CI's lower bounds to 0 (Methods and Supplementary Information). EFs include point-of-use emissions including non-CO_2 emissions and upstream emissions. Less than 1 (green shading) indicates under-crediting. Red shading indicates over-crediting, and yellow indicates accurate crediting.

this conversion used a value of 4.8 or higher. However, conversion efficiency is highly dependent on the specific location and charcoal production practices[40]. We do not use a firewood–charcoal conversion factor but instead use charcoal upstream and point-of-use EFs from Floess et al. [49].

Rebound effect

Households commonly increase their overall cooking energy consumption with access to an improved stove (for example, ref. 52). The improved stove lowers the 'cost' of cooking and provides another burner, allowing the household to increase their fuel consumption. Only projects that utilize KPTs capture this increase, which we confirm within our sample. We reduced our emission reduction estimation by 22% for projects that do not utilize KPTs, drawing on published literature that models or tracks the time stoves were used before and after the acquisition of an improved/clean stove through temperature sensors (Supplementary Information)[29,52–55].

Over/under-crediting analysis results

To find the total amount of over-crediting across our sampled portfolio, we estimate each project's over-crediting across analysed monitoring reports, then apply that to their total issued credits and compare our total ER estimates with their total issued credits. We estimate that our sample of cookstove projects are 9.2 times over-credited ((95% confidence interval (CI) 7.0, 11.5); Table 2 and Fig. 3). That is, the sample generated 26.7 million offset credits (as of May 2023), which is

https://doi.org/10.1038/s41893-023-01259-6

Table 3 | Recommendation for cookstove methodology reforms

To avoid over-crediting, new and current cookstoves methodologies should require, and until then, project developers should choose:	
Factor	**Recommendation**
fNRB	The 'Scenario B–low yield' value of Bailis et al. [2] at the lowest subnational level. Update to the Modelling Fuelwood Sustainability Scenarios value at the lowest subnational level as new research emerges.
Adoption, usage, stacking and rebound	One of the following options: 1. Meters or collect fuel purchase data for adoption, usage and stacking; a longitudinal survey or a conservative, literature-derived default for rebound; if a project has metered or fuel purchase data, this option is required 2. KPTs for usage and stacking, adjusted for the Hawthorne effect with a literature-derived default; robust longitudinal survey or conservative literature-derived default for adoption 3. Robust longitudinal surveys 4. Conservative literature-derived default values
Fuel consumption	Initial and update baseline KPTs and/or robust project-led surveys; enforce a reasonable range of 2–4 MJ-delivered per capita per day.
EFs	Upstream, point-of-use and non-CO_2 EFs for each cooking fuel from Floess et al., removing the need for a firewood–charcoal conversion factor IPCC's separate renewable/non-renewable global warming potentials for methane and nitrous oxide emissions, but continue to exclude black carbon pending future research

For full details on how to implement these recommendations, see our accompanying website[63].

over nine times our estimated carbon emission reductions of roughly 2.9 million tCO_2e.

Using the same approach, we extrapolate our estimates of over-crediting to the entire credit pool by methodology–stove combination. We find a total impact of roughly 5.2 million tCO_2e compared with the total 55.3 million VCM-issued credits.

We find that the average project in our sample is over-credited 27.6× (see Supplementary Information Section 6).

Respectively, fNRB, firewood–charcoal conversion, fuel consumption, adoption and usage produce the most over-crediting: 1.7, 1.5, 1.4, 1.4 and 1.4 times (Table 2). On average, only correcting the EFs resulted in under-crediting (0.6 times), while stove stacking and rebound minimally affects crediting amounts (1.1 and 1.0 times, respectively).

We find that all methodology–stove combinations over-credit (Fig. 3). AMS-II-G–firewood is the most over-credited project type from our sample (23.5 (0, 49.3)), stemming from specific project values (fNRB ~2.7× and consumption ~2.4×) and the methodology's approach (stacking ~2.5×, usage ~2.0×, adoption ~1.4× and rebound ~1.3×) that together have a multiplier effect. AMS-II-G–charcoal is the second most over-credited project type (21.0 (12.7, 29.4)) from the same sources, except their usage rates were closer to literature-derived values, while they had an additional source of over-crediting from the firewood–charcoal conversion (~1.3×). The CDM methodologies' weak monitoring approach overcomes the under-crediting from their use of the EF from a projected fossil fuel (~0.6–0.7×). GS-simplified–firewood (19.8 (2.5, 37.2)) is more over-credited than GS–firewood (8.9 (0, 26.9)) and GS–charcoal (8.6 (4.5, 12.8)), under GS-TPDDTEC, due to their less robust monitoring approach (that is, GS-simplified does not require KPTs). Compared with GS–firewood, GS–charcoal projects over-credited less from adoption, but over-credited from the firewood–charcoal conversion (~1.9×). GS–liquefied petroleum gas (LPG) over-credits by 5.9 (0, 16.3) times, from fNRB, adoption, usage and EFs. AMS-I-E–ethanol over-credits 5.4 (3.2, 7.6) times from adoption (1.6×), usage (1.9×), fNRB (2.9×) and rebound (1.3×), but under-credits from CDM's use of fossil fuel EFs (0.6×). GS-metered–pellets have the least over-crediting (1.5 (0.6, 2.4)), stemming only from fNRB and rebound, with slight under-crediting from EFs.

Over-crediting from fNRB stems from location-specific differences in the values of Bailis et al. [2] (Supplementary Fig. 4). Adoption, usage and stacking rates affect methodology–stove combinations based on the methodology's requirements (for example, meters, KPTs and surveys). GS–LPG, AMS-I-E–ethanol and GS-metered–pellets, on average, did not report fuel consumption values outside of a reasonable range, probably due to the use of KPTs, meters or sales data.

EF choices result in overall under-crediting (0.6×) from five methodology–stove combinations: CDM methodologies use the low EF of a projected fossil fuel as the baseline, GS–charcoal projects do not always include upstream emissions and GS-metered–pellets projects construct a weighted average baseline EF, which ultimately is lower than those in Floess et al. [49]. The EFs used by GS–firewood and GS–LPG for the baseline fuels are slightly higher than Floess et al. [49], leading to slight over-crediting, stemming from project-chosen values, not the switch to LPG.

Per stove-day, GS-metered–pellets and AMS-I-E projects reduce emissions by roughly 0.007 and 0.006 tCO_2e due to their renewable feedstocks, and thus minimal project emissions. They are followed, on a per stove-day basis, by GS–charcoal (0.003 tCO_2e), AMS-II-G–firewood, GS–firewood and LPG (0.001 tCO_2e), AMS-II-G–charcoal (0.0004 tCO_2e) and GS-simplified–firewood (0.0002 tCO_2e).

Discussion

We conservatively estimate that the total amount of over-crediting across our sample's issued credits is 9.2 (7.0, 11.5), stemming from misalignment across numerous, compounded factors.

The majority of over-crediting stems from lack of rigour and flexibility in how methodologies determine fNRB, adoption, usage, stacking and fuel consumption, despite periodic methodological updates. We provide recommendations for aligning methodologies with current science (Table 3). Regular updates will be needed to reflect future research advancements. Currently, project developers, who benefit financially from more credits, hire verifiers directly, possibly conflicting with the International Organization for Standardization (17029) that requires the verifier to be impartial (C5.3)[56]. The developers' incentives are evident, as robust fNRB values have been published for 8 years, yet all projects have opted to use higher CDM tool-derived or default values, and some projects track purchase data, yet fail to use it in reduction estimation. Eliminating the flexibility and requiring robust or conservative methods could reduce over-crediting easily, universally by 1.4–1.7 times for each factor.

Developers can apply these recommendations without incurring extra cost. For adoption, usage and stacking, while meters, longitudinal surveys and KPTs are the most accurate, they also can be costly depending on project infrastructure and size. For these factors, we include in our recommendations the option of literature-derived values that have no cost, and despite being less accurate, are likely to avoid over-crediting.

Additionally, increases in offset prices could make these needed reforms more affordable. There is a feedback loop—poor quality keeps offset prices too low to support accurately credited projects. Higher prices for accurately estimated reduction could incentivize and fund projects to promote behaviour change, increase awareness and address other market and behavioural barriers to cooking energy transformation[57].

In the current landscape, buyers are left confused about what constitutes quality, and often turn to rating companies. Similarly, for project co-benefits, some buyers are willing to pay more for projects with more co-benefits, but have been reported to care more about the number of SGDs than the quality of that contribution[58]. Project's claimed co-benefits are measured, unfortunately, alongside

https://doi.org/10.1038/s41893-023-01259-6

the adoption, usage and stacking rates, through single cross-sectional surveys, which are subject to the same biases our analysis outlines[58]. Low-quality tracking of both the carbon abatement and co-benefits leads to surface level, performative action, rather than meaningful, sustainable impact.

Our results are a call to action to overhaul offset programme design and the dominance of improved but not WHO-defined clean stoves. Prioritizing metered fuel switch projects and accurately quantifying their emission reductions would progress climate, energy and health SDGs. Our analysis indicates that these stoves currently offer the least over-credited credits and have the greatest abatement potential and health benefit. Further, they are often the most challenging for users to sustainably use, given the need for continuous fuel purchases, and thus are the cookstove project types that could most benefit from carbon finance. Our results further support Gill-Wiehl and Kammen's call for the VCM to exclusively fund WHO-defined clean stoves[59], and highlight the lost opportunity to use cookstove offsets to accelerate access to the cleanest stoves/fuels. Quality cookstove offsets could sustainably, instead of performatively, improve the health of people and the planet.

Methods

Due to the nature of this analysis, the results of our study of carbon accounting methods for cookstove projects are also the methods we used in our over/under-crediting analysis and inform our recommendations. Thus, our methods are summarized in the main text. Here, in the methods section, we include further explanation of how we adjusted factors, performed the MCM and estimated over/under-crediting, and discuss the limitations of our work. Further explanation and justification of our methods for each factor is provided in Supplementary Information.

Sample selection

We evaluate the quality of offset credits from the methodologies with the largest number of cookstove offset project activities on the VCM: GS-TPDDTEC, GS-simplified, CDM-AMS-II-G and CDM-AM-I-E. We also review the new GS-metered, released October 2021.

The methodologies deploy different project stoves. GS-TPDDTEC (previously GS's Methodology for Improved Cook-Stoves and Kitchen Regimes) is the most versatile methodology covering any thermal domestic technology switch that is less GHG intensive, including but not limited to improved biomass, heat retention, solar, LPG and electric stoves. CDM-AMS-I-E replaces non-renewable biomass with renewable energy (for example, renewable biomass, biogas, bioethanol and electric stoves). Designed for smaller projects, GS-simplified and CDM-AMS-II-G have limited scopes, only allowing for biomass efficiency projects (for example, traditional fuelwood stove to an improved fuelwood stove). GS-metered is designed for cookstoves with metered or other direct fuel monitoring (for example, purchase records) such as electric, LPG, biogas, bioethanol or advanced biomass pellet stoves.

Most cookstove projects are structured as programme of activities, in which multiple similar project activities (called voluntary project activities (VPAs) on the VCM and component project activities on the CDM) are bundled together to allow for rapid replication, only requiring a quick check from a validator and not a full registration procedure[60]. To reflect the diversity of projects on the VCM, we evaluated VPAs separately. CDM methodologies are used on both the CDM and the VCM, but we limited our scope to only VCM-registered projects (that is, those certified by GS or Verra).

In March 2021, we identified the 15 countries with the most credits from cookstove projects on the market and, for each country, selected the projects with the most credits for each methodology. For the GS-TPDDTEC, GS-simplified and CDM-AMS-II-G projects, we chose projects that posted at least one monitoring report and provided their exact calculations and the stove-days. There were very few projects under AMS-I-E and GS-metered and the only one that had been issued credits was also credited under AMS-I-I and so was not included in our sample. For these two methodologies, we selected all registered projects that provided enough information to recreate offset credit calculations on a stove-day basis for individual stove types. We included these methodologies because they offered different methods for monitoring stove usage and fuel consumption, and because of the greater potential emission reductions and health benefits from fuel switch projects that these protocols accommodate. We added additional projects as needed to ensure that our sample covered all types of fuel transitions, with the exception of electric stoves. There were no issued projects actively deploying an electric stove, and the only listed electric project under GS-metered had no files available. We do not include GS-metered's most recent methodology update, which allows for the participation of more complex cooking devices such as pressure cookers, in a new option called 'specific consumption' (Supplementary Information).

Additionally, we randomly selected ten small/medium-sized projects from GS-TPDDTEC (four), AMS-II-G (four) and GS-simplified (two) to ensure that our sample was representative of both large and small projects. We investigate the relationship between the amount of over/under-crediting and project size, and find a slight negative relationship between amount of over-crediting and total verified credits (evaluated on the log scale; Supplementary Fig. 1). This trend is not statistically significant and the R^2 is very low, but it indicates that our approach of focusing on large projects may have led to lower estimates of over-crediting.

This approach resulted in a sample of 51 projects, spanning 25 countries and 8 methodology–project type combinations: (1) GS–firewood, (2) GS-simplified–firewood, (3) GS–charcoal, (4) GS–LPG, (5) CDM-AMS-II-G–firewood, (6) CDM-AMS-II-G–charcoal, (7) CDM-AMS-II-E–ethanol and (8) GS-metered–pellet (WHO tier 4+ biomass pellet stove). Our sample covers 40% of all issued credits on the VCM from these methodologies (as of 10 May 2023). We have no reason to believe that these projects are not representative of the entire pool of cookstove credits on the VCM. The 31 GS projects in our sample represent 46% of the covered GS methodologies credits on the VCM. The 16 AMS-II-G contain 25% of that methodologies' credits.

Our sample of 51 projects tangentially represents 478 projects and 64% of total credits issued under the five studied methodologies as many projects are structured as largely identical VPAs under programme of activities.

Uncertainty

Quantification of emission reductions from offset programmes is inherently uncertain. Emission reductions must be estimated against an immeasurable counterfactual scenario. Other factors, notably fNRB, upstream emissions and leakage are also difficult to estimate, and with limited research so far, involve substantial uncertainty. Since offset credits often are used to 'offset' or trade with direct emission reductions, to maintain the integrity of an emission reduction claim, offset programmes are tasked with estimating programme impacts conservatively when there is uncertainty. Here, conservative means more likely to under-credit than to over-credit. Our analysis uses the most rigorous and up-to-date values from the literature when available (for example, fNRB). Instead of choosing conservative methods for all factors, we do not or minimally correct factors with little published research, notably additionality, leakage, non-permanence and overlapping claims, and instead recommend more research. In this way, we make methodological choices that probably underestimate the amount of over-crediting.

Methodology updates

All methodologies, except for recently released GS-metered, have undergone considerable updates over the years of credit generation that affect the methodological factors we study. Our recommendations and discussion below focus on the most recent version of each

https://doi.org/10.1038/s41893-023-01259-6

methodology and any updates proposed by the registry. However, most credits on the VCM, including those still available for purchase, are issued under previous methodology versions. Therefore, our quantitative over/under-crediting analysis assesses the credits generated regardless of the methodology version used. We note in the main text and detail in Supplementary Information where updated methodologies address over-crediting.

Adjusting factors

Using the values listed in the latest verified monitoring report or project documents of these 51 projects, we calculated the number of VERs on a per stove-day basis. We only included projects (or monitoring reports from projects) in our sample if we were exactly able to replicate the number of VERs either in total or on a per stove-time basis. Once we replicated the credits generated under the methodologies, we then adjusted all the identified factors contributing to over/under-crediting as described above. Then, we conducted analyses isolating each factor.

To make the factor analysis of EFs, firewood–charcoal conversion factor and consumption for GS-metered–pellet and AMS-I-E–ethanol comparable to all other methodology–stove combinations, we remove GS-metered and AMS-I-E's calculation approach and calculate the baseline emissions and project emissions separately. For example, we use the baseline and project consumption reported in their project documents to calculate the difference between baseline and project emissions instead of using their baseline conversion factor approach (see the 'Fuel consumption' section).

Finally, we conduct one analysis excluding adoption, usage and stacking rates, which are the only factors that are always monitored ex post. We do this for fair comparison with GS-metered–pellet and AMS-I-E–ethanol projects, which, as of the time of sampling, had generated no credits. In our main analysis, we use their ex ante values for adoption and stacking rates from the project documents rather than ex post values from monitoring reports as with all other projects.

In total, we have analyses in which (1) all factors are adjusted, (2) only adoption rates are adjusted, (3) only usage rates are adjusted, (4) only stacking rates are adjusted, (5) only fNRB values are adjusted, (6) only EFs (including upstream emissions) are adjusted, (7) only the firewood–charcoal conversion is adjusted, (8) only consumption (baseline and project) values are adjusted, (9) only rebound consumption is adjusted and (10) all factors are adjusted, except adoption, usage and stacking (Supplementary Information).

MCM

The MCM is a statistical framework that calculates possible outcomes when input parameters are randomly varied within a specified range using a given distribution[61]. When used for fNRB, adoption, usage and stacking rates, the MCM generates values within our defined limits, following the distribution defined in each factor's section, assuming independence (see 'Limitations' section). We specified the simulation to run 10,000 times, randomly generating new values for each of these factors and calculating an associated emission reduction. We acknowledge the inherent uncertainty within our factors and bound each one within a literature-derived range. We take this approach over other methods of error propagation given the inherent uncertainty and imprecision in the ranges within the literature. Johnson et al., for example, propagated error as they had direct field measurements for their study site for fNRB, EFs and fuel consumption. Without this level of precision for each carbon offset location, we take a higher level, although less precise approach. However, as discussed, we make methodological decisions that result in likely underestimation of the amount of over-crediting.

Estimating over/under-crediting

We estimate the over-crediting across our sampled portfolio in three ways. To estimate the total over-crediting of our sample, we estimate each project's over-crediting across analysed monitoring reports, apply that value to each project's total issued credits and compare our total ER estimates with their total issued credits (Fig. 3 and Supplementary Table 5). For the projects in our sample that have not generated credits (see the section on sampling), we use their estimated annual emission reductions from their project design documents. We then splice the results by methodology–project type combination (Fig. 3) and then by country (Supplementary Fig. 3).

Second, we average over-crediting by project across our sample (Supplementary Table 6). Finally, we take an average of our data points at the highest level of granularity, that is, at the level of the monitoring report or stove type within a monitoring report (Supplementary Table 7).

We construct CIs around the total amount of over-crediting by finding the standard deviation across the total over-crediting by project based off the average MCM for all and for the specific factor analysis. These CIs become larger within the subanalyses due to smaller sample sizes. Negative lower bounds of the CI are a function of large standard deviations due to specific project values and smaller sample sizes. Note that, within this over-crediting reporting framework, under-crediting is indicated by a value between 0 and 1, not negative. We thus limit the lower limit of CIs to zero.

To extrapolate to the entire cookstove market, we take the total rate of over-crediting for each methodology–stove combination found above, and then apply these rates to the total amount of credits issued for each methodology–stove combination. Thus, we find that the whole market is over-credited 10.6 times weighting by methodology–stove combination.

Commercial credits

A few of our sample projects included some stoves used for commercial purposes (restaurants, schools and so on), representing a small fraction of these projects' total credits. We do not adjust commercial stoves' adoption, usage or stacking rates, or baseline/project fuel consumption. There are still barriers to adoption, usage and ending stacking for commercial institutions; however, the literature on these rates is limited[62], and thus an area for future research.

Limitations

Our study has some limitations that must moderate our conclusions. This analysis does not cover 100% of projects under the five studied methodologies. We cover 40% of the market, and projects in 25 countries; however, we attempted to have a fully representative sample across methodologies, location and project type. We were limited to projects that were transparent enough to provide their exact calculations or stove-days within their monitoring or validation reports. All factors involve some amount of uncertainty, which we address with the MCM for some factors. We were limited by the details provided by the projects and the standards. For example, numerous projects did not specify the rural or urban setting or more specific administrative units, which is important for fNRB.

Finally, a key limitation in our work is that we assume that all factors are independent. This is an appropriate assumption for all factors, potentially except for adoption, usage and stacking. For example, there is no evidence in the literature that fNRB or EF is correlated with stove adoption; however, there could be correlation with stove adoption and usage. This correlation, however, would be highly context dependent and probably time variant (that is, a household's relationship with and use of an intervention has been shown to change over time). In creating the distributions for adoption, usage and stacking, we create ranges of uncertainty, since rates of adoption, usage and stacking have been reasonably well studied and there is an established literature that we draw from. Unfortunately, the correlation between these rates has not been well established and would require less grounded assumptions. This is also a reason that we pursue triangle distributions as we hesitate to make definitive claims on the underlying distributions, opting rather to simply present that the literature has established general ranges for these values as described above. Given this context, we therefore assume independence of all factors. This is a limitation of our work, but one that probably leads to more conservative findings. This is because incorporating the covariance between adoption, usage and stacking would further limit the input distribution of these factors and thus shrink our reported CIs. Thus, our reported ranges provide more coverage. We further feel comfortable with this methodological decision given the other areas that probably result in underestimation, as above.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All data and code are publicly available online at https://github.com/agillwiehl/GillWiehl_et_al_Pervasive_over_crediting.

Code availability

All data and code are publicly available online at https://github.com/agillwiehl/GillWiehl_et_al_Pervasive_over_crediting.

References

1. Abbafati, C. et al. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* **396**, 1223–1249 (2020).
2. Bailis, R., Drigo, R., Ghilardi, A. & Masera, O. The carbon footprint of traditional woodfuels. *Nat. Clim. Change* **5**, 266–272 (2015).
3. Defining clean fuels and technologies. *World Health Organization* https://www.who.int/tools/clean-household-energy-solutions-toolkit/module-7-defining-clean (2021).
4. So I., Haya, B. & Elias, M. Voluntary registry offsets database v.8. *University of California, Berkeley* https://gspp.berkeley.edu/faculty-and-impact/centers/cepp/projects/berkeley-carbon-trading-project/offsets-database (2023).
5. Stapp, J. et al. Little evidence of management change in California's forest offset program. *Commun. Earth Environ.* **4**, 331 (2023).
6. Haya, B. *Policy Brief: the California Air Resources Board's US Forest Offset Protocol Underestimates Leakage* (Goldman School of Public Policy, University of California, Berkeley, 2019); https://gspp.berkeley.edu/assets/uploads/research/pdf/Policy_Brief-US_Forest_Projects-Leakage-Haya_4.pdf
7. Haya, B. et al. Quality assessment of REDD+ carbon credit projects. Berkeley carbon trading project. *Goldman School of Public Policy, University of California, Berkeley* https://gspp.berkeley.edu/research-and-impact/centers/cepp/projects/berkeley-carbon-trading-project/REDD (2023).
8. West, T. A. P. et al. Action needed to make carbon offsets from forest conservation work for climate change mitigation. *Science* **381**, 873–877 (2023).
9. Cames, M. et al. How additional is the clean development mechanism? Analysis of the application of current tools and proposed alternatives. *DG CLIMA* https://doi.org/10.13140/RG.2.2.23258.54728 (2016).
10. Haya, B. K. *Carbon Offsetting: An Efficient Way to Reduce Emissions or to Avoid Reducing Emissions? An Investigation and Analysis of Offsetting Design and Practice in India and China*. PhD thesis, University of California, Berkeley (2010).
11. Bailis, R., Wang, Y., Drigo, R., Ghilardi, A. & Masera, O. Getting the numbers right: revisiting woodfuel sustainability in the developing world. *Environ. Res. Lett.* **12**, 115002 (2017).
12. Ramanathan, T. et al. Wireless sensors linked to climate financing for globally affordable clean cooking. *Nat. Clim. Change* **7**, 44–47 (2017).
13. Freeman, O. E. & Zerriffi, H. How you count carbon matters: implications of differing cookstove carbon credit methodologies for climate and development cobenefits. *Environ. Sci. Technol.* **48**, 14112–14120 (2014).
14. Sanford, L. & Burney, J. Cookstoves illustrate the need for a comprehensive carbon market. *Environ. Res. Lett.* **10**, 084026 (2015).
15. Simon, G. L., Bumpus, A. G. & Mann, P. Win–win scenarios at the climate–development interface: challenges and opportunities for stove replacement programs through carbon finance. *Glob. Environ. Change* **22**, 275–287 (2012).
16. Lee, C. M., Chandler, C., Lazarus, M. & Johnson, F. X. Assessing the climate impacts of cookstove projects: issues in emissions accounting. *Chall. Sustain.* **1**, 53–71 (2013).
17. Reduced emissions from cooking and heating—technologies and practices to displace decentralized thermal energy consumption (TPDDTEC). *The Gold Standard Foundation* https://globalgoals.goldstandard.org/407-ee-ics-technologies-and-practices-to-displace-decentrilized-thermal-energy-tpddtec-consumption/ (2021).
18. The Gold Standard simplified methodology for clean and efficient cookstoves. *The Gold Standard Foundation* https://globalgoals.goldstandard.org/408-ee-ics-simplified-methodology-for-efficient-cookstoves/ (2022).
19. AMS-II.G.: energy efficiency measures in thermal applications of non-renewable biomass version 12.0. *Clean Development Mechanism* https://cdm.unfccc.int/methodologies/DB/10PELMPDW951SVSW1B2NRCQEBAX96C (2022).
20. AMS-I.E.: switch from non-renewable biomass for thermal applications by the user version 12.0. *Clean Development Mechanism* https://cdm.unfccc.int/methodologies/DB/BLVN9ULDP1FRUVS2LYWW6WPYN9W78E (2021).
21. Methodology for metered and measured energy cooking devices. *The Gold Standard Foundation* https://globalgoals.goldstandard.org/news-methodology-for-metered-measured-energy-cooking-devices/ (2022).
22. Krumpal, I. Determinants of social desirability bias in sensitive surveys: a literature review. *Qual. Quant.* **47**, 2025–2047 (2013).
23. Kar, A., Brauer, M., Bailis, R. & Zerriffi, H. The risk of survey bias in self-reports vs. actual consumption of clean cooking fuels. *World Dev. Perspect.* **18**, 100199 (2020).
24. Wilson, D. L. et al. in *Technologies for Development* (eds Hostettler S. et al.) 211–221 (Springer, 2015).
25. Simons, A. M., Beltramo, T., Blalock, G. & Levine, D. I. Using unobtrusive sensors to measure and minimize Hawthorne effects: evidence from cookstoves. *J. Environ. Econ. Manage.* **86**, 68–80 (2017).
26. Shankar, A. V. et al. Everybody stacks: lessons from household energy case studies to inform design principles for clean energy transitions. *Energy Policy* **141**, 111468 (2020).
27. Hanna, R., Duflo, E. & Greenstone, M. Up in smoke: the influence of household behavior on the long-run impact of improved cooking stoves. *Am. Econ. J. Econ. Policy* **8**, 80–114 (2016).
28. Burwen, J. & Levine, D. I. A rapid assessment randomized-controlled trial of improved cookstoves in rural Ghana. *Energy Sustain Dev.* **16**, 328–338 (2012).
29. Beltramo, T., Blalock, G., Harrell, S., Levine, D. & Simons, A. M. The effects of fuel-efficient cookstoves on fuel use, particulate matter, and cooking practices: results from a randomized trial in rural Uganda. *UC Berkeley Center for Effective Global Action* https://escholarship.org/uc/item/365778pn (2019).
30. Rosa, G. et al. Assessing the impact of water filters and improved cook stoves on drinking water quality and household air pollution: a randomised controlled trial in Rwanda. *PLoS ONE* **9**, e91011 (2014).

31. Bensch, G. & Peters, J. The intensive margin of technology adoption—experimental evidence on improved cooking stoves in rural Senegal. *J. Health Econ.* **42**, 44–63 (2015).
32. Ruiz-Mercado, I., Masera, O., Zamora, H. & Smith, K. R. Adoption and sustained use of improved cookstoves. *Energy Policy* **39**, 7557–7566 (2011).
33. Islam, M. M. et al. Assessing the effects of stove use patterns and kitchen chimneys on indoor air quality during a multiyear cookstove randomized control trial in rural India. *Environ. Sci. Technol.* **56**, 8326–8337 (2022).
34. García-Frapolli, E. et al. Beyond fuelwood savings: valuing the economic benefits of introducing improved biomass cookstoves in the Purépecha Region of Mexico. *Ecol. Econ.* **69**, 2598–2605 (2010).
35. Agurto Adrianzén, M. Social capital and improved stoves usage decisions in the Northern Peruvian Andes. *World Dev.* **54**, 1–17 (2014).
36. Jeuland, M., Soo, J. S. T. & Shindell, D. The need for policies to reduce the costs of cleaner cooking in low income settings: implications from systematic analysis of costs and benefits. *Energy Policy* **121**, 275–285 (2018).
37. Ruiz-Mercado, I., Canuz, E., Walker, J. L. & Smith, K. R. Quantitative metrics of stove adoption using stove use monitors (SUMs). *Biomass Bioenergy* **57**, 136–148 (2013).
38. Pine, K. et al. Adoption and use of improved biomass stoves in rural Mexico. *Energy Sustain. Dev.* **15**, 176–183 (2011).
39. Pattanayak, S. K. et al. Experimental evidence on promotion of electric and improved biomass cookstoves. *Proc. Natl Acad. Sci. USA* **116**, 13282–13287 (2019).
40. MP88: meeting report/recommendations to the executive board. *CDM Methodologies Panel* https://cdm.unfccc.int/Panels/meth/index.html (2022).
41. Life cycle assessment of cooking fuel systems in India, China, Kenya, and Ghana. *US Environmental Protection Agency* https://cfpub.epa.gov/si/si_public_record_report.cfm?dirEntryId=339679&Lab=NRMRL&simplesearch=0&showcriteria=2&sortby=pubDate&timstype=Published+Report&datebeginpublishedpresented (2021).
42. Wathore, R., Mortimer, K. & Grieshop, A. P. In-use emissions and estimated impacts of traditional, natural- and forced-draft cookstoves in rural Malawi. *Environ. Sci. Technol.* **51**, 1929–1938 (2017).
43. Stockwell, T. et al. Estimating under- and over-reporting of drinking in national surveys of alcohol consumption: identification of consistent biases across four English-speaking countries. *Addiction* **111**, 1203–1213 (2016).
44. Ezzati, M., Martin, H., Skjold, S., Hoorn, S. V. & Murray, C. J. L. Trends in national and state-level obesity in the USA after correction for self-report bias: analysis of health surveys. *J. R. Soc. Med.* **99**, 250–257 (2006).
45. *Concept Note CDM-MP85-A07. Analysis and Options Regarding Caps Used in AMS-I.E, AMS-II.G and TOOL30 Version 01.0* (Clean Development Mechanism, 2013); https://cdm.unfccc.int/sunsetcms/storage/contents/stored-file-20210708220535947/MP85_EA07_Concept%20Note%20-%20Caps.pdf
46. Daioglou, V., van Ruijven, B. J. & van Vuuren, D. P. Model projections for household energy use in developing countries. *Energy* **37**, 601–615 (2012).
47. Tool 30: calculation of the fraction of non-renewable biomass (version 3). *Clean Development Mechanism* https://cdm.unfccc.int/methodologies/PAmethodologies/tools/am-tool-30-v1.pdf/history_view (2020).
48. Ghilardi, A. et al. Spatiotemporal modeling of fuelwood environmental impacts: towards improved accounting for non-renewable biomass. *Environ. Model Softw.* **82**, 241–254 (2016).
49. Floess, E. et al. Scaling up gas and electric cooking in low- and middle-income countries: climate threat or mitigation strategy with co-benefits? *Environ. Res. Lett.* **18**, 034010 (2023).
50. Whitman, T. L. & Lehmann, C. J. Systematic under and overestimation of GHG reductions in renewable biomass systems. *Clim. Change* **104**, 415–422 (2011).
51. Huang, Y. et al. Global radiative effects of solid fuel cookstove aerosol emissions. *Atmos. Chem. Phys.* **18**, 5219–5233 (2018).
52. Bailis R., et al. Enhancing clean cooking options in peri-urban Kenya: a pilot study of advanced gasifier stove adoption. *Environ. Res. Lett.* **15**, 084017 (2020).
53. Forum on natural capital accounting for better policy decisions: taking stock and moving forward. *World Bank Group* http://documents.worldbank.org/curated/en/904211580129561872/Forum-on-Natural-Capital-Accounting-for-Better-Policy-Decisions-Taking-Stock-and-Moving-Forward (2017).
54. Dufournaud, C. M., Quinn, J. T. & Harrington, J. J. A partial equilibrium analysis of the impact of introducing more-efficient wood-burning stoves into households in the Sahelian Region. *Environ. Plan. Econ. Space* **26**, 407–414 (1994).
55. Lambe, F. et al. Opening the black pot: a service design-driven approach to understanding the use of cleaner cookstoves in peri-urban Kenya. *Energy Res. Soc. Sci.* **70**, 101754 (2020).
56. ISO/IEC 17029:2019 general principles and requirements for validation and verification bodies. *ISO* https://www.iso.org/standard/29352.html (2019).
57. Khavari, B., Ramirez, C., Jeuland, M. & Fuso Nerini, F. A geospatial approach to understanding clean cooking challenges in sub-Saharan Africa. *Nat. Sustain.* **6**, 447–457 (2023).
58. Bakhtary, H., Tierney, M., Galt, H. & Gill-Wiehl, A. *More than just a Carbon Project: How Clean Cooking Projects Certified Under the Gold Standard Approach SDG Claims* (Climate Focus, 2023); https://mecs.org.uk/wp-content/uploads/2023/05/FINAL-SDG-Briefing-More-than-just-a-carbon-project.pdf
59. Gill-Wiehl, A. & Kammen, D. M. A pro-health cookstove strategy to advance energy, social and ecological justice. *Nat. Energy* **7**, 999–1002 (2022).
60. The handbook for programme of activities: practical guidance to successful implementation. *Climate Focus* https://climatefocus.com/publications/handbook-programmes-activities-practical-guidance-successful-implementation/ (2011).
61. Fitzpatrick, D. in *Analog Design and Simulation Using OrCAD Capture and PSpice* 2nd edn (ed Fitzpatrick, D.) 151–164 (Newnes, 2018).
62. Robinson, B. L., Clifford, M. J., Hewitt, J. & Jewitt, S. Cooking for communities, children and cows: lessons learned from institutional cookstoves in Nepal. *Energy Sustain. Dev.* **66**, 1–11 (2022).
63. Gill-Wiehl, A., Hogan, M. & Haya, B. A comprehensive quality assessment of cookstoves carbon credits. *Golman School of Public Policy, University of California, Berkeley* https://gspp.berkeley.edu/research-and-impact/centers/cepp/projects/berkeley-carbon-trading-project/cookstoves (2023).

