

Reduction of forest cover loss and associated CO₂ emissions in protected areas with German bilateral development assistance*

Yota Eilers,[†] Jochen Kluve,[‡] Johannes Schielein,[§] Melvin H.L. Wong[¶]

August 31, 2022

Abstract

We study how bilateal aid affects forest cover loss and associated CO₂ emissions in over 400 Latin American forest protected areas. We combine data from German bilateral aid data from 2004-2020, a spatial matching, and a Difference-in-Difference estimator to determine how supported protected areas effectively reduced forest cover loss and associated CO₂ emissions. This observation often contradicts monitoring data which often suggests ongoing forest and biodiversity losses without using a robust counter-factual scenario. **Keywords:** Development Aid, Forest Conservation, Latin America

JEL Codes: F35, O13, P45, Q23, Q51

Area-based conservation measures are crucial mechanisms to combat harmful global climate change and reduce the accelerating biodiversity crisis (Shukla et al., 2019; Díaz et al., 2019). The international High Ambition Coalition for Nature and People with almost 100 participating countries therefore called out the "30 by 30" target at the COP15 meeting of the Convention on Biological Diversity in 2021, which proposes to increase terrestrial protected area coverage from currently 17 % to 30 % by 2030.

More then USD 5 billion have been pledged by the coalition so far to increase protected area (PA) coverage and PA management effectiveness, especially for low- and middle-income countries where conservation relies heavily on foreign financial assistance. International public expenditure on biodiversity is estimated at USD 3.9-9.3 billion per year according to OECD/DAC, and Germany is amongst the largest bilateral donors with an ongoing portfolio of USD 2.6 Bio in 2021, and 602 supported PAs in 66 countries.

However, despite its economic and political importance, there has been no systematic portfolio-wide assessment of the effectiveness of bilateral assistance in reducing GHG emissions and biodiversity loss in PAs so far (as to our knowledge). Part of this might be due to the lack of comprehensive data on international conservation finance, especially on a disaggregated scale. In light of the problem's urgency, we think it is of utmost importance to study those effects in more detail to allocate financial resources effectively. To do so, KfW is creating a geospatial database of supported PAs which will enable us to systematically improve planning, monitoring and evaluation of our bilateral PA portfolio.

As a first analysis step, we study how forest cover loss and associated CO₂ emissions developed in almost 400 Latin American PAs, which German bilateral ODA financially assisted from 2004-2020. We use a spatial matching approach and a difference in difference estimator to predict that supported PAs effectively reduced forest cover loss and associated CO₂ emissions by at least 6.8 million tons in total. Nevertheless, effectiveness varies by year and country, giving meaningful insights to study heterogeneous treatment effects in more detail.

Our database currently contains 398 PAs in Latin America, which were supported between 2000 and 2021 by KfW Development bank. However, a global extension of the database is ongoing. We extracted names and countries of supported PAs from internal project documentation and matched them to The World Database on Protected Areas (WDPA) from IUCN (UNEP-WCMC, 2019). The WDPA contains georeferenced data for most of the world's PAs as polygon data, which can be intersected with other geospatial datasets such as the Global Forest Watch data. Our supported areas are divided into 337 terrestrial, 19 marine, and 42 partial marine/terrestrial protected areas. They cover a total surface of 0.993 Mio. km², which is about 2.8 times the size of Germany.

*The findings, interpretations, and conclusions expressed in this article are entirely those of the authors. They do not necessarily represent the views of KfW Development Bank and its affiliated organizations.

