*Environmental Data Science* (2026), 5: e28, 1–22
doi:10.1017/eds.2026.10060

CAMBRIDGE UNIVERSITY PRESS

APPLICATION PAPER

# A machine learning approach to detecting environmental crimes in Brazil

Matthew C. Ingram

Political Science, University at Albany, USA
Email: mingram@albany.edu

**Received:** 19 September 2025 UTC; **Revised:** 21 April 2026 UTC; **Accepted:** 17 June 2026 UTC

**Keywords:** Brazil; crime; environment; machine learning

## Abstract

This article contributes a machine learning approach to detecting environmental crimes in Brazil. Detecting these crimes is important for the health of the environment and public safety. The measurement of environmental crimes is challenging, in part due to efforts of those involved to remain hidden. However, administrative data on embargoes—legal liens on parcels of land—identify locations where the government documented some of these crimes. Using these data, the analysis develops a method for detecting yet-unidentified locations of these crimes. Embargo data are merged with a wide range of predictors that capture features of the physical environment, as well as demographic, socio-economic, political, and public safety attributes. After training a series of models on these data, the random forest (RF) model performs best (AUC = 0.92). Overall, a combination of changes in vegetation, homicide rates, the prevalence of young men, human development, employment, the presence of a poverty reduction program, and political factors that capture local electoral competition and participation are important predictors of illegal deforestation. These findings—some unsurprising and some surprising—show the value of multicausal, cross-disciplinary approaches to understanding and detecting environmental crime. The practical gains from integrating remote sensing, a wide range of administrative data, and machine learning help build more effective crime detection, enhancing the government's ability to conduct forest enforcement more safely and efficiently. Future research is promising as more and better data become available, and the approach can be extended to detect other illicit behavior, in Brazil and elsewhere.

## Impact Statement

This article demonstrates the value of integrating remote sensing, administrative data, interdisciplinary data from the physical and social sciences, and machine learning tools for the remote detection of illegal deforestation in Brazil, showing how data science approaches can assist government agencies to more safely and effectively prevent illegal deforestation and environmental degradation.

## 1. Introduction

Environmental crimes in Brazil include mining, the trafficking of plant and animal species, and illegal deforestation to convert land for cattle ranching, soy production, and logging for the black market in lumber (Boekhout van Solinge, 2010; Boekhout van Solinge, 2014; Skidmore et al., 2021; Ministério do

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

© The Author(s), 2026. 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.

https://doi.org/10.1017/eds.2026.10060 Published online by Cambridge University Press

Meio Ambiente e Mudança do Clima, 2025). Detecting these crimes is important for several reasons, including the health of the environment, especially the Amazon biome, as well as broader social harms that result from environmental degradation (Carpio-Domínguez, 2024). Most deforestation in the region is illegal (1, 272) and is increasing: the period 2016–2020 showed almost twice as much deforestation as in comparable time spans from 2001 to 2005 or 2006 to 2010 (da Silva et al., 2023a), and the first half of 2025 marked a record-breaking surge in deforestation (Berti, 2025). Environmental crimes are often closely connected to violence and criminal networks, including persistent evasion of authorities and oversight (Carvalho et al., 2019), so improving detection of these crimes is also important to combat corruption and organized crime, and to reduce lawlessness and insecurity in the region (Alston et al., 2000; Muggah, 2023; Muggah and Smith, 2024). Further, much of the violence and evasion that accompanies environmental crimes occurs in rural areas at the frontiers of government attention and regulation, so-called "brown areas" where the law is absent (O'Donnell, 2001; Carvalho et al., 2019), and thus where safety is a concern, so improvements to remote methods offer a way to monitor environmental degradation with lower personal risks.

Measuring environmental crimes is challenging, just as with other criminal behavior, in part due to the illicit, clandestine nature of the activity and the active efforts of those involved to remain hidden and hard to detect. Further, environmental crimes connected to deforestation, for example, illegal logging, are the kind of "gray" or "para-criminal" activity that is ingrained in regular, legal flows of economic activity that help camouflage the illegality (Vigh and Sausdal, 2025).

Yet, *embargoes*—legal liens on parcels of land for illegal environmental activity on the property—provide a way to identify locations where the government has already documented at least some of these crimes. Studies have leveraged these data to study compliance with embargoes (Silva et al., 2022) and changes in environmental enforcement over time (Nunes et al., 2024). However, while other administrative data have been incorporated into efforts to detect illegal deforestation (Ribeiro et al., 2024), to our knowledge, the data on embargoes have not. This study leverages information about these embargo locations to predict other, yet-unidentified locations of environmental crimes. This process complements existing efforts to improve government effectiveness in forest enforcement (Silva et al., 2022; Nunes et al., 2024).