Acknowledgements

The research that led to these results received funding from the Center for African Studies, University of California, Berkeley (A.G.-W.), Carbon Direct (B.K.H.), Google.org (D.K.), Katherine Lau Family Foundation (A.G.-W. and D.K.), the National Science Foundation's Graduate Researcher Fellowship Program (A.G.-W.), the Mulago Foundation/Better Cooking Company (A.G.-W.) and the Zaffaroni Family Foundation (D.K.). We thank A. Haya for help with initial data collection. We also thank Calyx Global and D. Lee. A.G.-W. thanks I. Ray and A. Hubbard for additional advising support.

Author contributions

A.G.-W. and B.K.H. co-led the research design. A.G.-W. compiled the data, conducted the analysis and co-led the write up of the paper. B.K.H. originated the idea and co-led the write up of the paper. D.K. contributed to the research design and write up, as well as funding the work.

Competing interests

A.G.-W. has received research support from Better Cooking Company Limited, whose leadership also provided comments on a draft of the manuscript. B.K.H. has received research support from Carbon Direct. D.K. declares no competeing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41893-023-01259-6.

Correspondence and requests for materials should be addressed to Annelise Gill-Wiehl.

Peer review information *Nature Sustainability* thanks Philippe Delacote, Ben Groom and Bruno Marino for their contribution to the peer review of this work.

nature portfolio

Corresponding author(s): Annelise Gill-Wiehl

Last updated by author(s): Dec 1, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used.
Data analysis	We used Python 3.0 to analyze the data. Figure 1 was created in R 4.2.1. All data and code are available at https://github.com/agillwiehl/GillWiehl_et_al_Pervasive_over_crediting

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All data and code are available at https://github.com/agillwiehl/GillWiehl_et_al_Pervasive_over_crediting

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description

We conduct a comprehensive, quantitative quality assessment of carbon offsets comparing cookstove offset methodologies and projects to published literature and our own analysis. We (1) discuss the accuracy of all estimation factors used (or not addressed) by the four most prominent cookstoves offset methodologies (GS-TPDDTEC17 , GS-Simplified18, CDM-AMS-II-G19, and CDM-AMS-I-E20) and the recent GS-Metered21 methodology (Table 1) (past and current versions) drawing from published literature and our own analysis (see methods). We then (2) recalculate the carbon emission reductions of a purposive sample of 51 cookstoves projects, addressing ranges of uncertainty using the Monte Carlo Method (MCM), and compare those results with actual credit issuance across eight methodology/stove type categories. We (3) suggest a specific set of methodological reforms to generate high-quality credits. In doing so, we (4) develop and demonstrate an over/under crediting analysis that can be used to systematically assess quality and inform methodology improvements across all offset project types

Research sample

Our sampling approach resulted in a sample of 51 projects spanning 25 countries and eight methodology-project type combinations: (1) GS-Firewood, (2) GS-Simplified-Firewood, (3) GS-Charcoal, (4) GS-LPG, (5) CDM AMS-II-G-Firewood, (6) CDM-AMS-II-G-Charcoal, (7) CDM-AMS-II-E-Ethanol, and (8) GS-Metered-Pellet (WHO Tier 4+ Biomass Pellet Stove). Our sample covers 40% of all issued credits on the VCM from these methodologies (as of May 10th, 2023). We have no reason to believe that these projects are not representative of the entire pool of cookstove credits on the VCM. The 31 GS projects in our sample represent 46% of the covered GS methodologies credits on the VCM. The 16 AMS-II-G contain 25% of that methodologies' credits.

Sampling strategy

In March 2021, we identified the 15 countries with the most credits from cookstove projects on the market and for each country selected the projects with the most credits for each methodology. For the GS-TPDDTEC, GS-Simplified, and CDM-AMS-II-G projects, we chose projects that posted at least one monitoring report and provided their exact calculations and the stove days. There were very few projects under AMS-I-E and GS-Metered and the only one that had been issued credits was also credited under AMS-I-I and so was not included in our sample. For these two methodologies, we selected all registered projects that provided enough information to recreate offset credit calculations on a stove-day basis for individual stove types. We included these methodologies because they offered different methods for monitoring stove usage and fuel consumption, and because of the greater potential emission reductions and health benefits from fuel switch projects that these protocols accommodate. We also added additional projects as needed to ensure that our sample covered all types of fuel transitions, with the exception of electric stoves. There were no issued projects actively deploying an electric stove, and the only listed electric project under GS Metered had no files available. We also do not include GS Metered's most recent methodology update which allows for the participation of more complex cooking devices such as pressure cookers, in a new option called "specific consumption". See supplemental methodology equation information.

Additionally, we randomly selected 10 small/medium sized projects from GS-TPDDTEC (4), AMS-II-G (4), and GS-Simplified (2) to ensure that our sample was representative of both large and small projects. We investigate the relationship between the amount of over/under crediting and project size and find a slight negative relationship between amount of over-crediting and total verified credits (evaluated on the log scale; see supplemental Figure S1). This trend is not statistically significant and the R-squared is very low, but it indicates that our approach of focusing on large projects may have led to lower estimates of over-crediting.

Data collection

The first author obtained all data from the publicly available databases from Gold Standard (SustainCert) and Verra's registry.

Timing and spatial scale	We selected the initial sample of projects and the respective documents from SustainCert in March 2021. We added 5 GS Simplified projects in March of 2023 after receiving feedback from Gold Standard. We added 10 projects after feedback from the review process.
Data exclusions	No data were excluded from the analysis once the sample was established.
Reproducibility	All attempts to repeat the analysis were successful.
Randomization	Randomization is not applicable to our study design.
Blinding	Blinding was not applicable to our study as we did not implement an experimental design.

Did the study involve field work? ☐ Yes ☒ No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study	
☒	☐	Antibodies
☒	☐	Eukaryotic cell lines
☒	☐	Palaeontology and archaeology
☒	☐	Animals and other organisms
☒	☐	Clinical data
☒	☐	Dual use research of concern

Methods

n/a	Involved in the study	
☒	☐	ChIP-seq
☒	☐	Flow cytometry
☒	☐	MRI-based neuroimaging

TYPE Original Research
PUBLISHED 21 March 2023
DOI 10.3389/ffgc.2023.958879

Check for updates

OPEN ACCESS

EDITED BY
Suzi Kerr,
Environmental Defense Fund, United States

REVIEWED BY
Luis Diaz-Balteiro,
Polytechnic University of Madrid, Spain
Ardalan Daryaei,
Tarbiat Modares University, Iran
Giorgio Vacchiano,
University of Milan, Italy

*CORRESPONDENCE
Barbara K. Haya
✉ bhaya@berkeley.edu

SPECIALTY SECTION
This article was submitted to Forest Management, a section of the journal Frontiers in Forests and Global Change

RECEIVED 01 June 2022
ACCEPTED 16 January 2023
PUBLISHED 21 March 2023

CITATION
Haya BK, Evans S, Brown L, Bukoski J, Butsic V, Cabiyo B, Jacobson R, Kerr A, Potts M and Sanchez DL (2023) Comprehensive review of carbon quantification by improved forest management offset protocols. *Front. For. Glob. Change* 6:958879. doi: 10.3389/ffgc.2023.958879

Comprehensive review of carbon quantification by improved forest management offset protocols

Barbara K. Haya[1,2]*, Samuel Evans[2,3], Letty Brown[2,4], Jacob Bukoski[2,5,6], Van Butsic[2,3], Bodie Cabiyo[2,7], Rory Jacobson[2,8], Amber Kerr[2,7], Matthew Potts[2,3] and Daniel L. Sanchez[2,3]

[1]Goldman School of Public Policy, University of California, Berkeley, Berkeley, CA, United States, [2]Carbon Direct Inc., Science Advisory Team, New York, NY, United States, [3]Department of Environmental Science, Policy, and Management, University of California, Berkeley, Berkeley, CA, United States, [4]Department of Integrative Biology, University of California, Berkeley, Berkeley, CA, United States, [5]Conservation International, Arlington, VA, United States, [6]Department of Forest Ecosystems and Society, Oregon State University, Corvallis, OR, United States, [7]Energy and Resources Group, University of California, Berkeley, Berkeley, CA, United States, [8]Yale School of the Environment, Yale University, New Haven, CT, United States

Improved forest management (IFM) has the potential to remove and store large quantities of carbon from the atmosphere. Around the world, 293 IFM offset projects have produced 11% of offset credits by voluntary offset registries to date, channeling substantial climate mitigation funds into forest management projects. This paper summarizes the state of the scientific literature for key carbon offset quality criteria—additionality, baselines, leakage, durability, and forest carbon accounting—and discusses how well currently used IFM protocols align with this literature. Our analysis identifies important areas where the protocols deviate from scientific understanding related to baselines, leakage, risk of reversal, and the accounting of carbon in forests and harvested wood products, risking significant over-estimation of carbon offset credits. We recommend specific improvements to the protocols that would likely result in more accurate estimates of program impact, and identify areas in need of more research. Most importantly, more conservative baselines can substantially reduce, but not resolve, over-crediting risk from multiple factors.

KEYWORDS
improved forest management, IFM, offsets, offset protocols, offset methodologies, forest carbon accounting, offset quality criteria

1. Introduction

Forests play a critical role in meeting greenhouse gas mitigation objectives with their potential to store large quantities of carbon and to act as an ongoing sink removing carbon from the atmosphere (Griscom et al., 2017; Fargione et al., 2018; Austin et al., 2020). Forest climate change mitigation activities generally fall into three broad categories: conserving existing forests; increasing forest extent through reforestation, afforestation, and agroforestry; and changing the management of existing forests to increase carbon in forests and forest products (improved forest management—IFM). Opportunities for increasing carbon sinks generally fall within the latter two categories, while forest conservation is focused on protecting existing forest carbon storage. Forest carbon activities can also have a range of ecosystem and societal co-benefits, including maintaining and enhancing biodiversity and providing forest products (Kremen and Merenlender, 2018; Asbeck et al., 2021).

Around the world, 293 carbon offset projects to date have channeled substantial carbon funding into improved forest management (So et al., 2023). Offsets are seen as a critical source of funds for IFM and an important alternative mitigation option to high-cost and hard-to-abate sources of emissions. This paper examines how well currently used IFM carbon offset protocols align with the scientific literature on carbon accounting, forest management, and land use change and how they can be amended to more accurately estimate program carbon benefits.

Studies suggest that IFM has the potential to increase carbon stocks by 0.2–2.1 Gt CO_2e/year globally (Griscom et al., 2017; Roe et al., 2019; Austin et al., 2020) without compromising the fiber and ecosystem co-benefits provided by managed forestlands. IFM includes a broad range of practices that increase carbon in forests and forest products (see Ontl et al., 2020; Ameray et al., 2021; Kaarakka et al., 2021 for detailed reviews of the range of IFM practices). For example, extending rotations can increase carbon stored on the landscape with continued or increased timber production for forests managed below maximum productivity (Sohngen and Brown, 2008; Foley et al., 2009; Nunery and Keeton, 2010). Reduced-impact logging in tropical forests can reduce forest degradation and increase or preserve soil carbon stocks, making forestry more sustainable and the conversion to agriculture less likely (Sasaki et al., 2016; Nabuurs et al., 2017; Ellis et al., 2019). Improved forest management can also make forests less susceptible to future carbon reversals from wildfire, drought, and pests (Anderegg et al., 2020).

In regulatory and voluntary carbon offset markets, carbon registries establish offset protocols that define project eligibility criteria and methods for monitoring and calculating the carbon impacts of each participating project. The registries also require third-party verification and issue offset credits. Each offset credit should represent one metric ton of carbon dioxide (tCO_2) emissions reduced or removed from the atmosphere. The protocols set the standard for the quality of the carbon offsets and their design allocates carbon financing toward eligible project types. Offset quality—the degree to which offset credits represent real emissions reductions and removals—is determined by protocol rules around additionality (would the project activities have occurred without the offset income?), counterfactual baselines (what would have happened without the offset income?), leakage (does the project cause increased emissions outside of project accounting boundaries?), durability (is the risk that stored carbon will be released back into the atmosphere managed and accounted for?), and carbon accounting (are the methods for monitoring and calculating carbon stocks, fluxes, and process emissions accurate and conservative?).

Peer-reviewed and non-peer-reviewed studies of IFM offset projects and protocols have shown evidence of over-crediting and non-conservative methodological rules. Studies of the California Air Resources Board (ARB) forest offset protocol found that the protocol is likely to significantly over-generate credits due to its methods for assessing project baselines (Badgley et al., 2022b; Coffield et al., 2022), leakage (Haya, 2019), and risk of reversal (Anderegg et al., 2020; Badgley et al., 2022a), as well as to create incentives counter to long term carbon stability in fire-prone areas (Herbert et al., 2022). Several peer reviewed and investigative case study analyses of projects using different IFM protocols identified substantial over-crediting (van Kooten et al., 2015; Elgin, 2020; Koberstein and Applegate, 2021).

Offset quality is essential for four main reasons. First, polluters often purchase offsets instead of directly reducing their own emissions. When used this way, offsets do not reduce emissions but rather trade where emissions reductions occur. When more offsets are generated than the program's actual climate benefits, they can reduce overall climate action. Second, when forest carbon is used to offset fossil fuel or other greenhouse gas emissions, offsets trade a known quantity of emissions with a much less certain and less durable quantity of reductions or removals (Haya, 2010; Haya et al., 2020). Third, the protocols send investment signals into the offset market sectors. If protocols result in over-crediting, climate mitigation funds will be over-allocated into less valuable activities. Fourth, over-crediting also creates a credibility problem for the offset market as a whole, undermining its ability to continue to direct private funds into effective climate mitigation. It is therefore critical that IFM protocols reflect current science and conservatively account for uncertainties.

To our knowledge, no study has yet comprehensively compared IFM offset protocols to the science of carbon accounting, forest management, and land use change to assess offset quality at the protocol level. The objective of this study is to qualitatively compare the IFM offset protocols against the scientific literature on quantifying IFM carbon impacts, with a particular focus on additionality, baselines, leakage, durability, and forest carbon accounting. Each section and our concluding discussion describe specific ways that the protocols can be improved to avoid over-crediting and to effectively support improved forest management practices that increase carbon storage in existing forests.

1.1. Background

Three voluntary offset market registries have generated the vast majority of IFM offset credits globally to date—American Carbon Registry (ACR), Climate Action Reserve (CAR), and Verified Carbon Standard (VCS). Each has offset protocols generating credits for voluntary use. All three also act as registries for the California Air Resources Board (ARB) offset program, hosting ARB-approved offset protocols and managing the monitoring, reporting, and verification processes for offset credits that can be used by California emitters to meet the state's cap-and-trade emissions targets.

Most IFM protocols were developed by interested stakeholders, including project developers, before the registry put them through a public vetting process. A list of the protocols reviewed for this study, along with the number of projects and credits issued by each, is shown in Table 1. We reviewed all IFM protocols with credits issued on voluntary market registries as of March 2022. While this analysis focuses on voluntary offset registries, governments also issue tradable credits from improved forest management projects, such as the UK Woodland Carbon Code.

Forest projects accounted for 30% of the total offset credits issued by voluntary registries in 2022 (Figure 1, top panel; So et al., 2023), mostly from REDD+ [Reducing Emissions from Deforestation and (forest) Degradation], which is the primary type of avoided deforestation offset (21% of 2022 credits), IFM (6%), and afforestation/reforestation (3%). IFM projects have generated 193 million offset credits since the first credits were issued in 2008. This represents 28% of the total forest-based offset credits and 11% of all

TABLE 1 IFM protocols reviewed.

Registry	Protocol	Number of projects	Credits issued to date	Countries
ARB	U.S. Forest Protocol	127	154,782,386	U.S.
ACR	IFM on Non-federal U.S. Forestlands	44	12,057,942	U.S.
CAR	CAR-U.S. Forest Protocol	29	13,549,474	U.S.
	CAR-Mexico Forest Protocol	90	1,099,403	Mexico
VCS	Conversion from Logged to Protected Forest (VM0010)	13	5,871,632	Australia, Canada, China, Malaysia, Romania, U.S
	IFM in Temperate and Boreal Forests (VM0012)	4	4,397,168	
	Rotation Extension (VM0003)	3	384,492	
	Conversion of Low Productivity to High Productivity Forest (VM0005)	1	509,540	
Total unique projects (counting projects that switch registries only once)		293	192,652,037	

Number of IFM projects that have been issued credits, credits issued, and countries hosting projects under each protocol, current through the end of 2022.

offset credits generated. While 293 IFM projects in seven countries have been issued offset credits, nearly all issued credits (94%) were in the United States and most (80%) are registered under the ARB compliance offset protocol (Figure 1, lower panel). Further, IFM projects generated close to half of all offset credits from projects in the United States.

To date, most IFM offset credits across all registries have been generated for reducing forest carbon losses by significantly reducing harvesting compared to the chosen baseline scenarios. While some projects support the types of activities highlighted in the literature as having high IFM potential—e.g., improving forest health for greater productivity and resilience, extended timber rotations, and reduced impact logging—so far the majority of credits are from activities that more resemble conservation and avoided degradation than IFM.

All protocols assess project impacts and the number of credits generated as the difference in carbon emissions and removals in the baseline scenario compared to actual levels. As relevant to the particular type of activity, all protocols take into account the major sources of carbon emissions and sinks affected by IFM projects—onsite carbon loss from logging and forest treatments, forest growth, process emissions (e.g., from equipment), and carbon held in harvested wood products. All protocols include procedures for reducing credits generated by an uncertainty deduction, and all set a proportion of credits aside in an insurance buffer pool which can be used to cover reversals such as from fire. Projects that reduce harvesting compared to the baseline also account for estimated displacement of timber harvesting to other lands (leakage). These carbon accounting factors are all discussed in the following section.

2. Review of quality criteria

2.1. Additionality and baselines

A project's baseline represents land management that most likely would have occurred in the absence of the offset program and is the scenario against which a project's carbon impact is measured. The "true" baseline (counterfactual) is inherently uncertain, because once a project takes place, the baseline cannot be observed. Baseline choice has a large effect on the number of credits issued, so baseline credibility and conservativeness are important to the quality of offset credits (Griscom et al., 2009).

For IFM projects, it is hard to distinguish additionality from baselines. Unlike most types of offset projects that involve a single action in time, such as building a landfill gas capture system, IFM involves a change in practice over the project lifetime. Additionality (would the project activities have occurred without the offset income?) and baselines (what would have happened without the offset income?) are closely related questions. ARB and CAR protocols combine them and treat all divergence from the baseline as additional, while ACR and VCS use separate baseline and additionality assessments (Table 2).

2.1.1. Summary of literature on IFM offset project baselines

Badgley et al. (2022b) documented that most ARB projects define their baseline at, or very close to, the minimum level allowed by the protocol. For most projects the minimum allowed baseline is the regional average carbon stock density for the forest type. Badgley et al. found that many participating projects are composed of species with greater carbon stocks than the regional/forest type average as defined by the protocol. Because carbon stocks often change gradually over space but the minimum baseline is defined regionally, there is a strong incentive to enroll lands with naturally higher carbon stocks than the regional average. Badgley et al. estimated that this has led to over-crediting of close to 30% across the study's projects compared to what would have been credited if a more refined method was used to determine the minimum allowed baseline. Coffield et al. (2022) used remote sensing-based datasets to compare the outcomes of 37 California-based ARB IFM offset projects with similar "control" lands. They found lack of evidence that the offset program influenced land management and therefore lack of project additionality. van

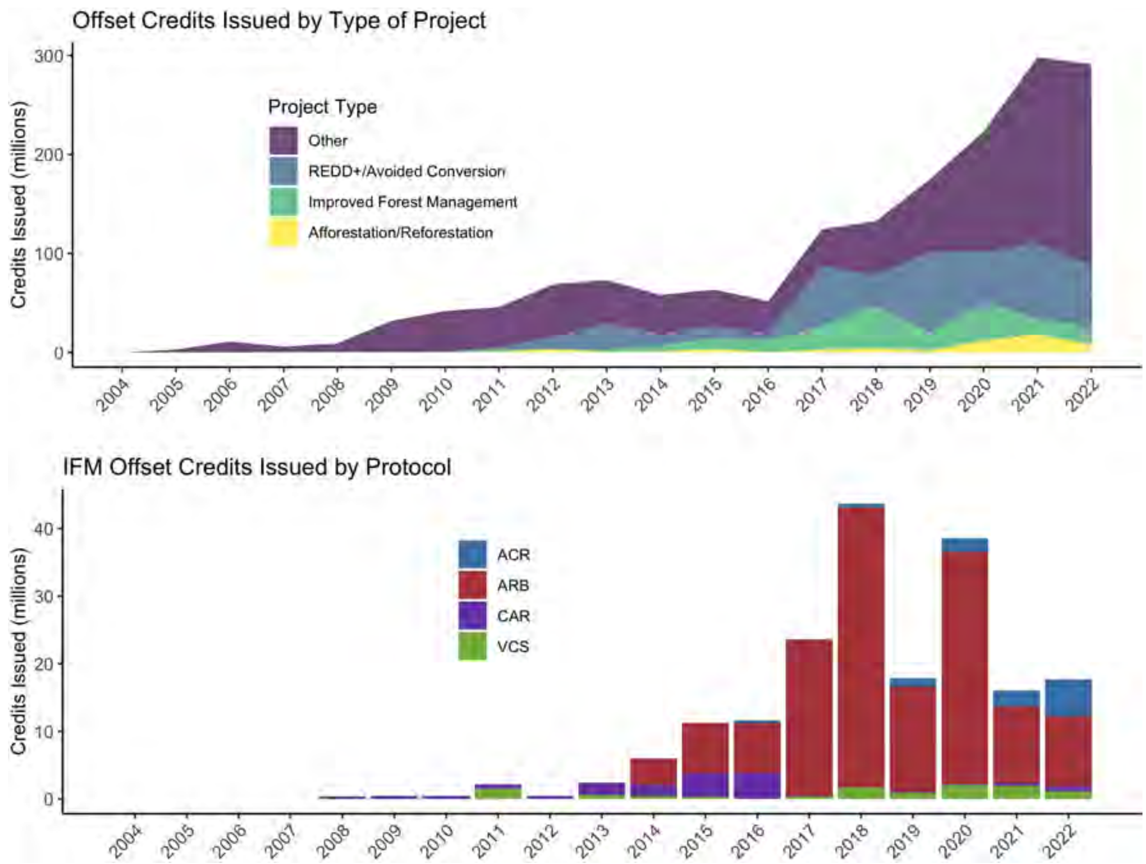


FIGURE 1
Current trends in forest-based carbon offset markets (based on data from So et al., 2023). The **upper** panel shows the breakdown of credits issued by project type for forest and non-forest carbon offset projects. The **lower** panel shows the trend in IFM credit issuances by program/registry.

Kooten et al. (2015) investigated a large VCS IFM project in British Columbia that assumed a "lumber liquidator" counterfactual—that an alternative forest owner would have aggressively logged the forest. van Kooten et al. found that in this case, the chosen baseline created substantially more carbon credits than would have been generated if a more likely sustainable management scenario was used as the baseline.

Qualitative research also has consistently identified problem areas in baseline setting. Several studies identified asymmetric information as a pervasive, inherent problem in baseline setting for IFM projects. Asymmetric information creates uncertainty for the program administrator and third-party verifier but not the project developer, who implements a project with full information (van Kooten et al., 2009; Asante and Armstrong, 2016; Gren and Aklilu, 2016). For example, one study highlighted the trend of pulp timberland acquisitions by real estate investment trusts (REITs) and timber investment management organizations (TIMOs), who aggressively harvest and then sell the land to carbon project developers (Gifford, 2020). The project developers can report a low baseline carbon stocking as a result of the recent harvesting. This is an example of how a complex management history and asymmetric information make accurate baseline-setting difficult.

One study documented how program administrators deflated baselines in order to reduce barriers to entry in IFM projects. The study quoted one project developer stating that "if baselines are set too high, many potential projects will not be viable for participation" (Ruseva et al., 2017). To the best of our knowledge, Anderson et al. (2017) is alone in finding "strong evidence of additionality" of projects under ARB's IFM protocol and suggests that baseline/additionality criteria may be too strict and may impede projects with "multiple desirable features." However, an expanded discussion by the authors suggested that they based their assessment on their observation that some rather than all projects are likely to be additional (Anderson and Perkins, 2017). Their survey of landowners with IFM projects showed that 5 of 17 (29%) self-reported that they were either not confident or unsure whether the offset credits generated by their projects "represent additional carbon sequestration that would not have happened without the forest offset program."

2.1.2. Description of the protocols

2.1.2.1. ARB and CAR-U.S. protocols

ARB and CAR-U.S. protocols define the baseline as the average onsite carbon stocks over a modeled 100-year baseline management scenario that should be no lower than the minimum baseline level allowed (Figure 2). Typical baselines are set at around 30% below initial carbon stocks (calculated from Badgley et al., 2021), and just above common practice (Badgley et al., 2022b). The ARB and CAR-U.S. protocols only require that the baseline scenario is financially

TABLE 2 Overview of how IFM protocols treat baselines and additionality.

Registry	Baseline setting	Additionality
ARB/CAR-U.S.	100-year baseline model - Aligned with legal and other obligations - Must be financially feasible - Not lower than common practice if initial stocks are above common practice - Otherwise, typically at initial stocks Standardized approach to additionality—any forest carbon above the baseline is considered additional.	
CAR-Mexico	Initial carbon stocks standardized approach to additionality—any forest carbon above initial carbon stocks is considered additional.	
ACR	Economic baseline: assume harvest to the level that maximizes net present value (NPV) over many rotations	Project-by-project: - financial barriers, - exceed common practice, - exceed regulation
VCS	Different baseline approaches (e.g., NPV and historical management)	Project-by-project: - not most financially beneficial option or experience other barriers, - exceed common practice

feasible and complies with all legal and contractual requirements. Further, the chosen baseline scenario does not need to be shown to be the most feasible or likely without offsets.

Setting the baseline below initial or historic carbon stocks raises an over-crediting concern. Instead of being credited for taking action, the forest owner is credited for not taking action that would have reduced the carbon stocks on their lands. In other words, the assumption is that in the absence of offset payments, the land owner would change their management practice in a way that releases carbon. Non-additional crediting has arguably been the most significant quality challenge for carbon offsets generally (Cames et al., 2016; Haya et al., 2020). For the majority of IFM projects with baselines below historical levels, additionality assessment is even more challenging because it is being tested for not taking an action.

In addition, timing of credit generation against the baseline is another quality concern for the majority of these projects. Although baselines are derived from modeled scenarios that are intended to represent realistic harvesting over time (decreasing solid orange line in Figure 2), in the 1st year of the project, project credits are issued against the 100-year-average baseline, which usually represents a sharp, unlikely drop from initial carbon stocks (flat dotted orange line in Figure 2). Thus, even in cases where the baseline is an accurate reflection of the true without-offsets scenario over decades, a large proportion of credits are generated in the 1st year of the project for reductions that will actually take place over a much longer period of time. In effect, this means that future reductions can be used to offset current emissions.

2.1.2.2. CAR-Mexico protocol

By using ton-year accounting, the CAR-Mexico protocol is structured differently from all other protocols discussed in this paper. Under this approach, the project developers decide on the length of time they commit to maintaining credited carbon stocks, ranging from one to 100 years. A chosen term of 100 years earns full credits without discounting. Any shorter commitment earns a fraction of the calculated carbon impact such that a 1-year commitment earns 1% of the calculated carbon benefits, and a term of 50 years earns 50%.

Using initial carbon stocks as the baseline is more conservative than other protocols and reduces over-crediting risk. However, flexibility in the term of the commitment increases risk of non-additional crediting. For example, terms that match rotation lengths can potentially earn offset credits without any change to harvest schedules.

2.1.2.3. ACR protocol

The ACR protocol uses net present value (NPV) to set the baseline. Project baselines are typically set to a 20-year crediting period and based on a 100-year NPV-maximizing harvest schedule. In general, the approach of setting the baseline as the scenario that maximizes NPV is sound for landowners who seek to maximize profit over a long term, like industrial forest owners who have access to reliable markets. However, this method may poorly predict the management decisions of other landowners who may manage for multiple goals like ecosystem or recreation benefits (Butler et al., 2016). Even where landowners wish to maximize long-term profit alone, irregular market demand may push them to shift their management away from what a simple NPV analysis would predict (Keegan et al., 2011). For example, small plantation owners in the U.S. Southeast currently have limited access to wood markets and, as a result, have older trees, on average, than is economically optimal (Grove et al., 2020). In addition, NPV calculations are based on internal costs, which can be difficult for verifiers to verify.

2.1.2.4. VCS protocols

The VCS IFM protocols use multiple approaches to baseline-setting, including historical baselines, legal baselines, common practice baselines, and baselines based on documented management activities. Therefore, there are multiple pathways for establishing a baseline within a single protocol, and these approaches can be applied with more or less rigor. Such flexibility is logical given the diversity of lands that might seek to enroll. However, they also allow project developers to pick the most advantageous baseline, which may lead to over-crediting. Such flexibility means potential offset credit buyers must conduct enhanced diligence to determine how appropriate the chosen baseline is.

VCS uses two additionality tools for its forestry projects which both closely mirror the Clean Development Mechanism (CDM) approach to additionality testing. Landowners must demonstrate that the project is not the most cost-effective land management approach or that other barriers would have prevented the landowner from carrying out the land management credited under the offset project. The land owner must also demonstrate that the credited land management approach is not common practice. In general, these tests have proven to be insufficient in ensuring the additionality of CDM projects (Haya, 2010; Cames et al., 2016), leaving additionality to be assessed primarily with baseline-setting as with the ARB and CAR-U.S. protocols.

2.1.3. Persistent issues and baseline recommendations

Where good data on forest harvest exists, baseline uncertainty can potentially be reduced and conservativeness increased by developing

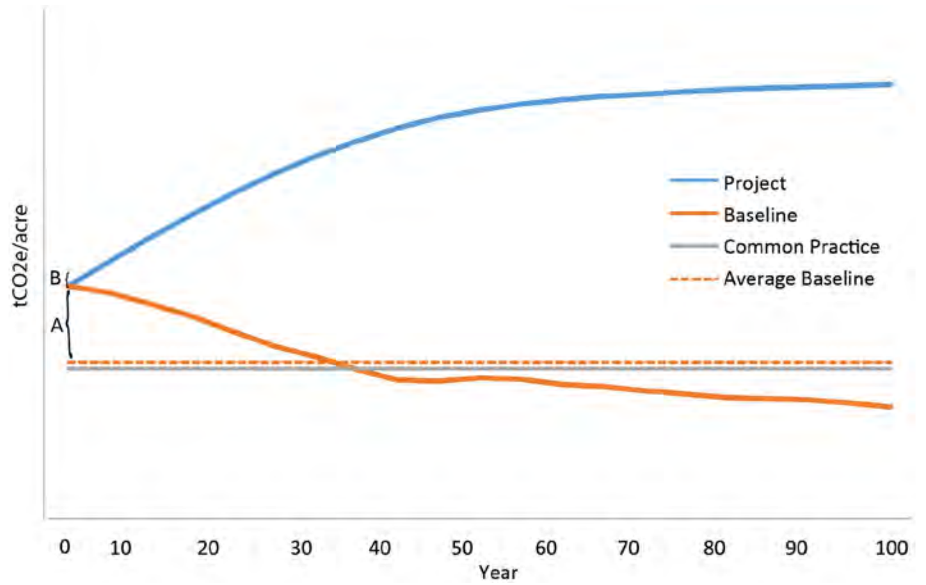


FIGURE 2
A sample ARB project and baseline scenario based on a project in Oregon (ACR 260). The pattern shown is similar in many other IFM offset projects. "A" represents the credits generated in the 1st year of the project from the difference in actual onsite carbon stocks compared to the 100-year-average baseline. "B" represents the credits generated in years 2 through 5 of the project from forest growth.

baselines on historical practice, initial carbon stocks, similar lands with "dynamic" baselines, and NPV for landowners where NPV is reasonably predictive with some restrictions.

When NPV is used as the baseline, project developers should describe their capacity to harvest at this level and also the market conditions and mill capacity to absorb this harvest. Project developers wishing to use NPV can justify their case by demonstrating that they have a strong history of harvesting on similar lands, or better yet, can demonstrate a history of NPV harvesting on that project property. For projects that cannot demonstrate NPV-type harvest schedules, NPV is likely inappropriate.

Baselines that reflect current carbon stocking of the participating parcel are usually more conservative than broad regional averages. Such baselines only credit removals through growth.

When past management actions are used as baselines, statistical land use models can be used to provide quantitative estimates on the likelihood of harvest given a project's characteristics (Lewis, 2010). Such models can be used to create credible baselines and importantly, these models can be used to simulate alternative baselines which might reflect different market conditions (Radeloff et al., 2012).

The use of dynamic baselines is similar to control plots in experimental science. In this system, properties similar to the offset property in past management, market conditions, ecosystem, landowner type, etc., can be used as the baseline for offset projects. Matching methods developed for causal inference can be used to create comparison sets (Andam et al., 2008; Ferraro and Hanauer, 2014). Each year, the carbon values of the offset and the baseline properties can be compared, and credits can be issued on the basis of this comparison.

An advantage of dynamic baselines is that by observing similar properties in each year, changing market conditions can be integrated into baselines. For example, consider an offset in an area where mill capacity falls dramatically. Under static baselines, the offset would continue to generate credits, even though in reality there may be no market for timber in the area. Conversely, if a new technology increases the profit of harvesting, more credits could be granted. Dynamic baselines solve this problem by accurately reflecting baseline conditions relative to the project in pre-defined time periods. Such baselines might be particularly useful in areas where markets are in rapid flux, where forest managers cannot show that they have historically managed for NPV, or where land use is rapidly changing.

With all of these options, adverse selection might still lead to over-crediting. Because landowners or project developers will always know more than registries and verifiers about what would have happened without the offset income, adverse selection is a persistent issue. Statistically, adverse selection can be thought of as an unobserved variable that is correlated with the treatment decision (project enrollment) and the outcome (forest harvest). If this unobserved variable is correlated with increased enrollment and decreased forest harvest, the baseline is an overestimate of the true counterfactual. For example, this might be the case where a landowner has a strong conservation ethic and prefers to preserve rather than cut down their trees. A case like this can lead to over-crediting, because such a landowner is unlikely to harvest, even in the absence of the program.

Using historical forest harvest data can help to control conditions that lead to adverse selection, especially if these conditions do not change over time. For example, in the case of a conservation-minded landowner, if they have held similar preferences in the past, a baseline that takes into account their historical harvest levels would not over-credit (even though we cannot measure their land management philosophy). At the same time, a baseline based on regional averages or NPV alone would likely over-credit.

However, while historical baselines can help to account for unobserved variables that do not change over time, they cannot account for cases where the unobserved variable is not static. An example of this could be when a property is inherited or purchased by a new landowner. The application of a historical baseline for a property that had been harvested, but was purchased by a conservation NGO and then later enrolled in an offset program could lead to over-crediting because the true counterfactual for the new landowner is different than from the past landowner.