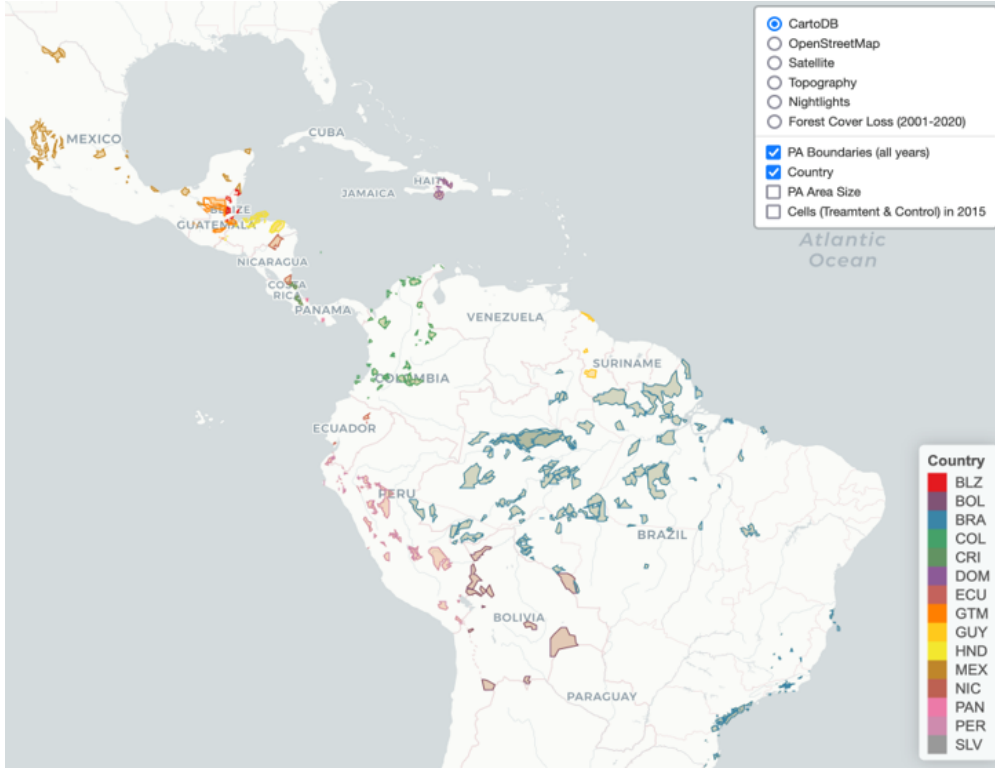
[†]yota.eilers@kfw.de

[‡]jochen.kluve@kfw.de

[§]johannes.schielein@kfw.de

[¶]melvin.wong@kfw.de

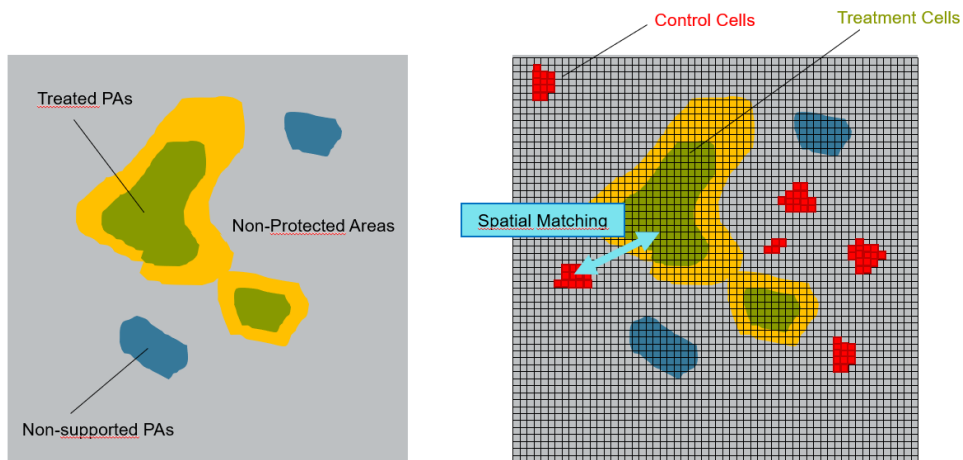
Figure 1: Overview supported Protected Areas



The database furthermore contains project metadata such as project number and yearly disbursements from 2000 to 2021 (see Appendix). To analyze our portfolio, we created a spatial grid covering all countries supported by KfW. The grid has a spatial resolution of 5 sqkm. We use grid cells as observation units and differentiate between "treated" and "non-treated" cells. We subsequently apply a matching procedure to find comparable "treatment" and "control" cells described in more detail below.

However, there are some cells that are omitted from the matching procedure. Omitted cells either cover other (non-supported) PAs or are located in the neighborhood around a supported PA (50 km buffer). We exclude the latter to avoid comparing treated cells to outside cells which might be affected by conservation leakage (i.e., forest area might be destroyed just outside the borders of a PA which would have been destroyed inside if the PA was not to exist there).

