

APPLICATION PAPER 

Bayesian state-space synthetic control method for deforestation baseline estimation for forest carbon credits

Keisuke Takahata¹ , Hiroshi Suetsugu¹, Keiichi Fukaya² and Shinichiro Shirota³

¹sustainacraft Inc., Tokyo, Japan

²Biodiversity Division, National Institute for Environmental Studies, Ibaraki, Japan

³Center for the Promotion of Social Data Science Education and Research, Hitotsubashi University, Tokyo, Japan

Corresponding author: Shinichiro Shirota; Email: shinichiro.shirota@gmail.com

Received: 28 February 2023; **Revised:** 28 September 2023; **Accepted:** 19 January 2024

Keywords: Bayesian modeling; causal inference; state-space modeling; carbon credit; REDD+

Abstract


Carbon credits from the reducing emissions from deforestation and degradation (REDD+) projects have been criticized for issuing junk carbon credits due to invalid ex-ante baselines. Recently, the concept of ex-post baseline has been discussed to overcome the criticism, while ex-ante baseline is still necessary for project financing and risk assessment. To address this issue, we propose a Bayesian state-space model that integrates ex-ante baseline projection and ex-post dynamic baseline updating in a unified manner. Our approach provides a tool for appropriate risk assessment and performance evaluation of REDD+ projects. We apply the proposed model to a REDD+ project in Brazil and show that it may have had a small, positive effect but has been overcredited. We also demonstrate that the 90% predictive interval of the ex-ante baseline includes the ex-post baseline, implying that our ex-ante estimation can work effectively.

Impact Statement

Our novel approach estimates deforestation baselines for forest carbon credits, addressing criticisms of junk carbon credit issued from REDD+ projects and problems of early financing simultaneously. By integrating ex-ante baseline projection and ex-post dynamic updating, we evaluate ex-ante risks and ex-post project performance in a unified manner. Our approach can contribute to sound allocation of funds for forest conservation projects with significant positive impacts on climate change action.

1. Introduction

Carbon credit is an incentive scheme to promote projects that have additional benefits for climate change mitigation, and is expected to play an important role in offsetting the gap from net zero emission after reduction efforts (Griscom et al., 2020). Reducing deforestation and forest degradation are considered to be one of the most effective approaches to reduce carbon emission and the reducing emissions from deforestation and degradation (REDD+) mechanism is a framework to promote such efforts through the issuance of carbon credit. However, carbon credits from REDD+ have been subject to several criticisms.

 This research article was awarded Open Materials badge for transparent practices. See the Data Availability Statement for details.

© The Author(s), 2024. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Credits issued for projects without actual positive effects on climate change mitigation are called “junk carbon credit”, and several studies have showed that many REDD+ projects may have produced junk carbon credits (e.g., West et al., 2020, 2023; Guizar-Coutiño et al., 2022).

Criticisms to carbon credit are mainly about the validity of ex-ante baseline, that is, a counterfactual scenario in the absence of a project that is estimated before the project starts. Considering this issue, the concept of ex-post (or dynamic) baseline has recently been discussed (Verra, 2022a). For example, the use of synthetic control method (SCM) has been studied in this context (West et al., 2020, 2023). In this framework, baseline is sequentially updated at every observation of the forest cover after intervention, allowing for the effects of changes in the external environment to be taken into account.

However, a financing issue still remains as result-based payment means project proponents must wait several years to obtain the first credit issuance. From an investor perspective, ex-ante baseline projection is needed to quantify the risk associated with projects for investment decisions (Verra, 2022b). With those in mind, we can find a need for the integration of both ex-ante baseline forecasting before intervention and ex-post baseline updating at each observation after intervention.

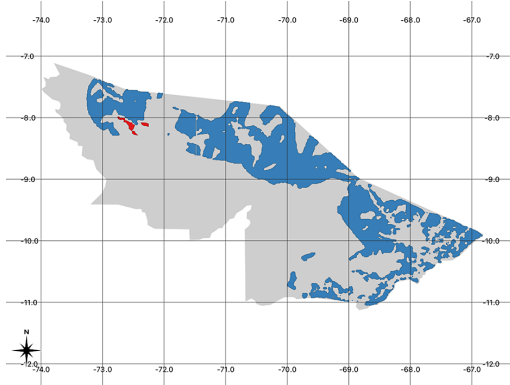
We propose a new model to address the above problem. First, we introduce a Bayesian state-space model that naturally integrates the estimation of ex-ante baseline and the dynamic updating of ex-post baseline. We achieve this by combining state-space modeling for forecasting and SCM for dynamic updating. Second, we consider covariate balancing in state-space modeling by using the general Bayesian updating method for valid causal inference. We also address a small sample problem, which often arises in time-series causal inference, by introducing the regularized Horseshoe prior. Finally, we apply the proposed model to a REDD+ project in Brazil and show that both ex-ante and ex-post baseline can be effectively estimated by our model. Our approach would enable appropriate ex-ante risk assessment and ex-post performance evaluation of forest conservation projects and contribute to the sound allocation of funds to projects that have significant positive impacts on climate change mitigation.

2. Preliminaries and related work

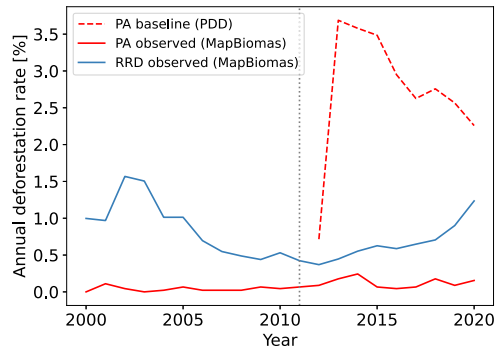
2.1. VM0007: A REDD+ methodology for emission reduction evaluation

VM0007 (Verra, 2020) is one of the major methodologies that defines the calculation of emission reductions in a REDD+ project. A key concept of VM0007 is the Reference Region for projecting the rate of deforestation (RRD), which is equivalent to a control unit in causal inference. The RRD is chosen to match the project area (PA) as closely as possible in terms of the following variables: deforestation drivers, landscape factors (e.g. forest types, soil types, elevation, etc.), and socioeconomic variables (access to infrastructures, policies, and regulations, etc.). For the 10–12 years preceding the implementation of a project, deforestation rates are aggregated over the RRD and projected as a baseline for the crediting period. The projection method can be a simple historical average or a predefined linear/nonlinear model, where the former is often used. After the projection, spatial modeling is applied for certain types of forest to adjust the baseline.

