Bayesian State-Space SCM for Deforestation Baseline Estimation for Forest Carbon Credit

Keisuke Takahata* **Hiroshi Suetsugu*** **Keiichi Fukaya**† **Shinichiro Shirota**‡

Abstract

In forest carbon credit, the concept of dynamic (or ex-post) baseline has been discussed to overcome the criticism of junk carbon credit, while an ex-ante baseline is still necessary in terms of project finance and risk assessment. We propose a Bayesian state-space SCM, which integrates both ex-ante and ex-post baseline estimation in a time-series causal inference framework. We apply the proposed model to a REDD+ project in Brazil, and show that it might have had a small, positive effect but had been over-credited and that the 90% predictive interval of the ex-ante baseline included the ex-post baseline, implying our ex-ante estimation can work effectively.

1 Introduction

Background 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 [1]. Reducing deforestation and forest degradation are considered to be one of the most effective approaches to reduce carbon emission and REDD+ is a framework to promote such efforts through the issuance of carbon credit. However, carbon credits from REDD+ have been subject to several criticisms. 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 [2].

Criticisms to carbon credit are mainly about the validity of baseline, i.e., a counterfactual scenario in the absence of a project. Considering this issue, the concept of dynamic baseline has recently been discussed [3, 4]. 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. Ex-post approach, e.g., the use of synthetic control method (SCM), has been investigated in this context [2].

However, there still remain a financing issue since result-based payment requires several years for project proponents to wait until they obtain the first credit issuance. From investor's perspective, ex-ante baseline projection is needed to quantify the risk of projects for their investment decision [5].

With those in mind, we can find a need for the integration of both ex-ante baseline prediction before intervention and ex-post dynamic baseline updating at each observation after intervention.

Summary of our contributions We propose a new model for solving the issue mentioned above. First, we introduce a Bayesian state-space model that naturally integrates the forecast of deforestation baseline before intervention and the dynamic updating of baseline after intervention. We achieve this by combining state-space modeling for forecasting and SCM for dynamic updating.

Tackling Climate Change with Machine Learning: workshop at NeurIPS 2022.

*sustainacraft Inc., Tokyo, Japan.
†National Institute for Environmental Studies, Ibaraki, Japan.
‡Center for the Promotion of Social Data Science Education and Research, Hitotsubashi University, Tokyo, Japan. Email: shinichiro.shirota@gmail.com

Second, we consider covariate balancing in state-space modeling by using the method of general Bayesian updating for a valid causal inference. Finally, we apply the proposed model to a REDD+ project in Brazil and show that both ex-ante and ex-post baseline by our model can work effectively. 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 to climate change action.

2 Preliminaries and Related Work

VM0007: A REDD+ methodology for emission reduction evaluation VM0007 [6] is one of the major methodologies that define how to calculate emission reduction in a REDD+ project. A key concept of VM0007 is Reference Region for projecting rate of Deforestation (RRD), which is used as a control unit for estimating a baseline. An RRD is chosen so that the following variables are as close as possible to those of the Project Area (PA): deforestation drivers, landscape factors (e.g. forest types, soil types, elevation, etc.), and socio-economic variables (access to infrastructures, policies, and regulations, etc.). For the past 10–12 years before intervention, 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 pre-defined linear/non-linear model, where the former is often used (see Appendix for an example of RRD and baseline setting). Several studies have reported that baselines set under VM0007 were overestimated because they failed to consider a counterfactual scenario or to eliminate the effect of external factors, e.g., policy changes [2].

Causal inference for time-series data Synthetic Control Method (SCM) [7] 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 a REDD+ project setting. Given that an RRD consists of multiple sub-units (hereinafter "control units"), SCM finds an optimal weight to match both pre-intervention 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 using observations after intervention. CausalImpact [8] 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, it can forecast a baseline with a little modification of the original work [8] since it is based on state-space modeling. In addition, CausalImpact does not include a covariate balancing mechanism between a treatment and synthetic controls.