The approach proceeds in three steps. First, using public, administrative data on embargoes, we identify documented locations of environmental crimes, focusing only on illegal logging and deforestation in the Legal Amazon, permitting us to isolate this one type of environmental crime within a key defined region. Second, we merge this information with data on a wide, interdisciplinary range of potential explanatory factors, including features of the physical environment from satellite imagery, as well as demographic, social, economic, political, and public safety attributes at each location. This varied combination of explanatory factors speaks to the call for interdisciplinary, multicausal, "multifactorial" approaches to illegal deforestation (Carpio-Domínguez, 2024). Third, we divide the data into training and testing subsets of the data, and train a series of models. The random forest (RF) model performs best, and we interpret this model in a series of steps, moving from a black box to a more understandable, explainable result consistent with interpretable machine learning (Doshi-Velez and Kim, 2017; Miller, 2019; Molnar, 2025). The combination of remote sensing, administrative and interdisciplinary data, and machine learning responds to calls for greater use of remote sensing in monitoring illegal deforestation (Thompson and Magrath, 2021; Ribeiro et al., 2024), and complements efforts to build data infrastructure on forests and pair these data with machine learning (Ouaknine et al., 2025). The approach can be extended to detecting other environmental crimes and illicit behavior, in Brazil and elsewhere.

Looking ahead, the findings show that a combination of changes in vegetation, levels of violence, changes in violence, the prevalence of young men, human development (an index of income, life expectancy, and education), employment, the prevalence of a poverty reduction program (Bolsa Familia), and political factors that capture local electoral competition and participation, are important predictors of illegal deforestation. These findings—some unsurprising and some surprising—endorse a multicausal, cross-disciplinary approach that draws on ecology, criminology, demography, sociology, and political science. Practically, integrating remote sensing, administrative data, a wide range of interdisciplinary

https://doi.org/10.1017/eds.2026.10060 Published online by Cambridge University Press

data, and machine learning enhances the effectiveness of environmental crime detection, helping the government to conduct forest enforcement more safely and efficiently.

Future research is promising, especially for collaboration across disciplines and with government agencies, and as more and better data become increasingly available. The conclusion returns to these issues, addressing limitations and future prospects.

## 2. Related work

This study sits at the intersection of literatures on environmental monitoring and environmental crime, and complements approaches in both that combine remote sensing, administrative data, and machine learning. The goal of environmental monitoring is to minimize harms to the ecosystem by documenting the natural environment, evaluating changes, and studying possible causes and consequences of those changes (Artiola et al., 2004; Olawade et al., 2024). Existing work includes spatiotemporal tracking of forest and other natural vegetation (Vancine et al., 2024), temperature changes (Jiménez-Muñoz et al., 2016; Rani et al., 2025), fire classification (Roy and Kumar, 2017), rainfall, floods, and droughts (Lima and AghaKouchak, 2017; Lima et al., 2021; Santos et al., 2018), the flow of waterways (Lappicy and Lima, 2023), groundwater contamination (de Oliveira et al., 2026), and tree species classification (Ramesh et al., 2025; Hou et al., 2025).

In the literature on environmental crime, studies examine illegal mining, the trafficking of plant and animal species, and illegal deforestation to convert forest land for cattle ranching, soy production, and timber (Boekhout van Solinge, 2010; Boekhout van Solinge, 2014; Skidmore et al., 2021; Ministério do Meio Ambiente e Mudança do Clima, 2025). Studies in Brazil also evaluate the effectiveness of environmental enforcement, including compliance with legal liens where environmental crimes have already been documented (Silva et al., 2022), and changes in environmental enforcement over time (Nunes et al., 2024). Existing research indicates that environmental crime "cannot be reduced to a single cause," and that a broad range of social, economic, and individual factors are relevant (Suprayogi et al., 2025; Ministério do Meio Ambiente e Mudança do Clima, 2025). Focusing on illegal deforestation, studies show this kind of environmental crime has multiple, varied explanatory factors, and therefore solutions require a cross-disciplinary, "multifactorial response" (Carpio-Domínguez, 2024), addressing environmental, social, economic, legal, and political drivers.