Figure 2: Grid Creation and units of observation for the empirical analysis



The significant challenge for identifying the conservation effects of protected areas is to account for selection bias. The protection of forest areas is not assigned randomly, but policymakers make conscious decisions to declare protected areas. For example, a major selection bias is the remoteness of an area. Forest areas distant from settlements have a lower probability of being deforested since the population to harvest

timber is absent. Likewise, machines, such as harvesters, have a greater challenge in reaching rugged forest areas. In contrast, areas with soil highly suitable for agriculture may be subject to increased land-cover pressures.

To overcome this selection bias, we use Coarsened Exact Matching (CEM) to create statistical twins across our units of observations (Iacus et al., 2012; Blackwell et al., 2009). CEM is a matching method that relies on pruning data so that each data point has a relevant counterpart. As an example, matching treatment and control units by travel time to the nearest settlement with CEM involves: First, defining range bins of the variable (e.g., 10, 30, 60, 90 minutes). Second, keeping only those observations in bins with treatment and control units. That is, if units in the bin of 90 minutes travel time and above consist only of control units or only of treatment units, the observations would be discarded from the analysis. Third, if there happens to be an unequal number of treatment and control units in a bin, weights are calculated. Finally, the pruned data may be used for further analysis, such as weighted regressions.

After the matching approach we construct a panel dataset to estimate conservation effectiveness to reduce forest cover loss in supported PAs using a difference-in-difference estimator:

$$loss_{c,t} = \beta_1 T_{c,t} + \delta_c + \tau_t + \epsilon_{c,t}, \quad (1)$$

where $loss_{c,t}$ is the change in forest cover in cell c , and year t . $T_{c,t}$ a dummy variable being 1 since the year of KfW disbursements, and 0 otherwise. All regressions include cell fixed effects and year fixed effects to account for time-invariant heterogeneity and time-variant covariates that are common across all units of observations, respectively. Moreover, the regressions are clustered on the level of protected areas, where the control group in each country forms a single cluster. For an overview of Variables used for matching please refer to Appendix 2.

Our estimates are based on multiple panel estimates, which we refer to as "matching frames." A matching frame is a common longitudinal data with the only difference that it includes observations only from the treatment and respective control group. Matching frames are constructed by selecting the start year of the intervention and performing the CEM for the treatment cells of that start year against all potential control cells. After that, all unmatched units of observations are discarded, and time-variant covariates are merged with the matched units of observations. This process is repeated for the following year. Matching frames are necessary, as the treatment cells vary each year and are, hence, matched to different control cells. In the following, the matching performance is illustrated for 2015. The regression results are shown for each matching frame with valid parallel trends for the cell's forest cover area in hectares.