Dynamic baselines cannot directly account for the problem of adverse selection. To the extent that similar properties also have similar unobserved variables, then matching may reduce the impact of adverse selection. However, there is limited empirical evidence for this. Indeed, using nearby non-enrolled parcels as "control plots" could actually increase the effect of unobserved variables: if some parcels enroll and others do not, then it may precisely be an unobserved variable that is influencing this self-selection, biasing the dynamic baseline in favor of over-crediting.

2.2. Leakage

The Intergovernmental Panel on Climate Change IPCC (2007) defines leakage as "the unanticipated increase or decrease in greenhouse gas (GHG) benefits outside of the project's accounting boundary as a result of the project activities." Three types of leakage are relevant for forest-based offset programs: activity leakage, output market leakage, and land market leakage (Meyfroidt et al., 2020). The latter two types of leakage are collectively referred to as market leakage. Activity leakage occurs when mobile factors of production (labor and capital) are no longer needed in the offset program area and are reallocated to similar activities outside of the program area. Output market leakage occurs when changes in harvesting inside the project area affect timber prices and change harvesting outside the project area by non-participating forest managers. Land market leakage occurs when changes in timber harvesting on offset project lands changes the value of timber land relative to other land uses and provides incentives for land conversion into managed timber land or from timber land into other uses.

There is no broad agreement on how offset registries should incorporate leakage into their IFM protocols. The approach taken by the protocols is to deduct credits from a project based on a specified leakage rate. The protocols differ in the leakage rate applied, when and how it is applied, and whether the protocols account for activity leakage explicitly. Each of these aspects of leakage is discussed below and summarized in Table 3.

2.2.1. Market leakage rate

All protocols have a mechanism for deducting leakage when timber harvesting is lower in a project relative to the baseline. All protocols use a leakage rate that reflects the assumed percent of onsite carbon loss (or gain) from a change in timber harvesting due to the offset projects that are lost (or gained) in other forests to which the harvesting is displaced.

ACR applies a 10% leakage rate if the project reduces harvesting by 5–25% compared to the baseline, and 40% if reduction in harvesting is more than 25% compared to the baseline. In the ARB, CAR-U.S., and CAR-Mexico protocols, leakage is deducted at a constant rate of 20%. Leakage rates used by all of the VCS protocols reviewed vary based on the carbon density, defined as the ratio of merchantable biomass to total biomass, of the forests where the displaced harvesting is assumed to occur compared to the forest enrolled in the carbon project. If harvesting is expected to shift to a forest with a ratio of merchantable biomass more than 15% lower than the project forest, a higher leakage rate (70%) is applied; if the destination forest produces more than 15% more merchantable biomass, relative to the project forest, a lower leakage rate applies (20%); if displacement occurs in a similar forest type, a 40% leakage rate is applied. VCS's extended rotation protocol (VM0003) also prescribes a 10% leakage rate if the rotation extension is <10 years and the harvest reduction over this time frame is <25%. VCS protocols exclude international leakage from their deduction formulas and allow for project-specific justifications for the application of a 0% leakage rate.

The academic literature has estimated forest carbon leakage using two general methods. Partial and general equilibrium models are complex optimization models based on economic theory of how markets function and calibrated to real-world data. Behavioral parameters, such as supply and demand elasticities, are drawn from the economic literature. These models are designed to capture the interconnectedness of different markets. General equilibrium models capture all economic flows within an economy, while partial equilibrium models usually focus in more detail on a subsection of the overall economy. Equilibrium models are generally used for ex-ante economic and policy analysis. Causal econometric models, which are an ex-post evaluation methodology that use statistical techniques to evaluate programs, have been utilized to assign causal attribution to leakage from other project types (e.g., Roopsind et al., 2019), but not IFM programs or projects. Challenges in applying causal inference methods to IFM include difficulty in observing a plausible harvesting counterfactual and the challenge of isolating program effects when so many IFM offset programs are currently being implemented with different rules.

Studies estimating leakage rates from reducing harvest activities have found a wide range of plausible leakage rates depending on different locations, spatial scales, time horizons, and methodological approaches. Some studies focused on national IFM programs (primarily in the United States), while others focused on global estimates. Studies in the United States context showed that leakage rates are generally higher than those commonly used in the protocols. In an econometric study of the effects of an 85% reduction in harvest on public lands in the Pacific Northwest of the United States during the 1990's, Wear and Murray (2004) found substantial evidence of output market leakage as softwood lumber prices increased by 15%. They estimated that nearly 84% of the timber harvest restriction shifted to unrestricted areas. Of that 84% leakage, they found that 43% in the region, 15% in other U.S. markets, and an additional 26% in Canadian markets. Using a general equilibrium model, Gan and McCarl (2007) estimated leakage rates from U.S. forest offset programs to be in the 75–78% range, including both domestic and global leakage.

One challenge in applying rates from the published literature to the protocols is that most, rather than quantifying leakage in units of carbon, estimate leakage of another metric like harvested

TABLE 3 Summary of leakage treatment in IFM protocols.

Protocol	Market leakage rate	Leakage timing	Can leakage be positive?	How leakage is deducted?	Accounts for market leakage?	Monitors for activity leakage?
ACR	0, 10, and 40%	Consistent	No	% reduction in total credits issued	Yes	Yes, landowner must demonstrate that there is no activity leakage beyond de minimis levels
ARB	20%	Inconsistent	No	% of difference between project and baseline onsite carbon harvested	Yes	No
CAR-U.S.	20%[a]	Inconsistent	Yes, but only to earn back past leakage deducted		Yes	No
CAR-Mexico	20%	Consistent			Yes	No
VCS—VM0010	0, 20, 40, and 70%	Consistent	No	% of net emissions from harvesting in the baseline	Yes	Yes, landowner must demonstrate that there is no activity leakage
VCS—VM0012[b]	0, 20, 40, and 70%	Consistent	No	% of difference between project and baseline emissions	Yes	
VCS—VM0003	0 and 10% for short extensions, 0, 20, 40, and 70% otherwise	Consistent	No		Yes	
VCS—VM0005	0, 20, 40, and 70%	Consistent	No	% of net carbon stock change in the baseline	Yes	No

[a]Except version 3.0 used a sliding 0–80% scale depending on how much harvesting was reduced compared to the baseline.
[b]VM0012 also allows projects in North America to apply the CAR leakage deduction (20%). Based on leakage formula in Murray et al. (2004).

wood products (Wear and Murray, 2004) or economic welfare (Gan and McCarl, 2007). Murray et al. (2004) and Murray et al. (2005) applied modeling frameworks that estimate carbon leakage directly. Murray et al. (2004) showed that domestic leakage rates (ignoring international leakage and focused on carbon instead of timber) for forest offset set-aside programs in the United States can vary from 16 to 68% depending on where the offset occurs in the country and carbon density of the protected forest. Murray et al. (2005) also conducted extensive carbon leakage analysis of forest sector carbon programs but did not focus explicitly on improved forest management is the focus of the protocols reviewed here.

Sun and Sohngen (2009) used a global economic optimization model and found that set-aside programs applied globally, which permanently reduced the land available for forest harvest, resulted in leakage rates of 47–52%, depending on the specific land taken out of production. Several studies in countries other than the United States showed significant variation in IFM leakage rates. Kallio and Solberg (2018) estimated leakage rates of 60–100% from harvest reduction projects in Norway. While the model had a relatively limited temporal and carbon accounting framework, it found that the variation in leakage rates is driven by the degree of harvest reduction, the type of forest product considered (e.g., pulpwood vs. sawlogs), and the forest product supply elasticity. By contrast, Sohngen and Brown (2004), estimated leakage rates of 2–38% for a Bolivian forest set-aside program. The country-to-country differences were likely driven by the country's integration into global wood product markets.

Based on findings from the literature and factors identified in Murray et al. (2004), leakage risk is likely to be highest in tight timber markets with responsive supply and in regions where non-participating land can produce similar timber products. One important caveat is that the economic equilibrium models used in the academic literature assumed that all actors have perfect information and as a result may slightly overestimate leakage risk in practice when markets are slower to adjust. More research is needed to update and refine understanding of leakage in IFM carbon projects. One particularly important area of future research is in leakage from short-term harvest deferrals.

2.2.2. Activity leakage

There is variation in how the protocols consider market vs. activity leakage. CAR and ARB do not distinguish between market and activity leakage; any activity leakage is effectively included in the 20% market leakage rate. ACR and VCS monitor activity leakage separately. Under both of these registries, if production declines by more than 5% relative to the baseline, the landowner must demonstrate that no leakage occurs on other lands they manage or operate outside of the offset project. Landowners can demonstrate that no activity leakage occurs with historical harvesting records, or forest management plans prepared at least 2 years prior to the start of the project showing no change in harvesting on non-project lands with the implementation of the offset project. ACR includes a third option where landowners can demonstrate that they are not engaging in activity leakage if all lands owned by the landowner are certified as sustainable, such as by the Forest Stewardship Council (FSC).

These requirements prevent the most flagrant violations of activity leakage, but there are plausible cases when activity leakage might still occur. For example, a landowner could write a forest management plan with increased levels of harvesting and then enroll part of their lands in a carbon project 2 years later. As another example, FSC certification does not prevent any increase in harvesting, and thus activity leakage could easily occur on FSC-certified land. On the other hand, cumbersome activity leakage rules may prevent timberland owners from being able to enroll portions of their forest holdings as carbon projects due to the inability to manage unenrolled lands in response to changing wood product markets.

2.2.3. Timing of the leakage deduction

In addition to market leakage rates, the timing of the leakage deduction can have large effects on the number of credits issued. Prior research found that the ARB and CAR-U.S. protocols tend to greatly over-credit at the start of each project, due to a timing mismatch in the construction of the baseline scenario (Haya, 2019; Haya and Stewart, 2019). Most ARB IFM projects start with carbon stocks far above estimated baseline levels; initial carbon stocks 40–50% higher than baseline levels are typical (Haya, 2019). This is based on the assumption that without the offset program, timber would be aggressively harvested, reducing onsite carbon stocks substantially. This initial onsite carbon above the 100-year-average baseline is credited in the first reporting period, promptly generating a large number of credits without requiring any change in land management.

However, the displacement of harvesting (leakage) associated with that large reduction in harvesting is not all deducted in the project's 1st year, but rather is deducted evenly over the 100-year life of the project. This results in over-crediting at the start of the project, which is gradually paid back over the project life. We are not aware of any academic literature that has examined the correct timing of harvest displacements in timber markets. A conservative approach would apply the leakage deduction in the year that harvest was assumed to occur in the baseline and is credited by the project. Haya (2019) estimated that this correction would reduce the number of credits generated by the ARB protocol by 35%, and if the correction were combined with a higher leakage rate of 40–80%, crediting would be reduced by 51–82%. Levels of over-crediting would be even higher if reversals were not adequately monitored and compensated for after the end of the final reporting period in which credits were issued (Haya, 2019). The CAR-Mexico, ACR, and VCS protocols do not have this timing issue.

Leakage can also result in positive carbon outcomes when the project increases timber harvesting, thus leading to less harvesting elsewhere. None of the protocols account for reverse leakage from increased harvesting compared to the baseline, which is a form of conservativeness built into the protocols. Only the CAR protocols allow for reverse leakage to be counted if cumulative leakage from the project start is positive. While accounting for leakage annually is more conservative, cumulative leakage accounting may create more incentive for forest owners to decrease harvesting temporarily and conduct thinning to enable increases in harvesting later from an older, better managed forest.

2.2.4. Recommendations on leakage

Leakage is a complex economic phenomenon that is both hard to quantify and likely varies considerably across many dimensions, including IFM project type, location, and supply and demand conditions. The risk of over-crediting due to leakage would be reduced considerably if baselines were set more conservatively as described above. More conservative baselines that involve no or little difference in harvesting compared to the project would involve lower estimates of leakage, and so uncertainty in the leakage rate would have less impact on the number of credits generated.

ARB and CAR-U.S. protocols, which attribute leakage evenly over 100 years, are likely to over-credit significantly in the 1st year of each project that chooses a baseline lower than initial carbon stocks (which is the case for most projects). This source of over-crediting can be easily removed if leakage were deducted at the same time that the onsite benefits of reducing harvest are credited.

Current literature does not provide much guidance on the appropriate leakage rate to apply in specific contexts. Generally, the literature supports higher leakage rates than are currently used, although there are only a few studies that are mostly decades old and based on national or global economic equilibrium models or statistical evidence from large policy changes. For projects that reduce harvesting permanently, a higher leakage rate than those used by current protocols would be conservative given the large uncertainties. However, there is a risk that large, immediate leakage deductions may discourage extended rotation projects with only temporary leakage risk. This may be partially remedied without over-crediting by assuming leakage plays out over several years. This would strike a balance between the ARB and CAR-U.S. protocols (which average baseline harvesting, and therefore leakage deductions, over 100 years), and ACR, VCS, and CAR-Mexico protocols (which deduct leakage immediately). In addition, assessing leakage cumulatively would better reflect the impact of projects that defer rather than reduce harvesting. Currently only the CAR protocols credit projects for reverse leakage when increased harvesting compared to the baseline is likely to cause less harvesting elsewhere. These credits can be earned if cumulative emissions from leakage over the project lifetime are still positive. Lastly, discretion for projects to choose the leakage rate, as offered by all VCS protocols reviewed, has the potential to lead to under-counting leakage impacts.

2.3. Durability

Carbon stored in ecosystems is inherently impermanent. Forest carbon can be released through natural occurrences like fire, drought, disease, and wind, and through human actions like harvesting and land use conversion. Protocols address these risks of reversal with commitments to maintain carbon storage over a designated period (the project term), incentives to design projects to reduce reversal risk, and recourse if reversals do occur.

The project term describes the length of time during which a project is contracted to maintain credited carbon stocks. Some protocols create incentives for forest management that reduces reversal risk. All registries host an insurance buffer pool to replace credits if a reversal does occur. Buffer pool contributions are designed to cover the calculated likelihood that those carbon stocks will be reversed, i.e., re-emitted to the atmosphere. Programs and projects vary widely across project term, risk of reversal, and reversal recourse.

The reviewed protocols have varied project terms that range from a year to a century (Table 4). The CAR-U.S. and ARB forest offset protocols have the longest project terms: 100 years from the date of credit issuance. By contrast, other protocols define the project term from the project start date rather than from the last credit issuance. For example, a VCS project with a term of 30 years may generate credits in year 20 that are only guaranteed for the remaining 10 years.

For large registries, buffer pools can be made up of a large, diverse pool of credits that offer significant risk mitigation for

TABLE 4 Durability terms and buffer pool contributions (± one standard deviation) across offset protocols.

Registry	Minimum term	Recourse
ACR	40 years*	23.5 ± 2% buffer pool, for both intentional and unintentional reversals
ARB	100 years	16.1 ± 2.8% buffer pool reversal risk assessment includes unintentional and intentional reversals; intentional reversals must be replaced with similar credits
CAR-U.S.	100 years	7.7 ± 2.6% buffer pool, intentional reversals must be replaced
CAR-Mexico	1 year	8% buffer pool, primarily for unintentional reversals but can be used at the discretion of CAR
VCS	20 years*	17.4 ± 11.4% buffer pool, for both intentional and unintentional reversals. Verra is the only registry that allows buffer pool credits to be returned to the salable credit pool as the risk of reversal within the project lifetime diminishes over time.

*From project start date, not the date credits are issued. Verra is considering extending the monitoring of reversals into the post-crediting period for compensation by the buffer pool.[1]

individual projects. Each protocol has a different approach to allocating buffer pool credits. Intentional reversals can include, for example, negligence on the part of the project developer or active harvesting. Unintentional reversals include natural reversals, like fire and disease, and human-caused reversals that are outside the control of the project operator. Notably, the ACR and VCS buffer pools can be used to cover both intentional and unintentional reversals, while ARB, CAR-U.S., and CAR-Mexico buffer pools can only be used to cover unintentional reversals. Under these protocols, intentional reversals must be replaced. VCS allows a portion of buffer pool credits to be returned to the salable credit pool if the risk of reversal within the project lifetime can be shown to decline over time.

2.3.1. Do the protocols adequately ensure durability?

Project terms are highly variable across protocols, but even the longest term (100 years) does not constitute a truly permanent offset equivalent to reducing fossil fuel emissions. Forest credits used to offset fossil fuel emissions convert carbon permanently stored as fossil fuels into carbon stored in trees in the short-term carbon cycle. If the end of a project term represents a reversal event, then non-permanent carbon storage (like all IFM projects) can more accurately be understood as delaying, not fully neutralizing, emissions (Herzog et al., 2003). Decisions about the appropriate duration of carbon storage fundamentally depend on assumptions about the future, and academics have called the default choice of 100 years "political" (Archer et al., 2009; Allen et al., 2016). In practice, project terms in IFM projects can range from 1 to 100 years, and there is not yet a widely adopted framework for comparing these different terms. Even taking for granted that these projects do not represent permanent offsets, questions remain about whether the current approach (relying on buffer pools) can achieve the promised durability.

Three key limitations of buffer pools could critically undermine their usefulness. First, none of the reviewed protocols take climate change into account in estimating buffer pool allocations and so may not reflect increasing risks of reversal over decadal time scales. For example, the ARB protocol for U.S.-based projects includes a buffer allocation of 2–4% for fire, 3% for biotic risks, and 3% for "other episodic catastrophic events" (e.g., drought). However, because annual acreage of forest fires in the United States is projected to quadruple by the end of the century even under a moderate emissions scenario (Anderegg et al., 2022), current buffer pool allocations may prove insufficient on the basis of wildfire risk alone. If recent wildfire trends continue in the United States, the entirety of the buffer pool for existing ARB projects will be consumed well before its intended lifetime is up (Badgley et al., 2022a). The ACR and VCS protocols have similarly low buffer allocations for natural disturbances, although no systematic assessment of these buffer pools have been conducted in the academic literature. A proposed VCS risk calculation tool may remedy this by using Climatic Impact Drivers (CIDs) to project increased risk.[2]

Second, some registries may not have a sufficiently diversified offset portfolio to effectively mitigate risk through the buffer pool mechanism. Such systemic risks may arise when a large proportion of projects in a registry are similar and/or exist in a constrained geographic area or ecological type. For example, the ARB compliance offset pool, which is composed mostly of IFM projects entirely in the United States (Badgley et al., 2022b), may be exposed to systemic forest risks that decrease the efficacy of the buffer pool as a risk mitigation tool.

Third, a buffer pool is defined by the quality of its constituent credits. Buffer pools composed of low quality credits have little value. Extensive work has shown systematic issues with additionality, baselines, leakage, and carbon accounting for land-based offset projects across protocols (e.g., Haya, 2019; West et al., 2020; Badgley et al., 2022b). Further, the ACR protocol allows project developers to put credits into the buffer pool from any ACR project (not just the project under consideration), which creates a perverse incentive to fill the buffer pool with low-value, potentially non-additional credits.

2.3.2. Recommendations on durability

Broadly, climate change is expected to push forest systems toward younger, shorter, less carbon-dense forests (McDowell et al., 2020). These future forests are expected to have higher rates of mortality due to climate-exacerbated disturbances, making the carbon they store less durable (Anderegg et al., 2020; McDowell et al., 2020). Many types of disturbances are expected to increase in both frequency and severity. Offset registries should incorporate these increasing risks into the rules defining buffer pool allocations. If possible, reversal risk should be defined in a spatially explicit way to reflect the fact that different types of risks vary tremendously depending on the location, species composition, and stand structure (Anderegg et al., 2020). Further, existing protocols give minimal incentive to reduce disturbance hazards and could be updated to more actively reward management activities like prescribed burning, species selection, and

1 https://verra.org/wp-content/uploads/2021/12/LTRMS-Public-Consultation.pdf

2 https://verra.org/wp-content/uploads/2019/02/Risk-Report-Calculation-Tool-Guidance_DRAFT_v0.1.pdf

thinning that increase resistance to reversals (Stephens et al., 2020; Herbert et al., 2022).

New time accounting frameworks have been proposed to clarify the value of shorter project terms. These fall into two broad categories: vertical and horizontal stacking of offset credits. Vertical stacking approaches, which include ton-year accounting like that used by the CAR-Mexico protocol, involve purchasing multiple short-term credits upfront to offset emitted CO_2. The multiple approaches to vertical stacking can have highly varied results depending on which assumptions are made (Levasseur et al., 2012; Groom and Venmans, 2022) and have been criticized for simply postponing climate impacts (Carton et al., 2021). Horizontal stacking, sometimes called offset rental or leasing, involves repeat purchasing of offset credits after they expire or after a reversal occurs (Herzog et al., 2003), which, if adequately enforced, could ameliorate some of the challenges of short durability terms.

2.4. Carbon accounting

Carbon accounting in the context of IFM protocols includes a variety of measurement and estimation techniques that attempt to accurately and precisely quantify carbon stocks in biomass and harvested wood products, as well as changes in these stocks that result from project activities (Table 5). Major sources of uncertainty in estimating onsite carbon stocks in the biomass pools fall into four categories: (i) accuracy of measurements in the field; (ii) choice of allometric models (including selection of wood density values and root:shoot ratios); (iii) sampling uncertainty related to plot size; and (iv) sampling uncertainty related to statistical representativeness of the plots within the whole landscape (Chave et al., 2004; Temesgen et al., 2015). For the soil and litter pools, substantial uncertainty exists around both the processes of organic carbon cycling, as well as accurately quantifying highly variable carbon stocks across space. Lastly, uncertainty surrounding carbon benefits from harvested wood products primarily relates to life cycle considerations, such as duration of use or potential climate benefits from product substitution.

All protocols include estimation of carbon stocks in aboveground and belowground biomass, with the exception of the VCS protocol for the Conversion of Logged to Protected Forests (VM0010), which presumes that root biomass is likely to remain constant or moderately increase. Typically, when a carbon pool is excluded from project-level carbon accounting, the decision is justified by an assumption that the change in the pool will be negligible under approved project activities, or will result in net carbon accumulation and thus can be excluded for conservative estimation. For example, in the context of the soil carbon pool, the stock is only estimated and included in project emissions to subtract losses from disruptive management activities or site preparation from a project's carbon benefit. Carbon pools with relatively smaller stocks compared to living tree biomass, such as standing or lying dead biomass or aboveground non-tree vegetation, are included or excluded on the basis of whether the activities eligible under the protocol are likely to have significant impacts on these stocks.

We discuss the protocol methods for estimating carbon in aboveground biomass, belowground biomass, soil carbon stocks, and harvested wood products in the following sections. Further, we identify several accounting practices that may be uncertain or yield systematic errors in carbon accounting.

2.4.1. Aboveground biomass

The protocols employ standardized approaches to measurement of aboveground carbon stock changes. High level-guidance from the IPCC tends to distinguish between "stock change" vs. "flux" approaches to measuring carbon sources and sinks. While "flux" approaches measure GHG exchanges to and from forested systems, "stock change" approaches quantify carbon stocks across pools as well as the changes in them. The protocols that we reviewed primarily use stock change approaches, which include plot-based inventories with extrapolation to the project area, field measurement of trees, and use of allometric equations (which describe non-linear relationships between a tree's biomass and its more easily measured parameters, such as its height and/or diameter).

The protocols tend to provide appropriately rigorous, high-level guidance on inventory design under a stock change approach that aligns with recommendations from the IPCC (2019). Forest structure and composition (and thus aboveground biomass) can be highly variable. The protocols allow flexibility in carbon accounting such that project developers can adapt methods to local conditions and efficiently conduct monitoring, reporting, and verification. Protocols allow either permanent or temporary sample plots (ACR, ARB) as well as stratified random or systematic random plot designs (CAR-U.S.). Both approaches can produce unbiased and precise estimates of aboveground carbon stocks, but will depend on local forest structure and composition as well as the field inventory design used. IFM projects in regions with fewer relevant datasets may use less appropriate allometric equations and thus less robust estimates of aboveground biomass (Yuen et al., 2016). Depending on the methods used, overestimation of aboveground carbon stocks can occur (Clough et al., 2016), but this is likely to be less consequential to the overall validity of a forest carbon project than other considerations (e.g., baselines and leakage).

Methods for quantifying forest carbon stocks and their changes are rapidly evolving, including through the integration of field-based methods and remote sensing. Although challenges associated with accurately measuring changes in below-canopy forest structure for some remote sensing types (e.g., optical imagery) may limit their application to IFM projects (Asbeck and Frey, 2021), we expect technological advances to improve its future utility. However, a full discussion of these future opportunities is out of scope of this study, and we refer the reader to other reviews of the topic (Goetz and Dubayah, 2011; Xiao et al., 2019).

2.4.2. Belowground biomass

Belowground biomass refers to living roots, typically comprising 15–25% of total living biomass in a forest (Jackson et al., 1996). The belowground biomass pool does not include soil carbon, microbial carbon, or dead roots (although living roots contribute directly to each of these other pools *via* complex processes including root death, root exudates, and interactions of mycorrhizal fungi). Belowground biomass estimation models vary widely across protocols. Because

TABLE 5 Summary of carbon pools in IFM protocols.

Protocol	Carbon pool							
	Aboveground tree biomass	Aboveground non-tree biomass	Belowground biomass	Standing dead	Lying dead	Litter pool	Soil carbon	Wood products (in-use)
ACR	Included	Included	Included	Included/ optional	Optional	Excluded	Excluded	Included
ARB	Included	Included	Included	Included	Excluded	Excluded	Included/ excluded[a]	Included
CAR-U.S.	Included	Excluded	Included	Included	Excluded	Excluded	Included[b]	Included
CAR-Mexico	Included	Included[c]	Included	Included	Excluded	Excluded	Excluded[d]	Excluded
VCS VM0010	Included	Excluded	Excluded	Excluded[g]	Excluded[f]	Excluded	Excluded	Included
VCS VM0012	Included	Excluded	Included	Included	Included	Excluded	Excluded	Included
VCS VM0003	Included	Excluded	Included	Conditional[e]	Conditional[f]	Excluded	Excluded	Conditional[f]
VCS VM0005	Included	Excluded	Optional	Included	Included	Excluded	Excluded	Included

[a]Soil carbon must be included in the Offset Project Boundary if (1) Site preparation activities involve deep ripping, furrowing, or plowing where soil disturbance exceeds 25% of the Project Area over the Project Life, or (2) mechanical site preparation activities are not conducted on contours.
[b]Included for estimating site preparation emissions.
[c]Included for estimating site preparation emissions.
[d]Site preparation with deep ripping techniques may require suspension of forest carbon credits for a number of crediting periods directly proportional to the area of the site impacted.
[e]Dead wood stocks can be excluded unless the project scenario produces greater levels of slash than the baseline and slash is burned as part of forest management. If slash produced in the project case is left in the forest to become part of the dead wood pool, dead wood may be excluded. Project proponents may elect to include the pool (where included the pool must be estimated in both the baseline and with project cases) as long as the dead wood pool represents $<$50% of total carbon volume on the site in any given modeled year.
[f]The protocol provides an approach for accounting for this pool, but also allows for exclusion of wood products if transparent and verifiable information can demonstrate that carbon stocks in wood products are rising faster in the project case than in the baseline or are decreasing faster in the baseline than in the project case.
[g]Dead wood from logging (slash) is included in the baseline.

empirical measurement of belowground biomass is difficult and time-consuming (requiring excavating, cleaning, sorting, and weighing roots), belowground biomass is estimated indirectly based on aboveground biomass measurements. The IFM protocols estimate belowground biomass using allometric equations or root:shoot ratios, which are inherently unable to capture detailed natural variation and, additionally, may introduce systematic errors by being inappropriately matched to the system in question (Ledo et al., 2018). Root:shoot ratios assume that belowground biomass occurs in a fixed ratio to aboveground biomass, whereas allometric equations allow for non-linear relationships.

VCS protocols tend to provide the greatest flexibility in ratio selection for belowground biomass estimation. VCS establishes basic criteria for eligible models, including peer-review, appropriate parameterization, and consistency with the original scope of the study. Regions with more abundant literature documenting root:shoot ratios enable developers to select estimates that produce the greatest number of credits. For example, VM0003 allows for use of the standard root:shoot ratios cited in Cairns et al. (1997), or any root:shoot value from research literature or national inventories with comparable climate and forest type. VM0012 is more stringent, requiring the use of the Cairns et al. ratios unless project-specific measurements have been taken. VM0010 is the only protocol that excludes belowground biomass entirely.

Both CAR and ARB require that projects in Washington, California, and Oregon use the Cairns et al. ratios. For other contiguous states, CAR and ARB protocols provide region-specific component ratio methods (which further divide aboveground and belowground biomass into subcompartments). ACR requires use of USFS merchantable volume equations tailored for region and species, which are then extrapolated to belowground biomass using ratios in Jenkins et al. (2003).

Because relatively little empirical belowground biomass data exists for validating either the allometric or root:shoot ratio approaches, it is not well-understood which of these approaches is preferable, what magnitude of error they may introduce, and whether they systematically over- or underestimate belowground biomass according to vegetation type, region, or climate regime (Xing et al., 2019). Across protocols, the Cairns et al. (1997) and Jenkins et al. (2003) reviews underpin nearly all belowground biomass estimates in IFM projects. Efforts to "spot-check" the validity of these simple modeling approaches have sometimes revealed large errors: for example, Xing et al. (2019) used empirical data to reveal that a root:shoot ratio approach overestimated belowground biomass in a Canadian poplar forest by between 18 and 42%.

2.4.3. Soil carbon

Soils comprise 56% of the carbon stock within managed ecosystems across the United States, and 80% of the terrestrial carbon pool globally (Lal, 2008; Domke et al., 2017). IFM protocols rarely require the measurement or estimation of soil organic carbon (SOC) stocks and fluxes due to the assumption that changes in the soil pool are negligible relative to credit volumes and due to the considerable expense and logistical challenge of measuring the soil carbon stock accurately and comprehensively (Paustian et al., 2019). ACR and VCS IFM protocols fail to account for advances in soil science, and potentially omit declines in SOC caused by certain IFM practices. In some instances this omission could enable over-crediting by neglecting substantial losses in soil organic matter that are likely not recuperated during the crediting period (Johnson and Curtis, 2001; Jandl et al., 2007; Noormets et al., 2015; Johnson and Henderson, 2018). A growing body of literature indicates that site preparation and ongoing management can cause significant disturbance to soil stocks, especially in litter, organic, and topsoil carbon pools, partially eroding the benefits of biomass stock increases (Jandl et al., 2007; Achat et al., 2015). In the crediting context, the primary consideration should be whether soil disturbance and SOC stock declines under IFM exceed the baseline.

Some IFM practices, such as extended rotation and retention of coarse woody material, are unlikely to yield significant or persistent changes in the soil carbon stock, and may prevent SOC losses that may have occurred under the baseline (Mayer et al., 2020). In contrast, mechanical site preparation, such as thinning, planting, removal of brush or shrubbery, or partial harvesting, may have significant and long lasting negative impacts on the SOC pool (Walmsley and Godbold, 2010; Zhang et al., 2018). The CAR and ARB protocols most appropriately and conservatively include these fluxes by requiring that projects with site preparation, harvesting, or treatment (deep ripping, furrowing, or plowing where soil disturbance exceeds 25% of project area or is not done on contours) estimate the loss of soil carbon as a product of biomass removal, mineral soil exposure, and frequency of disturbance. Estimated carbon stocks and losses are calculated using predetermined coefficients, which are determined by the soil order, harvesting intensity, disturbance frequency, site treatment, and tree type composition.

This is aligned with a growing body of evidence demonstrating that harvesting can yield losses between 8 and 11% in the top meter of soil (James and Harrison, 2016). Similarly, thinning and removal of dead biomass reduce organic matter inputs, compact topsoil, mix soil layers, and reduce the total SOC stock (Mayer et al., 2020; Kaarakka et al., 2021). These impacts are most substantial in the organic layer and topsoil (0–10 cm) even under conventional thinning practices, demonstrating losses of ∼25 and 5% of total SOC stock 10 years after management, respectively (Achat et al., 2015). SOC stocks are not homogenous and can be considered relatively recalcitrant or labile depending on the degree to which the carbon is mineral-associated or particle-associated organic matter (Lavallee et al., 2020). On average, the top 20 cm of forest soils in the United States contain ∼230 tCO_2/ha (Cao et al., 2019), thus a loss of 15% of this stock across only 20% of the project area may reduce total project credits on the order of 7 tCO_2/ha. For context, across the 74 projects reviewed by Badgley et al. (2022b), credit issuances averaged 73 tCO_2/ha, implying an average project could over-credit by 10% or more without violating CAR or ARB SOC stock estimation requirements. However, this is only relevant to crediting outcomes if the SOC stock under IFM declines more substantially than the baseline, which is unlikely in projects that involve a reduction in harvesting.

Only CAR and ARB allow for the inclusion of the SOC pool, and require it if the stock is likely to decline due to site preparation disturbances or other management activities. Appropriately, none of the IFM protocols include an option for additional crediting from increases in SOC. All VCS and ACR IFM protocols presume that impacts on soil carbon would be negligible or positive relative to the baseline. To rigorously incorporate the impact of SOC losses within IFM projects, protocols would need to quantify not only the impact of project management practices, but also the alternative impact to the soil carbon stock under the baseline scenario.

2.4.4. Harvested wood products (HWPs)

The harvest of biomass for use in wood products is included in all reviewed IFM protocols with the exception of the CAR-Mexico protocol, whose projects are not expected to significantly alter the production of wood products. The ARB, ACR, CAR-U.S., and VCS protocols all offer detailed methodologies for estimating the carbon stock stored in wood products. The methodologies require an estimate of the carbon stock for both baseline and project HWPs. In general, they follow a similar process where project proponents must estimate (a) the volume of timber removed in the project and baseline scenarios, (b) the merchantable carbon in these HWPs, the carbon loss due to mill processing, and (c) the decay of HWP carbon in final products and landfills over a 100 year horizon. This decay rate varies based on the lifetime of the product category.

For example, in ARB and CAR-U.S. projects, carbon in HWPs is annualized across a 100-year decay function to generate a HWP "storage factor." This means that each year, carbon flowing into the HWP pool is immediately discounted to its 100-year average value. In other words, a large portion of carbon reduced in the forest as a result of harvesting is assumed to instantaneously decay. Since much of that carbon is actually released over decades rather than immediately, for the first 50 years of the project, if the project harvests less than that projected in the baseline scenario carbon, which is the case for most IFM projects, benefits and credits are overestimated. ACR and VCS protocols use similar "storage factor" approaches for estimating carbon in HWPs.

All of the protocols we reviewed exaggerate the emissions associated with the production of HWPs by ignoring their displacement of other fossil-intensive alternatives. Substitution benefits are typically high for construction-based materials, such as steel or concrete (Smyth et al., 2017; Geng et al., 2019) and vary widely for energy products, such as biomass used to generate electricity and heat, based on the product displaced (Cabiyo et al., 2021). Ignoring these benefits results in some over-crediting and also shifts protocol incentives toward projects that reduce harvesting.