Several studies have reported that baselines set under VM0007 were overestimated because these baselines failed to consider a counterfactual scenario correctly or to eliminate the effect of external factors, such as policy changes (West et al., 2020, 2023). As an example, Figure 1 describes the spatial distribution and the deforestation trends of the PA and RRD of the Valparaiso project (Eaton et al., 2014), where the project started in 2011. The deforestation rates are calculated using MapBiomas Brazil (Souza et al., 2020), which will be used in the case study in Section 4. The baseline reported in the Project Design Document (PDD) is also included in the chart. We can see that the baseline (the red dashed line in Figure 1b) might have been overestimated considering the actual deforestation level in the PA. In addition, we can also see that the baseline might have failed to capture the change of trends where the deforestation rates in the RRD kept decreasing from 2002 to 2012 but changed to an upward trend again. The main cause of these changes may be the implementation and change of Brazil’s National Climate Change Plan (Ferrante and Fearnside, 2019; West et al., 2019, 2020). The current methodology cannot effectively reflect such impacts to the baseline.



(a) Spatial distribution of the PA (red) and RRD (blue) of the Valparaíso project in the state of Acre, Brazil (gray). The RRD is chosen to have similar characteristics as the PA in terms of deforestation risk (the image is made by the authors based on the PDD (Eaton et al., 2014).)



(b) Annual deforestation rates in the PA (red solid line) and RRD (blue solid line) calculated using MapBiomas, and the baseline by reported in the PDD (red dashed line). The project started at 2011 (vertical dashed line).

Figure 1. The Valparaíso project.

2.2. Related work

The SCM (Abadie et al., 2010) is one of the most popular methods for causal inference with time-series data. This method is designed for a case with one treatment unit and multiple control units, which is suited for the evaluation of a REDD+ project. Given that the RRD consists of multiple subunits (hereinafter “control units”), SCM finds an optimal weight to match both preintervention deforestation trends and covariates of the synthetic control (i.e., the weighted average of control units) to those of the PA. Note that a baseline estimated by SCM can reflect effects of external factors occurred after intervention while it cannot forecast the baseline, because it is calculated based on observations after intervention. CausalImpact (Brodersen et al., 2015) is another popular method based on Bayesian state-space models. It has several ideas in common with SCM, but there are several differences. In contrast to SCM, CausalImpact can forecast a baseline with a little modification of the original model since it is based on state-space modeling, while it does not include a covariate balancing mechanism between a treatment and synthetic controls.

Sparse estimation needs to be considered to deal with small sample problems when applying time-series causal inference models to REDD+ program evaluations. In these cases, the sample size is the number of time points of deforestation data before a project starts. Since deforestation data are often updated annually and available from the 1980s at the earliest, the sample size is usually around 30–40. On the other hand, a considerable number of control units can be necessary to get a good synthetic control, which may lead to a larger number of parameters than the sample size. Several studies consider the use of penalization methods in SCMs in the frequentist perspective (Doudchenko and Imbens, 2016; Abadie, 2021; Ben-Michael et al., 2021). In CausalImpact, the spike-and-slab prior (George and McCulloch, 1993) is employed to select influential samples on the baseline. Although the spike-and-slab prior is a natural solution to variable/sample selection problems in Bayesian modeling, it leads to substantial computational costs. The Horseshoe prior is a continuous analogue of the spike-and-slab prior and is computationally more tractable (Carvalho et al., 2010). If some prior knowledge is available about the degree of sparsity, it can be incorporated by specifying a prior on the global shrinkage parameter (Pironen and Vehtari, 2017).

In the context of general carbon crediting schemes including REDD+, the issue of junk carbon credit or overcrediting has been discussed for a long time (e.g., Haya, 2009; Aldy and Stavins, 2012; Badgley et al., 2022), and one of the sources has been identified to be baseline setting (Bento et al., 2016; Haya et al., 2020). SCM and CausalImpact have been applied to several studies evaluating REDD+ projects (Roopsind et al., 2019; Correa et al., 2020; Ellis et al., 2020; West et al., 2020, 2023) and other forest conservation activities (Sills et al., 2015; Rana and Sills, 2018; Simmons et al., 2018) in considering a counterfactual scenario. Many of the studies using causal inference methods have reported some degree of overcrediting. To the best of our knowledge, there is no study in the literature concerning ex-ante prediction and ex-post evaluation at once.

3. Model

3.1. Bayesian state-space SCM

We propose a Bayesian state-space SCM, leveraging the time-series modeling in CausalImpact and the covariate balancing in SCM. Consider a REDD+ project that has one PA and whose RRD is divided as multiple control units. We model the preintervention deforestation rates of the PA and the control units by the following state-space model with a local linear trend:

$$y_t = \tilde{z}_t' \beta + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_y^2), \quad (1)$$

$$z_t = \tilde{z}_t + e_t, \quad e_t \sim N(0, \sigma_z^2 I), \quad (2)$$

$$\tilde{z}_{t+1} = \tilde{z}_t + v_t + \eta_t, \quad \eta_t \sim N(0, \sigma_z^2 I), \quad (3)$$

$$v_{t+1} = v_t + \zeta_t, \quad \zeta_t \sim N(0, \sigma_v^2 I), \quad (4)$$

$$t = 1, \dots, T_0,$$

where y_t is an observed deforestation rate for the PA at time t , $z_t = (z_{1,t}, \dots, z_{J,t})'$ is a vector of observed deforestation rates of the control units at time t , J is the number of the control units, β is a weight to be applied to the control units defined on the simplex,

$$\beta \in S^{J-1} = \left\{ \beta \mid \sum_{j=1}^J \beta_j = 1, \beta_j \geq 0, j = 1, \dots, J \right\}, \quad (5)$$

$\tilde{z}_t = (\tilde{z}_{1,t}, \dots, \tilde{z}_{J,t})'$ is a latent state vector, and $v_t = (v_{1,t}, \dots, v_{J,t})'$ is a vector of slope components of the local linear trend. Time $t = 1$ denotes the year when observation started and $t = T_0$ the previous year before the intervention. Equations (1)–(2) link the observed data, y_t and z_t , to the latent state, \tilde{z}_t : it assumes that for the control units the deforestation rates are observed as the addition of the latent state and noise, while for the PA, the deforestation rate is written as the addition of the weighted sum of the latent states of the control units and noise (Brodersen et al., 2015). The latter relates this model to SCM except for covariate balancing. Equations (3)–(4) define the temporal evolution of the latent state by the simple local linear trend model, which enables us to forecast z_t , and thus the baseline.