In the Amazon, several existing projects focus on detecting illegal deforestation, including on-site and off-site approaches with various tradeoffs. On-site approaches include labor-intensive activities by researchers, private actors, and government investigators. Scholarly studies have deployed intensive fieldwork, including manual sampling and laboratory analysis (Olawade et al., 2024), as well as qualitative approaches and deep ethnographic studies (Boekhout van Solinge, 2010; Boekhout van Solinge, 2014). Private actors drive important participation in community-based environmental monitoring (Damonte et al., 2024). And the government has taken a variety of approaches. Among official authorities, the Brazilian Institute for the Environment and Natural, Renewable Resources (IBAMA) has Special Investigative Groups (Grupos Especializados de Fiscalização, GEF), which engage in multidisciplinary and even tactical operations on the ground to identify and control environmental crimes (Anderson, 2024; Socrates, 2025). The Priority Municipalities (*Municipios Prioritários*) program also allocates additional enforcement resources to municipalities with higher rates of deforestation, and this program has been found to be effective (Oliveira and Miranda, 2024). Ordinary citizens can also report suspected environmental crimes to IBAMA. These on-site techniques can provide close, in-person attention to environmental crimes. Similarly, in-person ethnographic research can provide a deeper understanding based on local voices (Boekhout van Solinge, 2014).

However, on-site techniques can be laborious, costly, and dangerous. Traditional sampling and laboratory approaches can be expensive (including training and labor), and have also been criticized for small study areas, delays, subjectivity, and environmental impact (e.g., of travel and on-site activity) (Ye et al., 2020; Olawade et al., 2024). Given the resource-intensive nature of these activities, it can be especially challenging when new governments enter office, political commitments and policies change,

https://doi.org/10.1017/eds.2026.10060 Published online by Cambridge University Press

and leaders are less supportive of institutions like IBAMA. This happened during the military dictatorship from 1964 to 1985, and more recently during the administration of Jair Bolsonaro (2019–2022) (Saad-Diniz and Gianecchini, 2021), when IBAMA was seriously undermined and its personnel reduced from around 6,000 to around 2,000 (Mendes, 2025). Further, on-site approaches are restricted to smaller, predetermined geographic areas; that is, you must already know where to look. And in rural or remote areas, like deep parts of the Amazon and especially in some indigenous territories, there is a long history of violence between (a) environmental activists and/or local inhabitants and (b) extractivists (Boekhout van Solinge, 2014; Garnelo et al., 2023; Dal'Asta et al., 2025; Comissão Pastoral da Terra, 2025). Violence and security concerns pose safety risks to in-person monitoring efforts. These risks also hold for projects aimed at detecting other kinds of illicit behavior, for example, clandestine graves (Jessee and Skinner, 2005). For instance, 2023 was one of the most violent years for the Guajajara people in the Arariboia territory in the northern state of Maranhão (Mendes, 2024).

Off-site approaches include aerial or satellite collection of data, including a variety of machine learning approaches to assess deforestation in these data (Bragagnolo et al., 2021). Aerial approaches (also called "sub-orbital" or low-altitude) employ aircraft or drones to collect photographic images or sensor data, and satellite-based approaches (also called "orbital") leverage optical or radar sensors on a wide range of satellites. Aerial, sub-orbital approaches can cover a larger area than on-site approaches, like the helicopters and drones of IBAMA's GEF (Anderson, 2024; Socrates, 2025), but these approaches are inappropriate for large regions, and one still needs to know generally where to look. One only knows where to look after a crime has occurred; thus, it is infeasible to collect systematic data retrospectively.

In contrast, orbital images can be obtained from several continuous, ongoing satellite programs. Thus, once an event of interest occurs, it is feasible to obtain images of the area (at various levels of spatial and temporal resolution) retrospectively, that is, looking back after the fact, leading up to an event of interest and extending beyond it. Orbital approaches of this kind include the System for Detecting Deforestation in Real Time (Sistema de Detecção de Desmatamento em Tempo Real, or DETER), and the Program to Calculate Deforestation in the Amazon (Programa de Cálculo do Desflorestamento da Amazônia, or PRODES) (Ribeiro et al., 2024). DETER uses satellite imagery to provide daily alerts on changes in forestry and vegetation, and is meant to be an early-warning system for deforestation. From 2004 to 2017, DETER used the MODIS sensor from the TERRA/NASA satellites, with a spatial resolution of 250 m (250 m × 250 m cells); since 2018, DETER has used the WFI sensor on the CBERS/INPE satellite, with a spatial resolution of 60 m (Instituto Nacional de Pesquisas Espaciais, 2025a). PRODES also uses satellite imagery to measure deforestation, but this program offers data on an annual basis. PRODES initially made substantial use of Landsat satellites with a spatial resolution of 20–30 m, but has increasingly used other satellites, including the Sino-Brazilian CBERS satellites, India's IRS satellites, and the United Kingdom's UK-DMC2 satellite. Regardless of the imagery used, PRODES identifies deforestation locations with a minimum area of 6.5 hectares, which corresponds with a 250 m cell (Instituto Nacional de Pesquisas Espaciais, 2025b). Both DETER and PRODES are operated by Brazil's National Institute for Space Research (Instituto Nacional de Pesquisas Espaciais, INPE), and report their data to IBAMA. PRODES was designed to detect primary forest loss, so complementary efforts have also focused on detecting nonforest (NF) loss (Messias et al., 2024).