Related Work In the context of general carbon crediting schemes including REDD+, the issue of junk carbon credit or over-crediting has been discussed for a long time (e.g. [9, 10, 11]) and one of the sources has been identified to be baseline setting [12, 13]. SCM and CausalImpact have been applied to several studies evaluating REDD+ projects [2, 14, 15, 16] and other forest conservation activities [17, 18, 19] to consider a counterfactual scenario and set a reasonable baseline. Many of those studies using causal inference methods have reported that there was an over-crediting to some extent.

3 Bayesian State-Space SCM

We propose a Bayesian state-space SCM, leveraging the time-series modeling in CausalImpact and the covariate balancing in SCM. We model the pre-intervention deforestation rates of PA and RRD by the following state-space model with a local linear trend:

$$\begin{bmatrix} y_{1,t} \\ z_t \end{bmatrix} = \begin{bmatrix} \beta' \\ I \end{bmatrix} \tilde{z}_t + \epsilon_t, \quad \epsilon_t \sim N(0, Q_t), \tag{1}$$

$$\tilde{z}_{t+1} = \tilde{z}_t + v_t + \eta_t, \quad \eta_t \sim N(0, R_t), \tag{2}$$

$$v_{t+1} = v_t + \xi_t, \quad \xi_t \sim N(0, S_t), \tag{3}$$

where $y_{j,t}$ is an observed deforestation rate for the unit j at t ($j = 1$ for PA and $j = 2, \ldots, J+1$ for control units), β is a weight to be applied to control units in RRD, $z_t = (y_{2,t}, \ldots, y_{J+1,t})'$ is a vector of the observed deforestation rates of control units at t, and $\tilde{z}_t = (\tilde{z}_{2,t}, \ldots, \tilde{z}_{J+1,t})'$ is a latent state vector. Equation (1) links the observed data, $y_{1,t}$ and z_t, to the latent state, $\tilde{z}_t$: it assumes that for control units the deforestation rates are observed as the addition of the latent state and noise, while for the treatment unit (PA) the deforestation rate is written as the weighted sum of the latent states of control units and noise [8]. The latter relates this model to SCM except for covariate balancing. Equations (2)–(3) define the temporal evolution of the latent state by a simple local linear trend model, which enables us to forecast z_t, and thus the baseline. Details of the model specifications can be found in Appendix.

For covariate balancing, we apply the method of general Bayesian updating [20]. With this, we can reflect the distance of the covariates between PA and synthetic controls as a covariate-dependent prior on β: $p(\beta \mid \{x_j\}_{j=1}^{J+1}) \propto \exp(-wL(\beta; \{x_j\}_{j=1}^{J+1}))p(\beta)$, where x_j is a $K \times 1$ vector of the covariates, L is a loss function that measures a distance between PA and synthetic controls, and w is a tuning parameter. Here we choose L to be a SCM-like loss function for covariate balancing: $L(\beta; \{x_j\}_{j=1}^{J+1}) = 1/(2J) \cdot (x_1 - X_0\beta)'V(x_1 - X_0\beta)$, where $X_0 = (x_2, \ldots, x_{J+1})$ and V is the inverse of the covariance matrix of X_0. Combining equations (1)–(3), we obtain the full posterior as

$$
\begin{aligned}
&p(\beta, \{u_t\}_{t=1}^{T_0}, \{\Sigma_t\}_{t=1}^{T_0} \mid \{z_t\}_{t=1}^{T_0}, \{y_{1,t}\}_{t=1}^{T_0}, \{x_j\}_{j=1}^{J+1}, w) \\
&\propto \prod_{t=1}^{T_0} f(y_{1,t}, z_t, u_t \mid u_{t-1}, \beta, \Sigma_t) \cdot \exp(-wL(\beta; \{x_j\}_{j=1}^{J+1}))p(\beta)p(u_0)p(\{\Sigma_t\}_{t=1}^{T_0}),
\end{aligned} \tag{4}
$$

where T_0 is the number of periods before intervention, f is the density function of the model (1)–(3), $u_t = \{\tilde{z}_t, v_t\}$, and $\Sigma_t = \{Q_t, R_{t-1}, S_{t-1}\}$. 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