We consider a shrinkage prior on β to select important control units for more interpretable and stable results. Sparse estimation is needed because the sample size T_0 tends to be small due to data availability, while we need to take a satisfactory number of the control units J into account so that there exists some weight that well approximates the PA. Here we introduce the regularized Horseshoe prior proposed by Piironen and Vehtari (2017): for $j = 1, \dots, J$,

$$\begin{aligned}
 p(\beta_j | \lambda_j, \tau, c) &\propto \frac{1}{2\tilde{\lambda}_j\tau} \exp\left(-\frac{\beta_j^2}{2\tilde{\lambda}_j^2\tau^2}\right), \\
 \lambda_j &\sim C^+(0, 1), \quad \tilde{\lambda}_j^2 = \frac{c^2\lambda_j^2}{c^2 + \tau^2\lambda_j^2}, \quad c^2 \sim \text{Inv-Gamma}(1/2, 1/2), \\
 \tau | \sigma_y &\sim C^+(0, \tau_0), \quad \tau_0 = \frac{m_0}{J - m_0} \frac{\sigma_y}{\sqrt{T_0}},
 \end{aligned}
 \tag{6}$$

where $C^+(0, 1)$ is the standard half-Cauchy distribution, τ is a global shrinkage parameter, λ_j is a local shrinkage parameter, and m_0 is the effective number of nonzero coefficients in β , which is given as a hyperparameter from prior knowledge. The global shrinkage parameter shrinks all coefficients toward zero, while the local shrinkage parameter controls the shrinkage level of each coefficient. Although the global shrinkage parameter has large impacts on the results, the systematic way of specifying a prior for τ based on the information about the sparsity has not been well discussed. Piironen and Vehtari (2017) provide the systematic way of specifying a prior for τ based on prior knowledge of the sparsity. We can control the number of control units that show significantly large positive weights through the effective sample size m_0 .

For covariate balancing, we rely on the method of general Bayesian updating (Bissiri et al., 2016). With this, we reflect the distance of the covariates between the PA and synthetic controls as a covariate-dependent prior on β . Together with (6), we set the prior on β as

$$p(\beta | \{x_j\}_{j=1}^{J+1}, \lambda, \tau, c) \propto \exp(-wL(\beta; \{x_j\}_{j=1}^{J+1})) \prod_{j=1}^J p(\beta_j | \lambda_j, \tau, c),
 \tag{7}$$

where x_j is a $K \times 1$ vector of the covariates with $j = J + 1$ for the PA, L is a loss function that measures a distance between the PA and synthetic controls, and w is a tuning parameter. Here we choose L to be a SCM-like loss function for covariate balancing:

$$L(\beta; \{x_j\}_{j=1}^{J+1}) = 1/2 \cdot (x_{J+1} - X_0'\beta)'V(x_{J+1} - X_0'\beta),
 \tag{8}$$

where $X_0 = (x_1, \dots, x_J)'$ and V is the inverse of the covariance matrix of X_0 .

Combining equations (1)–(8), we obtain the full posterior as

$$\begin{aligned}
 &p(\beta, \{u_t\}_{t=1}^{T_0}, \Sigma, \lambda, \tau, c | \{z_t\}_{t=1}^{T_0}, \{y_t\}_{t=1}^{T_0}, \{x_j\}_{j=1}^{J+1}, w) \\
 &\propto \prod_{t=1}^{T_0} f(y_t, z_t, u_t | u_{t-1}, \beta, \Sigma) \cdot p(\beta | \{x_j\}_{j=1}^{J+1}, \lambda, \tau, c) p(\lambda, \tau, c) p(u_0) p(\Sigma),
 \end{aligned}
 \tag{9}$$

where f is the density function of the model (1)–(4), $u_t = (\tilde{z}'_t, v'_t)'$, and $\Sigma = (\sigma_y, \sigma_z, \sigma_{\tilde{z}}, \sigma_v)'$. After the observation at $t = T_1 (\geq T_0)$, we can obtain the posterior predictive distribution of the baseline up to the target period $t = T_2 (\geq T_1)$ as

$$\begin{aligned}
 &p(\{y_t^{\text{bsl}}\}_{t=T_0+1}^{T_2} | \{z_t\}_{t=1}^{T_1}, \{y_t\}_{t=1}^{T_1}, \{x_j\}_{j=1}^{J+1}, w) = \int \prod_{t=T_0+1}^{T_2} f(y_t^{\text{bsl}}, z_t, u_t | u_{t-1}, \beta, \Sigma) \\
 &\cdot p(\beta, \{u_t\}_{t=1}^{T_0}, \Sigma, \lambda, \tau, c | \{z_t\}_{t=1}^{T_0}, \{y_t\}_{t=1}^{T_0}, \{x_j\}_{j=1}^{J+1}, w) \cdot d\beta \cdot d\Sigma \cdot d\lambda \cdot d\tau \cdot dc \cdot \prod_{t=1}^{T_2} du_t \cdot \prod_{t=T_1+1}^{T_2} dz_t,
 \end{aligned}
 \tag{10}$$

where $\{y_t^{\text{bsl}}\}_{t=T_0+1}^{T_2}$ is the estimated baseline of the PA from $t = T_0 + 1$ to $t = T_2$. As the project proceeds, the baseline can be updated in a unified manner for ex-ante baseline ($T_1 < t \leq T_2$) and for ex-post baseline ($T_0 < t \leq T_1$). Note that the estimation of β is based on the data up to $t = T_0$, because β represents the relation between the PA and the control units without intervention.

3.2. Estimation algorithm

In this section, we discuss an efficient Markov chain Monte Carlo (MCMC) algorithm for sampling the parameters from (9). Although we fit our model using Stan (Stan Development Team, 2016) for the simplicity of implementation in the empirical example below (Section 4), it would be worth discussing the structures of the full conditional distributions and the relevant literature. By the construction of the model, the full conditional distribution of each parameter reduces to a standard distribution, which enables us to obtain a Gibbs sampler easily, except for the weight β and the parameters regarding the regularized Horseshoe prior. Therefore, we focus on the sampling from the full conditional distribution of these parameters.