2.4.5. Recommendations on forest carbon accounting

The accuracy and precision of estimating forest carbon stocks within IFM protocols should improve over time as measurement technologies, inventories, allometric equations, and root:shoot ratios improve. IFM protocols generally provide appropriate selection criteria for plot distribution, measurement, and carbon stock estimation and distribution methods. The accuracy of a given site's carbon stock estimate is likely to be most significantly impacted by the availability of regionally tailored and species-specific allometric equations and root:shoot ratios to approximate the impact of IFM practices on biomass distribution. Accounting for carbon in harvested wood products is more straightforward than estimating carbon in the ecosystem, and unnecessary over-crediting in the early decades of a project could easily be avoided by modeling HWPs in a temporally realistic way instead of immediately discounting them to their 100-year "storage factor." Lastly, the protocols should account for potentially significant and lasting losses in soil carbon pools as a result of disruptive site preparation and management methods. While CAR and ARB have already incorporated literature-driven methods to account for reductions in the soil carbon stock of a project, more research is needed to understand how specific practices, species, and soil types respond to interventions.

3. Discussion and conclusions

Carbon offsets have the potential to direct substantial funds into improved forest management, helping realize the potential for forest management to sequester carbon and achieve a range of other environmental and societal benefits. Carbon offset quality matters. Offsets are designed to compensate for known GHG emissions, reducing the overall cost of meeting an emissions target. If they generate more credits than their actual impact, they can reduce and obscure the efficacy of climate change mitigation efforts. In this paper, we compare the offset protocols that have generated offset credits from IFM globally with literature on quantifying carbon impacts from IFM activities. Focusing on all major elements of carbon accounting—baselines, additionality, leakage, durability, and carbon pool quantification—we document shortcomings of each protocol, and suggest specific ways they could be improved to reduce the risk of over-crediting.

The most important area for reducing over-crediting is changing the way baselines are determined. All protocols, except for CAR-Mexico, offer substantial flexibility in setting project baselines. When there is flexibility, project developers have a financial incentive to choose the option that generates the most credits. ARB and CAR-U.S. allow the developer discretion to use any modeled baseline that is financially, legally, and contractually feasible, and not below the minimum allowed baseline, which is defined as the regional average for most projects. With that discretion, most developers choose baselines at or very close to minimum allowed levels (Badgley et al., 2022b).

Similarly, for the ACR protocol, baselines are defined as the scenario with the highest net present value (NPV) for the landowner. While NPV is a conceptually accurate way to predict land management for industrial forest owners, it is not a good predictor for many landowners seeking to manage for multiple uses, like recreational or ecosystem benefits. Further, it can be difficult for verifiers to assess NPV claims due to information asymmetries. All four VCS protocols provide developers with flexibility in choosing the baseline scenarios. Only the CAR-Mexico protocol prohibits baselines below initial carbon stocks, but the ability for project developers to choose any crediting period between 1 and 100 years increases the risk of non-additional crediting.

In the current market, flexible baseline setting rules have resulted in a large portion of credits being generated from claims that projects prevent forest carbon loss with large reductions in timber harvesting. These projects look more similar to conservation or avoided degradation projects than to improved forest management. While these baselines might be accurate for some projects with potential for real climate benefit, the flexibility all protocols give can lead to significant over-crediting.

Several changes to the protocols could result in more accurate and conservative baselines. Baselines set at current levels or past practice for the particular parcel (not for a broad regional average) or with dynamic baselines or NPV for some forest lands are more conservative than current methods that have systematically resulted in aggressive harvesting baselines. Choosing baselines at or close to initial carbon stocks, and avoiding the deep baselines currently used allows landowners to be credited for changing their land management practice (compared to the past, present, or other similar lands dynamically), rather than for not changing it. Avoiding aggressive harvesting baselines would also lessen over-crediting from leakage and harvested wood product accounting and improve the effectiveness of reversal buffer pools by improving the quality of the credits in them.

NPV baselines are justifiable for industrial timberland owners who can show a history of management consistent with NPV and who have steady access to contract labor and mills. Dynamic baselines, while unproven in the market, offer a number of advantages because they can adjust to market conditions over time. However, until dynamic baselines are applied to real-world settings, their strengths and weaknesses may not be completely understood.

All of these baseline setting methods still risk over-crediting due to adverse selection. Adverse selection can occur because landowners that do not need to change their forest management practice to earn offset credits are the most likely to participate and earn credits against standardized rules, undetected due to information asymmetries.

While setting more conservative baselines is likely to remedy a large portion of over-crediting risk under current protocols, we identified several other areas where the current protocols could be better aligned with the scientific literature.

One important correction to the ARB and CAR-U.S. protocols is to fix a contradiction in the baseline scenario. Currently, in the 1st year of a project, landowners are rewarded for the difference in onsite carbon stocks between actual onsite carbon stocks and the often much lower baseline level, while deductions for leakage and carbon in harvested wood products in that year are based on 100-year average harvest rates. A straightforward correction is to assume levels of harvesting in the baseline that match any assumed drop in onsite carbon stocks. In order to avoid discouraging projects that extend rotations by reducing harvesting for short periods, leakage deductions could be applied over several years, and all protocols could account for positive leakage cumulatively rather than annually when harvesting is larger in the project than in the baseline scenario. Similarly, protocols could avoid over-crediting by crediting against temporally explicit HWP decay functions rather than using static HWP "storage factors" for a given time period.

The science on leakage is not yet robust enough to develop rules that satisfactorily address leakage risk from projects that reduce harvesting. The protocols have opted to apply low leakage rates, which are generally inconsistent with the scant literature available. It would be prudent to apply higher leakage rates until new data and methods can be developed to support a more refined approach.

The protocols likely under-allocate credits to the buffer pool, in large part because they do not adequately address the increasing risk of reversal due to climate change. Larger buffer pool deductions along with regularly updating the protocols based on the latest science would help to address this issue. Protocols may also consider incentivizing, and avoid dis-incentivizing, practices that reduce carbon in the short run but increase resilience in the long-run, like thinning and fuels treatments that reduce the risk of catastrophic wildfire (Hurteau et al., 2011; North and Hurteau, 2011; Herbert et al., 2022).

Finally, methods for estimating onsite carbon stocks in the protocols allow for a great deal of flexibility. If implemented properly, current rules are sufficient to ensure high integrity. However, this flexibility also allows for less accurate carbon accounting, including through the use of reference literature for allometric equations and root:shoot ratios that may not be appropriate or conservative for the project under development. While the implications of this flexibility have not yet been systematically studied in the context of IFM projects, it appears to be relatively less consequential than the baseline, leakage, and durability issues identified above.

These changes will significantly reduce the risk of over-crediting and bring protocols more in line with the scientific literature. Still, we highlight one persistent challenge with ensuring the quality of IFM offset credits: uncertainty in the true baselines. Our recommendations reduce, but do not eliminate, the risk of over-crediting from baseline choices. Due to the inherent uncertainty in true baselines, baseline setting rules necessarily involve a tradeoff between false positives and false negatives (Trexler et al., 2006). If baselines err on the side of inclusiveness, allowing projects to choose baselines well below initial carbon stocks can accommodate worthwhile projects on lands at risk of degradation or conversion, but this flexibility also allows lands not at high risk of being degraded or converted to choose similar baselines, leading to over-crediting (false positives). Choosing more conservative baselines as we recommend means that some valuable projects will earn fewer credits than their true climate impact and some opportunities for real climate mitigation will be missed (false negatives). The greater the baseline uncertainty, the greater the tradeoff between false negatives and false positives. Setting the baseline at the average, given uncertainty, is not sufficient to avoid over-crediting because of information asymmetry and adverse selection.

Another potential solution to the inevitability of adverse selection (and more broadly, the incentive for project developers to take advantage of flexible rules to choose the option that results in the most credits), is to build more sources of under-crediting into the protocols so that if over-crediting occurs for any particular project, the integrity of the portfolios of projects under a protocol as a whole is not compromised.

If a higher burden of evidence for quality was required across the whole offset market, the number of credits generated by each project would shrink, and the price would go up. Poor quality of IFM and other project types keeps offset prices lower than what is needed to effectively drive mitigation without over-crediting. Expected growing demand in the voluntary market and constrained supply will likely push carbon prices higher in the future allowing offsets to play a larger role in driving real change with more accurate protocols.

IFM has a large potential to reduce emissions and sequester carbon through forest restoration, conservation of ecologically important forests, increased stand productivity through changed management, extended rotations of working forestlands, restoration of degraded forests, and reduced-impact logging. Carbon offsetting has the potential to create meaningful incentives to achieve this potential. This study identified ways to bring the IFM protocols better in line with the literature on carbon accounting and forest management to significantly reduce the risk of over-crediting. Most importantly, more conservative baselines that avoid the assumption of significantly increased harvesting can substantially reduce over-crediting risk, but does not resolve it due to persistent uncertainty and adverse selection. Better aligning protocol rules with current understanding of carbon accounting practices will help re-allocate carbon financing toward projects that can have meaningful climate impact.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

BH and MP designed the research. All authors performed the research, wrote the original manuscript, and created figures and tables.

Acknowledgments

We thank Anna Escuer, Kelli Wright, and Rebekah Braswell for helpful discussion and feedback. We also thank Google and Land Life for support for an early draft of this analysis.

Conflict of interest

All authors have been compensated for performing consultant advisory services for Carbon Direct Inc. JB was employed by Conservation International.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Achat, D. L., Fortin, M., Landmann, G., Ringeval, B., and Augusto, L. (2015). Forest soil carbon is threatened by intensive biomass harvesting. *Sci. Rep.* 5, 15991. doi: 10.1038/srep15991

Allen, M. R., Fuglestvedt, J. S., Shine, K. P., Reisinger, A., Pierrehumbert, R. T., and Forster, P. M. (2016). New use of global warming potentials to compare cumulative and short-lived climate pollutants. *Nat. Clim. Change* 6, 773–776. doi: 10.1038/nclimate2998

Ameray, A., Bergeron, Y., Valeria, O., Montoro Girona, M., and Cavard, X. (2021). Forest carbon management: A review of silvicultural practices and management strategies across boreal, temperate and tropical forests. *Curr. For. Rep.* 7, 245–266. doi: 10.1007/s40725-021-00151-w

Andam, K. S., Ferraro, P. J., Pfaff, A., Sanchez-Azofeifa, G. A., and Robalino, J. A. (2008). Measuring the effectiveness of protected area networks in reducing deforestation. *Proc. Natl. Acad. Sci. U. S. A.* 105, 16089–16094. doi: 10.1073/pnas.0800437105

Anderegg, W. R. L., Chegwidden, O. S., Badgley, G., Trugman, A. T., Cullenward, D., Abatzoglou, J. T., et al. (2022). Future climate risks from stress, insects and fire across US forests. *Ecol. Lett.* 25, 1510–1520. doi: 10.1111/ele.14018

Anderegg, W. R. L., Trugman, A. T., Badgley, G., Anderson, C. M., Bartuska, A., Ciais, P., et al. (2020). Climate-driven risks to the climate mitigation potential of forests. *Science* 368, eaaz7005. doi: 10.1126/science.aaz7005

Anderson, C., and Perkins, J. (2017). *Counting California Forest Carbon Offsets: Greenhouse Gas Mitigation Lessons from California's Cap-and-Trade U.S. Forest Compliance Offset Program (submitted as public comment to the California Air Resources Board on the 2017 Scoping Plan Update)*. Stanford School of Earth, Energy amd Environmental Sciences. Available online at: https://www.arb.ca.gov/lists/com-attach/app-zip/89-scopingplan2030-UyBWM1Q6BSZWOQBu.zip (accessed February 01, 2023).

Anderson, C. M., Field, C. B., and Mach, K. J. (2017). Forest offsets partner climate-change mitigation with conservation. *Front. Ecol. Environ.* 15, 359–365. doi: 10.1002/fee.1515

Archer, D., Eby, M., Brovkin, V., Ridgwell, A., Cao, L., Mikolajewicz, U., et al. (2009). Atmospheric lifetime of fossil fuel carbon dioxide. *Annu. Rev. Earth Planet. Sci.* 37, 117–134. doi: 10.1146/annurev.earth.031208.100206

Asante, P., and Armstrong, G. (2016). Carbon sequestration and the optimal forest harvest decision under alternative baseline policies. *Can. J. For. Res.* 46, 656–665. doi: 10.1139/cjfr-2015-0222

Asbeck, T., and Frey, J. (2021). Weak relationships of continuous forest management intensity and remotely sensed stand structural complexity in temperate mountain forests. *Eur. J. Forest Res.* 140, 721–731. doi: 10.1007/s10342-021-01361-4

Asbeck, T., Sabatini, F., Augustynczik, A. L. D., Basile, M., Helbach, J., Jonker, M., et al. (2021). Biodiversity response to forest management intensity, carbon stocks and net primary production in temperate montane forests. *Sci. Rep.* 11, 1625. doi: 10.1038/s41598-020-80499-4

Austin, K. G., Baker, J. S., Sohngen, B. L., Wade, C. M., Daigneault, A., Ohrel, S. B., et al (2020). The economic costs of planting, preserving, and managing the world's forests to mitigate climate change. *Nat. Commun.* 11, 5946. doi: 10.1038/s41467-020-19578-z

Badgley, G., Chay, F., Chegwidden, O. S., Hamman, J. J., Freeman, J., and Cullenward, D. (2022a). California's forest carbon offsets buffer pool is severely undercapitalized. *Front. For. Glob. Change* 5, 930426. doi: 10.3389/ffgc.2022.930426

Badgley, G., Freeman, J., Hamman, J. J., Haya, B., and Cullenward, D. (2021). *California Improved Forest Management Offset Project Database (1.0) [Data Set]*. Zenodo.

Badgley, G., Freeman, J., Hamman, J. J., Haya, B., Trugman, A. T., Anderegg, W. R. L., et al. (2022b). Systematic over-crediting in California's forest carbon offsets program. *Glob. Change Biol.* 28, 1433–1445. doi: 10.1111/gcb.15943

Butler, B. J., Hewes, J. H., Dickinson, B. J., Andrejczyk, K., Butler, S. M., and Markowski-Lindsay, M. (2016). Family forest ownerships of the United States, 2013: Findings from the USDA forest service's national woodland owner survey. *J. For.* 114, 638–647. doi: 10.5849/jof.15-099

Cabiyo, B., Fried, J. S., Collins, B. M., Stewart, W., Wong, J., and Sanchez, D. L. (2021). Innovative wood use can enable carbon-beneficial forest management in California. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2019073118. doi: 10.1073/pnas.2019073118

Cairns, M. A., Brown, S., Helmer, E. H., and Baumgardner, G. A. (1997). Root biomass allocation in the world's upland forests. *Oecologia* 111, 1–11. doi: 10.1007/s004420050201

Cames, M., Harthan, R., Füssler, J., Lee, C., Erickson, P., and Spalding-Fecher, R. (2016). *How Additional Is the Clean Development Mechanism? Analysis of the Application of Current Tools and Proposed Alternatives*. Freiberg: Oeko Institute. Available online at: https://climate.ec.europa.eu/system/files/2017-04/clean_dev_mechanism_en.pdf (accessed February 01, 2023).

Cao, B., Domke, G. M., Russell, M. B., and Walters, B. F. (2019). Spatial modeling of litter and soil carbon stocks on forest land in the conterminous United States. *Sci. Tot. Environ.* 654, 94–106. doi: 10.1016/j.scitotenv.2018.10.359

Carton, W., Lund, J. F., and Dooley, K. (2021). Undoing equivalence: Rethinking carbon accounting for just carbon removal. *Front. Clim.* 3, 664130. doi: 10.3389/fclim.2021.664130

Chave, J., Condit, R., Aguilar, S., Hernandez, A., Lao, S., and Perez, R. (2004). Error propagation and scaling for tropical forest biomass estimates. *Phil. Trans. R. Soc. Lond. B* 359, 409–420. doi: 10.1098/rstb.2003.1425

Clough, B. J., Russell, M. B., Domke, G. M., and Woodall, C. W. (2016). Quantifying allometric model uncertainty for plot-level live tree biomass stocks with a data-driven, hierarchical framework. *For. Ecol. Manag.* 372, 175–188. doi: 10.1016/j.foreco.2016.04.001

Coffield, S. R., Vo, C. D., Wang, J. A., Badgley, G., Goulden, M. L., Cullenward, D., et al. (2022). Using remote sensing to quantify the additional climate benefits of California forest carbon offset projects. *Glob. Change Biol.* 2022, gcb.16380. doi: 10.1111/gcb.16380

Domke, G. M., Perry, C. H., Walters, B. F., Nave, L. E., Woodall, C. W., and Swanston, C. W. (2017). Toward inventory-based estimates of soil organic carbon in forests of the United States. *Ecol. Appl.* 27, 1223–1235. doi: 10.1002/eap.1516

Elgin, B. (2020). *JPMorgan, Disney, Blackrock Buy Nature Conservancy's Useless Carbon Offsets*. Bloomberg News. Available online at: https://www.bloomberg.com/features/2020-nature-conservancy-carbon-offsets-trees/ (accessed February 1, 2023).

Ellis, P. W., Gopalakrishna, T., Goodman, R. C., Putz, F. E., Roopsind, A., Umunay, P. M., et al. (2019). Reduced-impact logging for climate change mitigation (RIL-C) can halve selective logging emissions from tropical forests. *For. Ecol. Manag.* 438, 255–266. doi: 10.1016/j.foreco.2019.02.004

Fargione, J. E., Bassett, S., Boucher, T., Bridgham, S. D., Conant, R. T., Cook-Patton, S. C., et al. (2018). Natural climate solutions for the United States. *Sci. Adv.* 4, eaat1869. doi: 10.1126/sciadv.aat1869

Ferraro, P. J., and Hanauer, M. M. (2014). Quantifying causal mechanisms to determine how protected areas affect poverty through changes in ecosystem services and infrastructure. *Proc. Natl. Acad. Sci. U. S. A.* 111, 4332–4337. doi: 10.1073/pnas.1307712111

Foley, T. G., Richter, D., de, B., and Galik, C. S. (2009). Extending rotation age for carbon sequestration: A cross-protocol comparison of North American forest offsets. *For. Ecol. Manag.* 259, 201–209. doi: 10.1016/j.foreco.2009.10.014

Gan, J., and McCarl, B. A. (2007). Measuring transnational leakage of forest conservation. *Ecol. Econ.* 64, 423–432. doi: 10.1016/j.ecolecon.2007.02.032

Geng, A., Chen, J., and Yang, H. (2019). Assessing the greenhouse gas mitigation potential of harvested wood products substitution in China. *Environ. Sci. Technol.* 53, 1732–1740. doi: 10.1021/acs.est.8b06510

Gifford, L. (2020). "You can't value what you can't measure": A critical look at forest carbon accounting. *Clim. Change* 161, 291–306. doi: 10.1007/s10584-020-02653-1

Goetz, S., and Dubayah, R. (2011). Advances in remote sensing technology and implications for measuring and monitoring forest carbon stocks and change. *Carbon Manag.* 2, 231–244. doi: 10.4155/cmt.11.18

Gren, I.-M., and Aklilu, A. Z. (2016). Policy design for forest carbon sequestration: A review of the literature. *For. Pol. Econ.* 70, 128–136. doi: 10.1016/j.forpol.2016.06.008

Griscom, B., Shoch, D., Stanley, B., Cortez, R., and Virgilio, N. (2009). Sensitivity of amounts and distribution of tropical forest carbon credits depending on baseline rules. *Environ. Sci. Pol.* 12, 897–911. doi: 10.1016/j.envsci.2009.07.008

Griscom, B. W., Adams, J., Ellis, P. W., Houghton, R. A., Lomax, G., Miteva, D. A., et al. (2017). Natural climate solutions. *Proc. Natl. Acad. Sci. U. S. A.* 114, 11645–11650. doi: 10.1073/pnas.1710465114

Groom, B., and Venmans, F. (2022). The social value of offsets. *Res. Squ. [Preprint]*. doi: 10.21203/rs.3.rs-1515075/v1

Grove, P. M., Conrad, J. L., Harris, T. G., and Dahlen, J. (2020). Consulting forester timber sale practices in the US South. *For. Sci.* 66, 221–229. doi: 10.1093/forsci/fxz068

Haya, B. (2010). *Carbon Offsetting: An Efficient Way to Reduce Emissions or to Avoid Reducing Emissions? An Investigation and Analysis of Offsetting Design and Practice in India and China*. Available online at: https://escholarship.org/content/qt7jk7v95t/qt7jk7v95t.pdf (accessed February 01, 2023).

Haya, B. (2019). *The California Air Resources Board's U.S. Forest Offset Protocol Underestimates Leakage*. University of California, Berkeley. Available online at: https://gspp.berkeley.edu/assets/uploads/research/pdf/Policy_Brief-US_Forest_Projects-Leakage-Haya_4.pdf (accessed February 1, 2023).

Haya, B., Cullenward, D., Strong, A. L., Grubert, E., Heilmayr, R., Sivas, D. A., et al. (2020). Managing uncertainty in carbon offsets: Insights from California's standardized approach. *Clim. Pol.* 20, 1112–1126. doi: 10.1080/14693062.2020.1781035

Haya, B., and Stewart, W. (2019). *Response to Comments by the California Air Resources Board on Policy Brief: The California Air Resources Board's U.S. Forest Offset Protocol Underestimates Leakage*. University of California, Berkeley. Available online at: https://gspp.berkeley.edu/assets/uploads/research/pdf/Response_to_comments_by_ARB_on_leakage_under_forest_protocol_2.pdf (accessed February 01, 2023).

Herbert, C., Haya, B., Stephens, S. L., and Butsic, V. (2022). Managing nature-based solutions in fire-prone ecosystems: Competing management objectives in California forests evaluated at a landscape scale. *Front. For. Glob. Change*. 2022, 957189. doi: 10.3389/ffgc.2022.957189

Herzog, H., Caldeira, K., and Reilly, J. (2003). An issue of permanence: Assessing the effectiveness of temporary carbon storage. *Clim. Change* 59, 293–310. doi: 10.1023/A:1024801618900

Hurteau, M. D., Stoddard, M. T., and Fulé, P. Z. (2011). The carbon costs of mitigating high-severity wildfire in southwestern ponderosa pine: Carbon costs of mitigating wildfire. *Glob. Change Biol.* 17, 1516–1521. doi: 10.1111/j.1365-2486.2010.02295.x

IPCC. (2007). "Summary for policymakers," in *Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, eds B. Metz, O. R. Davidson, P. R. Bosch, R. Dave, and L. A. Meyer (Cambridge; New York, NY: Cambridge University Press).

IPCC. (2019). "Forest land," in *2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories*. (Vol. 4): Agriculture, Forestry and Other Land Use, eds E. Calvo Buendia, K. Tanabe, A. Kranjc, J. Baasansuren, M. Fukuda, S. Ngarize, A. Osako, Y. Pyrozhenko, P. Shermanau, and S. Federici (Cham: IPCC).

Jackson, R. B., Canadell, J., Ehleringer, J. R., Mooney, H. A., Sala, O. E., and Schulze, E. D. (1996). A global analysis of root distributions for terrestrial biomes. *Oecologia*. 108, 389–411. doi: 10.1007/BF00333714

James, J., and Harrison, R. (2016). The effect of harvest on forest soil carbon: A meta-analysis. *Forests.* 7, 308. doi: 10.3390/f7120308

Jandl, R., Lindner, M., Vesterdal, L., Bauwens, B., Baritz, R., Hagedorn, F., et al. (2007). How strongly can forest management influence soil carbon sequestration? *Geoderma* 137, 253–268. doi: 10.1016/j.geoderma.2006.09.003

Jenkins, J. C., Chojnacky, D. C., Heath, L. S., and Birdsey, R. A. (2003). National-scale biomass estimators for United States tree species. *For. Sci.* 49, 12–35. doi: 10.1093/forestscience/49.1.12

Johnson, D. W., and Curtis, P. S. (2001). Effects of forest management on soil C and N storage: meta analysis. *For. Ecol. Manag.* 140, 227–238. doi: 10.1016/S0378-1127(00)00282-6

Johnson, D. W., and Henderson, P. (2018). "Effects of forest management and elevated carbon dioxide on soil carbon storage," in *Soil Management and Greenhouse Effect*, eds R. Lal, J. M. Kimble, E. Levine, and B. A. Stewart (Boca Raton, FL: CRC Press), 12. doi: 10.1201/9780203739310-12

Kaarakka, L., Cornett, M., Domke, G., Ontl, T., and Dee, L. E. (2021). Improved forest management as a natural climate solution: A review. *Ecol. Solut. Evid.* 2, 12090. doi: 10.1002/2688-8319.12090

Kallio, A. M. I., and Solberg, B. (2018). Leakage of forest harvest changes in a small open economy: Case Norway. *Scand. J. For. Res.* 33, 502–510. doi: 10.1080/02827581.2018.1427787

Keegan, C. E., Sorenson, C. B., Morgan, T. A., Hayes, S. W., and Daniels, J. M. (2011). Impact of the great recession and housing collapse on the forest products industry in the Western United States. *For. Prod. J.* 61, 625–634. doi: 10.13073/0015-7473-61.8.625

Koberstein, P., and Applegate, J. (2021). *Native Alaskan Company Involved in Controversial Carbon Offsets Project to Quit Logging. Earth Island Journal.*

Kremen, C., and Merenlender, A. M. (2018). Landscapes that work for biodiversity and people. *Science* 362, eaau6020. doi: 10.1126/science.aau6020

Lal, R. (2008). Carbon sequestration. *Philos. Trans. Royal Soc. B* 363, 815–830. doi: 10.1098/rstb.2007.2185

Lavallee, J. M., Soong, J. L., and Cotrufo, M. F. (2020). Conceptualizing soil organic matter into particulate and mineral associated forms to address global change in the 21st century. *Glob. Change Biol.* 26, 261–273. doi: 10.1111/gcb.14859

Ledo, A., Paul, K. I., Burslem, D. F. R. P., Ewel, J. J., Barton, C., Battaglia, M., et al. (2018). Tree size and climatic water deficit control root to shoot ratio in individual trees globally. *New Phytol.* 217, 8–11. doi: 10.1111/nph.14863

Levasseur, A., Lesage, P., Margni, M., Brandão, M., and Samson, R. (2012). Assessing temporary carbon sequestration and storage projects through land use, land-use change and forestry: Comparison of dynamic life cycle assessment with ton-year approaches. *Clim. Change* 115, 759–776. doi: 10.1007/s10584-012-0473-x

Lewis, D. J. (2010). An economic framework for forecasting land-use and ecosystem change. *Resour. Energy Econ.* 32, 98–116. doi: 10.1016/j.reseneeco.2009.11.007

Mayer, M., Prescott, C. E., Abaker, W. E. A., Augusto, L., Cécillon, L., Ferreira, G. W. D., et al. (2020). Tamm review: Influence of forest management activities on soil organic carbon stocks: A knowledge synthesis. *For. Ecol. Manag.* 466, 118127. doi: 10.1016/j.foreco.2020.118127

McDowell, N. G., Allen, C. D., Anderson-Teixeira, K., Aukema, B. H., Bond-Lamberty, B., Chini, L., et al. (2020). Pervasive shifts in forest dynamics in a changing world. *Science* 368, eaaz9463. doi: 10.1126/science.aaz9463

Meyfroidt, P., Börner, J., Garrett, R., Gardner, T., Godar, J., Kis-Katos, K., et al. (2020). Focus on leakage and spillovers: Informing land-use governance in a tele-coupled world. *Environ. Res. Lett.* 15, e090202. doi: 10.1088/1748-9326/ab7397

Murray, B. C., McCarl, B. A., and Lee, H.-C. (2004). Estimating leakage from forest carbon sequestration programs. *Land Econ.* 80, 109–124. doi: 10.2307/3147147

Murray, B. C., Sohngen, B. L., Sommer, A. J., Depro, B., Jones, K., McCarl, B. A., et al. (2005). *Greenhouse Gas Mitigation Potential in U.S. Forestry and Agriculture*. Washington, DC: United States Environmental Protection Agency.

Nabuurs, G.-J., Delacote, P., Ellison, D., Hanewinkel, M., Hetemäki, L., and Lindner, M. (2017). By 2050 the mitigation effects of EU forests could nearly double through climate smart forestry. *Forests* 8, 484. doi: 10.3390/f8120484

Noormets, A., Epron, D., Domec, J. C., McNulty, S. G., Fox, T., Sun, G., et al. (2015). Effects of forest management on productivity and carbon sequestration: A review and hypothesis. *For. Ecol. Manag.* 355, 124–140. doi: 10.1016/j.foreco.2015.05.019

North, M. P., and Hurteau, M. D. (2011). High-severity wildfire effects on carbon stocks and emissions in fuels treated and untreated forest. *For. Ecol. Manag.* 261, 1115–1120. doi: 10.1016/j.foreco.2010.12.039

Nunery, J. S., and Keeton, W. S. (2010). Forest carbon storage in the northeastern United States: Net effects of harvesting frequency, post-harvest retention, and wood products. *For. Ecol. Manag.* 259, 1363–1375. doi: 10.1016/j.foreco.2009.12.029

Ontl, T. A., Janowiak, M. K., Swanston, C. W., Daley, J., Handler, S., Cornett, M., et al. (2020). Forest management for carbon sequestration and climate adaptation. *J. For.* 118, 86–101. doi: 10.1093/jofore/fvz062

Paustian, K., Collier, S., Baldock, J., Burgess, R., Creque, J., DeLonge, M., et al. (2019). Quantifying carbon for agricultural soil management: From the current status toward a global soil information system. *Carb. Manag.* 10, 567–587. doi: 10.1080/17583004.2019.1633231

Radeloff, V. C., Nelson, E., Plantinga, A. J., Lewis, D. J., Helmers, D., Lawler, J. J., et al. (2012). Economic-based projections of future land use in the conterminous United States under alternative policy scenarios. *Ecol. Appl.* 22, 1036–1049. doi: 10.1890/11-0306.1

Roe, S., Streck, C., Obersteiner, M., Frank, S., Griscom, B., Drouet, L., et al. (2019). Contribution of the land sector to a 1.5 °C world. *Nat. Clim. Chang.* 9, 817–828. doi: 10.1038/s41558-019-0591-9

Roopsind, A., Sohngen, B., and Brandt, J. (2019). Evidence that a national REDD+ program reduces tree cover loss and carbon emissions in a high forest cover, low deforestation country. *Proc. Natl. Acad. Sci. U. S. A.* 116, 24492–24499. doi: 10.1073/pnas.1904027116

Ruseva, T., Marland, E., Szymanski, C., Hoyle, J., Marland, G., and Kowalczyk, T. (2017). Additionality and permanence standards in California's Forest Offset Protocol: A review of project and program level implications. *J. Environ. Manag.* 198, 277–288. doi: 10.1016/j.jenvman.2017.04.082

Sasaki, N., Asner, G. P., Pan, Y., Knorr, W., Durst, P. B., Ma, H. O., et al. (2016). Sustainable management of tropical forests can reduce carbon emissions and stabilize timber production. *Front. Environ. Sci.* 4, 50. doi: 10.3389/fenvs.2016.00050

Smyth, C., Rampley, G., Lemprière, T. C., Schwab, O., and Kurz, W. A. (2017). Estimating product and energy substitution benefits in national-scale mitigation analyses for Canada. *GCB Bioenergy* 9, 1071–1084. doi: 10.1111/gcbb.12389

So, S. I., Haya, B. K., and Elias, M. (2023). *Voluntary Registry Offsets Database, v7.1. University of California Berkeley, Berkeley Carbon Trading Project*. Available online at: https://gspp.berkeley.edu/research-and-impact/centers/cepp/projects/berkeley-carbon-trading-project/offsets-database (accessed February 1, 2023).

Sohngen, B., and Brown, S. (2004). Measuring leakage from carbon projects in open economies: A stop timber harvesting project in Bolivia as a case study. *Can. J. For. Res.* 34, 829–839. doi: 10.1139/x03-249

Sohngen, B., and Brown, S. (2008). Extending timber rotations: Carbon and cost implications. *Clim. Pol.* 8, 435–451. doi: 10.3763/cpol.2007.0396

Stephens, S. L., Westerling, A. L., Hurteau, M. D., Peery, M. Z., Schultz, C. A., and Thompson, S. (2020). Fire and climate change: Conserving seasonally dry forests is still possible. *Front. Ecol. Environ.* 18, 354–360. doi: 10.1002/fee.2218

Sun, B., and Sohngen, B. (2009). Set-asides for carbon sequestration: Implications for permanence and leakage. *Clim. Change* 96, 409–419. doi: 10.1007/s10584-009-9628-9

Temesgen, H., Affleck, D., Poudel, K., Gray, A., and Sessions, J. (2015). A review of the challenges and opportunities in estimating above ground forest biomass using tree-level models. *Scand. J. For. Res.* 2015, 1–10. doi: 10.1080/02827581.2015.1012114

Trexler, M. C., Broekhoff, D. J., and Kosloff, L. H. (2006). A statistically-driven approach to offset-based GHG additionality determinations: What can we learn? *Sustain. Dev. Law Pol.* 6, 30–40. Available online at: https://digitalcommons.wcl.american.edu/sdlp/vol6/iss2/8/ (accessed January 29, 2023).

van Kooten, G. C., Bogle, T. N., and de Vries, F. P. (2015). Forest carbon offsets revisited: Shedding light on darkwoods. *For. Sci.* 61, 370–380. doi: 10.5849/forsci.13-183

van Kooten, G. C., Laaksonen-Craig, S., and Wang, Y. (2009). A meta-regression analysis of forest carbon offset costs. *Can. J. For. Res.* 39, 2153–2167. doi: 10.1139/X09-139

Walmsley, J. D., and Godbold, D. L. (2010). Stump harvesting for bioenergy - A review of the environmental impacts. *Forestry* 83, 17–38. doi: 10.1093/forestry/cpp028

Wear, D. N., and Murray, B. C. (2004). Federal timber restrictions, interregional spillovers, and the impact on US softwood markets. *J. Environ. Econ. Manag.* 47, 307–330. doi: 10.1016/S0095-0696(03)00081-0

West, T. A. P., Börner, J., Sills, E. O., and Kontoleon, A. (2020). Overstated carbon emission reductions from voluntary REDD+ projects in the Brazilian Amazon. *Proc. Natl. Acad. Sci. U. S. A.* 117, 24188–24194. doi: 10.1073/pnas.2004334117

Xiao, J., Chevallier, F., Gomez, C., Guanter, L., Hicke, J. A., Huete, A. R., et al. (2019). Remote sensing of the terrestrial carbon cycle: A review of advances over 50 years. *Remot. Sens. Environ.* 233, 111383. doi: 10.1016/j.rse.2019.111383

Xing, D., Bergeron, J. A. C., Solarik, K. A., Tomm, B., Macdonald, S. E., Spence, J. R., et al. (2019). Challenges in estimating forest biomass: Use of allometric equations for three boreal tree species. *Can. J. For. Res.* 49, 1613–1622. doi: 10.1139/cjfr-2019-0258

Yuen, J. Q., Fung, T., and Ziegler, A. D. (2016). Review of allometric equations for major land covers in SE Asia: Uncertainty and implications for above- and below-ground carbon estimates. *For. Ecol. Manag.* 360, 323–340. doi: 10.1016/j.foreco.2015.09.016

Zhang, X., Guan, D., Li, W., Sun, D., Jin, C., Yuan, F., et al. (2018). The effects of forest thinning on soil carbon stocks and dynamics: A meta-analysis. *For. Ecol. Manag.* 429, 36–43. doi: 10.1016/j.foreco.2018.06.027

News and intelligence on carbon markets, greenhouse gas pricing, and climate policy

Americas > ANALYSIS: Do offset registry revenue models offer perverse incentives to over-credit?