$$
\begin{aligned}
p(\{y_{1,t}^{\mathrm{bsl}}\}_{t=T_0+1}^{T_2} \mid \{z_t\}_{t=1}^{T_1}, \{y_{1,t}\}_{t=1}^{T_1}, \{x_j\}_{j=1}^{J+1}, w) = \int \prod_{t=T_0+1}^{T_2} f(y_{1,t}^{\mathrm{bsl}}, z_t, u_t \mid u_{t-1}, \beta, \Sigma_t) \\
\cdot\, p(\beta, \{u_t\}_{t=1}^{T_0}, \{\Sigma_t\}_{t=1}^{T_0} \mid \{z_t\}_{t=1}^{T_0}, \{y_{1,t}\}_{t=1}^{T_0}, \{x_j\}_{j=1}^{J+1}, w) \cdot d\beta \cdot \prod_{t=1}^{T_2} \Sigma_t \cdot \prod_{t=1}^{T_2} du_t \cdot \prod_{t=T_1+1}^{T_2} dz_t,
\end{aligned} \tag{5}
$$

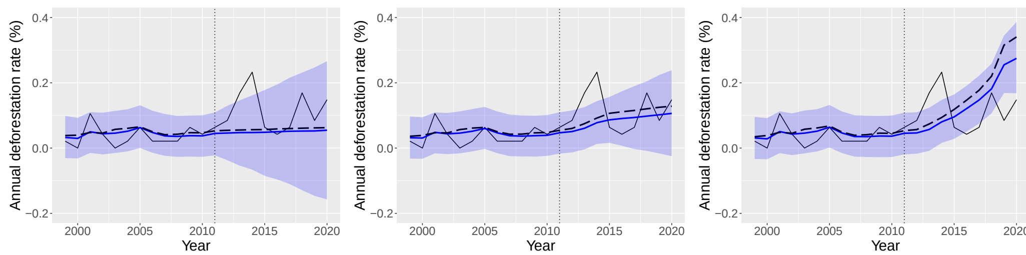
where $\{y_{1,t}^{\mathrm{bsl}}\}_{t=T_0+1}^{T_2}$ is the estimated baseline of PA from $t = T_0 + 1$ to $t = T_2$. As a project proceeds, the baseline can be updated in a unified manner for ex-ante baseline ($T_1 < t \leq T_2$) and for ex-post baseline ($T_0 < t \leq T_1$). Note that the estimation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

4 Case studies

We apply the proposed model to the Valparaiso project [21] to demonstrate the performance of the model. The Valparaiso project is a REDD+ project run in Acre state, Brazil, whose main purpose is to avoid unplanned deforestation (e.g. illegal logging or conversion of forest to agricultural land). The intervention consists of multiple activities, including community outreach and the employment of local community members as forest guards or other project staff. We obtain forest cover data from MapBiomas Brazil [22] and follow [2] for preprocessing. Considering that deforestation is caused by different drivers [23, 24], we include the following covariates into our model: elevation from FABDEM [25], (pixel-based Euclidean) distance to road, distance to urban centers, and distance to recently deforested pixels [26]. For each control unit, deforestation rates and covariates are aggregated over a polygon called CAR, which is a georeferenced property organized by Brazil's Rural Environmental Registry, while aggregated over the project boundary found in the registry [21] for the PA. For model estimation we use Stan [27] and obtain 6000 MCMC samples where the first 1000 samples are discarded as warm-up.

Figure 1 shows the result of the estimation. Comparing Figures 1a (1b) and 1b (1c) (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 1c), we can see that the project had no effect during the first 4 years (2011–2015), 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 [28, 29], while the PA was protected from that trend. Figure 1 also includes the posterior mean of the estimated baseline without the covariate balancing (i.e. $w = 0$), which is estimated separately.

Although the difference in the posterior mean is negligible when the observed deforestation rates were close to zero (Figure 1a), the baseline without the covariate balancing became higher as the rates went up (Figure 1c). This would imply the importance of the covariate balancing in the baseline estimation.



(a) Pre-intervention ($T_1 = 2010$) **(b)** 5 years after intervention ($T_1 = 2015$) **(c)** 9 years after intervention ($T_1 = 2019$)

Figure 1: 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 $T_2 = 2020$.)