Figure 3: Overview of matching variable balancing

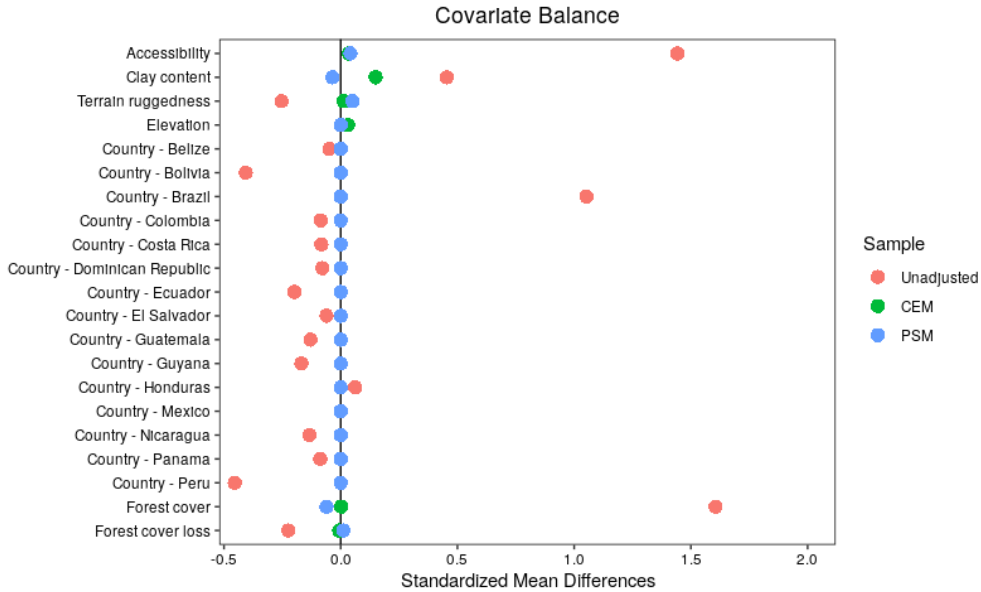
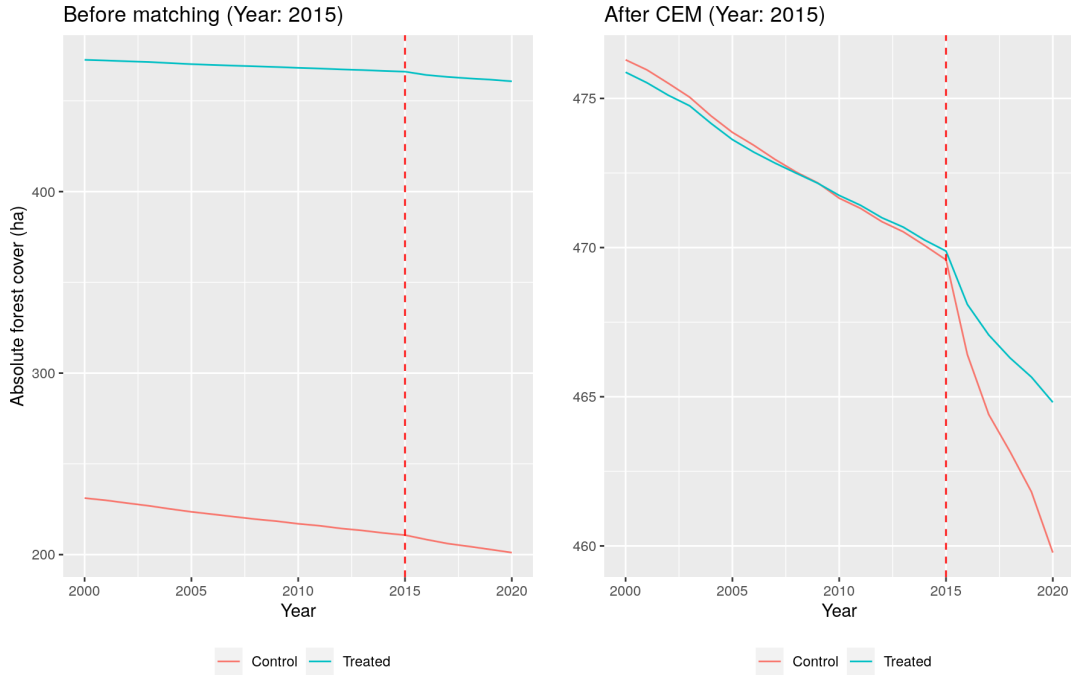


Figure 3 shows that matching establishes a balanced dataset as the mean differences between the treatment and control groups are significantly reduced and partially eliminated. The treatment and control groups become exceptionally comparable after matching with regard to travel distance to the nearest settlement and forest cover. Moreover, Figure 4 demonstrates that matching treatment and control groups reduces the differences between them while establishing a parallel trend up to the time of the intervention.

Figure 4: Parallel trends for the 2015 matching frame



Bilateral disbursements significantly reduce forest cover loss in protected areas in the treatment group compared to the control group. Table 1 shows the coefficient estimates for each year where protected areas started to receive financing, that is, for each matching frame. The total number of cells refers to the number of individual 500 hectares cells. Hence, a total number of treated cells of 300 corresponds to a treatment area of 1,500 hectares. The expectation is that finances of KfW and its partners reduce the anthropogenic pressure on forest cover. Hence coefficient should be negative. This is the case for estimates with matching frames that exhibit a parallel trend, such as 2015 (see figure 4). In some instances, the matching approach was insufficient to establish a parallel trend between the treatment and control groups. Hence estimates may become positive. This has been the case for all other estimates in the table. A back-of-the-envelope calculation shows that financing which started in 2004 avoided about 588 hectares, which corresponds to roughly 280,000 tons of CO₂. Thus, the valid estimates show that financial support for protected areas reduces forest cover loss in these areas.