ANALYSIS: Do offset registry revenue models offer perverse incentives to over-credit?

Published 13:03 on March 20, 2023 / Last updated at 10:17 on March 21, 2023 / Joan Pinto / Americas, International, Nature-based, US, Voluntary

Recent allegations of widespread over-crediting in the voluntary carbon market (VCM) have raised questions of whether the dominant revenue model of offset standards bodies financially incentivises those involved to maximise issuance while compromising integrity.

Recent allegations of widespread over-crediting in the voluntary carbon market (VCM) have raised questions of whether the dominant revenue model of offset standards bodies financially incentivises those involved to maximise issuance while compromising integrity.

The environmental integrity of carbon offsets in the VCM has come into focus in 2023 after media articles and academic papers alleged numerous instances of over-crediting within Verra-registered REDD+ avoided deforestation undertakings (https://carbon-pulse.com/188182/) and Gold Standard (GS) and CDM-developed clean cookstove projects (https://carbon-pulse.com/195650/) .

This has caused some VCM participants to halt credit sales (https://carbon-pulse.com/189380/) from affected projects and deepened an ongoing loss of buyer confidence (https://carbon-pulse.com/181420/) in REDD+ offsets – as reflected in lower credit prices (https://carbon-pulse.com/195450/) and activity (https://carbon-pulse.com/195192/) – though registries have contested the accuracy of the investigations (https://carbon-pulse.com/189728/) , and Verra has pledged to undergo (https://carbon-pulse.com/191924/) a sweeping revamp of its practices that predates recent media reports.

A common thread underlying offset issuances across the VCM is registries' revenue models, wherein self-regulated firms get paid through their business of issuing verified emissions reduction (VERs) credits based on methodologies for which they charge fees to approve.

Most registries that function as non-profit organisations issue carbon credits while operating on a $/credit revenue model, ratcheting up issuance fees as they award VERs.

The more credits the certifiers pump out, the more money comes back to them in fees, boosting their coffers and potentially providing a financial incentive to overlook problems that could interrupt that income stream, in turn risking irreputable damage to efforts many see as crucial in tackling emissions.

"The $/credit fee structures encourage maximum credit issuance and financially discourage efforts to restrict over-crediting, so this revenue model – whether for-profit or non-profit – contributes to core structural problems in the voluntary carbon markets," Danny Cullenward, research fellow with American University's Institute for Carbon Removal Law and Policy, told Carbon Pulse.

Pricing per credit can vary, and publicly advertised fees are garnered through the life of an offset project that uses registry-approved protocols and methodologies to calculate the number of VERs eventually doled out.

Carbon Pulse has assessed registry issuance fees for the four main VCM standards bodies – Verra, GS, American Carbon Registry (ACR), and Climate Action Reserve (CAR) – along with related fees to review methodologies, validate, verify, and convert units (see table below).

ISSUES WITH ISSUANCES?

Elias Ayrey, co-founder and head scientist of offsets rating agency Renoster, estimated VCM over-crediting to sit at around 75% of total issuance.

A total of nearly 1.5 billion VCM credits have been issued thus far according to the Climate Focus dashboard, with Verra's Verified Carbon Units (VCUs) making up nearly 1.1 bln – or 71.7% – of total offsets distributed.

GS follows with a 16.5% share of the VCM, ACR with a 6.1% market share, and CAR with 5.2%.

Verra's net assets nearly quadrupled (https://carbon-pulse.com/194038/) between 2019 and 2021 as carbon credit issuances spiked, with levies from the organisation's various carbon a[illegible]nefits standards amounting to $37.5 mln for the firm in 2021 – some 92% of its revenue – according to financial statements.

In the US, registered non-profits are required to reinvest profits at the end of the year into growth, not pay dividends to shareholders, and lack an executive compensation scheme typically seen in for-profit organisations.

“Non-profit status is a tax status, not proof of ethical behaviour or aligned incentives,” Cullenward noted.

While increased revenue might not directly factor into worker pay-cheques or shareholder dividends, a healthy income from maximum issuance would still be welcome to a non-profit in securing business, while offering salaries capable of attracting and retaining top talent.

Maximising revenues could also be tempting as these entities seek to invest in developing and enhancing their offerings, or look to maintain and grow their market share and influence in an increasingly competitive marketplace.

Analysis firm AlliedOffsets tracked 17 registries that list offset projects spanning 150 countries, signalling the heightened competition that the historic ‘big four’ have come under of late.

Over-crediting issues aren’t just recent developments, however. SourceMaterial’s recent investigation identified 2011 filings in Verra’s registry showing how developer South Pole had estimated 52 mln tCO2 abatement from its Kariba REDD project in Zimbabwe (VCS ID 902), with Verra ultimately signing off on a revised figure of 197 mln (https://carbon-pulse.com/189380/) .

In response to Carbon Pulse inquiries, Verra CEO David Antonioli said that the firm had not considered switching away from the present $/credit revenue model.

“If someone were to show up and write me a cheque to run the organisation on a yearly basis for many years to come, certainly we’d consider it, but the reality is that you’ve got to keep the lights on, and you’ve got to be able to hire good people,” Antonioli added.

Jamie Ballantyne, Gold Standard director of marketing and communications, similarly defended the offset certifier’s revenue model, citing the absence of any philanthropic, government, or private actor that would fund private companies towards a commercial market mechanism.

“The standard model for certifying commodities is a levy per unit,” Ballantyne said in an extensive emailed statement to Carbon Pulse.

Ballantyne detailed that Gold Standard saw no evidence that its current fee model incentivised over-crediting, while outlining safeguards put in place such as independent approval of methodologies by a non-remunerated Technical Advisory Committee (TAC) and independent certification decision making.

Ballantyne also highlighted Gold Standard’s membership of ISEAL Alliance – a leading voice on governance in sustainability systems.

Mary Grady, president and CEO of the Environmental Resources Trust (ERT) at Winrock International – ACR’s operator – also referred to the $/credit model as being common practice across global carbon markets.

“While we are always open to exploring new approaches, we are not aware of any alternative revenue models that would be acceptable in the market, nor do we believe that an alternative would be free from similar criticism of incentivising certain outcomes,” Grady said, refuting the notion that $/credit revenue models were to blame in over-crediting.

ACR does not earn activation fees on the offset volumes associated with credit deductions for uncertainty, leakage, or buffer pool contributions, but rather only on emission reduction or removals credits that are issued and eventually activated for transfer or retirement, Grady noted.

Additionally, all ACR methodologies are approved through a blind, scientific peer-review, independent process to ensure the rigour in the carbon accounting, and all projects are independently verified by an accredited validation and verification body (VVB) for conformance with the ACR standard and applicable methodology, Grady reiterated.

While ACR states it issues offsets for free, the firm’s $0.15/offset activation fee for ACR-issued credits, in addition to $0.02/offset in transfer fees and $0.02/offset in retirement fees, are essentially $/credit fees generated through the issuance/retirement process.

CAR has not responded to Carbon Pulse questions by time of press on whether the $/credit issuance model facilitates over-crediting.

“A RACE TO THE BOTTOM”

Although registries point to numerous safeguards built in the system to suppress the perverse incentive to over-credit, critics see deeper flaws with the process.

“I think in either a [for]-profit model or a non-profit model, if your payment depends on something that requires your own assessment that’s just a basic problem,” said Peter Riggs, director at non-profit climate justice advocacy Pivot Point.

The current revenue model enables and is what drives what Riggs called a “race to the bottom”, which he believes reveals companies that are complicit in that kind of behaviour, but also showcases companies that avoid holding questionable, “crap credits”, instead opting for higher quality offsets.

Thomas Day, researcher at think-tank NewClimate Institute, echoed Riggs’ sentiment that the $/credit revenue model facilitates the race to the bottom mentality.

“The business model of $/credit for offsetting claims was supposed – in theory – to lead to the maximum implementation of mitigation projects in the most cost-effective manner,” Day said.

“In reality, it creates a race to the bottom for the integrity of carbon credits and the credibility of offsetting claims; over-crediting is just one of the many ways in which this [issue] manifests,” he added.

Incentives lie not just with project developers creating protocols and methodologies, the non-profit registries remunerated by approval of protocols and offset issuance revenues, but everybody involved in the system getting paid for certifying offsets, Renoster’s Ayrey argued.

“Project developers, put forward their own numbers, verifiers considered independent, get paid, of course, by the project developers to do this verification, and there’s no real incentive for them to find flaws,” he said.

In Cullenward’s view, the absence of any meaningful VCM regulatory framework or enforcement agenda was more important, but the $/credit business model was still a factor that encouraged this race to the bottom.

“I see the methodologies and revenue models as symptoms of broader quality problems, not the root causes of those problems. That said, a system in which financially inte arket proponents write their own rules, subject to ‘oversight’ from registries that take per-credit fees, is one in which no one has an incentive to protect the atmosphere,” Cullenward stated.

However, Franck Gbaguidi, senior analyst, energy, climate & resources at political risk consultancy Eurasia Group, held a slightly differing opinion that the $/credit revenue model was not a factor per se causing over-crediting.

"The underlying issue is threefold: unreasonable and imperfect baseline assumptions that are triggering over-crediting, lack of methodology standardisation among VCMs that is further fuelling over-crediting, and insufficient review and monitoring of current methodologies," Gbaguidi told Carbon Pulse in an emailed statement.

The Integrity Council for the Voluntary Carbon Market (ICVCM) this month intends to publish (https://carbon-pulse.com/188196/) its Core Carbon Principles (CCPs) and Assessment Framework for identifying high-quality carbon credits.

Existing offset standards bodies will need to receive accreditation both at the programme level and crediting level in order to receive CCP certification.

POTENTIAL SOLUTIONS

While the $/credit issuance model has dominated over the lifespan of the VCM, several observers offered potential fixes that could reduce the risk of over-crediting going forward.

Laurie Wayburn, president at non-profit conservation land trust Pacific Forest Trust, said the offset registries that provided accounting to compliance systems had some level of oversight that was the basis for being more reliable.

For example, California regulator ARB has its own set of compliance protocols to vet registry offset credits (ROCs) from ACR and CAR for eligibility under the state's WCI-linked cap-and-trade programme, though researchers recently found substantial over-crediting (https://carbon-pulse.com/194477/) within the scheme's improved forest management methodology.

Wayburn proposed a system similar to the US "ENERGY STAR" programme, which would likely be developed and maintained by the US Environmental Protection Agency (EPA), even for international projects.

The US EPA awards ENERGY STAR certification to products that meet certain criteria of energy efficiency, scoring 75 or higher on EPA's 1–100 scale.

This would mean registry-approved methodologies would have to pass a quality test prior to credit distribution, and have a layer of oversight on issuance levels.

Riggs suggested minimising the role of intermediaries in the VCM issuance process as a means to reduce the incentives that promoted over-crediting.

The key term, Riggs stressed, was "disintermediation", which contrasted against the $/credit revenue model of offset certifiers.

Riggs believed the frontier for innovation was in formulating agreements with community monitoring protocols, and that part of the reporting was not just about carbon but about livelihoods.

Carbon indicators, biodiversity indicators, and livelihood indicators, all in one package was a solution Riggs floated, with the involvement of Indigenous people to co-design and co-manage these products in ways that were consistent and enhancing of their livelihoods.

"The scale needs to tip towards the Article 6.8 non market approaches, with a global registry, all transactions taking place on a public registry, no dark trades, and better limits on retirement or encouraging retirement, but also the subsequent use of credits," Riggs suggested.

Article 6.8 is a yet-to-be-operationalised provision that recognises non-market cooperative approaches among nations to promote mitigation and adaptation, with specifications for a web-based platform to record and exchange information on non-market cooperation.

In this scenario, Riggs was of the view that there would be less need for registries' skill in external verification of carbon, but the standards' community engagement and land management competence would remain relevant.

Steve Suppan, senior policy analyst at non-profit advocacy Institute for Agriculture and Trade Policy (IATP), believed a shift to jurisdictional crediting with no more individual offset issuances – but rather an averaging of credits – would drive more standardisation in the VCM.

A more robust use of buffer accounts could also have some impact on balancing instances of over-issuance, Suppan added.

To Renoster's Ayrey, the solution to reduce VCM over-crediting was simple.

"Somebody needs to be paid throughout the process to be financially incentivised to find issues with these projects or issues with the protocols" Ayrey said.

As an example, Ayrey pointed to Google's bug bounty programme that paid hackers to submit flaws in the software giant's applications.

"I think the fundamental question is whether incumbent registries are capable of raising the bar, given their long history in, and financial commitment to, problematic market segments, or whether it will take new entrants to change the dynamics," Cullenward concluded.

Gbaguidi said there was still a large consensus among VCM market participants that non-profits were best placed to run VCM registries given their legal status, longstanding advocacy efforts, and internal know-how.

"But their credibility will be undermined if overestimation issues continue to arise," Gbaguidi cautioned.

"This threat will lead many registries to pause, regroup, and work toward more high-quality, high-integrity credit schemes. They will do so because their very existence is at stake."

Non-exhaustive registry fee breakdown:

Fee Category	Verra (https://verra.org/wp-content/uploads/Program-Fee-Schedule_v4.1.pdf)	GS (https://globalgoals.goldstandard.org/fees/)	ACR (https://americancarbonregistry.org/how-it-works/membership/acr-fee-schedule/acr-fee-schedule-january-2021.pdf)	CAR (https://www.climateactionreserve.org/how/program-resources/program-fees/)
Account fees	$500 account opening $0.10/credit, max	$1,000 account annual $2,500 account reactivation	$500 account opening $500 annual account	$500 account set up $500 reactivation

	$10,000 registration			$500 annual maintenance $200 project account set-up $200 project account maintenance annual
Issuance fees	$0.025/credit >10 mln $0.05/credit <10,000 $0.14/credit 10,001–1 mln $0.05/credit conversion $1,500/event retroactive label	$0.15/credit GS VER $0.30/credit GS VER subsequent issuance $0.05/credit GS CER $0.10/credit GS CER subsequent issuance $0.15/credit renewal period annual avg.	'free' issuance $0.15/offset activation $0.02/offset transfer $0.03/offset cancellation $0.02/offset retirement	$0.19/credit issuance $0.03/credit transfer $0.03/credit cancellation
Project/Methodology fees	$2,000 methodology application, non-refundable $13,000 methodology processing $1,500 methodology revision application $6,000 methodology revision processing	$500–$3,500 preliminary review $0.05–$0.15/credit project design review $650–$4,500 performance review $0.10/credit or $500 design change (whichever greater) $1,000 deviation reconsideration	$2,500 review new methodology $7,500 review new & modified ACR methodology $750 ARB CCO $1,500 ARB ODS second verification $1,000 project listing ACR methodology $0.08/offset project transfer	$700 ARB CCO $500 Reserve protocol $1,350 project variance review $500 project transfer
Expert review fees	$375	$50/hour additional review $2,500 expedited review	TBD scientific peer-review	
Validation Verification Body fees	$2,500 annual	$5,000–$20,000 validation $2,500 additional VPA validation $1,500–$2,500 annual verification		

Source: Verra, GS, ACR, CAR; excluding discounts and rebates.

By Joan Pinto – joan@carbon-pulse.com

Harvard
Business
Review

Business And Society

What Every Leader Needs to Know About Carbon Credits

by Varsha Ramesh Walsh and Michael W. Toffel

December 15, 2023

Tom Penpark/Getty Images

Summary. Many companies have begun to look into credits to offset their emissions as a way to support their net zero goals as their target years get closer and closer. As it stands, the carbon credit market is too small to bear the brunt of reducing companies' impacts on the... **more**

In the absence of government regulations requiring dramatic reductions of greenhouse gas (GHG) emissions that are causing climate change, a growing number of companies are adopting "net zero" targets. More than one third of the world's 2,000 largest publicly held companies have declared net zero targets according to Net Zero Tracker, a database compiled by a collaboration of academics and nonprofits. These targets typically entail public commitments to reduce GHG emissions through measures such as process modification, product reformulation, fuel switching, shifting to renewable power, investing in carbon removal projects — and a pledge to zero-out their remaining emissions by purchasing carbon offsets, also known as carbon credits. Carbon credits are financial instruments where the buyer pays another company to take some action to reduce its greenhouse gas emissions, and the buyer gets credit for the reduction.

As companies creep closer to their net zero target years, many have already begun purchasing carbon credits. The market for carbon credits is projected to grow 50-fold within a decade, from nearly $2 billion in 2022 to nearly $100 billion by 2030, and as much as $250 billion by 2050, according to Morgan Stanley. But navigating the world of carbon credits creates brand risk because the market remains immature and complex, with wide variation in project types, developers, location, and cost, resulting in unclear quality, transparency, and credibility.

Companies routinely choose to purchase rather than produce goods and services that other companies can create more inexpensively, and this decision doesn't often attract the attention of activists or the media. Not so for carbon mitigation: Activists are vocal about how companies choose to meet their net-zero goals. Corporate carbon mitigation plans viewed as overly reliant on buying carbon credits rather than making carbon reductions to their own operations and supply chains risk being accused of not being sufficiently serious about decarbonization and seeking to "buy their way out" of meaningfully achieving their goals. In part, this is because the carbon credit market is far too small to accommodate the dramatic carbon reductions necessary to meet companies' net-zero goals or for the world to reduce GHG emissions by 45% by 2030 and reach net zero by

2050 that the UN claims is necessary to avoid the worst effects of climate change by limiting the global average increase to 1.5 degrees Celsius. Questions about credits' credibility abound, including whether they deliver on their promise to reduce GHGs, whether any such reductions will endure, and whether the project would have occurred even without the sale of carbon credits. From John Oliver's claim that "offsets are bullshit" to the *Guardian* calling some carbon credits purchased by Disney, Gucci, and Shell "largely worthless," some offsets receive charges of "greenwashing" — environmental performance claims that outstrip reality. That's hardly the reputation boost firms seek.

Yet, the voluntary carbon market has the potential to drive billions of dollars over the coming decade into climate solutions, creating along the way an estimate of cost-of-carbon in goods and services. What's the best way to participate in the market when carbon credits claiming to avoid or remove one metric ton of GHG range in price from nearly $2 per ton to $1,800 per ton? Which types of credits are considered to be the highest quality, and thus least likely to lead your company to be named and shamed? Despite the emergence of standards and registries meant to inject confidence in the market, many quality concerns remain. Leaders need guidance to apply due diligence to decisions regarding the carbon credit market.

What projects create carbon credits?

Carbon credits are created from projects that avoid the generation of GHG emissions or that remove GHGs from the atmosphere. These projects include "nature-based solutions," such as reforestation and regenerative agriculture efforts, and "engineered solutions," such as combusting methane emitted from landfills to generate electricity and direct air capture.

Examples of Carbon Credit Projects

This table illustrates the differences between nature-based and engineered solutions for both carbon emissions avoidance and carbon removal.

	Carbon emissions avoidance	**Carbon removal**
NATURE-BASED SOLUTIONS	Preservation of forest land to avoid its conversion into farmland	Regenerative agriculture practices that sequester (embed) atmospheric carbon into soils and vegetation
ENGINEERED SOLUTIONS	Carbon capture and storage of GHGs from smokestacks at coal- and natural-gas-fueled power plants and other types of factories New solar-and wind-power plants that substitute for fossil fuel electricity Combustion of stockpiles of ozone-depleting substances that would otherwise leak into the atmosphere Combustion of methane emissions from landfills	Direct air capture of GHGs from the atmosphere with deep-well storage

Source: Varsha Ramesh Walsh and Michael W. Toffel

Some companies focus only on some of these types. Microsoft, for example, invests only in carbon removals. Others create a portfolio across the spectrum, such as Delta's $137 million investment in carbon credits that include REDD+ (reducing emissions from deforestation and forest degradation) avoidance credits, avoidance credits from solar and wind-power projects, and removal credits including afforestation and carbon capture and storage.

Who are the players?

Unlike with stock exchanges, carbon credits lack widely adopted standards and large centralized marketplaces. This makes it difficult to find, understand, and compare carbon credit projects.

Instead, leaders have to navigate a maze of various standards and players with frustratingly overlapping roles. There are numerous carbon credit registries and standards bodies that provide minimum requirements for various project attributes and in some cases list projects that meet their own standards.

- **Carbon credit verifiers**, also known as validation/verification bodies (VVBs), assess whether projects meet certain standards. They range from global companies to niche players that focus on just one type of project.
- **Carbon credit brokers and marketplaces** connect buyers with project developers. Some list projects they helped finance and develop, raising the potential for conflicts of interest.
- **Carbon credit ratings agencies** assess carbon credit projects along various dimensions, including but not limited to the attributes featured in standards. They tend to sell their ratings via a subscription model to prospective credit buyers. These ratings agencies provide much-needed transparency and convey key attributes of the projects they rate.

With so many players and many standards, it's no wonder companies find it difficult to navigate the landscape. The Voluntary Carbon Markets Integrity Initiative or Oxford Net-Zero Aligned Offsetting Principles provide holistic carbon credit and offsetting principles and are a great place for leaders to start, but even these are updated periodically to keep pace with the changing landscape.

Examples of players in the carbon credit marketplace

- **Carbon credit registries and standards bodies**: VERRA's Verified Carbon Standard (VCS), Gold Standard, Climate Action Reserve, American Carbon Registry (ACR), Puro.Earth, and Isometric
- **Carbon credit and offsetting principles**: Integrity Council for the Voluntary Carbon Market's Core Carbon Principles and Oxford Net Zero-Aligned Offsetting Principles
- **Carbon credit verifiers**: SCS Global Services and DNV
- **Carbon credit brokers and marketplaces**: 3Degrees, Cloverly, Lune, Patch, South Pole Group, and Terrapass
- **Carbon credit ratings agencies**: BeZero, Calyx, and Sylvera

What makes for a high-quality carbon credit?

It's crucial to understand the key dimensions that differentiate the quality and brand risk associated with different types of credits. While the most reputationally sensitive organizations might choose to pursue high quality along all dimensions, others are more willing to make tradeoffs. Some buyers prefer credits related to their industry — such as food companies that prefer nature-based approaches in agriculture — and are willing to accommodate higher permanence risk, or the risk that a project's climate benefits might only be temporary rather than permanent. Some companies prefer credits generated from newer technologies, hoping that early purchases encourage them to continue developing their technology to scale up and reduce their costs.

While standards bodies and registries have attempted to clarify what quality means, they have not kept up with the pace of innovation and explosive growth of the carbon credit market. Some of the older methodologies have accumulated limited proof of climate impact, and projects with novel technology are meanwhile trying to measure and communicate their climate impact with scientific rigor. The proliferation of standards has created conflict and confusion.

Below, we describe five attributes that constitute high-quality carbon credits and discuss the co-benefits.

Additionality

This refers to the idea that the carbon credit project would not have occurred without the expected revenue from selling the carbon credit. For a project to be additional, the revenues from selling the carbon credits must play a decisive ("make or break") role in the project developer's decision to implement the project. Evaluating additionality is often subjective: Projects are reviewed by registries, scientists via the peer review process, and third-party evaluators to determine if the project can demonstrate financial, technological, or institutional additionality.

For an example where additionality is clear, consider startup Heirloom Carbon Technologies. It uses limestone to extract carbon dioxide from the atmosphere and relies on the promise of carbon credit revenues to fund its capital-intensive process. Stripe Climate and other leaders in the carbon dioxide removal space have expressed high confidence that Heirloom's carbon credits exhibit additionality. For some companies, such as Google, additionality is the main attribute they look for when selecting carbon credits. "If Google hadn't invested in an offset project, it wouldn't have been built. We want to make sure that those tons of carbon wouldn't have been avoided or sequestered hadn't it been for our investment," said Anna Escuer, the lead for water and carbon at Google. She also noted that additionality enables companies to claim “that it was [their] money that really made that project happen.”

One caution: Additionality can be a moving target. Innovations that reduce technology costs and public policies such as new subsidies can mean some types of projects to no longer meet additionality tests. For example, a particular project developed a few years ago that passed additionality tests and thus created carbon credits might, if launched today, become profitable even without selling credits — meaning it would not be additional, thus would not be able to issue carbon credits.

The takeaway: Pay close attention to additionality claims. Carbon removal projects are more likely to be additional because their carbon credits more clearly depend on credit revenue.

Quantification

This refers to the method a project uses to determine the volume of carbon dioxide reduced, avoided, or removed. Carbon credits are supposed to represent one metric ton of carbon dioxide equivalent gas. How do we know that the quantity reported by a project is accurate? Project methodologies define standard quantification approaches, and these are used to establish both a baseline scenario — how much GHG emissions would be released without the project — and to estimate how much less GHG emissions result due to the project, which determines how many carbon credits the project generates.

Often, project developers describe in their calculations any deviations they made from those standard approaches to reflect unique aspects of their process. Even the most rigorous of projects tend to require some degree of estimation. Projects address this by highlighting potential discrepancies and building adequate buffer pools. Carbon credit buyers can also mitigate the risk of overestimating carbon benefits by purchasing more credits than they need in order to build in a buffer for reduction claims such as achieving net zero targets.

The takeaway: Understand the quantification risks that your carbon credit projects entail, including any deviations and discrepancies from standard quantification approaches, and develop strategies to mitigate those risks, such as overpurchasing or building in a buffer pool.

Leakage

This is the risk that emissions avoided or removed by a project are pushed outside the project boundary. For example, if you are paying to avoid deforestation, does the project's cordoning off some land simply shift the deforestation pressure to adjacent land? Do projects that entail the destruction of stored ozone-depleting chemicals spur the manufacture of new ozone-depleting chemicals that wouldn't have otherwise been produced? High-quality projects account for many potential sources of leakage and quantify their impact into the number of credits they generate.

The takeaway: Similar to quantification, look for projects that take leakage into account and develop a sound approach to incorporating it into your quantification estimates. Assessing these project-specific details requires technical expertise that credit buyers can invest in or can ascertain from assessments provided by carbon credit rating agencies.

Permanence risk

This refers to the possibility that the GHG emissions avoided or removed from the atmosphere as a result of a carbon credit project might be temporary rather than permanent benefits, with a "reversal" resulting in such carbon being released into the atmosphere. For example, consider carbon credits sold based on the emissions avoided because a project prevented some hectares of forest from being cut down. It can be quite difficult to estimate the probability that the forest will never be cut down — or perhaps catch fire — in the future, either of which would undo the project's carbon benefits. On the other hand, carbon credits created from a project that utilizes direct air capture and stores that carbon in a sealed well are very likely to remain there indefinitely.

Many standards set the permanence target as 100 years, with Frontier Climate setting the most stringent standard: 1,000 years. Carbon credits with very high confidence of permanence — that is, that exhibit low permanence risk — exist but are in short supply and tend to be quite expensive. Due diligence efforts should assess the likelihood that policymakers might authorize and even encourage policy changes that can override today's permanence claims, such as future policies that foster the development of lands previously conserved via carbon credit projects.

To manage permanence risk, some carbon credit providers include a form of insurance: They set aside a percentage of the carbon credits generated by a project to hold in reserve and offer as compensation if reversals occur. Alternatively, buyers can self-insure by purchasing more credits than they currently need.

The takeaway: When purchasing carbon credits with permanence risk, buyers can manage reversal risk — that carbon benefits might turn out to be temporary — by purchasing insurance or by self-insuring, and should include this insurance cost when comparing carbon credits.

Vintage

This refers to the year that credits were issued by the project. Buying credits that were recently issued and/or that were generated from projects that were recently launched increases the odds that your funds will go towards organizations that are actively innovating and launching projects. In addition, these credits are more likely to have met more stringent voluntary certification standards, which in general have become more rigorous over time as scientific knowledge has evolved. In fact, some crediting schemes — such as the aviation sector's CORSIA program — only allow for project vintages starting at 2016.

The takeaway: Recent vintage carbon credits can take advantage of improved standards and can reward and encourage active project developers focusing on high-quality credits.

Co-benefits

Some carbon credit projects not only avoid or remove GHG emissions; they also provide co-benefits such as reducing health problems, enhancing biodiversity and water quality, and creating jobs. For example, the Yagasu Project that restored mangrove forests in Sumatra, Indonesia to sequester carbon also created a village revolving fund and empowered women to serve in management roles and to participate in town hall meetings to discuss local business development.

Such co-benefits can provide tangible publicity benefits to credit buyers. For example, Infosys describes the co-benefits of investing in biogas and cookstove carbon credit projects in India as generating "over 2,400 jobs" that benefit "more than 102,000 families."

Purchasing carbon credits with co-benefits can also provide confidence that your funding had social *and* environmental impact. Equinor, which recently bought carbon credits in partnership with Sylvera, strategically chose a portfolio of different types of carbon credit projects, prioritizing those with co-benefits.

The takeaway: Carbon credit projects that produce co-benefits can offer enhanced reputational benefits and diversify the impact of the project — but don't eliminate the need to understand carbon credit quality.

...

With the enormous growth in companies' use of carbon credits to meet their sustainability goals and carbon footprint reduction targets, there is an increasingly wide range of projects that are generating carbon credits from which to choose. Companies should choose the carbon credit strategy that best meets their objectives. Purchasing high-quality credits reduces the risk of negative publicity and greenwashing charges and bolsters the odds that the carbon you think you are avoiding or removing is *actually* being avoided and removed — in both the short and long terms.

Editor's note (12/22/23): CarbonPlan, which does not create or sell ratings of credits or offset projects, has been removed from the list of carbon credit ratings agencies.

VW

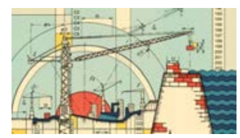
Varsha Ramesh Walsh is the Founder and CEO of Offstream, a platform for seamless carbon compliance management, and has worked on carbon removal solutions at Indigo Agriculture and Patch.

Michael W. Toffel is the Senator John Heinz Professor of Environmental Management at Harvard Business School, host of the HBS Climate Rising podcast, and co-lead of the HBS Online Business and Climate Change course.

Recommended For You

How ENGOs Can Support Corporate Climate Change Efforts

Accounting for Carbon Offsets



Creating High-Impact Coalitions

PODCAST
Can Remote Surgeries Digitally Transform Operating Rooms?

Received: 21 January 2022 | Accepted: 28 February 2022

DOI: 10.1002/aepp.13254

SUBMITTED ARTICLE

AAEA Agricultural & Applied Economics Association WILEY

Challenges to voluntary Ag carbon markets

Oranuch Wongpiyabovorn[1] | **Alejandro Plastina**[1] | **John M. Crespi**[1,2]

[1]Department of Economics, Iowa State University, Ames, Iowa, USA

[2]Center for Agricultural and Rural Development, Iowa State University, Ames, Iowa, USA

Correspondence
Alejandro Plastina, Department of Economics, Iowa State University, 478E Heady Hall, 518 Farm House Lane, Ames, IA 50011-1054, USA.
Email: plastina@iastate.edu

Funding information
Center for Agricultural and Rural Development at Iowa State University; USDA National Institute of Food and Agriculture, Grant/Award Number: Hatch Project IOW04099

Editor in charge: Daniel Petrolia

Abstract

The credibility of agricultural carbon credits will play a critical role in the determination of payments received by farmers through voluntary carbon markets. This article analyzes the major challenges from both the demand and supply sides to voluntary agricultural carbon credit programs and serves as a resource to researchers, producers, policymakers, and other stakeholders who seek a comprehensive analysis of the challenges that still face this market despite recent positive developments in global agreements.

KEYWORDS

agriculture, carbon credits, carbon market, carbon removal, greenhouse gasses

JEL CLASSIFICATION

Q20, Q28, Q58

In a market economy, prices carry signals for both producers and consumers. When greenhouse gas (GHG) emissions are costless to the emitter, there is little incentive to incorporate the environmental and social costs imposed by the emissions (the externalities) into the emitter's decision-making process (internalize the externalities). The World Bank's Carbon Pricing Dashboard (World Bank Group 2021) lists 65 carbon[1] pricing programs around the world, including mandatory and voluntary emissions trading systems (ETSs), GHG taxes, and combinations of ETS and taxes, covering 21.5% of GHG emissions worldwide and generating \$53 billion in revenue. In the United States, two mandatory ETSs are currently in place (California's cap-and-trade program and the Regional Greenhouse Gas Initiative), and one is scheduled to launch in 2023 (Washington's cap-and-invest program). Additionally, 11 voluntary agricultural carbon credit programs have started or are in the pilot period (Plastina & Wongpiyabovorn, 2021).

Arguably, the successes of voluntary carbon markets thus far have ranged from modest to none. An example of the latter is the Chicago Climate Exchange[2], which was closed in 2010 after 7 years of activity due to lack of trading volume and very low carbon prices. An example of the former is the Kyoto Protocol,[3] in legal effect for 37 nations since 2005 and stalled since 2012 when the Doha Amendment could not be approved. The Kyoto Protocol was plagued with uncertainties on how to avoid double counting of emissions reductions and on how participating nations could generate and use certified emission reductions (CERs)[4] to meet the Kyoto target.

Nevertheless, voluntary carbon markets received a strong boost from the leaders of almost 200 nations in December 2021 through the Glasgow Climate Pact,[5] subscribed at COP26 and intended to initiate the transition from the old Kyoto Protocol regime to the instruments of the 2015 Paris Agreement.[6] In particular, the Pact proposes a new carbon crediting mechanism based on standardized modalities, procedures, and guidelines (MPGs) to issue and report the use of carbon credits from emission-reducing activities (UK Government 2021). This new mechanism, the United Nations Framework Convention on Climate Change (UNFCCC), is expected to increase international cooperation and transparency, and help countries achieve their national climate action plans under the Paris Agreement via trade of international carbon credits and linking of existing emission trading systems (ETSs).[7] The bodies that supervise CERs and the UNFCCC mechanism will meet in 2022 to start to deliver the rules for the new projects and the processes for transitioning existing projects into the new system under the Paris Agreement (UNFCCC, 2021).