5 Discussion and limitations

Our results suggested that the baseline set under VM0007 might be overestimated (see Appendix), while they had small, positive effects on mitigating deforestation. In particular, the implementation of dynamic baseline would lead to more reasonable baseline estimates because it could reflect the effects of policy changes as noted in [28, 29]. Our results agree with [26] qualitatively (i.e. small, positive impacts), but there are some differences in the magnitude of effects. One reason for this may be that the unit of our analysis is CAR while theirs is pixels. Our analysis follows [2] in the sense that the unit of their analysis is CAR and that they concluded the baselines were overestimated. However, their results are more negative about the additionality of the projects than ours. One reason for this difference could lie in the difference of covariates considered in the model. In particular, distance to deforestation edge is considered to be an important driver [23, 24], and the fact that we take this into the model may lead to the difference in the estimated weights.

Although we assumed that all covariates, except for distance to deforestation edge, are constant over time, it is also important to consider them as time-series, especially for socio-economic covariates. For example, distance to road is known to be an important driver; indeed, the background of the Valparaiso 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 plots by introducing a transition matrix in the system equation (2), which would reduce error in estimation.

Acknowledgments and Disclosure of Funding

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

References

[1] Bronson W. Griscom, Jonah Busch, Susan C. Cook-Patton, Peter W. Ellis, Jason Funk, Sara M. Leavitt, Guy Lomax, Will R. Turner, Melissa Chapman, Jens Engelmann, Noel P. Gurwick, Emily Landis, Deborah Lawrence, Yadvinder Malhi, Lisa Schindler Murray, Diego Navarrete, Stephanie Roe, Sabrina Scull, Pete Smith, Charlotte Streck, Wayne S. Walker, and Thomas Worthington. National mitigation

potential from natural climate solutions in the tropics. Philosophical Transactions of the Royal Society B, 375(1794):20190126, 2020.

[2] Thales A. P. West, Jan Börner, Erin O. Sills, and Andreas Kontoleon. Overstated carbon emission reductions from voluntary REDD+ projects in the Brazilian Amazon. Proceedings of the National Academy of Sciences, 117(39):24188–24194, 2020.

[3] Verra. Methodology for Afforestation Reforestation and Revegetation Projects. https://verra.org/methodology/methodology-for-afforestation-reforestation-and-revegetation-projects/.

[4] Verra. Methodology for Improved Forest Management. https://verra.org/methodology/methodology-for-improved-forest-management/.

[5] Verra. Public Consultation: Projected Carbon Units. https://verra.org/public-consultation-projected-carbon-units/, 3 May 2022.

[6] Verra. VM0007 REDD+ Methodology Framework (REDD+ MF). https://verra.org/methodology/vm0007-redd-methodology-framework-redd-mf-v1-6/.

[7] Alberto Abadie, Alexis Diamond, and Jens Hainmueller. 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, 2010.

[8] Kay H. Brodersen, Fabian Gallusser, Jim Koehler, Nicolas Remy, and Steven L. Scott. Inferring causal impact using Bayesian structural time-series models. The Annals of Applied Statistics, 9(1):247–274, 2015.

[9] Barbara Haya. Measuring Emissions Against an Alternative Future: Fundamental Flaws in the Structure of the Kyoto Protocol's Clean Development Mechanism. SSRN Electronic Journal, 2009.

[10] Joseph E. Aldy and Robert N. Stavins. The Promise and Problems of Pricing Carbon. The Journal of Environment & Development, 21(2):152–180, 2012.

[11] Grayson Badgley, Jeremy Freeman, Joseph J. Hamman, Barbara Haya, Anna T. Trugman, William R. L. Anderegg, and Danny Cullenward. Systematic over ‐ crediting in California's forest carbon offsets program. Global Change Biology, 28(4):1433–1445, 2022.

[12] Antonio Bento, Ravi Kanbur, and Benjamin Leard. On the importance of baseline setting in carbon offsets markets. Climatic Change, 137(3-4):625–637, 2016.