The model (1) is a Gaussian sparse linear regression model, and if the prior is the basic (without-regularization) Horseshoe prior the full conditional distribution of β becomes also a Gaussian distribution and the Gibbs sampler can be derived by using data augmentation techniques (Carvalho et al., 2010; Bhadra et al., 2019). However, our formulation has two problems: (1) since we consider the regularized Horseshoe prior, it is hard to find a closed form expression for the full conditional distribution of the shrinkage parameters, $\lambda_j (j = 1, \dots, J)$ and τ and (2) since the weight β is defined on the simplex space, the full conditional is not a simple Gaussian.

For the first problem, we need to rely on a general sampling algorithm that can be applied densities without a closed form, e.g., Hamiltonian Monte Carlo (HMC) and Langevin Monte Carlo methods (Neal, 2011). For the second problem, we can refer to the literature of the sampling from truncated Gaussian distributions. It is known that sampling from a Gaussian distribution defined on the simplex S^{J-1} reduces to sampling from certain truncated Gaussian distribution on \mathbb{R}^{J-1} (Altmann et al., 2014). In our formulation, the covariate balancing term in (7) can be incorporated in this framework because the choice of the quadratic loss function in (8) makes the full conditional distribution have also a Gaussian kernel:

$$p(\beta | \cdot) \propto \exp\left(-\frac{1}{2\sigma_y^2}(\beta - A^{-1}b)'A(\beta - A^{-1}b)\right), \quad \beta \in S^{J-1}, \tag{11}$$

where

$$A = \tilde{Z}_0 \tilde{Z}'_0 + w\sigma_y^2 X_0 V X'_0 + D, \quad D = \text{diag}\left(\tilde{\lambda}_1^{-2}, \dots, \tilde{\lambda}_p^{-2}\right) / \tau^2, \\ b = \tilde{Z}'_0 y + w\sigma_y^2 X'_{J+1} V X'_0, \quad \tilde{Z}_0 = (\tilde{z}_1, \dots, \tilde{z}_{T_0})', \quad y = (y_1, \dots, y_{T_0})'$$

Equation (11) indicates that the full conditional distribution of β is the Gaussian $N(A^{-1}b, \sigma_y^2 A^{-1})$ constrained by the linear constraint $\beta \in S^{J-1}$, which reduces to a truncated normal distribution on \mathbb{R}^{J-1} .

To sample from a truncated Gaussian distribution, we may use the exact Hamiltonian Monte Carlo (EHMC) method (Pakman and Paninski, 2014). A traditional approach to this problem is the combination of the Gibbs sampling from a Gaussian and the rejection of samples that are outside of the domain, but it is known to be inefficient, particularly in a case of high dimension. On the other hand, Pakman and Paninski (2014) propose an efficient HMC algorithm for sampling from a multivariate Gaussian that is defined on a constrained space by linear and quadratic inequalities (or products thereof). In this special case, their algorithm achieves a hundred percent acceptance rate, while HMC methods in general need to reject a proposed sample with some small probability.

In Stan, we implement the following variable transformation so that β is contained in the simplex S^{J-1} :

$$\tilde{\beta}_j = \alpha_j \cdot \tilde{\lambda}_j \cdot \tau, \quad \beta_j = \tilde{\beta}_j / \sum_{j=1}^J \tilde{\beta}_j, \quad j = 1, \dots, J,$$

where α_j is defined as a non-negative scalar. In this formulation, α is sampled based on the full posterior (9) and then transformed into β . We use a truncated normal distribution as the prior for α .

4. Case study

We apply the proposed model to the Valparaíso project (Eaton et al., 2014) to demonstrate the performance of the model. The Valparaíso project, implemented in the state of Acre, Brazil, is a REDD+ project with the main objective of avoiding unplanned deforestation, such as illegal logging or the conversion of forest to agricultural land. Since 2011, the project has implemented several countermeasures against deforestation. These include community outreach and employing local community members as forest guards or in other project staff roles.

To monitor deforestation, we use MapBiomas Brazil (collection 6) (Souza et al., 2020), which is a landcover data estimated from Landsat satellite imagery, and convert it to a 0–1 forest map at 250 m resolution from 1995 to 2020. We followed West et al. (2020) for preprocessing of the data. Figure 2 illustrates the progress of deforestation around the PA between 2000 and 2020. We can see that between 2000 and 2010 the deforestation area approached the PA, especially between two boundaries. Our focus is to evaluate how much deforestation would have occurred within the PA after 2011 in the absence of the project.

Given that deforestation is caused by different drivers (Laurance et al., 2011; Busch and Ferretti-Gallon, 2017), we consider the following covariates into our model to find control units with similar deforestation risks to the PA: elevation from FABDEM (Hawker et al., 2022), (pixel-based Euclidean) distance to road, and annual deforestation rates in the 20 km buffer area around the boundary before the project started. The last covariate is considered to express the risk of deforestation surging to the area. Since the selection of covariates would affect the accuracy of the model, we carry out covariate selection based on the predictive performance within the preproject period. As a result, in addition to distance to road and elevation, we include the annual deforestation rate in the buffer area in 2010 (one year prior to the project implementation). See Appendix A for details on covariate selection.

For the boundary of control units, we use the forest district polygon called CAR (Cadastro Ambiental Rural),¹ which is a georeferenced property organized by Brazil's Rural Environmental Registry. For each control unit, deforestation rates and covariates are aggregated over a CAR polygon, while for the PA, they are aggregated over the project boundary found in the project registry². To reduce the pool size of CARs, we query them where the historical average of the deforestation rates between 1999 and 2010 (i.e., 12 years prior to the intervention) are close to the PA both in the boundary (from 0.0% to 0.2%) and in the 20 km buffer area (from 0.0% to 0.4%), obtaining 57 CARs. Figure 3 describes the spatial distribution and the deforestation rate of the PA and the selected CARs.

We fit the model using Stan (Stan Development Team, 2016). We obtain 3000 MCMC samples where the first 1000 samples are discarded as warm-up. The default setting in Stan is used for prior specifications except for the ones described in Section 3. For the interpretability of results, we set the effective number of nonzero parameters m_0 as 4. This may also depend on the sample size of the data. If the sample size is limited, one may consider a smaller number as m_0 . We need to fine-tune w for covariate balancing. We adjust it so that the balanced covariates become close enough to the result by SCM, and set $w = 100$ to get the following results.