If nations are now mostly in agreement on the MPGs for generating carbon credits, a goal that eluded the parties to the 2015 Paris Agreement, voluntary carbon markets may overcome much of their previous participation disincentives. Standardization and agreed-upon metrics for practices and outcomes could translate into an opportunity for the agricultural sector worldwide, as agricultural practices could be adapted to reduce GHG emissions (e.g., using nitrogen stabilizers or nitrification inhibitors to avoid the release of nitrous oxide from fertilization applications into the atmosphere in crop production) and to remove GHGs from the atmosphere (e.g., switching from conventional tillage to no-till crop production, or using cover crops) in conformity with the new UNFCCC mechanism.

While the current supply of agricultural carbon credits is very limited, the number of voluntary programs to generate carbon credits in the agricultural sector grew rapidly in recent years, in tandem with "net zero" GHG emissions pledges from nations and corporations (Black et al., 2021). However, since carbon credits and offsets are credence goods,[8] scaling up voluntary agricultural carbon markets faces multiple challenges, even if the standardized MPGs are adopted by all voluntary carbon programs. This article analyzes the major challenges from both the demand and supply sides to voluntary agricultural carbon markets and presents four possible scenarios for those markets in the United States.

The remainder of the article is organized into four sections describing demand and supply challenges, parallelisms between carbon markets and organic markets, and alternative scenarios for agricultural carbon, followed by brief concluding comments.

DEMAND-SIDE CHALLENGES

Issues that can undermine a market for credence goods are well known in economics. Where labels or certification are used to verify a claim on a credence good, markets fail in the presence of difficult-to-verify claims, misunderstood or poorly worded labels, the lack of clear, consistent, and uniform guidelines across certifying parties, a lack of trust for certifiers (especially when these are not independent third-parties), and label proliferation (the existence of too many labels in a market or on a good leading to confusion about competing claims). Economists already know much about these issues, as they have examined them in other areas. Giannakas (2002) and Bonroy and Constantatos (2015) examine information asymmetries in the organics markets and conclude that a viable market must have viable certification and undermining of the labels could do great damage to the industry. When Bithas and Latinopoulos (2021) elicited consumers' willingness to pay for carbon sequestration in a stated preference experiment of forest product consumption, they asserted to the respondents that the carbon truly was being sequestered, something that may only be inferred in a real market. In the absence of verification, adverse selection (Akerlof, 1970) may lead to a market failure over a carbon sequestration claim. As is seen in the variety of third-party certifiers in the carbon sequestration market today, the need for verification is already understood.

Consumers would likely not trust the manufacturer to correctly self-report carbon sequestration because it is arduous for consumers to detect whether a firm's suppliers follow carbon sequestration processes—search costs to verify a label are indeed large barriers (Teisl & Roe, 1998). Certification agents (public or private) who specialize in such detection are necessary in cases where the labels signal the production methods, regional sourcing, environmental impacts, safety, or quality of a good. The absence of the label for a desirable attribute creates a "lemons problem" (Akerlof, 1970) where consumers who have a higher willingness to pay for a carbon credit cannot detect the attribute in the absence of a label and will not believe it in the absence of certifier credibility. The market can fail not because of a lack of demand but because of a lack of information. Caswell and Mojduszka (1996) and Marette and Roosen (2011) delve into this issue in the case of food labeling, Crespi and Marette (2003) and Crespi and Marette (2005) examine the issue in the case of public labels and eco labels, respectively, while Roe and Sheldon (2007) and Roe et al. (2014) examine the literature on credence good labels in general.

Without government-backed standards, we should expect questionable carbon claims and an increase in competing claims, so-called "label proliferation." Kiesel and Villas-Boas (2013) and Marette (2014) explore this issue, which arises when products and markets contain multiple labeled attributes. The concern here is a different type of market failure where consumers become so overwhelmed by competing messages that they lower their willingness to pay for an attribute because of the noise. Label proliferation leads to a "crowding out" of desirable attributes similar to Akerlof's lemons problem. In short, in the absence of standards and verification, buyers of carbon credits and the downstream consumers of credit buyers' products or services may be reticent to assign much value to a GHG sequestration or emission reduction claim.

Another challenge in voluntary carbon markets is that entities promising net zero emissions or specific GHG emissions targets usually place the target date a decade or more into the future. While such behavior makes sense from a planning perspective, it also allows those entities to commit some investments at the time of the initial announcement and then postpone further investments until near the target date. The disconnect between long-term voluntary goals and short-term annual purchases of carbon credits or investments in carbon credit generation could result in pent-up demand in years of large announcements, followed by years of low demand and prices, and high demand again in target years. Such cyclicality, combined with the multi-year processes required to produce agricultural credits, could generate incentives to discontinue carbon sinking practices and disrupt the supply of carbon credits prior to the target years.

Although not currently a barrier to the development of agricultural carbon markets, the carbon footprint of the whole system involved in generating carbon credits, including issuance and tracking of the serial numbers for each project in the carbon registries, along with financing projects and trading credits, could become a concern for consumers of carbon credits or the end products or services where carbon credits are applied to reduce their carbon footprint. For example, West and Marland (2002) find that the carbon stored in soil organic matter by reduced-tillage is offset by the GHG emissions into the atmosphere through increased production, transportation, and application of chemicals. Another example is that an afforestation program under carbon markets in a specific region could result in net losses in stored carbon because of the intensification of agricultural production in unregulated regions (Haim et al., 2016). Carbon programs that use energy-intensive accounting and verification systems (e.g., "proof of work" consensus systems in blockchain technology) might generate net positive carbon emissions, and could become less desirable than carbon programs with smaller GHGs footprints. High-quality carbon credits require that leakage (i.e., the increase in GHG emissions outside of the carbon project as a consequence of the generation of a carbon credit) be prevented in the process that leads to their issuance and use.

SUPPLY-SIDE CHALLENGES

Related to the credence attribute of carbon credits, farmers may be reticent to change production practices in order to generate carbon credits of unknown value. Likewise, in the face of an uncertain market, lending institutions may be reticent to fund producers who possibly need specific assets for the production methods applied in the generation of carbon credits.

Accurate measurement and verification of carbon credits from agricultural and forestry activities are typically difficult and costly (van Kooten, 2008). Collecting soil samples and measuring soil organic carbon is currently the most accurate way to gauge the amount of carbon stored in the soil, but it is too costly and time-consuming to be widely used (Castagné et al., 2020). Data collection from satellite mapping and remote technologies may provide an accurate calculation of soil carbon at a lower cost. However, this method is still lacking in terms of roughness, soil moisture, and vegetation cover, which would lead to less robust estimation although advances to the systems are being developed (Angelopoulou et al., 2019).

Voluntary carbon programs currently follow different protocols based on different models to calculate how much carbon is sequestered through the implementation of agricultural practices (Plastina, 2021). For example, while CIBO Impact uses the System Approach to Land Use Sustainability model to calculate carbon credits, Nori and the Soil and Water Outcomes Fund use the COMET-farm model, and Ecosystem Services Market Consortium (ESMC) uses the DeNitrification-DeComposition model and the Operational Tillage Information System model to calculate carbon credits. The complexity involved in comparing potential carbon credits generated by one specific practice in a particular farm across programs could discourage objective technical comparisons of programs and result in farmers choosing programs with the best customer service rather than the highest potential profitability.

Non-additionality is one of the major risks making conservation programs cost-ineffective. Agricultural conservation practices are considered to yield additional environmental gains only if they would not have been adopted without payment. Estimating additionality for selected agricultural practices, Claassen et al. (2018) conclude that the adoption of three off-field structural practices (filter strips, riparian buffers, and field borders) and the elimination of fall application of nitrogen fertilizer was highly additional, while the adoption of conservation tillage was only moderately additional. Sawadgo and Plastina (2021) estimate that cover crops were moderately additional and that over half of farmland in cost-share programs funded cover crop acreage would not have been planted without payments. The 11 voluntary agriculture carbon credit programs analyzed by Plastina and Wongpiyabovorn (2021) require additionality to generate a carbon credit. However, not all programs require that farmers change their production practices since programs use a wide array of benchmarks to determine what is additional or different: some programs require a change of practices with respect to past practices on the same field, while others require that practices in the field be different from common practices in the area (even if the same practices have been implemented for many years in the field under consideration).

Permanence is a major driver of carbon credit quality. Carbon credits generated from Land Use, Land-Use Change, and Forestry (LULUCF) face natural risks such as fire, disease, pest outbreaks, and other natural disasters. In the past, some issued offsets were lost because of emissions reversal. For example, the California Air Resources Board (CARB) retired 178,642 credits from a U.S. Forest project due to unintentional reversal on September 12, 2017 (CARB, 2022). Generating high-quality credits with long-lived carbon storage in the soil is a costly process due to the required changes in farming practices that sometimes reduce productivity—even if temporarily—and the costs to verify and certify the carbon sequestration. For example, no-till could reduce crop productivity, particularly in cooler and/or wetter climatic conditions due to surface residues and lower soil temperatures (Ogle et al., 2012). According to Gramig and Widmar (2018), farmers in Indiana who have never adopted conservation tillage or no-till would require almost a $40 per acre increase in net revenue to implement no-tillage, while individuals who previously used conservation tillage would be willing to adopt with no payment. Gramig and Widmar (2018) also find that an additional $10.57 per acre is needed to enter the program with a multi-year contract that does not allow them to change their tillage practices during the contract term. Having a carbon project certified to generate high-quality carbon credits according to the Gold Standard registry can cost $5000 in one-time validation fees and $3500 per year in annual verification and registry fees (Gold Standard, 2021); large fixed costs could unintentionally impact the market structure of carbon credits (Crespi & Marette 2022). Furthermore, Plastina and Wongpiyabovorn (2021) report that when contracted practices are temporarily discontinued due to factors external to the farm (e.g., weather), some voluntary agricultural carbon programs impose penalties associated with skipping payments for the discontinued practices until reinstated (Soil and Water Outcomes Fund, CIBO Impact) or until additional gains in carbon sequestration are observed (ESMC, Indigo), while at least two initiatives do not report having any penalties for permanent disadoption (Gradable, Bayer).

In the present environment of burgeoning agricultural carbon programs, little attention is paid to the potential effects of alternating adoption, opportunistic adoption, and partial adoption on the total area under conservation practices (Pannell & Claassen, 2020), let alone their limiting effects on the development of voluntary carbon markets. Carbon reversal from disadoption of conservation practices can occur when a participant of a carbon program stops using the contracted practice when the contract expires. Jackson-Smith et al. (2010) studied a single watershed in Utah from 1992 to 2006 finding that 66% of crop production practices implemented were still maintained in 2007, while 32% of the discontinued practices were driven by farmers exiting farming or selling land for nonfarm development. Using county-level data from the 2012 and 2017 U.S. Censuses of Agriculture, Sawadgo and Plastina (2022) evaluate regional patterns of adoption and disadoption of conservation practices in the United States. They estimate that national disadoption rates in cover crops and no-till averaged 15.60% and 39.38%, respectively, between censuses. Plastina and Sawadgo (2021) report that 11% and 33% of the counties in Iowa, Illinois, and Indiana disadopted cover crops and no-till, respectively, reducing their areas in those conservation practices by 25% and 13% between 2012 and 2017. If these percentages are indicative of the probability that farmers participating in voluntary carbon programs could temporarily discontinue contracted practices and trigger penalties from carbon programs, those findings suggest that farmers planting cover crops and using no-till would face non-trivial probabilities of being penalized over the life of a multi-year carbon contract.

Even within a credible verification and certification system mitigating uncertainty in the conversion of agricultural practices into carbon credits, suppliers of agricultural carbon credits will face competition from other suppliers of carbon credits generated in forestry, geological carbon sequestration, ethanol production with carbon capture and sequestration, landfill methane capture and destruction, and multiple other sources. The quality of credible agricultural carbon credits, dependent mostly on the degree of additionality and permanence of the carbon sequestration, will play a critical role in the determination of payments received by farmers (via direct sale of credits to end users and brokers or indirectly via carbon programs that sell credits to investors).

The cyclicality in demand for carbon credits due to strategic behavior by entities with voluntary GHG emissions targets could, as explained above, generate price signals in the early stages of the cycle incentivizing farmers to enroll in multi-year carbon programs, generating an oversupply of credits and a decline in credit prices when demand drops in the middle of the cycle.

Although outside the context of carbon programs, multiple studies examine barriers to adoption of conservation practices and suggest that a diverse combination of economic and agronomic factors, social norms, perceptions of government programs, farm characteristics, land tenure factors, and knowledge-related factors pose barriers to conservation adoption (Nowatzke & Arbuckle, 2018; Prokopy et al., 2008, 2019; Ranjan et al., 2019).

A further barrier to participation in carbon programs is the lack of transparency in the price discovery mechanism for participating farmers. Farmers and ranchers interested in carbon programs are currently being offered anywhere between $10 and $40 per acre to implement practices that will generate carbon credits, but prices will be subject to market fluctuations beyond pilot programs (Plastina & Wongpiyabovorn, 2021). In March 2020, the CME Group began trading CBL Global Emission Offset (GEO) futures contracts. The aim of these futures contracts is to help manage risk in carbon prices and establish a global pricing benchmark for the voluntary emissions offset market (CME Group, 2021a). In August 2021, the CME Group also started trading futures contracts for offsets generated from agriculture, forestry, and other land use, called Nature-Based GEO (N-GEO). To ensure the transparency of N-GEO futures, only the offsets from Verra's Verified Carbon Standard for Agriculture, Forestry, and Other Land Use projects and/or the Climate, Community, and Biodiversity Standards are accepted for trading (CME Group, 2021b). As of January 19, 2022, the prices of GEO and N-GEO futures expiring in December 2022 were $8.16 and $16.07 per metric ton of carbon dioxide-equivalent (MtCO2e), respectively. Trading volumes in December 2021 averaged 102 and 172 contracts per day (equivalent to 0.1 and 0.17 million MtCO2e) for GEO and N-GEO futures, respectively, with an open interest of 2460 and 6792 contracts at the end of the month. The lack of "hard" caps on GHG emissions in voluntary programs and the small number of carbon credits traded, the cyclical pattern of demand for carbon credits, and the resulting lack of volatility to attract speculators that inject liquidity in the market are major reasons to be skeptical about the ability of GEO and N-GEO futures to serve as a pricing benchmark for voluntary agricultural offsets (Wongpiyabovorn et al., 2021).

Conservation practices not only sequester carbon and reduce GHG emissions, but they also benefit farmers by reducing soil erosion, improving water infiltration, soil water storage, and soil quality. In addition, cover crops and proper nutrient management could improve water quality by reducing nitrate leaching and phosphorous runoff to nearby water bodies. However, the co-benefits from adopting these practices are uncertain and take time to develop. For example, while no-till can be profitable depending on crop rotation and location (Al-Kaisi et al., 2015, 2016; Pendell et al., 2007) and might increase the incentives to use corn stover for ethanol production (Petrolia, 2008), its adoption might take more than 5 years to yield reduced soil erosion and sediment loss to water and wind and increase water-storage capacity (Toliver et al., 2012). If policymakers choose to incentivize farmers' participation in carbon and ecosystem services programs through subsidies or cost-share programs, it is important to keep in mind that uniform payments across geography and/or based on adopted practices are not cost-effective to deliver desirable environmental outcomes (Khanna, 2017). Secchi and Jones (2021) propose that the government use subsidies to support long-term or permanent practices, such as land retirement and reforestation due to their associated water quality and habitat co-benefits, rather than investing in carbon capture and storage projects at ethanol plants.

As long as buyers of agricultural carbon credits perceive differences in the quality of credits generated through alternative protocols, it can be expected that some programs will gain market share, affecting systemic risks for farmers and credit buyers (Plastina & Wongpiyabovorn, 2021). The risk to farmers could be partially mitigated through the standardization of equivalences for carbon farming practices across programs, and the introduction of transferable partial and full credits across protocols. However, the risk of a shorter-than-expected permanency of a carbon credit triggered in the event that a program exits the market and farmers who sold credits through that program discontinue the practices before the expiration of the retention period is only partially mitigated in a few programs through retained carbon credits. Credit reversals are a liability for which there is no insurance policy currently available.

Finally, since geopolitical borders are less relevant for the global atmospheric balance of GHGs than for trade, agricultural producers in different countries will face different incentives even if the law of one price holds for carbon credits. Carbon farming will tend to attract more attention in countries with weaker exchange rates and lower real incomes, and with more carbon-intensive agricultural practices. For example, the carbon program Boomitra –supported by the CGIAR, Yara International, and Chevron, among others– has the mission to remove GHGs from the atmosphere at scale and at the lowest cost, by focusing most of its projects on developing countries: farmers under the international poverty line burning crop residues in open fire pits for cooking and heating are not only able to generate carbon credits from changing agricultural practices but also from adopting "clean cooking" practices (i.e., using stoves and modern fuels).[9] These farmers could not only remove or avoid comparatively more units of CO2e emissions per acre than farmers in developed countries but also do so at a lower cost in U.S. dollars and with a relatively larger boost to their real household incomes. Depending on the relative costs of measuring, verification, and reporting systems across countries, farmers in developing countries could capture a substantial market share of the global market for agricultural carbon credits at the expense of farmers in developed countries.

A WAY FORWARD FOR VOLUNTARY AG CARBON PROGRAMS

A textbook example of overcoming a market failure for credence goods is the case of U.S. organic markets before and after certification. Prior to specific standards for production, the market for organics was very small with lenders reluctant to finance operations. Once standards were set and claims were verified, many farmers overcame their reluctance to join the industry, consumers overcame their distrust of product claims, wholesalers overcame their reticence to broker the goods, retailers devoted space to the items, and lenders had a greater understanding of the needs of producers in this new market (Giannakas, 2002; Jones et al., 2015; Klonsky & Smith, 2002; Kostandini et al., 2011).

In the international arena, the Glasgow Climate Pact described in the introduction intends to set the standard for the international trade of carbon credits. In the United States, a major piece of legislation in support of increasing transparency and standardization in voluntary agricultural carbon programs is the Growing Climate Solutions Act of 2021 (GCSA), passed by the U.S. Senate on June 24, 2021. If ratified by the U.S. House of Representatives, the GCSA will assist farmers, ranchers, and private forest landowners with participating in voluntary carbon markets and adopting conservation practices. Particularly, the legislation will provide the U.S. Department of Agriculture (USDA) authority to create a GHG Technical Assistance Provider and a Third-party Verifier Certification Program. Although the bill does not specify any details about carbon markets, it instructs the Secretary of Agriculture to provide necessary definitions of the markets and determine the rules for the certification program (Crespi & Tidgren, 2021). An effort to standardize or create equivalencies to the amount of carbon credit generated by the same practice in the same farm across private programs would add transparency and reduce systemic risks for participants.

ALTERNATIVE SCENARIOS FOR AG CARBON

Considering the functioning of voluntary carbon markets and the challenges described in the previous sections, we propose four possible scenarios for the future of voluntary agricultural carbon credits, based on the level of corporate demand for and the value of agricultural carbon credits received by farmers. Depending on the international implementation of MRV systems, the four basic scenarios can be combined to represent mixes of vertically differentiated (high vs. low quality) agricultural carbon credits in the global carbon market.

Scenario 1: Carbon farming is the next cash crop

If corporate demand for carbon credits is high and sustained, and agricultural carbon credits are traded at high values, then the carbon market will generate a valuable and stable source of revenue for participating farmers. A credible MRV system for agricultural carbon credits is necessary to achieve this scenario, as well as limited competition from industrial carbon sinks, forestry, and other sources (either via limited quantities at similar prices, or via a segmented market for carbon credits with different prices).

This scenario assumes large-scale adoption of practice changes according to production protocols that generate high-quality credits, and puts the agricultural sector at the forefront of global warming mitigation. Sustained demand for agricultural carbon credits and widespread farmer participation would result in liquid markets with moderate price volatility supported by robust financing and adequate risk-management services for farmers and purchasers of credits. Scenario 1 would be reinforced by the development of complementary value chains for low-carbon commodities that trade at a premium over conventional commodities, as well as by articulated protocols that would allow producers to migrate across carbon programs.

Scenario 2: Low-hanging fruit only

If corporate demand for carbon credits is high but the perceived quality of agricultural carbon credits is low, then agricultural carbon markets will likely be small and underdeveloped. A necessary condition for Scenario 2 to exist is that competition from other sources of low-value carbon credits be limited. Scenario 2 is likely to occur in the absence of a credible MRV system for agricultural carbon credits, resulting in participants implementing only the least-cost practices to generate carbon credits and most likely those practices that would be implemented even in the absence of carbon payments. Market liquidity would be low, with high volatility around low average prices with limited financing and risk-management services for farmers and purchasers of credits.

Scenario 3: Taxpayers fund carbon farming

If corporate demand for carbon credits is low but participation in voluntary carbon programs is highly subsidized (directly through cost-share programs to implement certain practices, or indirectly through crop insurance premium deductions, tax credits, or green financing backed with public bonds), to the extent that market prices for carbon credits become of secondary importance to farmers, then an inefficient market for agricultural carbon would develop, funded by present and future taxpayers. The focus of participating farmers and suppliers would turn to comply with regulations to receive government payments or subsidies (rent-seeking behavior), and the cost of administering carbon programs would be largely absorbed by the sponsoring government agencies.

A low corporate demand for carbon credits could stem from a weak MRV system or high competition from other sources of carbon credits. Market liquidity would be low, with high volatility around low average prices, and limited private financing and risk-management services for farmers and purchasers of credits. Scenario 3 would be unsustainable in the long run.

Scenario 4: Unsustainable carbon farming

If corporate demand for carbon credits is low and the perceived quality of agricultural carbon credits is low, resulting in low credit prices and possibly, but not necessarily, including adverse selection or moral hazard in the marketplace, then agricultural carbon markets will likely collapse. A low corporate demand for carbon credits could stem from a weak MRV system or high competition from other sources of carbon credits. A limited adoption of conservation practices will likely generate high volatility around low average agricultural credit prices and steer farmers away from carbon markets. There would be limited private financing and risk-management services for farmers and purchasers of credits. Scenario 4 would be unsustainable in the short run.

CONCLUDING COMMENTS

This article discusses the current state of voluntary agriculture carbon markets (with a focus on the United States), analyzes the major challenges from both the demand and supply sides to voluntary agricultural carbon credit programs and provides an assessment of four possible scenarios for the future of agricultural carbon. It serves as a resource to researchers, producers, policymakers, and other stakeholders who seek a comprehensive analysis of the challenges that still face this market despite recent positive developments in global agreements.

ACKNOWLEDGMENTS

This research was supported in part by the USDA National Institute of Food and Agriculture Hatch Project IOW04099 and the Center for Agricultural and Rural Development at Iowa State University. Plastina appreciates the opportunities to conduct research on carbon markets provided by the Farm Foundation through the 2021 Agricultural Economics Fellowship. Open access funding provided by the Iowa State University Library. [Correction added on 18 May 2022, after first online publication: ISUL funding statement has been added.]

ENDNOTES

[1] Because GHGs have different global warming potentials, all pricing programs express the covered GHGs in carbon dioxide-equivalent (CO2e) units. For example, nitrous oxide (the most prevalent GHG emission in crop production) has a global warming potential equivalent to 298 times the global warming potential of carbon dioxide (Intergovernmental Panel on Climate Change, 2007).

[2] Although the CCX was designed to have binding emission targets for its participants, participation itself and the choice of the baseline emission level were voluntary (Intercontinental Exchange, 2011). By contrast, the California Air Resources Board (CARB) regulates power plants' participation and GHG emissions caps in the fully mandatory California cap and trade system (CARB, 2015).

[3] As in the CCX, the Kyoto Protocol mandated target emission levels for each participating nation, but participation itself was voluntary. For example, the United States of America did not participate in the Kyoto Protocol.

[4] The Clean Development Mechanism (CDM) of the Kyoto Protocol allowed participating nations to invest in emission reduction projects in developing countries and earn CERs that could be used to offset emissions in the investing nation and meet its Kyoto target. The CDM was the first global environmental investment and credit scheme for GHG emissions reduction.

[5] The Pact sets the basis to develop new market and non-market mechanisms to help countries achieve their national climate action plans under the Paris Agreement (United Nations Climate Change, 2021). The market mechanisms include (a) guidance to recognize the bilateral transfer of greenhouse gas (GHG) emission reductions between countries, and (b) the adoption of modalities, procedures, and guidelines (MPGs) for a new mechanism to issue carbon credits from emission-reducing activities. The non-market mechanism consists of a work program to help countries and their institutions to cooperate in climate change mitigation and adaptation activities, as well as on sustainable development and poverty reduction.

[6] The Paris Agreement to the United Nations Framework Convention on Climate Change (UNFCCC) attempts to limit the global average temperature increase to 2°C (3.6°F) above pre-industrial levels, with a preferable goal of 1.5°C (2.7°F), by reducing international GHG emissions (UNFCCC, 2015). In 2020, Parties to the Agreement embodied their updated national climate action plans in their Nationally Determined Contributions (NDCs) for climate mitigation.

[7] Currently, the California ETS is linked with Québec's ETS, and Switzerland's ETS is linked to the EU ETS. Following Brexit, the UK implemented its own UK ETS, and the new UNFCCC Mechanism would incentivize its linkage with the EU ETS.

[8] Credence goods are goods with qualities that cannot be ascertained by consumers even after consumption (Darby & Karni, 1973). A carbon credit or offset based on a claim that GHGs have been sequestered from the atmosphere or emissions have been avoided through certain processes is a credence good.

[9] Personal communication with Mr. Josh Shaeffer, Boomitra Head of Global Partnerships.

REFERENCES

Akerlof, George. 1970. "The Market for "Lemons": Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics* 84: 488–500.

Al-Kaisi, Mahdi M., Sotirios V. Archontoulis, and David Kwaw-Mensah. 2016. "Soybean Spatiotemporal Yield and Economic Variability as Affected by Tillage and Crop Rotation." *Agronomy Journal* 108: 1267–80.

Al-Kaisi, Mahdi M., Sotirios V. Archontoulis, David Kwaw-Mensah, and Fernando Miguez. 2015. "Tillage and Crop Rotation Effects on Corn Agronomic Response and Economic Return at Seven Iowa Locations." *Agronomy Journal* 107: 1411–24.

Angelopoulou, Theodora, Nikolaos Tziolas, Athanasios Balafoutis, George Zalidis, and Dionysis Bochtis. 2019. "Remote Sensing Techniques for Soil Organic Carbon Estimation: A Review." *Remote Sensing* 11(6): 676. https://doi.org/10.3390/rs11060676

Bithas, Kostas, and Dionysis Latinopoulos. 2021. "Managing Tree-Crops for Climate Mitigation. An Economic Evaluation Trading-Off Carbon Sequestration with Market Goods." *Sustainable Production and Consumption* 27: 667–78.

Black, Richard, Kate Cullen, Byron Fay, Thomas Hale, John Lang, Saba Mahmood and Steve M. Smith. 2021. Taking Stock: A Global Assessment of Net Zero Targets. Energy & Climate Intelligence Unit and Oxford Net Zero. https://bit.ly/3rxQCGD. Last accessed October 6, 2021.

Bonroy, Olivier, and Christos Constantatos. 2015. "On the Economics of Labels: How Their Introduction Affects the Functioning of Markets and the Welfare of All Participants." *American Journal of Agricultural Economics* 97: 239–59.

CARB. 2015. Overview of ARB Emissions Trading Program. California Environmental Protection Agency. https://bit.ly/3rgneGs. Last accessed January 19, 2022.

CARB. 2022. ARB Offset Credit Issuance Table. https://ww2.arb.ca.gov/our-work/programs/compliance-offset-program. Last Accessed January 19, 2022.

Castagné, Manon, Sara Lickel, Tara Ritter and Gilles Dufrasne 2020. Carbon Markets and Agriculture: Why Offsetting is Putting Us on the Wrong Track. Institute for Agriculture and Trade Policy. https://bit.ly/3FJi8G8. Last accessed August 15, 2021.

Caswell, Julie, and Eliza M. Mojduszka. 1996. "Using Information Labelling to Influence the Market for Quality in Food Products." *American Journal of Agricultural Economics* 78: 1248–53.

Claassen, Roger, Eric N. Duquette, and David J. Smith. 2018. "Additionality in U.S. Agricultural Conservation Programs." *Land Economics* 94(1): 19–35.

CME Group. 2021a. CME Group Announces First Trades of Global Emissions Offset (GEO) Futures. https://bit.ly/3qG5Av6. Last accessed December 27, 2021.

CME Group 2021b. CME Group Announces First Trades of Nature-Based Global Emissions Offset (N-GEO) Futures. https://bit.ly/3IeNuWS. Last accessed December 27, 2021.

Crespi, John M., and Stephan Marette. 2003. "Some Economic Implications of Public Labelling." *Journal of Food Distribution Research* 34: 83–94.

Crespi, John M., and Stephan, Marette. 2005. "Ecolabeling Economics: Is Public Involvement Necessary?" In *Environment, Information and Consumer Behavior*, edited by Clifford S. Russell and Signe Krarup. Northampton: Edward Elgar Publishing.

Crespi, John M, and Kristine Tidgren. 2021. "The First Legal Step for an Agricultural Carbon Market is in the Growing Climate Solutions Act of 2021." Policy Brief 21-PB 33. Center for Agricultural and Rural Development, Iowa State University.

Crespi, John M., and Marette, Stephan. 2022. *How Carbon Credits are Certified Could Change the Market Structure*, Policy Brief 22-PB 37. Center for Agricultural and Rural Development, Iowa State University.

Darby, Michael R., and Edi Karni. 1973. "Free Competition and the Optimal Amount of Fraud." *Journal of Law and Economics* 16(1): 67–88.

Giannakas, Konstantinos. 2002. "Information Asymmetries and Consumption Decisions in Organic Food Product Markets." *Canadian Journal of Agricultural Economics* 50: 35–50.

Gold Standard. 2021. Gold Standard for the Global Goals Fee Schedule. https://globalgoals.goldstandard.org/fees/. Last accessed August 20, 2021.

Gramig, Benjamin M., and Nicole J. O. Widmar. 2018. "Farmer Preferences for Agricultural Soil Carbon Sequestration Schemes." *Applied Economic Perspectives and Policy* 40(3): 502–21.

Haim, David, Eric M. White, and Ralph J. Alig. 2016. "Agriculture Afforestation for Carbon Sequestration Under Carbon Markets in the United States: Leakage Behavior from Regional Allowance Programs." *Applied Economic Perspectives and Policy* 38(1): 132–51.

Intercontinental Exchange. 2011. CCX Fact Sheet. https://bit.ly/31bJ7vN. Last accessed January 19, 2022.

Intergovernmental Panel on Climate Change (IPCC). 2007. "Climate Change 2007." In *Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by Core Writing Team, R. K. Pachauri, and A. Reisinger, 104. Geneva: IPCC.

Jackson-Smith, Douglas B., Michael Halling, Ernesto de la Hoz, Jamie P. McEvoy, and Jeffery S. Horsburgh. 2010. "Measuring Conservation Program Best Management Practice Implementation and Maintenance at the Watershed Scale." *Journal of Soil and Water Conservation* 65(6): 413–23.

Jones, Ghangela, Cesar Escalante, and Hofner Rusiana. 2015. "Reconciling Information Gaps in Organic Farm Borrowers' Dealings with Farm Lenders." *Agricultural Finance Review* 75(4): 469–83.

Khanna, Madhu. 2017. "Nexus between Food, Energy and Ecosystem Services in the Mississippi River Basin: Policy Implications and Challenges." *Choices Magazine* 32(4): 1–9.

Kiesel, Kristin, and Sofia B. Villas-Boas. 2013. "Can Information Costs Affect Consumer Choice? Nutritional Labels in a Supermarket Experiment." *International Journal of Industrial Organization* 31: 153–63.

Klonsky, Karen, and Martin D. Smith. 2002. "Entry and Exit in California's Organic Farming Sector." In *Economics of Pesticides, Sustainable Food Production, and Organic Food Markets. Advances in the Economics of Environmental Resources*, Vol 4, edited by Darwin C. Hall and L. Joe Moffitt, 139–65. Bingley: Emerald Group Publishing Limited.

Kostandini, Genti, Elton Mykerezi, and Eftila Tanellari. 2011. "Viability of Organic Production in Rural Counties: County and State-Level Evidence from the United States." *Journal of Agricultural and Applied Economics* 43(3): 443–51.

Marette, Stephan. 2014. "Economic Benefits Coming from the Absence of Labels Proliferation." *Journal of Agricultural and Food Industrial Organization* 12: 1–9.

Marette, Stephan, and Jutta Roosen. 2011. "Bans and Labels with Controversial Food Technologies." In *The Oxford Handbook of the Economics of Food Consumption and Policy*, edited by J. L. Lusk, J. Roosen, and J. E. Shogren, 499–519. Oxford: Oxford University Press.

Nowatzke, Laurie, and J. Gordon Arbuckle. 2018. *Iowa Farmers and the Iowa Nutrient Reduction Strategy: Survey Results from the Missouri-Little Sioux Watershed*. Ames: Iowa State University Extension and Outreach.

Ogle, Stephen M., Amy Swan, and Keith Paustian. 2012. "No-till Management Impacts on Crop Productivity, Carbon Input and Soil Carbon Sequestration." *Agriculture, Ecosystems and Environment* 149: 37–49.

Pendell, Dustin L., Williams, Jeffery R., Boyles, Scott B., Rice, Charles W., and Nelson, Richard G.. 2007. "Soil Carbon Sequestration Strategies with Alternative Tillage and Nitrogen Sources under Risk." *Applied Economics Perspectives and Policy* 29(2): 247–68.

Pannell, David J., and Roger Claassen. 2020. "The Roles of Adoption and Behavior Change in Agricultural Policy." *Applied Economic Perspectives and Policy* 42(1): 31–41.

Petrolia, Daniel R. 2008. "An Analysis of the Relationship between Demand for Corn Stover as an Ethanol Feedstock and Soil Erosion." *Applied Economic Perspectives and Policy* 30(4): 677–91.

Plastina, Alejandro. 2021. How do Data and Payments Flow through Ag Carbon Programs? Ag Decision Maker File A1-77. Iowa State University Extension and Outreach. https://www.extension.iastate.edu/agdm/crops/pdf/a1-77.pdf

Plastina, Alejandro and Wendiam Sawadgo. 2021. "Cover Crops and No-till in the I-States: Non-Permanence and Carbon Markets." Agricultural Policy Review, Fall 2021. Center for Agricultural and Rural Development, Iowa State University. www.card.iastate.edu/ag_policy_review/article/?a=133. Last accessed January 19, 2022.