[13] Barbara Haya, Danny Cullenward, Aaron L. Strong, Emily Grubert, Robert Heilmayr, Deborah A. Sivas, and Michael Wara. Managing uncertainty in carbon offsets: insights from California's standardized approach. Climate Policy, 20(9):1–15, 2020.

[14] Anand Roopsind, Brent Sohngen, and Jodi Brandt. 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, 2019.

[15] Juliano Correa, Elías Cisneros, Jan Börner, Alexander Pfaff, Marcelo Costa, and Raoni Rajão. Evaluating REDD+ at subnational level: Amazon fund impacts in Alta Floresta, Brazil. Forest Policy and Economics, 116:102178, 2020.

[16] Edward A. Ellis, José Antonio Sierra-Huelsz, Gustavo Celestino Ortíz Ceballos, Citlalli López Binnqüist, and Carlos R. Cerdán. Mixed Effectiveness of REDD+ Subnational Initiatives after 10 Years of Interventions on the Yucatan Peninsula, Mexico. Forests, 11(9):1005, 2020.

[17] Erin O. Sills, Diego Herrera, A. Justin Kirkpatrick, Amintas Brandão, Rebecca Dickson, Simon Hall, Subhrendu Pattanayak, David Shoch, Mariana Vedoveto, Luisa Young, and Alexander Pfaff. Estimating the Impacts of Local Policy Innovation: The Synthetic Control Method Applied to Tropical Deforestation. PLoS ONE, 10(7):e0132590, 2015.

[18] Pushpendra Rana and Erin O. Sills. 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, 2018.

[19] B. Alexander Simmons, Raymundo Marcos-Martinez, Elizabeth A. Law, Brett A. Bryan, and Kerrie A. Wilson. Frequent policy uncertainty can negate the benefits of forest conservation policy. Environmental Science & Policy, 89:401–411, 2018.

[20] P. G. Bissiri, C. C. Holme, and S. G. Walker. A general framework for updating belief distributions. Journal of the Royal Statistical Society. Series B (Statistical Methodology), 78(5):1103–1130, 2016.

[21] The Valparaiso Project. https://registry.verra.org/app/projectDetail/VCS/1113.

[22] MapBiomas Brazil. https://mapbiomas.org/en.

[23] William F. Laurance, José L.C. Camargo, Regina C.C. Luizão, Susan G. Laurance, Stuart L. Pimm, Emilio M. Bruna, Philip C. Stouffer, G. Bruce Williamson, Julieta Benítez-Malvido, Heraldo L. Vasconcelos, Kyle S. Van Houtan, Charles E. Zartman, Sarah A. Boyle, Raphael K. Didham, Ana Andrade, and Thomas E. Lovejoy. The fate of Amazonian forest fragments: A 32-year investigation. Biological Conservation, 144(1):56–67, 2011.

[24] Jonah Busch and Kalifi Ferretti-Gallon. What Drives Deforestation and What Stops It? A Meta-Analysis. Review of Environmental Economics and Policy, 11(1):3–23, 2017.

[25] Laurence Hawker, Peter Uhe, Luntadila Paulo, Jeison Sosa, James Savage, Christopher Sampson, and Jeffrey Neal. A 30 m global map of elevation with forests and buildings removed. Environmental Research Letters, 17(2):024016, 2022.

[26] Alejandro Guizar‐Coutiño, Julia P.G. Jones, Andrew Balmford, Rachel Carmenta, and David A. Coomes. A global evaluation of the effectiveness of voluntary REDD+ projects at reducing deforestation and degradation in the moist tropics. Conservation Biology, 2022.

[27] Stan Development Team. Stan modeling language users guide and reference manual (ver. 2.30). https://mc-stan.org, 2016.

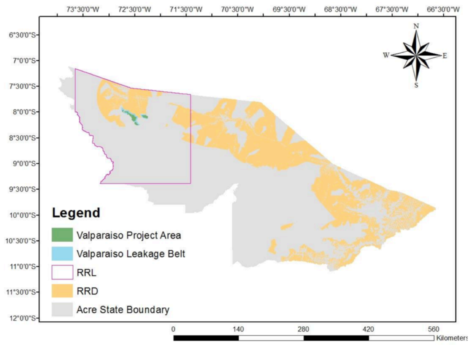
[28] Lucas Ferrante and Philip M Fearnside. Brazil's new president and "ruralists" threaten Amazonia's environment, traditional peoples and the global climate. Environmental Conservation, 46(4):261–263, 2019.