Figure 4 shows the result of the estimated baseline. Our baseline is much smaller than the one by the project developer (see Figure 1b). Given that the our baseline is naturally continuous with the preproject trend, it would be more reasonable than the one by the project developer. Comparing Figure 4a–c (respectively), we can find that the 90% interval of the ex-ante baseline includes the posterior mean of the ex-post baseline at least up to three years forward, implying that our ex-ante estimation worked to some extent. Looking at the ex-post baseline at 2019 (Figure 4c), we can see that the project had no effect during the first 4 years (2011–2014), but then gradually started to have a small positive effect after 2015. This may be because the baseline was lifted by the upward trend over Brazil since 2012 (Ferrante and Fearnside, 2019; West et al., 2019), while the PA was protected from that trend. Figure 4 also includes

¹ <https://www.car.gov.br/publico/imoveis/index>

² <https://registry.verra.org/app/projectDetail/VCS/1113>

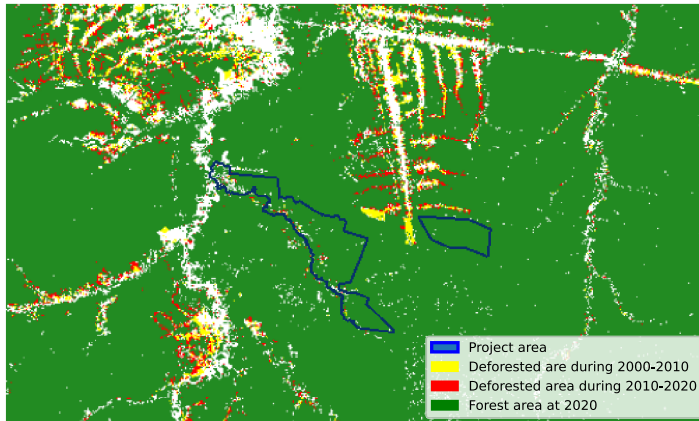


Figure 2. Forest transition map calculated from MapBiomass Brazil (Forest area: green, Nonforest areas: white, deforested area during 2000–2010: yellow, deforested area during 2010–2020: red, PA boundary: blue).

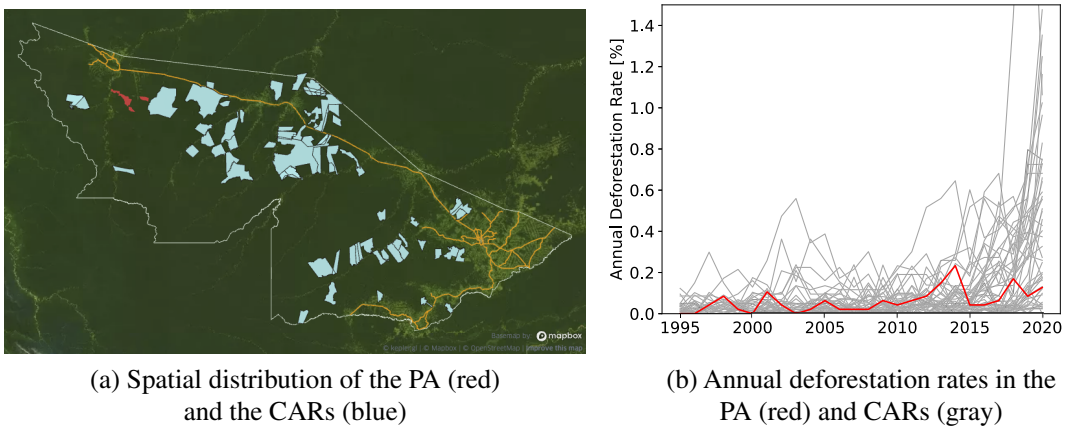


Figure 3. Spatial distribution and deforestation rate of the PA and all the CARs used in analysis.

the posterior mean of the estimated baseline without the covariate balancing (i.e., $w = 0$), which is estimated separately. We can see that the baseline without covariate balancing is generally lower than the baseline with covariate balancing. This may be because the former underestimated the risk of deforestation that is approaching to the PA, which can be represented by deforestation rate in the 20 km buffer area. This would imply the importance of the covariate balancing in the baseline estimation.

Table 1 shows the mean of the covariates for the PA and the synthetic control, where the synthetic control is evaluated by the posterior mean of β . We can see that the synthetic control is closer to the PA than the simple average of CARs in terms of the similarity in covariates. This suggests that the risk of deforestation to the PA is better approximated by the synthetic control rather than by the simple average of CARs.

We also compare the proposed approach with two popular methods, SCM (Abadie et al., 2010) and CausalImpact (Brodersen et al., 2015). We run these models using the R packages Synth (version 1.1-8) (Abadie et al., 2011) and CausalImpact (version 1.3.0) (Brodersen et al., 2015) respectively. We use the same set of covariates for SCM, while using no covariates for CausalImpact because it cannot consider covariate balancing. In addition, since these two methods do not consider the prediction of deforestation rates of control units, we only compare ex-post results at $T_1 = 2019$. Figure 5 shows the estimated baseline results from SCM and CausalImpact, along with the result of the proposed method with $w = 100$.