Plastina, Alejandro and Oranuch Wongpiyabovorn 2021. How to Grow and Sell Carbon Credits in US Agriculture. Ag Decision Maker File A1-76. Iowa State University Extension and Outreach. https://www.extension.iastate.edu/agdm/crops/pdf/a1-76.pdf

Prokopy, Linda S., Kristin Floress, J. Gordon Arbuckle, Sarah P. Church, Francis R. Eanes, Yuling Gao, Benjamin M. Gramig, Pranay Ranjan, and Ajay S. Singh. 2019. "Adoption of Agricultural Conservation Practices in the United States: Evidence from 35years of Quantitative Literature." *Journal of Soil and Water Conservation* 74: 520–34. https://doi.org/10.2489/jswc.74.5.520

Prokopy, Linda S., Kristin Floress, Denise Klotthor-Weinkauf, and Adam Baumgart-Getz. 2008. "Determinants of Agricultural Best Management Practice Adoption: Evidence from the Literature." *Journal of Soil and Water Conservation* 63: 300–11. https://doi.org/10.2489/jswc.63.5.300

Ranjan, Pranay, Sarah P. Church, Kristin Floress, and Linda S. Prokopy. 2019. "Synthesizing Conservation Motivations and Barriers: What Have We Learned from Qualitative Studies of Farmers' Behaviors in the United States?" *Society and Natural Resources* 32: 1171–99. https://doi.org/10.1080/08941920.2019.1648710

Roe, Brian E., and Ian Sheldon. 2007. "Credence Good Labelling. The Efficiency and Distributional Implications of Several Policy Approaches." *American Journal of Agricultural Economics* 89: 1020–33.

Roe, Brian E., Mario F. Teisl, and Corin R. Deans. 2014. "The Economics of Voluntary versus Mandatory Labels." *Annual Review of Resource Economics* 6: 1–21.

Sawadgo, Wendiam, and Plastina, Alejandro 2022. "The Invisible Elephant: Disadoption of Conservation Practices in the United States." *Choices Magazine* 37(1): 1–13.

Sawadgo, Wendiam, and Alejandro Plastina. 2021. "Do Cost-share Programs Increase Cover Crop Use? Empirical Evidence from Iowa." *Renewable Agriculture and Food Systems* 36(6): 527–35. https://doi.org/10.1017/S1742170521000132

Secchi, Silvia & Chris Jones 2021. "Opinion: With Iowa's and the Ag Industry's Carbon Efforts, Plenty of Reason to Question Motives." Des Moines Register, July 4, 2021. https://bit.ly/3qGFwjk. Last accessed August 15, 2021.

Teisl, Mario F., and Brian E. Roe. 1998. "The Economics of Labeling: An Overview of Issues for Health and Environmental Disclosure." *Agricultural and Resource Economics Review* 27: 140–50.

Toliver, Dustin K., James A. Larson, Roland K. Roberts, Burton C. English, Daniel G. De La Torre Ugarte, and Tristram O. West. 2012. "Effects of No-till on Yields as Influenced by Crop and Environmental Factors." *Agronomy Journal* 104(2): 530–41.

UNFCCC. 2015. Paris Agreement to the United Nations Framework Convention on Climate Change. T.I.A.S. No. 16-1104. https://bit.ly/3fCZvcD. Last accessed January 19, 2022.

UNFCCC. 2021. COP26 Outcomes: Market Mechanisms and Non-Market Approaches (Article 6). https://bit.ly/3FIKJvi. Last accessed December 22, 2021.

United Kingdom Government. 2021. COP26: The Negotiations Explained. https://ukcop26.org/wp-content/uploads/2021/11/COP26-Negotiations-Explained.pdf. Last accessed December 22, 2021.

United Nations Climate Change. 2021. COP26 Outcomes: Market Mechanisms and Non-Market Approaches (Article 6). https://bit.ly/3rw2dG7. Last accessed December 21, 2021.

van Kooten, G. Cornelis. 2008. "A Perspective on Carbon Sequestration as a Strategy for Mitigating Climate Change." *Choices Magazine* 23(1): 24–7.

West, Tristram O., and Gregg Marland. 2002. "A Synthesis of Carbon Sequestration, Carbon Emissions, and Net Carbon Flux in Agriculture: Comparing Tillage Practices in the United States." *Agriculture, Ecosystems and Environment* 91(1–3): 217–32.

Wongpiyabovorn, Oranuch, Alejandro Plastina & Sergio H. Lence 2021. "Futures Market for Ag Carbon Offsets under Mandatory and Voluntary Emission Targets." Agricultural Policy Review, Fall 2021. Center for Agricultural and Rural Development, Iowa State University. www.card.iastate.edu/ag_policy_review/article/?a=127

World Bank Group. 2021. Carbon Pricing Dashboard. https://carbonpricingdashboard.worldbank.org/map_data. Last accessed October 11, 2021.

How to cite this article: Wongpiyabovorn, Oranuch, Alejandro Plastina, and John M. Crespi. 2023. "Challenges to voluntary Ag carbon markets." *Applied Economic Perspectives and Policy* 45(2): 1154–1167. https://doi.org/10.1002/aepp.13254 nature communications

Perspective

https://doi.org/10.1038/s41467-023-44425-2

The principles of natural climate solutions

Received: 12 May 2023

Accepted: 13 December 2023

Published online: 23 January 2024

Check for updates

Peter Woods Ellis [1] ✉, Aaron Marr Page [2], Stephen Wood [1], Joseph Fargione [1], Yuta J. Masuda [3], Vanessa Carrasco Denney [1], Campbell Moore [1], Timm Kroeger [1], Bronson Griscom [4], Jonathan Sanderman [5], Tyson Atleo [6], Rane Cortez [1], Sara Leavitt [1] & Susan C. Cook-Patton [1]

Natural climate solutions can mitigate climate change in the near-term, during a climate-critical window. Yet, persistent misunderstandings about what constitutes a natural climate solution generate unnecessary confusion and controversy, thereby delaying critical mitigation action. Based on a review of scientific literature and best practices, we distill five foundational principles of natural climate solutions (nature-based, sustainable, climate-additional, measurable, and equitable) and fifteen operational principles for practical implementation. By adhering to these principles, practitioners can activate effective and durable natural climate solutions, enabling the rapid and wide-scale adoption necessary to meaningfully contribute to climate change mitigation.

While the general mechanisms by which plants affect climate have been understood for over a century[1–3], in 2017 scientists and conservation practitioners framed the holistic concept of 'natural climate solutions' (NCS) to adapt existing knowledge and experience to climate action. As originally defined, NCS are deliberate human actions (NCS pathways) that protect, restore, and improve management of forests, wetlands, grasslands, oceans, and agricultural lands to mitigate climate change[4]. NCS were also defined as having no net negative impact on food and fiber supply and no net harm to biodiversity, while ensuring actions are implemented in socially and culturally responsible ways[4,5].

In the past six years, interest in NCS has increased dramatically. The conversation has tripled in size, from < 2% to > 6% of climate-related social media traffic (see Supplementary Methods), and funding commitments have doubled[6]. However, this pace must accelerate exponentially if we are to succeed[7]. The Intergovernmental Panel on Climate Change (IPCC) emphasizes that the rapid deployment of NCS (which the IPCC calls Agriculture, Forestry and Other Land Use [AFOLU] mitigation measures) is essential to reach net zero emissions and avoid catastrophic warming, but if deployed carefully and appropriately it can deliver a third of the climate mitigation needed by 2030. However, investments will require > $400 billion per year[8], which is over nine times the amount being spent today[9].

Accompanying this uptick in interest has been a concomitant rise in confusion and controversy. In some instances, the excitement around NCS has led to well-intentioned but hastily and poorly designed tree planting programs, which have rightly catalyzed heated dialog around considerations for implementing NCS programs[10]. In other instances, NCS have been dismissed as greenwashing because they are "vulnerable to exploitation by companies that want to appear at the vanguard of climate action[11]." A similar misperception portrays NCS as predominantly carbon offsetting mechanisms promoted by energy intensive industries[12]. Another confusion arises from the overlap between NCS and carbon dioxide removal (CDR); some NCS (for example reforestation) do indeed remove carbon dioxide (CO_2) from the atmosphere, but others (avoided peatland conversion) avoid CO_2 or other greenhouse gas emissions. Furthermore, NCS are often conflated with other terms such as nature-based solutions (NbS), nature-based climate solutions (NbCS) and AFOLU. Compared with NCS, NbS refer to a much broader set of actions that address a range of societal challenges beyond only climate mitigation (Fig. 1)[13–15]; NbCS are nearly identical to NCS but include some additional activities in engineered ecosystems (for example, macroalgae farming[16]) that have been moved further from their natural state (Principle 1.2 below)[17,18]. NCS are synonymous with AFOLU mitigation measures as defined by the IPCC[8].

Perhaps another source of confusion is driven by the fact that the NCS concept builds upon a long history of conservation science and practice but focuses the framework on measurable climate change mitigation. For example, the United Nations' reducing emissions from deforestation and forest degradation in developing countries program (REDD+), payment for ecosystem services,

[1]The Nature Conservancy, Arlington, VA, USA. [2]Forum Nobis, Iowa City, IA, USA. [3]Paul G. Allen Family Foundation, Seattle, WA, USA. [4]Conservation International, Arlington, VA, USA. [5]Woodwell Climate Research Center, Falmouth, MA, USA. [6]Nature United, Victoria, BC, Canada. ✉e-mail: pellis@tnc.org

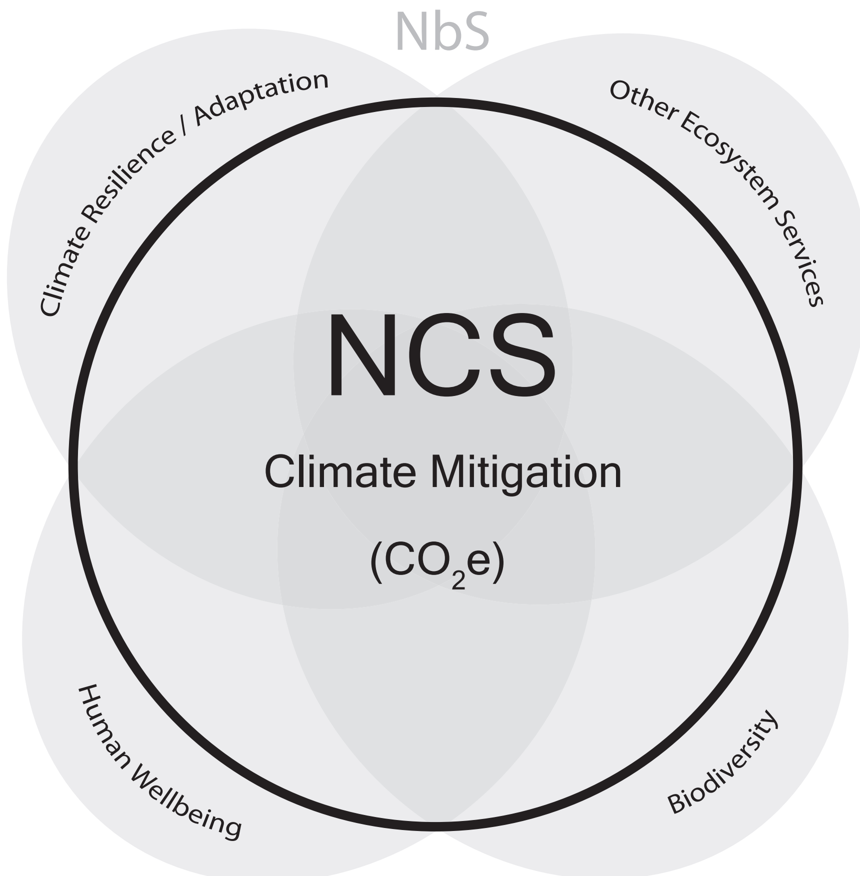


Fig. 1 | Overlap of natural climate and nature-based solutions. Conceptual diagram showing the overlap between nature-based solutions (NbS) and natural climate solutions (NCS). While NCS focus on have a single outcome (CO_2 equivalents; CO_2e), NbS can be defined by multiple outcomes with multiple metrics.

community-based conservation programs, national government conservation incentive programs, and other well-established and studied interventions or programs all presumably fall under the umbrella of NCS and provide a rich and diverse body of evidence on advances and challenges in NCS implementation. The original NCS study[4] was powerful because it explicitly estimated the climate change mitigation potential of known and long-studied conservation interventions, providing insight on whether, which, and to what extent natural solutions could feasibly advance global climate change mitigation goals. One challenge that emerged from the analysis was that there is no guarantee that past interventions and programs were truly climate additional even though they plausibly fit within the NCS umbrella. Meanwhile, several studies have been conducted that provide insights on the determinants of successful NCS projects which can inform future action[19–22].

Finally, significant confusion arises when untested NCS interventions are promoted widely without a robust scientific basis. However, recent publications have provided a mechanism for differentiating well-tested 'ready-to implement' NCS pathways from emerging or nascent pathways that require further research[17,23].

While confusion and controversy are to be expected in a rapidly growing field, without a clear framework for productive conversation and reliable action, there is a concern that momentum for NCS implementation will stall. Avoiding climate catastrophe requires immediate NCS action[24,25], but to be effective, NCS must be implemented equitably and sustainably. When initiatives claim to implement NCS but fail to achieve sustainable, equitable climate mitigation, they divert resources away from legitimate climate solutions and undermine public support for true NCS.

Many have commented on the risks of poorly-informed NCS action[18,19,21,26–28], but few have offered a clear path forward for real, fair and well-informed NCS action. If the NCS movement is to scale rapidly, the evolving NCS conversation needs normative criteria to help practitioners, policymakers, researchers, and the public evaluate whether NCS options are tangible, viable, and appropriate.

In this Perspective, we outline these normative criteria as a set of NCS principles (Box 1 and Fig. 2) that can be used to identify NCS actions worthy of support. These principles address problems that contribute to unproductive confusion and controversy around climate change mitigation. NCS foundational principles are criteria (nature-based, sustainable, climate-additional, measurable, and equitable) that provide a working definition of NCS based on the existing literature. This working definition enables more effective and productive NCS action by clarifying the boundaries of the NCS conversation. NCS operational principles guide NCS implementation by specifying how to apply NCS foundational principles to real-world action. These fifteen operational principles provide guidance to policymakers and practitioners so that risks can be navigated intelligently without impeding action. Taken together, the NCS principles orient NCS activities to the appropriate scope (Principle 1), ensure positive climate benefits (Principle 4) and avoid negative impacts (Principles 2, 4, and 5). While we hope the normative nature of these principles facilitates implementation by mitigating against confusion and controversy, we acknowledge that, as with all normative criteria, they are idealized, and the real-world complexities of conservation action will continue to expose new uncertainties that should allow the concept to evolve through active and adaptive action. Over time, it may be necessary to add or adjust principles.

Fig. 2 | The wheel of natural climate solutions. Foundational principles are shown along the outer edge of the wheel, while operational principles are the 'spokes' inside the wheel. See Box 1 and main text for the full definition of each principle.

Foundational principle 1: NCS are nature-based

Principle 1.1: NCS result from the human stewardship of ecosystems

NCS involve active management decisions that affect human stewardship of ecosystems and result in net climate mitigation. An ecosystem is broadly defined to include all the living organisms, including humans, and their interrelationships within a physical environment. Notably, this definition includes both natural and working lands. For example, agroforestry can lead to carbon sequestration in soil and woody plants[29] and occurs within an ecosystem involving farmers, crops, soil, and the soil microbial community. Principle 1.1 acknowledges that, despite common misconceptions of nature as something devoid of humans, humans have shaped natural lands and waters for millennia[30].

Changes in human stewardship of ecosystems can be triggered locally by supply-side decisions (for example, a rancher adopting silvopastoral practices) or demand-side decisions (for example, an urban resident deciding to stop eating beef). While we acknowledge the influence of demand-side climate solutions such as diet change, shifts to longer-lived wood products, reduced food waste, or biofuel usage, it is beyond the scope of this paper to trace these supply and demand-side interactions for each NCS Pathway. IPCC refers to the interacting aggregate of supply and demand-side actions as AFOLU mitigation measures[8]. Conservation International groups AFOLU mitigation measures into people-centered (rather than ecosystem centered) NCS Action Tracks[31].

Principle 1.2: NCS do not move ecosystems further from their natural state

Being nature-based also means that NCS stewardship does not move ecosystems further away from their natural state than they are already. Elements of structure, composition, and function of the unmodified, naturally occurring system must be considered, as well as the current land use. For example, while tree planting may be perceived as a positive activity, replacing a natural forest by planting a plantation of non-native tree species, even if faster-growing, would not be considered an NCS, because the natural structure and function of the forest would be diminished as a result[32]. Similarly, artificial fertilization of ocean water to stimulate algal bloom[33], ocean alkalinization[34], and kelp/seaweed afforestation[35] would not be considered NCS because human intervention moves the ecosystem farther from its unmodified state.

In many cases, advocates and practitioners will need a nuanced and dynamic understanding of an ecosystem's natural state to decide whether an intervention adheres to Principle 1.2. The structure and composition of many ecosystems are transitioning as climate changes, and some need human assistance to become more climate resilient. Therefore, careful assessment of actions is often needed to determine what movement away from, or towards, a natural, climate-resilient state means in practice. For example, if there is clear evidence that a forest ecosystem is transitioning into a grassland ecosystem in response to climate change, then restoration NCS should be aligned with these dynamics.

 https://doi.org/10.1038/s41467-023-44425-2

BOX 1

The principles of natural climate solutions

Foundational principles define natural climate solutions (NCS). operational principles (for example, Principle 1.1, 2.1) guide NCS implementation by specifying how to operationalize foundational principles.

Foundational Principle 1: NCS are Nature-based.
Principle 1.1: NCS result from the human stewardship of ecosystems.
Principle 1.2: NCS do not move ecosystems further from their natural state.

Foundational Principle 2: NCS are Sustainable.
Principle 2.1: NCS sustain biodiversity.
Principle 2.2: NCS sustain food production.
Principle 2.3: NCS sustain fiber and wood production.
Principle 2.4: NCS sustain climate adaptation services.

Foundational Principle 3: NCS are Climate-additional.
Principle 3.1: NCS provide additional climate mitigation that would not happen without human intervention.
Principle 3.2: NCS provide durable mitigation.
Principle 3.3: NCS are not used to compensate for readily abatable emissions.

Foundational Principle 4: NCS are Measurable.
Principle 4.1: NCS are quantified in terms of cumulative effects on radiative forcing.
Principle 4.2: NCS accounting is conservative.
Principle 4.3: NCS with uncertainty ranges greater than the estimated climate mitigation should be flagged as emerging.
Principle 4.4: NCS accounting avoids double-counting.

Foundational Principle 5: NCS are Equitable.
Principle 5.1: NCS respect human rights.
Principle 5.2: NCS respect Indigenous self-determination.

Foundational principle 2: NCS are sustainable

Principle 2.1: NCS sustain biodiversity

Activities that transition an ecosystem further away from its unmodified state would also fail to qualify as NCS because NCS should sustain biodiversity. NCS can have neutral near-term local impacts, but must avoid reductions in alpha, beta, and/or gamma diversity as is described in the Convention on Biological Diversity[4,36]. For example, turning agricultural residue into soil biochar stores carbon in the soil without harming biodiversity[37]. In contrast, adding trees to native grasslands may increase carbon sequestration at the expense of native grassland diversity and thus would not qualify as an NCS[38].

Principle 2.2: NCS sustain food production

NCS should sustain food production. Climate solutions will not be durable unless food security can be provided alongside farmer and fisher livelihoods[39]. Maintaining food security in the face of a growing human population and changing diets is a complex challenge. Extensive reforestation on cultivated lands and constraints on agricultural inputs (for example, fertilizer use) can negatively impact food security[40]. However, various combinations of maintaining existing cropland[4], improving fertilizer management in ways that reduce the cost of crop production[41], implementing silvopasture to increase livestock productivity in existing pastoral systems[42], limiting bioenergy production and associated land demand[43,44], and adopting more climate-friendly diets enable NCS implementation while increasing food security[45].

Principle 2.3: NCS sustain fiber and wood production

NCS should sustain fiber and wood production as climate solutions will not be durable unless rising wood product demand can be met while maintaining forest-based livelihoods[46]. Wood-based materials are particularly important to the success of sustainable climate mitigation because, when well-sourced through selective logging and existing forest plantations, they have much smaller carbon footprints than building alternatives such as steel and concrete. However, unsustainable timber exploitation that involves deforestation or that harvests forests above a scientifically defined level of sustained yield[47] would undermine this principle and would therefore not be considered an NCS.

Principle 2.4: NCS sustain climate adaptation services

Finally, while NCS are focused on climate mitigation outcomes, at a minimum they sustain climate adaptation services through the ecosystems in which they are implemented. A rich literature identifies the multitude of ecosystem and climate adaptation services that NCS can provide, such as attenuation of floods, soil erosion, landslides, storm surges, and the resulting human benefits[48–53]. However, existing ecosystems often already provide adaptation services. Any implementation of NCS should at sustain existing levels of adaptation services to ensure adaptation is co-produced alongside mitigation.

Climate change is a significant threat to people, biodiversity, and other ecosystem services, and NCS offer real near-term climate mitigation that alleviates this threat. Thus, over longer time scales, NCS deliver positive benefits to ecosystems and people, either directly (by providing net positive local effects) or indirectly (by sustaining the climate stability upon which these benefits rely). A comprehensive approach to climate change must consider both mitigation and adaptation, and projects that provide both, as many NCS projects do, are particularly valuable in this regard.

Note that Principles 2.1–2.4 are sensitive to scale. For example, a 100-hectare cropland reforestation NCS project may fail to sustain food production at the project scale, but a state-wide reforestation program may not, if any lost cropland is either marginal or accommodated through intensification.

While NCS does not by definition require additional non-climate benefits, on-the-ground NCS projects are frequently implemented to achieve multiple co-benefits. Indeed, the promise of NCS co-benefits is often a primary motivating factor for decision-makers, especially when juggling multiple conservation and climate priorities alongside other sustainable development goals. Therefore, it is important for researchers to continue compiling evidence for decision-makers to understand whether and to what extent NCS provide co-benefits (or involve tradeoffs).

Foundational principle 3: NCS are climate-additional

Principle 3.1: NCS provide additional climate mitigation that would not happen without human intervention

NCS provide additional climate mitigation that would not happen without human intervention. Additionality is traditionally assessed in reference to a 'business as usual' baseline scenario[54]. For example, establishing a forest reserve in a remote landscape with high carbon stocks would not count as NCS unless that landscape was threatened by human disturbance. If that landscape is not threatened, then reserve status does not alter the fate of the carbon stocks stored within it. Similarly, in locations experiencing land abandonment and natural recovery of native ecosystems, it would not be appropriate to count pre-existing recovery as NCS, because that recovery is part of the

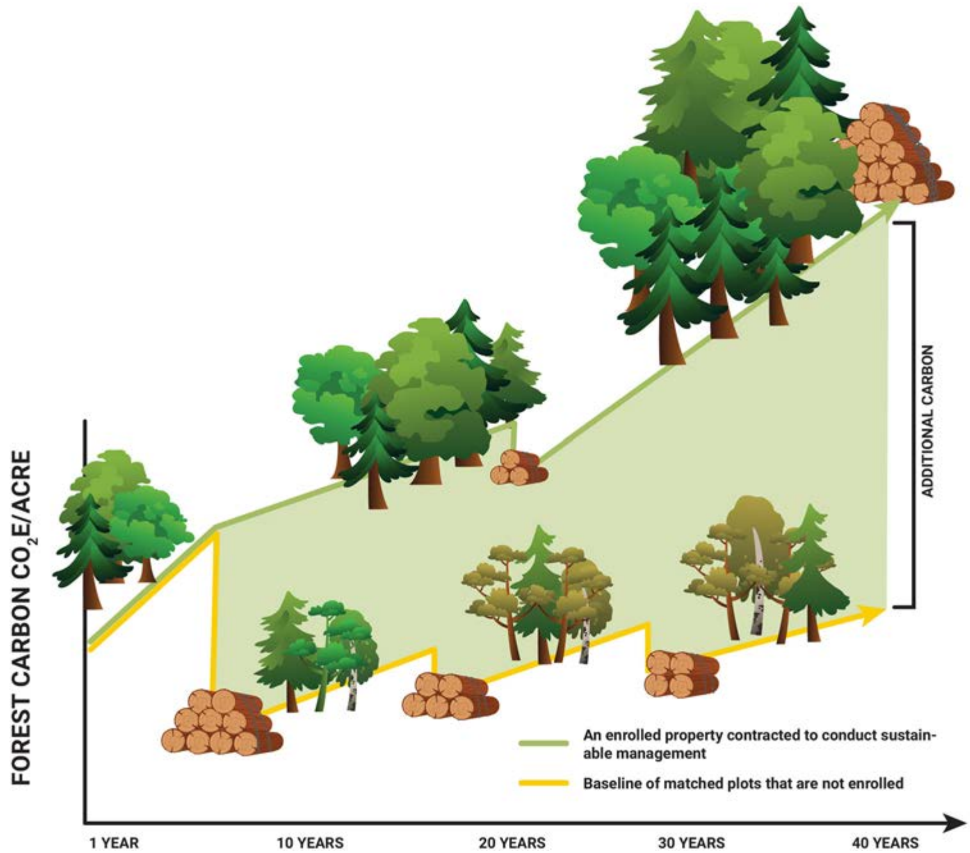


Fig. 3 | Schematic representation of additionality calculations. Figure shows how additionality (Principle 3.1) is calculated for improved forest management projects within the Family Forest Carbon Program (FFCP)[108,109]. Increasing stocks over time represent long term increases in timber and carbon yields predicted under FFCP practices, ensuring sustainability (Principle 2.3). Periodic harvests show that FFCP is committed to maintaining working forests (Principle 1.1). Adapted from the Dynamic Baselines infographic with permission from The Nature Conservancy.

baseline condition and not specifically associated with human intervention. However, cleared land that is expected to remain cleared (for example, as a pasture), and where a deliberate choice is made to instead allow natural recovery, would count as NCS because human intervention changes the trajectory of the land use. Note that this principle applies to both avoided emissions (for example, avoided forest conversion) and CDR (for example, reforestation) NCS.

Methods for demonstrating additionality are varied, and each approach has strengths and weaknesses. The scientific community that supports NCS action should commit to continuous improvement focused on methodological transparency, accuracy, and efficiency. For example, traditional baseline approaches to forest NCS measurement are being updated with dynamic global monitoring systems that can compare projects to matched control sites and track benefits as they accrue using coordinated networks of field measurements and improved remote-sensing technologies (see Fig. 3)[55].

Principle 3.2: NCS provide durable mitigation

NCS must also provide durable mitigation, meaning that additional climate benefits persist over time. Durability is defined in terms of different greenhouse gas (GHG) residence times in different pools (for example, soil or aboveground biomass) at different scales (for example, national or global), but should avoid binary classifications (that is, permanent versus impermanent)[56,57]. The key consideration is whether the NCS activity will provide mitigation for a long enough period to deliver measurable, additional, net positive climate benefits[58]. For example, peatland rewetting initiatives might not qualify as NCS if the restored peatland is not maintained long enough to counterbalance the adverse warming effect of the methane emissions pulse from rewetting[59].

Many ecosystems stewarded through NCS have proven extremely durable. For example, Australia's Daintree Rainforest has effectively stored carbon for 135 million years, and is currently protected as a national park[60]. Other ecosystems are less durable, either due to human or environmental threats, such as conversion for urban development or increased wildfire risk due to a warming climate[61,62]. To adequately account for this variation, NCS should be considered in the context of both the imperative to act now, and the imperative to simultaneously build robust systems to ensure durability over time. Enabling conditions are needed to ensure NCS durability at adequate spatial and temporal scales; for example, a global NCS monitoring system is needed to detect and quantify reversals, and long-term insurance and financial systems are needed to 'pay back' the atmosphere in the event of reversals. Carbon market standards are in the process of developing these systems, but more research is needed to calibrate these systems with quantifiable risks of reversal[58]. As with principle 3.1, durability considerations are equally important to CDR and avoided emissions NCS.

Principle 3.3: NCS are not used to compensate for readily abatable emissions

It is imperative that NCS not be used to compensate for readily abatable emissions. Non-NCS strategies that reduce emissions from fossil fuels will need to deliver the majority of the mitigation required to meet Paris Agreement goals[60,61]. NCS must proceed in parallel to these strategies. In most cases, this is unproblematic. Governments and other climate mitigation actors may choose to invest in or focus on NCS for many legitimate reasons, including cost efficiency and the multiple layers of co-benefits offered by some NCS. But when ecosystem-based mitigation is used to offset readily abatable emissions in another sector it provides no real climate benefit.

Application of principle 3.3 requires nuance and consideration of the different venues in which climate progress is pursued (for example, public sector commitments or voluntary carbon markets (VCM)). Different actors will have different views on complex questions of what emissions are 'readily abatable', but establishing clear best practice across different sectors and contexts is necessary to maintain societal buy-in. This principle can accommodate such differences while also serving as a foundation to resist non-credible claims. At a minimum, parties seeking to use NCS carbon credits to compensate for unabated emissions should be able to have a record of and target for abatement, articulate why the emissions to be compensated are not currently abatable, and locate the credits on a path to reach a legitimate mitigation goal under the Paris Agreement or applicable private sector standards. Notably, many efforts are underway that can help corporations and other entities align their emissions reduction goals with a science-based path to limit global warming and identify 'residual' emissions that would not have been abated even under ambitious decarbonization scenarios[62–64]. These demand-side integrity programs require further development to ensure true 'atmospheric additionality' of mitigation but are both necessary and sufficient to allow the purchase and trade of NCS credits in a way that will accelerate rather than delay global climate progress[65].

It is possible that NCS carbon market projects will be developed without knowing—or being able to control—how generated credits will later be used by purchasers. This means that we need to push for systems that enable greater transparency (for example, the demand-side integrity programs mentioned above) so that the stakeholders who generate the credits can better choose to whom they sell credits. Principle 3.3 speaks directly to purchasers as key participants in the overall solution.

Foundational principle 4: NCS are measurable

There are multiple potential NCS actions that can occur in a given landscape and quantifying the overall magnitude of opportunity can help to focus efforts on the actions that can offer the largest mitigation returns. However, appropriate accounting is required to ensure that NCS potential is consistently and clearly quantified.

Principle 4.1: NCS are quantified in terms of cumulative effects on radiative forcing

NCS mitigation is quantified in terms of cumulative effects on radiative forcing due to changes in stewardship of ecosystems. For consistent comparison, radiative forcing ($W\ m^{-2}$) is converted to CO_2 equivalents (CO_2e) and, for GHGs, considered as a function of GHG flux into or out of the atmosphere[3]. Avoided emissions and removals NCS are treated equally, as their effect on the quantity of GHGs in the atmosphere is the same. GHGs to be considered include CO_2, methane (CH_4), and nitrous oxide (N_2O) fluxes. Ideally, NCS mitigation estimates also incorporate biophysical factors that affect top-of-atmosphere radiative forcing, including black carbon deposited from particulate matter, changes in albedo resulting from changes in land cover, and changes in water vapor[66]. These first two non-GHG factors have the potential to significantly alter the net climate impact of NCS and thus shape where, when, and how NCS implementation occurs. Although these factors have been difficult to directly quantify, there is opportunity for them to be included as the science improves. Failing to consider all climate forcing agents could lead to adoption of NCS that have little climate benefit. For example, tree planting in dry forest areas may appear to be net climate positive when quantifying carbon, but actually might be net negative after considering changes in albedo[67].

To better facilitate comparisons, it is possible to translate non-CO_2 climate pollutants into CO_2e using established Global Warming Potentials (GWP)[8](see section 7.6 of Ref. 8). How to make these comparisons depends on whether one is accounting for near-term or long-term climate impacts. To appropriately consider both near-term and long-term climate impacts, we recommend accounting for short-lived and long-lived climate pollutants using different GWP conversions. For example, following IPCC guidance, GWP_{100} is appropriate to calculate the climate impact of NCS over the long term and for long-lived climate pollutants such as CO_2 and N_2O[68]. However, conversion equations such as GWP* are better suited to account for the climate impact of short-lived climate pollutants, such as CH_4 and black carbon[69]. In this way, targets and monitoring systems for long and short-lived climate pollutants can be accounted for separately, thereby enabling incentives to be developed appropriate to the timing of their atmospheric impacts[56].

Principle 4.2: NCS accounting is conservative

NCS accounting adheres to the convention of conservativeness[70], whereby data[56] are considered only when sufficient evidence exists to support their inclusion. Buma et al.[17]. have recently assessed the strength of the science underlying different NCS actions and show how some NCS pathways (for example, avoided forest conversion) have well-constrained estimates of mitigation, but others (such as avoided benthic disturbance) include unresolved or incomplete accounting. Sufficient evidence means that additional NCS mitigation potential estimates are significantly different from zero with medium or greater confidence according to the IPCC[68]. New NCS pathways and activities can be added as the science evolves and estimates improve. For example, data on rates of urban tree cover loss in Canada enabled estimates of the mitigation potential of maintenance of urban tree cover in that country[71], allowing this pathway to be incorporated into NCS mitigation assessments[72]. Some pathways may also delineate new activities as more refined data becomes available. For example, global estimates of climate-smart forestry NCS (originally termed natural forest management) were initially calculated in aggregate[4], but more recent research has begun to parse this into specific activities such as reduced-impact logging for climate mitigation[73] and liana removal[74].

Principle 4.3: NCS with uncertainty ranges greater than the estimated climate mitigation should be flagged as emerging

Determining whether a candidate NCS pathway will provide significant mitigation requires estimating the uncertainty of the quantified NCS mitigation. NCS with uncertainty ranges greater than the estimated climate mitigation should be flagged as emerging. Robust and complete assessments of uncertainty can and should affect decision-making. For example, Griscom et al.[4]. identify reforestation as the single largest NCS pathway based on biophysical potential (10.1 $PgCO_2\ yr^{-1}$), nearly equal to all other NCS mitigation potentials combined. However, the 95% confidence interval is 74% of the mean estimate (7.5 $PgCO_2\ yr^{-1}$). If reforestation offered only 2.6 $PgCO_2\ yr^{-1}$, it would dramatically alter the implications for prioritizing reforestation globally. Quantifying and reporting NCS uncertainty at the appropriate scale is important to ensure that incentives are focused on actionable NCS, and research is focused on emerging NCS[17,23].

Principle 4.4: NCS accounting avoids double-counting

NCS accounting should make all attempts to avoid double-counting. For example, when estimating NCS opportunity, a given pasture could be reforested or could have improved grazing management activities, but not both. In contrast, some NCS pathways can be deployed in parallel, such as agroforestry practices and biochar application on the same agricultural land. Careful consideration of which pathways can and cannot be applied on the same area at the same time is necessary to avoid overstating (see Principle 4.2) or understating the potential of NCS on a given landscape. The original NCS publication established a series of hierarchical rules to eliminate double counting, described in its supporting information appendix (see Supplementary Note 1)[4]. Similarly, when generating carbon credits or reporting toward country-level or other jurisdictional targets, NCS projects should follow accounting best practices to ensure that the mitigation they provide is not double counted within emission inventories or carbon market schemes. To avoid double counting, it is helpful to use the NCS framework to categorize activities into biomes, conservation actions, and pathways (see Supplementary Note 1). Note that conservation pathways related to protection measure avoided emissions, while those related to restoration measure additional sequestration. In contrast, improved management pathways can be a mix of both increased sequestration and avoided emissions. For example, agroforestry increases C sequestration, but biochar and no tillage (contrary to some misconceptions)[75] avoid emissions via decomposition of vegetation or avoided oxidation of soil carbon.