Our analysis shows that conservation finance in PAs effectively conserves forests and should be further supported. This observation is often contradicted by monitoring data suggesting ongoing forest and biodiversity losses without using a robust counter-factual scenario. However, we can observe different effects in different years, which indicates that we should further scrutinize our results. There are different improvements that we seek to implement in the upcoming weeks. Those include more control variables to separate anthropogenic from natural forest cover loss (data about the occurrence of hurricanes, droughts, and wild-fires). In addition, we will look at the effects on the level of individual projects and countries. Furthermore, our database is steadily growing, allowing us to analyze KfWs' global conservation portfolio. On extension to this database will also be the categorization of treatment types according to an internationally recognized standard (Conservation Action Types). This will allow us to analyze heterogeneous treatment effects in more detail, ultimately allowing us to improve the design of new conservation projects.

Table 1: Dependent variable: Forest cover loss

	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Financial support	-1.959*** (0.205)	-0.284 (0.221)	-0.198*** (0.037)	-0.462** (0.142)	-0.122 (0.422)	0.703+ (0.303)	0.998*** (0.029)	-0.217 (0.599)	0.009 (0.068)	-0.440 (1.455)	0.493*** (0.000)	-0.767*** (0.124)	-1.685*** (0.138)	-0.006 (0.076)
Num.Obs.	4656920	5475260	3013860	12937100	5202440	286000	1722400	1379020	6576260	1292640	1041160	19885180	12063180	4905880
R2 Adj.	0.081	0.082	0.193	0.054	0.142	0.147	0.034	0.123	0.162	0.084	0.056	0.058	0.063	0.176
FE: .assetid	X	X	X	X	X	X	X	X	X	X	X	X	X	X
FE: year	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Total num. cells	232846	273763	150693	646855	260122	14300	86120	68951	328813	64632	52058	994259	603159	245294
–treated	300	933	11398	21862	9034	518	1738	3007	11398	478	128	122433	2571	7875
–control	232546	272830	139295	624993	251088	13782	84382	65944	317415	64154	51930	871826	600588	237419

Standard errors clustered by zones of forest protected areas (WDPA).