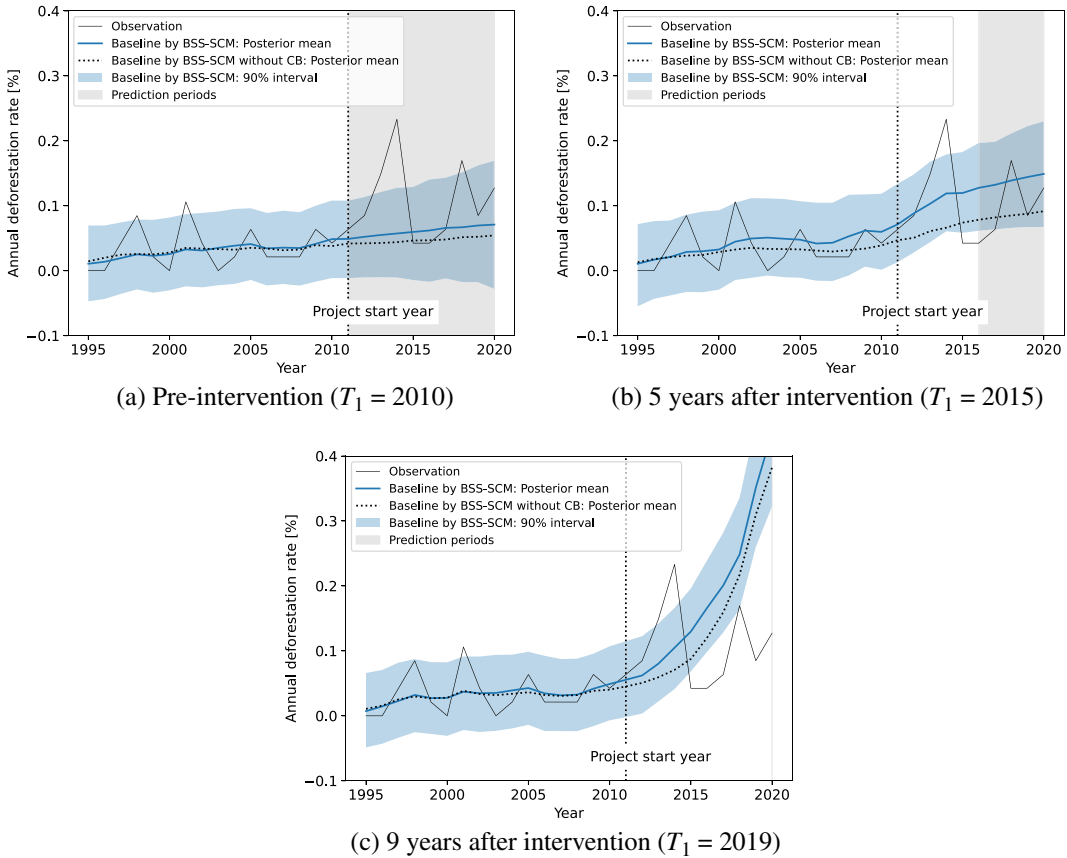


Figure 4. Estimated ex-ante and/or ex-post baseline for the Valparaiso project (x-axis: year; y-axis: annual deforestation rate; dotted vertical line: the time when the intervention started (2011); solid line (black): the observed deforestation rate; solid line (blue): the posterior mean of the estimated baseline; blue area: the 90% credible interval of the estimated baseline; dashed line (black): the posterior mean of the estimated baseline without the covariate balancing (i.e., $w = 0$); throughout (a)–(c), $T_0 = 2010$ and $T_2 = 2020$).

Table 1. Comparison of covariates

	PA	RRD: Synthetic	RRD: Average
Distance to road [km]	41.745	41.584	49.329
Elevation [m]	201.216	202.630	219.040
Deforestation rate in buffer area in 2010 [%]	0.132	0.131	0.0739

In comparison with SCM the proposed method yielded a similar baseline after the ex-post data is observed. It may imply that the proposed method can be regarded as a natural extension of SCM in terms of the capability of ex-ante prediction. On the other hand, the baseline estimated by CausalImpact remained around zero. This may be explained by the fact that CausalImpact cannot include covariates and fit the weight parameter only by minimizing the error in annual deforestation rates. Therefore, it might underestimate the deforestation risk of PA. This would show the importance of introducing the covariate-dependent prior as in Eq. (7).

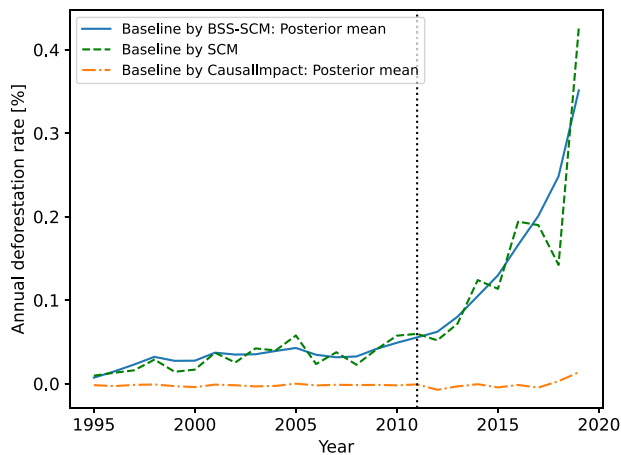


Figure 5. Comparison of baseline estimates between different methods (proposed method: solid blue; SCM: dashed green; CausalImpact: dash-dotted orange).

5. Discussion and limitations

Our results suggested that the baseline set under VM0007 might be overestimated in the Valparaíso project, while the project had small, positive effects on reducing deforestation. The implementation of dynamic baseline updating would lead to more reasonable baseline estimates because it could reflect the effects of policy changes as noted in Ferrante and Fearnside (2019) and West et al. (2019). Our results were qualitatively consistent with Guizar-Coutiño et al. (2022) (i.e., small, positive impacts), but there are some differences in the magnitude of the effects. One reason for this may be that the unit of our analysis is CAR while theirs is pixels. Our analysis followed West et al. (2020) in the sense that the unit of their analysis is CAR and they concluded that the baselines were overestimated. However, their results are more negative about the additionality of the project than ours. One reason for this difference could lie in the difference of the covariates considered in the model. In particular, distance to deforestation edge is considered to be an important covariate (Laurance et al., 2011; Busch and Ferretti-Gallon, 2017), and the fact that we included the deforestation rate in a buffer area as a proxy variable for it may lead to the difference in the estimated weights.

Although we assumed that all covariates, except for deforestation rate in the 20 km buffer area, are constant over time, it is also important to consider them as time-series, especially for socioeconomic covariates. For example, distance to road is known to be an important driver; indeed, the background of the Valparaíso project is to stop deforestation accelerated by the road development. Given the limited update frequency of public data, the monitoring and/or modeling of covariate growth using, e.g., remote sensing would be necessary. As for the modeling, we can consider many different extensions. One possible way would be to consider spatial correlation between control units by introducing a transition matrix in the system equation (3), which would reduce error in estimation.

Finally, we should mention the technical report published in January 2023 by Verra (Verra, 2023). In the report, Verra criticized West et al. (2020, 2023) for several reasons³. Here we discuss two of their arguments in relation to the present paper. First, they claim that the synthetic controls were constructed “by looking at only a small set of superficial physical characteristics such as initial forest cover, slope, and proximity to state capitals, while excluding the key determinants of deforestation such as forest type, agricultural practices, and in fact any socioeconomic factors whatsoever.” This may be partly true. However, if the effects of the key determinants are reflected in proxy variables that are directly linked to

³ They also criticized Guizar-Coutiño et al. (2022) and an article by Britain’s Guardian, but here we focus on West et al.’s studies in terms of proximity to the present study.