Foundational principle 5: NCS are equitable

The original NCS publication emphasized that "work also remains to refine methods for implementing NCS pathways in socially and culturally responsible ways"[4]. This statement was a call for additional research and action, but also a recognition that NCS must be equitable. The need for a social equity lens is driven by historical and ongoing injustices associated with management of natural resources[76,77]. Vulnerable populations such as Indigenous Peoples, local communities, farmers, forest managers, coastal communities, conservationists, women, and other marginalized groups are often the least responsible for historic emissions, bear the greatest costs and impacts of climate change[78], and yet are often the most active and effective NCS stewards[79–83].

Despite best practices in conservation around rigorous review procedures and community consultation, recent scholarship has identified deeper equity issues in the field that persist and that have been explored through a number of frameworks such as the capabilities approach[84], the 'Just Sustainabilities' concept[85], human rights[86], and Indigenous Peoples' rights[87]. This critical lens reveals a new set of challenges: while equality of participation and material outcomes remains important, true equity may require reimagining underlying root concepts to consider previously marginalized and excluded interests and experiences. Environmentalism, conservation, sustainability, and similar root concepts must be understood as culturally embedded and linked to particular social identities and political choices, rather than as abstract, inherent, and universal.

More concretely, two studies present an analytical framework that recognizes multiple dimensions of equity, including procedural (involvement and inclusiveness of all rightsholders and interested parties), distributive (fair allocation of costs, benefits, burdens, and rights), recognitional (respect for knowledge systems, values, social norms, and rights of all rightsholders and interested parties), and contextual (attention to power dynamics and the social conditions that affect ability to advocate for equity on the other dimensions)[88,89]. Many social equity safeguard frameworks for use with NCS have already emerged and are being piloted[90–92]. While NCS can work toward social equity in myriad ways, the following principles can help to deliver socially and culturally responsible NCS implementation and meet the basic commitment to equity.

Principle 5.1: NCS respect human rights

NCS should respect human rights. This means that NCS activities comply with national laws and international human rights law, as reflected in the International Bill of Human Rights[93,94], the International Labor Organization Fundamental Principles and Rights at Work[95], the United Nations Declaration on the Rights of Indigenous Peoples (UNDRIP)[87], and other key conventions and sources. For example, the use of financial and legal resources to acquire land used customarily by subsistence farmers who lack the resources to acquire legal title might be allowed by a national legal system but would be considered a 'land grab' linked to violation of internationally recognized human rights, and thus would not be acceptable to advance NCS[96].

NCS projects should be able to demonstrate respect for human rights. This usually means a policy foundation and a 'due diligence' or assessment practice that helps an NCS project identify potential human rights impacts. Drawing from existing practice under the UN Guiding Principles on Business and Human Rights, NCS proponents should strive for human rights policies[97] that acknowledge human rights frameworks, identify priority risk areas and corresponding safeguards, and explain how people who experience a rights violation related to NCS implementation can bring it to the attention of jurisdictional authorities and NCS project managers. NCS proponents should strive for due diligence practices[98] that define and continuously assess key indicators of human rights risk and invite people directly affected to co-create any needed mitigation strategies, thus supporting the multiple dimensions of procedural, distributional, and recognitional equity. Further, NCS projects should mainstream gender equity considerations[99] in all design and implementation processes and should focus due diligence and mitigation efforts on vulnerable groups and identities.

Consistent with Principle 4.2 and the convention of conservativeness, NCS activities should not proceed in the face of allegations or concerns about specific human rights impacts until a due diligence system is in place to demonstrate how the impacts have been considered and addressed in a manner consistent with the multiple dimensions of equity.

Principle 5.2: NCS respect indigenous self-determination

As a subset of human rights that particularly relates to NCS implementation, NCS should respect indigenous self-determination, including governance, knowledge, and spirituality. As such, NCS projects should aim to enhance local leadership and decision-making for both Indigenous Peoples and local communities generally.

Self-determination is a multi-dimensional collective right, most clearly articulated and protected in the UNDRIP (Art. 3)[87]. It includes enumerated articles recognizing specific indigenous collective rights, including the right to autonomy or self-government in internal or local matters (Art. 4), the right to participate in decision-making (Art. 18), the right to determine and develop priorities and strategies (Art. 23), the right to territories and resources (Art. 32), the right to give or withhold Free Prior and Informed Consent (FPIC) (Arts. 19, 32.2), the right to protect and strengthen histories, languages, oral traditions (Art. 13), cultures (Art. 15), spiritual and religious traditions (Art. 12), distinctive spiritual relationships with lands (Art. 25), traditional knowledge and cultural expressions (Art. 31), and institutional structures and practices (Art. 34)[87]. The ways in which different NCS actions might contribute to or detract from the various social, economic, and political dynamics and processes related to self-determination are deeply context specific[100]. In most cases, respect for self-determination will require promoting indigenous leadership or deep collaboration in decision-making throughout the design, implementation, monitoring, and benefit-sharing of any project or program affecting Indigenous People.

Principle 5.2 requires that all NCS actors respect FPIC rights for Indigenous Peoples, consistent with the UNDRIP. NCS actors should also ensure FPIC for any local community that could be significantly affected. Numerous tools are publicly available to help NCS actors understand and ensure FPIC is carried out[101–105]. The self-determination of local communities can and should be amplified by preserving or increasing local decision-making and control over key priorities and strategies.

It is particularly important that NCS implementation not increase security threats faced by Indigenous People or local communities, nor result in dispossession or increased pressure on communities that use land on a customary but legally insecure basis. NCS can best avoid such outcomes by embedding respect for human rights and self-determination into project design and implementation activities. A failure to demonstrate FPIC or address human rights risks could make an NCS project ineligible to register carbon credits on the compliance or voluntary carbon markets, or under the anticipated Article 6 of the UN Paris Agreement[106], and therefore could undermine the ability to achieve national climate goals. Strengthened self-determination can activate critical local knowledge and add valuable local experience to the global NCS learning and science community. NCS projects that demonstrate respect, responsibility, and equity will be more resilient, will inspire action rather than controversy, and will better advance the climate solutions that we so urgently need.

Monitoring and policy considerations

For NCS to effectively contribute to climate change mitigation, it is critical that there exist robust systems for measuring, monitoring, reporting, and verifying (MMRV) net emissions changes as a result of NCS implementation. There are widespread efforts to advance these systems, such as science and technology measurement systems for corporate supply chain inventory accounting methods (for example, Greenhouse Gas Protocol) and efforts to establish best practice MMRV through global initiatives like the Integrity Council for Voluntary Carbon Markets (IC-VCM). These initiatives are a good start and appropriately recognize the need for continual improvement. There remain important accounting uncertainties, principally in scientific best practice for establishing the baseline/counterfactual scenario against which NCS progress is measured, and leakage. Advanced remote sensing, machine learning, and impact evaluation methods from other disciplines offer rich near-term opportunities to establish a new high bar of NCS accounting. The scientific community should strive for consensus in best practices to give markets and policymakers the certainty needed to support NCS implementation at large scales. However, there remain critical uncertainties and gaps in these systems, such as whether outcomes can be accurately quantified at large scales, or how to align accounting across scales without double counting from project to value chain and national inventory. It is not a foregone conclusion that we will be able to adequately achieve the ambition of developing high quality global MMRV systems for NCS, but if we are to succeed in realizing 11 Pg of cost-effective global NCS potential, a diverse and concerted effort to accelerate the development of high-quality global MMRV systems for NCS is needed[4,24].

In this vein, these NCS principles could be used to inform efforts to achieve high quality NCS in multiple fora. In the VCM, initiatives like the IC-VCM (focused on supply side quality) and the Voluntary Carbon Markets Integrity Initiative (VCMI) (focused on demand side integrity) should align updates of their rules with these principles. Carbon buyers and investors in the VCM should ensure their market activity aligns as well. Similarly, crediting protocols for regulatory or compliance carbon markets should be modified to calculate project credits based on the change in total radiative forcing, characterize uncertainty around mitigation, demonstrate compliance with human rights due diligence practices and indigenous self-determination, and align with global best practice on use of carbon credits for compliance purposes. In short, NCS credits can be used to close the gap between readily abatable emissions and the ambition needed to meet the Paris Agreement. But NCS credits should only be used for residual emissions. This approach will require defining what counts as 'residual' in each industry, which will need to be based either on unit abatement cost (preferred but difficult to verify) or technology (suboptimal, but readily verifiable). There currently exist >30 compliance carbon markets ranging in jurisdictional scale from subnational (for example, California's Compliance Offset Projects) to supra-national (the European Union's Emissions Trading Scheme)[107], so the effort required to promote NCS principle adoption by even a sizable share of these schemes will be substantial. Another potentially powerful mechanism would be the incorporation of these principles into a country's Nationally Determined Contributions (self-defined national climate pledges under the Paris Agreement). This would likely take the form of voluntary individual country commitments unless the Paris Agreement signatories make compliance with the NCS principles mandatory for AFOLU commitments in Nationally Determined Contributions.

NCS exemplar

The Family Forest Carbon Program (FFCP) is an NCS initiative launched by the American Forest Foundation and The Nature Conservancy in early 2020[108,109]. In an effort to solve the inequitable market access of existing forest carbon projects, the FFCP was specifically designed to deliver measurable, additional climate mitigation (Principles 3 and 4) in managed natural forests (Principle 1) while maintaining a sustainable supply of timber (Principle 2) and equitable carbon revenue for small landholders in the United States (Principle 5). Since its inception, the FFCP team has validated an improved forest management methodology through Verra's Verified Carbon Standard (VCS)[110], and enrolled > 400 small landowners in climate-smart forestry practice agreements on nearly 60,000 acres. The FFCP is currently undergoing VCS project validation and initial verification, intending to deliver a first tranche of credits to vetted buyers in early 2024. We believe that FFCP adheres to the spirit and practice of the NCS principles, while continuing to evolve and improve (see Supplementary Note 2).

Outstanding challenges

While we have attempted to resolve some of the persistent confusion and controversy around NCS through the articulation of foundational and operational principles, many real issues remain where critiques and debates will be fruitful. First, much more work is needed to understand and address the feasibility constraints (inputs, markets, behaviors and attitudes, institutions, policies, and governance) that limit NCS action. Second, many ecosystem stewards view NCS benefits in the context of a broader set of benefits (for example, biodiversity, water, air, soil, human well-being, climate resilience); but more work is needed to quantify where and how NCS action can deliver these co-benefits, and where and how there are real trade-offs. Third, continuous effort is needed to ensure that NCS are indeed additional, especially to the extent that NCS activities contribute credits to carbon markets. Fourth, additionality is notoriously difficult to prove in areas with high carbon stocks and low historic rates of disturbance (such as high forest cover, low deforestation zones), despite real increasing future threats; the degree and timing of risk that these forests face needs to be better quantified to determine the relative additionality of ongoing actions to protect them[111,112]. Fifth, additional research is needed to ensure that NCS mitigation remains durable to future disturbances, especially the droughts and wildfires that are expected to increase with climate change. Sixth, NCS science has, to date, largely focused on measuring NCS opportunity, but to be successful, we need consistent, compatible NCS monitoring systems to accurately quantify impacts and learn adaptively. Seventh, as we expand the scale of NCS action in an adaptive management cycle, we need a rapid global learning network to replicate successes and prevent repeating mistakes. Finally, the NCS community as a whole needs to demonstrate a commitment to equity by creatively and continuously seeking ways to recognize and integrate the leadership and, with their consent, the knowledge and experience of Indigenous Peoples and other NCS stewards.

Conclusion

The coauthors of this paper believe it is time for NCS action. Debate and discussion are a healthy component of applied science, and NCS is no exception. But the urgency of our climate predicament requires human society to adopt a culture of adaptive management, in which climate solutions (natural and otherwise) can adapt rapidly and transparently, in concert with their widespread adoption. The foundational and operational principles outlined in this paper are intended to help resolve confusion to expedite action while also fostering discussion and learning focused on important outstanding questions. Many fora are emerging for this type of productive action-oriented NCS learning and conversation (for example, naturebase.org and restor.eco). We hope that these principles facilitate urgent, productive, and collective action toward the widespread adoption of robust NCS projects.

References

1. Arrhenius, S. *Ueber den Einfluss des atmosphärischen Kohlensäuregehalts auf die Temperatur der Erdoberfläche*. (PA Norstedt & Söner, 1876).
2. Dyson, F. Can we control the Carbon dioxide in the atmosphere? *Energy* **2**, 287–291 (1977).
3. Houghton, J. et al. *Climate change 2001: the scientific basis*. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. pp 285–310 (2001).
4. Griscom, B. W. et al. Natural climate solutions. *Proc. Natl. Acad. Sci. USA* **114**, 11645–11650 (2017).
5. Leavitt, S. et al. *Natural Climate Solutions Handbook: A Technical Guide for Assessing Nature-based Mitigation Opportunities in Countries*. https://www.nature.org/content/dam/tnc/nature/en/documents/TNC_Natural_Climate_Solutions_Handbook.pdf (2021).
6. Climate Policy Initiative. *Landscape of Climate Finance for Agriculture, Forestry, Other Land Use and Fisheries: Preliminary Findings*. https://www.climatepolicyinitiative.org/wp-content/uploads/2022/11/Landscape-of-Climate-Finance-for-Agriculture-Forestry-Other-Land-Uses-and-Fisheries.pdf (2022).
7. United Nations Environment Programme. *State of Finance for Nature. Time to Act: Doubling Investment by 2025 and Eliminating Nature-Negative Finance Flows*. https://wedocs.unep.org/20.500.11822/41333 (2022).
8. IPCC. *Climate Change 2022, Mitigation of Climate Change: Working Group III Contribution to the Sixth Assessment Report*. 2030 https://www.ipcc.ch/report/ar6/wg3/ (2022).
9. Buchner, B. et al. *Global Landscape of Climate Finance 2023*. https://www.climatepolicyinitiative.org/publication/global-landscape-of-climate-finance-2023/ (2023).
10. Fleischman, F. et al. Pitfalls of tree planting show why we need people-centered natural climate solutions. *BioScience* (2020). https://doi.org/10.1093/biosci/biaa094.
11. Jones, B. No One Knows What "Nature-based Solutions" are. *Vox* https://www.vox.com/down-to-earth/2022/11/17/23460525/nature-based-solutions-climate-change-cop27 (2022).
12. IPES-Food. *Smoke and Mirrors: Examining Competing Framings of Food System Sustainability: Agroecology, Regenerative Agriculture, and Nature-Based Solutions*. https://ipes-food.org/_img/upload/files/SmokeAndMirrors.pdf (2022).
13. Cohen-Shacham, E., Walkers, G., Janzen, C., & Maginnis, S. (eds). *Nature-Based Solutions to Address Global Societal Challenges*. 1–97 (2016).
14. IUCN. *Global Standard for Nature-based Solutions: A User-friendly Framework for the Verification, Design and Scaling Up of NbS*. https://doi.org/10.2305/IUCN.CH.2020.08.en (2020).
15. United Nations Environment Assembly of the United Nations Environment Programme. *Nature-Based Solutions for Supporting Sustainable Development*. *UNEP/EA.5/Res.5* (2022).
16. Duarte, C. M., Wu, J., Xiao, X., Bruhn, A. & Krause-Jensen, D. Can seaweed farming play a role in climate change mitigation and adaptation? *Front. Mar. Sci.* **4**, 100 (2017).
17. Buma, B. et al. Nature-based climate solutions for quantifiable climate mitigation - some are ready, some are not. *Nat. Clim. Chang.* IN REVIEW.
18. Novick, K. A. et al. Informing nature-based climate solutions for the United States with the best-available science. *Glob. Chang. Biol.* **28**, 3778–3794 (2022).
19. Elmqvist, T. et al. Benefits of restoring ecosystem services in urban areas. *Curr. Opin. Environ. Sustain.* **14**, 101–108 (2015).
20. Hajjar, R. et al. A global analysis of the social and environmental outcomes of community forests. *Nat. Sustain.* **4**, 216–224 (2021).
21. Salzman, J., Bennett, G., Carroll, N., Goldstein, A. & Jenkins, M. The global status and trends of payments for ecosystem services. *Nat. Sustain.* **1**, 136–144 (2018).
22. Weatherley-Singh, J. & Gupta, A. Drivers of deforestation and REDD+ benefit-sharing: A meta-analysis of the (missing) link. *Environ. Sci. Polic.* **54**, 97–105 (2015).
23. Howard, J. et al. Blue carbon pathways for climate mitigation: Known, emerging and unlikely. *Mar. Polic.* **156**, 105788 (2023).
24. Roe, S. et al. Contribution of the land sector to a 1.5 °C world. *Nat. Clim. Change* **9**, 817–828 (2019).
25. Girardin, C. A. J. et al. Nature-based solutions can help cool the planet — if we act now. *Nature* **593**, 191–194 (2021).
26. Seddon, N. et al. Getting the message right on nature-based solutions to climate change. *Glob. Change Biol.* **27**, 1518–1546 (2020).
27. Seddon, N. et al. Understanding the value and limits of nature-based solutions to climate change and other global challenges. *Philos. Trans. R. Soc. B* **375**, 20190120 (2020).
28. Schulte, I., Eggers, J., Nielsen, J. Ø. & Fuss, S. What influences the implementation of natural climate solutions? A systematic map and review of the evidence. *Environ. Res. Lett.* **17**, 013002 (2022).
29. Poeplau, C. & Don, A. Carbon sequestration in agricultural soils via cultivation of cover crops – a meta-analysis. *Agric Ecosyst. Environ.* **200**, 33–41 (2015).
30. Ellis, E. C. et al. People have shaped most of terrestrial nature for at least 12,000 years. *Proc. Natl Acad. Sci.* **118**, e2023483118 (2021).
31. Wolosin, M. *Exponential Roadmap for Natural Climate Solutions*. https://www.conservation.org/priorities/exponential-roadmap-natural-climate-solutions (2022).
32. Bremer, L. L. & Farley, K. A. Does plantation forestry restore biodiversity or create green deserts? A synthesis of the effects of land-use transitions on plant species richness. *Biodivers. Conserv.* **19**, 3893–3915 (2010).
33. Strong, A., Chisholm, S., Miller, C. & Cullen, J. Ocean fertilization: time to move on. *Nature* **461**, 347–348 (2009).
34. Burns, W. & Corbett, C. R. Antacids for the sea? artificial ocean alkalinization and climate change. *One Earth* **3**, 154–156 (2020).
35. Ross, F., Tarbuck, P. & Macreadie, P. I. Seaweed afforestation at large-scales exclusively for carbon sequestration: Critical assessment of risks, viability and the state of knowledge. *Front. Mar. Sci.* **9**, 1015612 (2022).
36. United Nations. Convention on Biological Diversity. http://www.cbd.int/doc/legal/cbd-en.pdf (1992).
37. Bass, A. M., Bird, M. I., Kay, G. & Muirhead, B. Soil properties, greenhouse gas emissions and crop yield under compost, biochar and co-composted biochar in two tropical agronomic systems. *Sci. Total Environ.* **550**, 459–470 (2016).
38. Veldman, J. W. et al. Tyranny of trees in grassy biomes. *Science* **347**, 484–485 (2015).

39. Worm, B. & Paine, R. T. Humans as a hyperkeystone species. *Trends Ecol. Evol.* **31**, 600–607 (2016).
40. Fujimori, S. et al. Land-based climate change mitigation measures can affect agricultural markets and food security. *Nat. Food* **3**, 110–121 (2022).
41. Sela, S. et al. Adapt-N outperforms grower-selected nitrogen rates in northeast and midwestern United States strip trials. *Agron. J.* **108**, 1726–1734 (2016).
42. Tölgyesi, C. et al. Maintaining scattered trees to boost carbon stock in temperate pastures does not compromise overall pasture quality for the livestock. *Agric Ecosyst. Environ.* **351**, 108477 (2023).
43. Searchinger, T. et al. Use of U.S. croplands for biofuels increases greenhouse gases through emissions from land-use change. *Science* **319**, 1238–1240 (2008).
44. Fargione, J. E., Plevin, R. J. & Hill, J. D. The ecological impact of biofuels. *Annu. Rev. Ecol. Evol. Syst.* **41**, 351–377 (2010).
45. Tilman, D. & Clark, M. Global diets link environmental sustainability and human health. *Nature* **515**, 518–522 (2014).
46. Radkau, J. *Wood: A History* (Polity, Cambridge, UK, 2012).
47. Putz, F. E. et al. Sustained timber yield claims, considerations, and tradeoffs for selectively logged forests. *PNAS Nexus* **1**, 1–7 (2022).
48. Bradshaw, C. J. A., Sodhi, N. S., Peh, K. S. H. & Brook, B. W. Global evidence that deforestation amplifies flood risk and severity in the developing world. *Glob. Change Biol.* **13**, 2379–2395 (2007).
49. Beck, M. W. et al. The global flood protection savings provided by coral reefs. *Nat. Commun.* **9**, 2186 (2018).
50. Huang, L., Shao, Q. & Liu, J. Forest restoration to achieve both ecological and economic progress, Poyang Lake basin, China. *Ecol. Eng.* **44**, 53–60 (2012).
51. Parsons, L. A. et al. Tropical deforestation accelerates local warming and loss of safe outdoor working hours. *One Earth* **4**, 1730–1740 (2021).
52. Wolff, N. H. et al. The effect of deforestation and climate change on all-cause mortality and unsafe work conditions due to heat exposure in Berau, Indonesia: a modelling study. *Lancet Planet Heal* **5**, e882–e892 (2021).
53. Zeppetello, L. R. V. et al. Consistent cooling benefits of silvopasture in the tropics. *Nat. Commun.* **13**, 708 (2022).
54. Gillenwater, M. *What Is Additionality? Part 2: A framework for a more precise definition and standardized approaches* (2012). Discussion Paper No. 002 (version 3).
55. Cook-Patton, S. C., Shoch, D. & Ellis, P. W. Dynamic global monitoring needed to use restoration of forest cover as a climate solution. *Nat. Clim. Change* **11**, 366–368 (2021).
56. Wood, S. A. et al. Mitigating near-term climate change. *Environ. Res. Lett.* **18**, 1–4 (2023).
57. Leifeld, J. Carbon farming: Climate change mitigation via non-permanent carbon sinks. *J. Environ. Manag* **339**, 117893 (2023).
58. Matthews, H. D., Zickfeld, K., Koch, A. & Luers, A. Accounting for the climate benefit of temporary carbon storage in nature. *Nat. Commun.* **14**, 5485 (2023).
59. Zou, J. et al. Rewetting global wetlands effectively reduces major greenhouse gas emissions. *Nature Geosci.* (2022) https://doi.org/10.1038/s41561-022-00989-0.
60. Anderson, V. & Gough, W. A. A typology of nature-based solutions for sustainable development: an analysis of form, function, nomenclature, and associated applications. *Land* **11**, 1072 (2022).
61. Griscom, B. W. et al. We need both natural and energy solutions to stabilize our climate. *Glob. Change Biol.* **25**, 1889–1890 (2019).
62. Science-Based Targets Initiative. *SBTi Corporate Net Zero Standard, ver 1*. 64 https://sciencebasedtargets.org/resources/files/Net-Zero-Standard.pdf (2001).
63. Natural Climate Solutions Alliance. *Natural Climate Solutions for Corporates*. https://www3.weforum.org/docs/WEF_NCSA_NCS_for_Corporates_2021.pdf (2021).
64. Voluntary Carbon Markets Integrity Initiative. *VCMI Provisional Code of Practice: Draft For Public Consultation and Corporate Road Testing*. https://vcmintegrity.org/wp-content/uploads/2022/06/VCMI-Provisional-Claims-Code-of-Practice.pdf (2022).
65. Goldstein, A. *Buying In: Taking Stock of the Role of Offsets in Corporate Carbon Strategies*. https://www.forest-trends.org/publications/buying-in/ (2016).
66. Seymour, F., Wolosin, M. & Gray, E. *Not Just Carbon: Capturing all the Benefits of Forests for Stabilizing the Climate from Local to Global Scales*. https://doi.org/10.46830/wrirpt.19.00004. (2022).
67. Rohatyn, S., Yakir, D., Rotenberg, E. & Carmel, Y. Limited climate change mitigation potential through forestation of the vast dryland regions. *Science* **377**, 1436–1439 (2022).
68. IPCC. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. https://report.ipcc.ch/ar6/wg1/IPCC_AR6_WGI_FullReport.pdf (2021).
69. Lynch, J., Cain, M., Pierrehumbert, R. & Allen, M. a means of reporting warming-equivalent emissions that captures the contrasting impacts of short- and long-lived climate pollutants. *Environ. Res. Lett.* **15**, 044023 (2020).
70. Neag, R. & Maşca, E. Identifying accounting conservatism – a literature review. *Proc. Econ. Financ* **32**, 1114–1121 (2015).
71. Drever, C. R. et al. Natural climate solutions for Canada. *Sci. Adv.* **7**, eabd6034 (2021).
72. Fargione, J. E. et al. Natural climate solutions for the United States. *Sci. Adv.* **4**, eaat1869 (2018).
73. Ellis, P. W. et al. Reduced-impact logging for climate change mitigation (RIL-C) can halve selective logging emissions from tropical forests. *For. Ecol. Manag.* **438**, 255–266 (2019).
74. Finlayson, C., Roopsind, A., Griscom, B. W., Edwards, D. P. & Freckleton, R. P. Removing climbers more than doubles tree growth and biomass in degraded tropical forests. *Ecol. Evol.* **12**, 1–13 (2022).
75. IPCC. *AR6 WGIII CDR Factsheet: Carbon Dioxide Removal*. 2 https://www.ipcc.ch/report/ar6/wg3/downloads/outreach/IPCC_AR6_WGIII_Factsheet_CDR.pdf (2022).
76. Jones, M. W. et al. National contributions to climate change due to historical emissions of carbon dioxide, methane, and nitrous oxide since 1850. *Sci. Data* **10**, 155 (2023).
77. Weiss, E. B. Chapter 5: Implementing intergenerational equity. in *Research Handbook on International Environmental Law* (eds Fitzmaurice, M. Ong, D. M. & Merkouris, P.) 100–116 (Edward Elgar Publishing, Inc., 2010).
78. King, A. D. & Harrington, L. J. The inequality of climate change From 1.5 to 2 °C of global warming. *Geophys. Res. Lett.* **45**, 5030–5033 (2018).
79. Brondizio, E., Diaz, S., Settele, J. & Ngo, H. T. *Global Assessment Report on Biodiversity and Ecosystem Services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*. https://doi.org/10.5281/zenodo.3831673 (2019).
80. Frechette, A. et al. *A Global Baseline of Carbon Storage in Collective Lands: Indigenous and Local Community Contributions to Climate Change Mitigation*. https://rightsandresources.org/wp-content/uploads/2018/09/A-Global-Baseline_RRI_Sept-2018.pdf (2018).
81. FAO & FILAC. *Forest Governance by Indigenous and Tribal Peoples. An Opportunity for Climate Action in Latin America and the Caribbean*. https://doi.org/10.4060/cb2953en (2021).
82. Garnett, S. T. et al. A spatial overview of the global importance of Indigenous lands for conservation. *Nat. Sustain* **1**, 369–374 (2018).

83. Schuster, R., Germain, R. R., Bennett, J. R., Reo, N. J. & Arcese, P. Vertebrate biodiversity on indigenous-managed lands in Australia, Brazil, and Canada equals that in protected areas. *Environ. Sci. Policy* **101**, 1–6 (2019).
84. Nussbaum, M. C. *Women and Human Development* (Cambridge University Press, 2000).
85. Agyeman, J., Bullard, R. D. & Evans, B. *Just Sustainabilities: Development in an Unequal World* (MIT Press, 2003).
86. United Nations. *The Human Right to a Clean, Healthy and Sustainable Environment*. A/RES/76/300. https://digitallibrary.un.org/record/3983329?ln=en (2022).
87. United Nations. *United Nations Declaration on the Rights of Indigenous Peoples*. A/RES/61/295. https://www.un.org/en/genocideprevention/documents/atrocity-crimes/Doc.18_declaration%20rights%20indigenous%20peoples.pdf (2007).
88. Pascual, U. et al. Social equity matters in payments for ecosystem services. *BioScience* **64**, 1027–1036 (2014).
89. McDermott, M., Mahanty, S. & Schreckenberg, K. Examining equity: a multidimensional framework for assessing equity in payments for ecosystem services. *Environ. Sci. Policy* **33**, 416–427 (2013).
90. United Nations Framework Convention on Climate Change. REDD + Safeguards:Reducing Emissions from Deforestation and Forest Degradation in Developing Countries. https://redd.unfccc.int/fact-sheets/safeguards.html (2015).
91. Green Climate Fund. Environmental and Social Safeguards | Green Climate Fund. https://www.greenclimate.fund/projects/safeguards/ess (2023).
92. Verra. *Climate, Community & Biodiversity Standards Third Edition, v3.1*. https://verra.org/wp-content/uploads/CCB-Standards-v3.1_ENG.pdf (2017).
93. United Nations. *International Covenant on Economic, Social and Cultural Rights*. resolution 2200A (XXI). https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-economic-social-and-cultural-rights (1966).
94. United Nations General Assembly. *International Covenant on Civil and Political Rights*. resolution 2200A (XXI). https://www.ohchr.org/sites/default/files/ccpr.pdf (1966).
95. International Labour Organization. *ILO Declaration on Fundamental Principles and Rights at Work*. https://www.ilo.org/wcmsp5/groups/public/---ed_norm/---declaration/documents/normativeinstrument/wcms_716594.pdf (2022).
96. Holland, M. B., Masuda, Y. J. & Robinson, B. E. *Land Tenure Security and Sustainable Development*. https://doi.org/10.1007/978-3-030-81881-4 (2022).
97. Amis, L. et al. *A Guide for Business: How to Develop a Human Rights Policy*. https://www.ohchr.org/sites/default/files/Documents/Issues/Business/guide-business-hr-policy.pdf (2015).
98. United Nations Human Rights Council. *Working Group on the Issue of Human Rights and Transnational Corporations and Other Business Enterprises*. https://documents-dds-ny.un.org/doc/UNDOC/GEN/N18/224/87/PDF/N1822487.pdf?OpenElement (2018).
99. UN Women. *Gender Mainstreaming: A Global Strategy for Achieving Gender Equality and the Empowerment of Women and Girls* https://www.unwomen.org/sites/default/files/Headquarters/Attachments/Sections/Library/Publications/2020/Gender-mainstreaming-Strategy-for-achieving-gender-equality-and-empowerment-of-women-girls-en.pdf (2020).
100. The Nature Conservancy. *The Voice, Choice, and Action Framework: A Conservation Practitioner's Guide to Indigenous and Community-Led Conservation, Version 2.0*. https://tncvoicechoiceaction.org/wp-content/uploads/TNC-VCAFramework-2.0-220502-FRLR.pdf (2022).
101. The Nature Conservancy. *The Nature Conservancy's Human Rights Guide for Working with Indigenous Peoples and Local Communities*. https://www.tnchumanrightsguide.org/ (2020).
102. Kennedy, T., Martin, T., Lee, M. & RESOLVE. *The Practice of FPIC: Insights from the FPIC Solutions Dialogue*. http://www.fpicdialogue.org (2021).
103. Equitable Origin, Roundtable on Sustainable Biomaterials & Coordinator of Indigenous Organizations. FPIC-360°. https://fpic360.org/.
104. World Wildlife Fund. *WWF Environmental and Social Safeguards Framework*. https://wwfint.awsassets.panda.org/downloads/wwf_environmental_and_social_safeguards_framework_consultation_draft.pdf (2019).
105. Mchelhinny, V., Degawan, M., Dunne, P. & Cruz, A. *Indigenous Negotiations Resource Guide*. https://www.conservation.org/docs/default-source/publication-pdfs/ci-indigenous-negotiations-resource-guide.pdf?sfvrsn=8ef3136a_3 (2021).
106. Eisen, N. et al. *Rights, Carbon, Caution: Upholding Human Rights Under Article 6 of the Paris Agreement*. https://www.ciel.org/wp-content/uploads/2021/02/Rights-Carbon-Caution.pdf (2021).
107. World Bank. *State and Trends of Carbon Pricing*. https://openknowledge.worldbank.org/entities/publication/58f2a409-9bb7-4ee6-899d-be47835c838f (2023).
108. Galik, C. S., Baker, J. S., Daigneault, A. & Latta, G. Crediting temporary forest carbon: Retrospective and empirical perspectives on accounting options. *Front. Glob. Change* **5**, 933020 (2022).
109. American Forest Foundation. Family Forest Carbon Program. https://www.forestfoundation.org/what-we-do/increase-carbon-storage/family-forest-carbon-program/ (2023).
110. Shoch, D. et al. *VM0045 Methodology for Improved Forest Management Through Reduced Impact Logging v1.0 - Verra*. https://verra.org/methodology/vm0035-methodology-for-improved-forest-management-through-reduced-impact-logging-v1-0/ (2022).
111. Funk, J. M. et al. Securing the climate benefits of stable forests. *Clim. Policy* **19**, 1–16 (2019).
112. Griscom, B., Shoch, D., Stanley, B., Cortez, R. & Virgilio, N. Sensitivity of amounts and distribution of tropical forest carbon credits depending on baseline rules. *Environ. Sci. Policy* **12**, 897–911 (2009).

Acknowledgements
The authors wish to thank the Bezos Earth Fund for their support to write this paper; Lucy Almond for inspiration to write it; Frances Seymour for her wisdom, guidance, careful review, and deep input into its framing; Kemen Austin for her excellent guidance and review; Amardeep Sidhu from Onalytica for providing NCS Twitter monitoring statistics; Doria Gordon, Mark Moroge, and Eric Holst at the Environmental Defense Fund for their partnership and review; Ali Surdoval, Chris Zganjar, Kelley Hamrick, Melissa Gallant, Ayla Norris-Smith, and Lindsay Smart for review and comments; Julia Kim for help with figures and graphics; and Christine Cadigan and Ethan Belair for their help articulating the Family Forest Carbon Program NCS Exemplar.

Author contributions
P.W.E., B.G., and R.C. conceived the idea of NCS Principles, P.W.E. designed the paper and coordinated all co-author input. P.W.E., A.M.P., S.A.W., J.F., Y.J.M., V.C.D., C.M., B.G., T.K., J.S., T.A., R.C., and S.C-P. directly contributed to drafting and review of the paper. S.C-P. provided guidance as senior author. S.L. provided comprehensive review for clarity, accuracy, and consistency.

Competing interests
The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-023-44425-2.

Correspondence and requests for materials should be addressed to Peter Woods Ellis.

Peer review information *Nature Communications* thanks Kyle Hemes and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.