[29] Thales A. P. West, Jan Börner, and Philip M. Fearnside. Climatic Benefits From the 2006–2017 Avoided Deforestation in Amazonian Brazil. Frontiers in Forests and Global Change, 2:52, 2019.

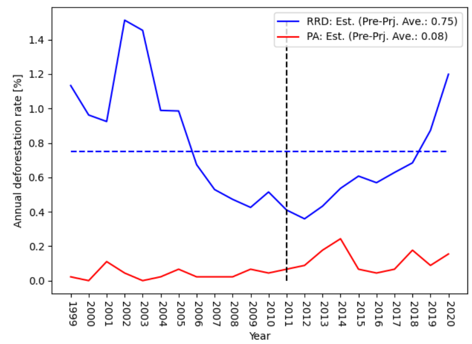
Appendix

Example of VM0007: The Valparaiso project

Figure 2a and 2b describe the RRD and the deforestation trends in the PA and the RRD of the Valparaiso project [21], with the baseline set by the simple historical average approach following VM0007 [6]. The project started at 2011. We can see that the baseline (the blue dashed line in Figure 2b) might have failed to capture the change of trends and have resulted in overestimation. Note that the baseline shown here is simplifed compared with the one by the project in the sense that the latter applied spatial [6] after the projection.



(a) RRD and PA (source: [21])



(b) Deforestation trends in PA and RRD and the baseline by the simple historical average approach (horizontal dashed line).

Figure 2: The Valparaiso project

Details of the model formulation and estimation

Let $g(x \mid \mu, \Phi)$ be the probability density function (pdf) of the multivariate normal distribution with the mean μ and the covariance matrix Φ. Then the pdf of the model (1)–(3) can be written as

$$\begin{aligned} & f(y_{1,t}, z_t, u_t \mid u_{t-1}, \beta, \Sigma_t) \\ =& g(\begin{bmatrix} y_{1,t} \\ z_t \end{bmatrix} \mid \begin{bmatrix} \beta_t' \\ I \end{bmatrix} \tilde{z}_t, Q_t) \cdot g(\tilde{z}_t \mid \tilde{z}_{t-1} + v_{t-1}, R_{t-1}) \cdot g(v_t \mid v_{t-1}, S_{t-1}). \end{aligned}$$

In the case study in Section 4, we used the time-invariant, diagonal covariance matrices as follows:

$$Q_t = \mathrm{diag}(\sigma_y^2, \underbrace{\sigma_z^2, \ldots, \sigma_z^2)}_{\text{the same } J \text{ elements}}, \quad R_t = \sigma_{\tilde{z}}^2 I, \quad S_t = \sigma_v^2 I \quad (t = 1, \ldots, T_2),$$

where I is a $J \times J$ identity matrix.

For the prior distributions of $p(\beta)$, $p(u_0)$, and $p(\{\Sigma_t\}_{t=1}^{T_0})$, we used non-informative priors and followed the default settings in Stan [27]. For the tuning parameter of covariate balancing, we used $w = 300$.

Result of the covariate balancing

Table 1 shows the mean of the covariates for the PA and synthetic controls, where synthetic controls are evaluated with the posterior mean of β. We can see that the synthetic control with covariate balancing ("CB") has the closest covariates to the PA.

	Dist to Road [km]	Dist to Urban [km]	Elevation [m]	Dist to DF Edge [km]
PA	41.75	33.17	201.22	7.01
CB	40.66	45.76	207.85	9.59
Non-CB	31.87	53.15	209.13	8.23
Ave	31.31	53.84	207.58	7.53

Table 1: Mean of the covariates for PA and synthetic controls. PA: Project Area, CB: Synthetic control with covariate balancing ($w = 300$), Non-CB: Synthetic control without covariate balancing ($w = 0$) Ave: Simple average over control units.