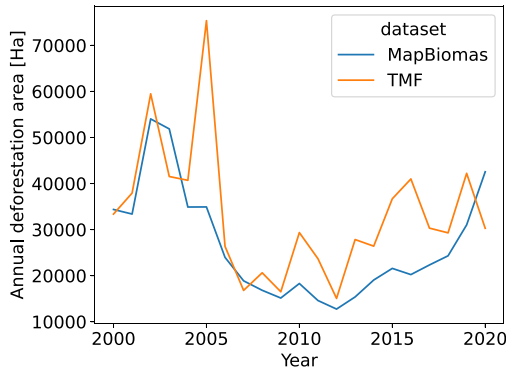


Figure 6. Annual deforestation rates for the RRD of the Valparaiso project calculated by MapBiomass and TMF.

the deforestation risk, such as deforestation rate in the buffer area, then the resulting synthetic control would not underestimate the deforestation risk. Furthermore, regardless of how a baseline is estimated, it should be continuous with the preproject deforestation level to some extent. At least, the fact that the official baseline considers those key determinants would not justify the substantial differences from the observed deforestation level in the PA/RRD as in Figure 1. In this sense, we think SCM-type approaches, including the proposed method in the paper, are effective in evaluating REDD+ projects.

Second, Verra argues that the use of satellite imagery can be helpful for deforestation estimation, but it can be problematic to depend too heavily on it. Deforestation or landcover data estimated from satellite imagery, including MapBiomass, often yield inconsistent results with survey data or visual inspection. Besides, the result is sensitive to the way it is preprocessed and/or at which resolution it is evaluated. Therefore, it needs to be handled carefully before used for REDD+ project assessment. We agree on these points. In fact, we can observe that deforestation trends differ depending on datasets in the Valparaiso project. Figure 6 compares the annual deforestation trends calculated by MapBiomass and TMF (Vancutsem et al., 2021) for the RRD, showing that there is a significant difference between them. Considering this point, we do not argue our results are a definitive judgment on the effectiveness of the project. Nonetheless, we still think that the substantial discrepancies observed between the official baseline, West et al.'s estimates, and our own, cannot solely be attributed to errors inherent in satellite-based products, and further discussions are needed toward more reliable carbon credits.

Author contribution. Conceptualization: K.T., H.S.; Data curation: K.T.; Formal analysis: K.T.; Funding acquisition: H.S.; Investigation: K.T., H.S., K.F., S.S.; Methodology: K.T., S.S.; Software: K.T.; Supervision: K.F., S.S.; Validation: K.T.; Visualization: K.T.; Writing—original draft: K.T.; Writing—review and editing: K.T., H.S., K.F., S.S.

Competing interest. The authors declare none.

Data availability statement. The codes for this work will be available to the public upon publication at the following GitHub repository: <https://github.com/stkhat/bssscm>.

Ethics statement. The research meets all ethical guidelines, including adherence to the legal requirements of the study country.

Funding statement. This paper is based on results obtained from a project, JPNP14012, subsidized by the New Energy and Industrial Technology Development Organization (NEDO).

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/eds.2024.5>.

References

Abadie A (2021) Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature* 59(2), 391–425.

- Abadie A, Diamond A and Hainmueller J** (2010) Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association* 105(490), 493–505.
- Abadie A, Diamond A and Hainmueller J** (2011) Synth: An R package for synthetic control methods in comparative case studies. *Journal of Statistical Software* 42(13), 1–17.
- Aldy JE and Stavins RN** (2012) The promise and problems of pricing carbon. *The Journal of Environment & Development* 21(2), 152–180.
- Altmann Y, McLaughlin S and Dobigeon N** (2014) Sampling from a multivariate Gaussian distribution truncated on a simplex: A review. In *2014 IEEE Workshop on Statistical Signal Processing (SSP)*, pp. 113–116, Gold Coast, QLD, Australia.
- Badgley G, Freeman J, Hamman JJ, Haya B, Trugman AT, Anderegg WRL and Cullenward D** (2022) Systematic over-crediting in California's forest carbon offsets program. *Global Change Biology* 28(4), 1433–1445.
- Ben-Michael E, Feller A and Rothstein J** (2021) The augmented synthetic control method. *Journal of the American Statistical Association* 116(536), 1789–1803.
- Bento A, Kanbur R and Leard B** (2016) On the importance of baseline setting in carbon offsets markets. *Climatic Change* 137(3–4), 625–637.
- Bhadra A, Datta J, Polson NG and Willard B** (2019) Lasso meets horseshoe: A survey. *Statistical Science* 34(3), 405–427.
- Bissiri PG, Holme CC and Walker SG** (2016) A general framework for updating belief distributions. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* 78(5), 1103–1130.
- Brodersen KH, Gallusser F, Koehler J, Remy N and Scott SL** (2015) Inferring causal impact using Bayesian structural time-series models. *The Annals of Applied Statistics* 9(1), 247–274.
- Busch J and Ferretti-Gallon K** (2017) What drives deforestation and what stops it? A meta-analysis. *Review of Environmental Economics and Policy* 11(1), 3–23.
- Carvalho CM, Polson NG and Scott JG** (2010) The horseshoe estimator for sparse signals. *Biometrika* 97(2), 465–480.
- Correa J, Cisneros E, Börner J, Pfaff A, Costa M and Rajão R** (2020) Evaluating REDD+ at subnational level: Amazon fund impacts in Alta Floresta, Brazil. *Forest Policy and Economics* 116, 102178.
- Doudchenko N and Imbens GW** (2016) Balancing, regression, difference-in-differences and synthetic control methods: A synthesis. Technical report. National Bureau of Economic Research.
- Eaton J, Dickson R, McFarland B, Freitas P and Lopes MB** (2014) The Valparaíso Project: A Tropical Forest Conservation Project in Acre, Brazil. Available at <https://registry.verra.org/app/projectDetail/VCS/1113>. Accessed 16 Feb. 2024.
- Ellis EA, Sierra-Huelsz JA, Ceballos GCO, Binnqüist CL and Cerdán CR** (2020) Mixed effectiveness of REDD+ subnational initiatives after 10 years of interventions on the Yucatan peninsula, Mexico. *Forests* 11(9), 1005.
- Ferrante L and Fearnside PM** (2019) Brazil's new president and "ruralists" threaten Amazonia's environment, traditional peoples and the global climate. *Environmental Conservation* 46(4), 261–263.
- George EI and McCulloch RE** (1993) Variable selection via gibbs sampling. *Journal of the American Statistical Association* 88(423), 881–889.
- Griscom BW, Busch J, Cook-Patton SC, Ellis PW, Funk J, Leavitt SM, Lomax G, Turner WR, Chapman M, Engelmann J, Gurwick NP, Landis E, Lawrence D, Malhi Y, Murray LS, Navarrete D, Roe S, Scull S, Smith P, Streck C, Walker WS and Worthington T** (2020) National mitigation potential from natural climate solutions in the tropics. *Philosophical Transactions of the Royal Society B* 375(1794), 20190126.
- Guizar-Coutiño A, Jones JP, Balmford A, Carmenta R and Coomes DA** (2022) A global evaluation of the effectiveness of voluntary REDD+ projects at reducing deforestation and degradation in the moist tropics. *Conservation Biology* 36(6), e13970.
- Hawker L, Uhe P, Paulo L, Sosa J, Savage J, Sampson C and Neal J** (2022) A 30 m global map of elevation with forests and buildings removed. *Environmental Research Letters* 17(2), 024016.
- Haya B** (2009) Measuring emissions against an alternative future: Fundamental flaws in the structure of the Kyoto Protocol's clean development mechanism. *University of California, Berkeley Energy and Resources Group Working Paper, ERG09-001*.
- Haya B, Cullenward D, Strong AL, Grubert E, Heilmayr R, Sivas DA and Wara M** (2020) Managing uncertainty in carbon offsets: Insights from California's standardized approach. *Climate Policy* 20(9), 1–15.
- Laurance WF, Camargo JL, Luizão RC, Laurance SG, Pimm SL, Bruna EM, Stouffer PC, Williamson GB, Benítez-Malvido J, Vasconcelos HL, Houtan KSV, Zartman CE, Boyle SA, Didham RK, Andrade A and Lovejoy TE** (2011) The fate of Amazonian forest fragments: A 32-year investigation. *Biological Conservation* 144(1), 56–67.
- Neal RM** (2011) MCMC using Hamiltonian dynamics. In *Handbook of Markov Chain Monte Carlo*, Vol. 2, p. 2. Chapman and Hall/CRC: New York.
- Pakman A and Paninski L** (2014) Exact Hamiltonian Monte Carlo for truncated multivariate Gaussians. *Journal of Computational and Graphical Statistics* 23(2), 518–542.
- Piironen J and Vehtari A** (2017) Sparsity information and regularization in the horseshoe and other shrinkage priors. *Electronic Journal of Statistics* 11(2), 5018–5051.
- Rana P and Sills EO** (2018) Does certification change the trajectory of tree cover in working forests in the tropics? An application of the synthetic control method of impact evaluation. *Forests* 9(3), 98.
- 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. *Proceedings of the National Academy of Sciences* 116(49), 24492–24499.