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix

Figure 5: KfW Disbursements for Protected Areas in Latin America

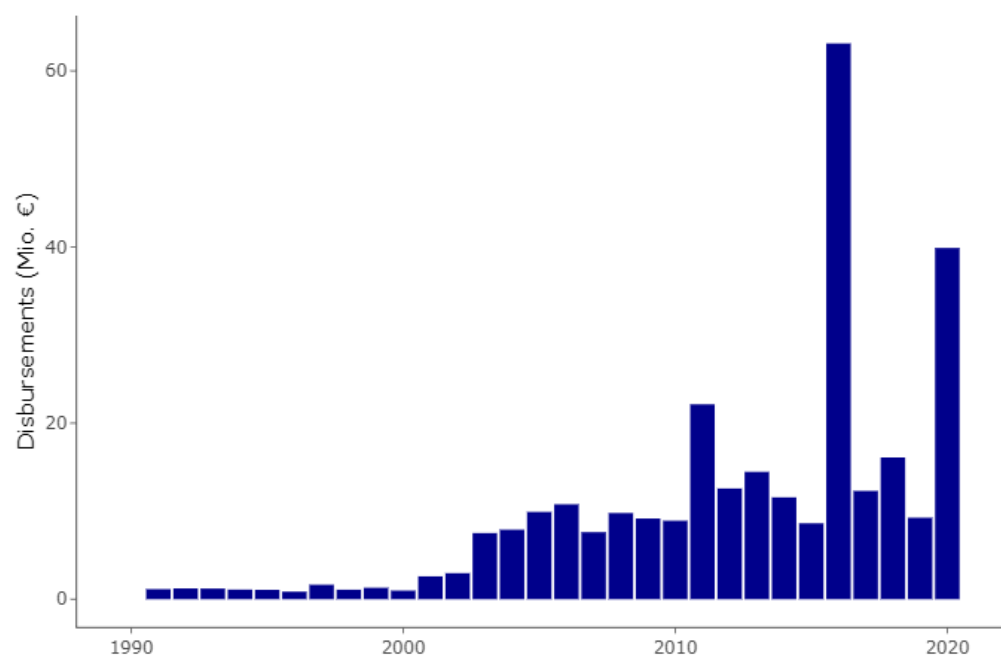


Table 2: Overview of matching variables

Covariate	Data source	Data description	Rationale
Travel time to next city (population >5,000)	Weiss et al. (2018)	Accessibility is the ease with which larger cities can be reached from a certain location. This resource represents the travel time to major cities in the year 2015. Encoded as minutes, representing the time needed to reach that particular cell from nearby city of target population range. We use the travel time to cities with a population of 5,000 to 110 million.	Proxy for infrastructure accessibility (main determinant of anthropogenic forest cover loss/deforestation)
Forest cover	Hansen et al. (2013)	“Tree cover in the year 2000, defined as canopy closure for all vegetation taller than 5m in height. Encoded as a percentage per output grid cell, in the range 0–100.”	Compare cells with similar level of forest cover and exclude non-forest areas (e.g. cities)
Forest cover loss (aggregated over three years before financing)	Hansen et al. (2013)	“Forest loss during the period 2000–2020, defined as a stand-replacement disturbance, or a change from a forest to non-forest state. Encoded as either 0 (no loss) or else a value in the range 1–20, representing loss detected primarily in the year 2001–2020, respectively.”	Captures pre-treatment forest cover loss dynamics
Terrain ruggedness	Farr et al. (2007); Riley et al. (1999)	This index is calculated using the same data as the elevation variable. The elevation difference between the center pixel and its eight immediate pixels are squared and then averaged and its square root is taken to get the TRI value. This function allows to efficiently calculate terrain ruggedness index (TRI) statistics for polygons. For each polygon, the desired statistics (mean, median or sd) is/are returned.	Agricultural mechanization suitability
Elevation	Farr et al. (2007)	The layer represents the 30m global terrestrial digital elevation model from the NASA Shuttle Radar Topographic Mission (SRTM), available for download as 5 degree x 5 degree tiles. It is encoded as meter, representing the elevation at the particular grid cell.	Proxy for agricultural climate suitability
Soil clay content	Hengl et al. (2017)	Proportion of clay particles < 0.002 mm in the fine earth fraction (g/100g)	Agricultural soil suitability
Country		Corresponding country of the gridcell.	Political and regulatory framework

References

- Blackwell, M., Iacus, S., King, G., and Porro, G. (2009). cem: Coarsened exact matching in stata. *The Stata Journal*, 9(4):524–546.
- Díaz, S. M., Settele, J., Brondízio, E., Ngo, H., Guèze, M., Agard, J., Arneth, A., Balvanera, P., Brauman, K., Butchart, S., et al. (2019). The global assessment report on biodiversity and ecosystem services: Summary for policy makers.
- Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., et al. (2007). The shuttle radar topography mission. *Reviews of geophysics*, 45(2).
- 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(6160):850–853.
- Hengl, T., Mendes de Jesus, J., Heuvelink, G. B., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangquan, W., Wright, M. N., Geng, X., Bauer-Marschallinger, B., et al. (2017). Soilgrids250m: Global gridded soil information based on machine learning. *PLoS one*, 12(2):e0169748.
- Iacus, S. M., King, G., and Porro, G. (2012). Causal inference without balance checking: Coarsened exact matching. *Political analysis*, 20(1):1–24.
- Riley, S. J., DeGloria, S. D., and Elliot, R. (1999). Index that quantifies topographic heterogeneity. *intermountain Journal of sciences*, 5(1-4):23–27.
- Shukla, P. R., Skeg, J., Buendia, E. C., Masson-Delmotte, V., Pörtner, H.-O., Roberts, D., Zhai, P., Slade, R., Connors, S., Van Diemen, S., et al. (2019). Climate change and land: an ipcc special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems.
- UNEP-WCMC (2019). User manual for the world database on protected areas and world database on other effective area-based conservation measures: 1.6.
- Weiss, D. J., Nelson, A., Gibson, H., Temperley, W., Peedell, S., Lieber, A., Hancher, M., Poyart, E., Belchior, S., Fullman, N., et al. (2018). A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature*, 553(7688):333–336.