- Sills EO, Herrera D, Kirkpatrick AJ, Brandão A, Dickson R, Hall S, Pattanayak S, Shoch D, Vedoveto M, Young L and Pfaff A** (2015) Estimating the impacts of local policy innovation: The synthetic control method applied to tropical deforestation. *PLoS ONE* 10(7), e0132590.
- Simmons BA, Marcos-Martínez R, Law EA, Bryan BA and Wilson KA** (2018) Frequent policy uncertainty can negate the benefits of forest conservation policy. *Environmental Science & Policy* 89, 401–411.
- Souza CM, Shimbo JZ, Rosa MR, Parente LL, Alencar AA, Rudorff BFT, Hasenack H, Matsumoto M, Ferreira LG, Souza-Filho PWM, de Oliveira SW, Rocha WF, Fonseca AV, Marques CB, Diniz CG, Costa D, Monteiro D, Rosa ER, Vélez-Martin E, Weber EJ, Lenti FEB, Paternost FF, Pareyn FGC, Siqueira JV, Viera JL, Neto LCF, Saraiva MM, Sales MH, Salgado MPG, Vasconcelos R, Galano S, Mesquita VV and Azevedo T** (2020) Reconstructing three decades of land use and land cover changes in Brazilian biomes with Landsat archive and Earth Engine. *Remote Sensing* 12(17), 2735. <https://doi.org/10.3390/rs12172735>.
- Stan Development Team** (2016) Stan modeling language users guide and reference manual (ver. 2.30). Available at <https://mc-stan.org>.
- Vancutsem C, Achard F, Pekel J-F, Vieilledent G, Carboni S, Simonetti D, Gallego J, Aragão LEOC and R. Nasi. Long-term (1990–2019)** (2021) Monitoring of forest cover changes in the humid tropics. *Science Advances* 7(10), eabe1603.
- Verra VM0007 REDD+ Methodology Framework (REDD+ MF), v1.6** (2020). Available at <https://verra.org/methodology/vm0007-redd-methodology-framework-redd-mf-v1-6/>.
- Verra** (2022a) Methodology for Improved Forest Management Using Dynamic Matched Baselines from National Forest Inventories, v1.0. Available at <https://verra.org/methodology/methodology-for-improved-forest-management/>.
- Verra** (2022b) Public Consultation: Projected Carbon Units. Available at <https://verra.org/public-consultation-projected-carbon-units/>.
- Verra** (2023) Technical review of West et al. 2020 and 2023, Guizar-Coutiño 2022, and coverage in Britain's Guardian. Technical report. Available at <https://verra.org/technical-review-of-west-et-al-2020-and-2023-guizar-coutino-2022-and-coverage-in-britain-guardian/>.
- West TAP, Börner J and Fearnside PM** (2019) Climatic benefits from the 2006–2017 avoided deforestation in Amazonian Brazil. *Frontiers in Forests and Global Change* 2, 52.
- West TAP, Börner J, Sills EO and Kontoleon A** (2020) Overstated carbon emission reductions from voluntary REDD+ projects in the Brazilian Amazon. *Proceedings of the National Academy of Sciences* 117(39), 24188–24194.
- West TAP, Wunder S, Sills EO, Börner J, Rifai SW, Neidermeier AN, Frey GP and Kontoleon A** (2023) Action needed to make carbon offsets from forest conservation work for climate change mitigation. *Science* 381(6660), 873–877.