Notably, detecting environmental change generally, or deforestation specifically, can include a wide range of phenomena, including natural disasters (e.g., wildfires prompted by lightning), and human behavior, including accidents, legal conduct, and illicit activities. Thus, differentiating illegal deforestation from other patterns is challenging. DETER tries to identify human, or "anthropic," deforestation by differentiating geometric from disorganized forest loss—where more structured or geometric forest loss indicates possible human behavior (Instituto Nacional de Pesquisas Espaciais, 2025a). However, a limiting factor is that when IBAMA receives data from either DETER or PRODES, the agency still needs to distinguish licit from illicit behavior, which requires the on-site approaches discussed earlier. Still, one study estimated that DETER had a meaningful impact on reducing violence by deterring deforestation (Araujo et al., 2025).

https://doi.org/10.1017/eds.2026.10060 Published online by Cambridge University Press

Existing work highlights the importance of remote sensing and machine learning as transformative tools for environmental monitoring and assessment. Remote sensing enhances transparency by making data more accessible, including publicly available satellite images and other data like the data employed here (Ribeiro et al., 2024), though some are also cautious about political misuses of this transparency (Oliveira and Siqueira, 2022). Remote sensing data can be paired with information in administrative data like Brazil's Registry of Rural Areas (Cadastro de Areas Rurais, CAR) to improve both remote monitoring and administrative data collection (Ribeiro et al., 2024). The CAR has been criticized for creating new opportunities for deforestation through built-in exceptions to limits and for the ease with which it can be subverted (Carvalho et al., 2019), yet there are still promising opportunities for collaboration among scholars and public agencies to improve forest monitoring and oversight.

Machine learning offers complementary advantages. In place of intensive, on-site fieldwork and analysis, machine learning offers a less costly, faster, more nimble, and less subjective alternative. A wide range of applications using linear regression, maximum likelihood, support vector machines, decision-trees, random forests, and neural networks constitute meaningful advances in environmental monitoring (Zhou et al., 2023; Chowdhury, 2024; Olawade et al., 2024).

This project builds on and extends these existing studies to offer an alternative solution to detecting environmental crime. This study takes an off-site approach, leveraging satellite imagery and the ability to collect data retrospectively on locations of interest, and combines this with machine learning. The off-site nature of the approach also means it does not carry any of the resource intensity or security risks associated with placing people in potentially unsafe locations in the field. The project also combines the remote sensing data with administrative data on embargoes from IBAMA, much like the combination of satellite imagery and CAR noted above (Ribeiro et al., 2024). Specifically, IBAMA provides local, expert details about location boundaries for specific, known sites of illegal deforestation, and we integrate this information with a larger dataset and our machine learning approach to predict yet-unknown locations of other illegal deforestation. The integration of IBAMA data promises to improve the ability to separate illegal deforestation from deforestation more broadly. Lastly, the project also integrates these data with a broad range of other explanatory predictors, responding to calls for "multifactorial" approaches to illegal deforestation (Carpio-Domínguez, 2024).

## 3. Method

This study builds on the environmental monitoring literature that employs remote methods using orbital satellite imagery and machine learning, and draws on the environmental crime literature by taking a broad, multicausal approach to the drivers of this kind of crime. The research design anticipates that the outcome of interest is influenced by a wide range of local features, including the physical environment, demography, economics, and politics. The predictive, machine learning approach is appropriate in that it allows for interactions and nonlinearities among the included features.

### *3.1. Data*

The outcome of interest is the predicted location of an embargo for illegal logging or deforestation. According to the Brazilian Institute for the Environment and Natural, Renewable Resources (IBAMA, 2026), there are more than 80,000 of these embargoes from 1988 to 2024. We use data on a subset of embargoes from the year 2011, focusing only on illegal logging and deforestation, to capture the outcome of interest. The year 2011 was selected because a wide range of predictors are organized in an existing data set covering 2000–2010 (see below), thus facilitating a multicausal analysis of illegal deforestation. The approach serves as a kind of proof of concept. If it is effective with these earlier data, the project can be extended to broader and more recent data, and is more promising with the increased temporal and spatial resolution of more recent data.

Key predictors include 5-year and 10-year changes in vegetation constructed from annual vegetation indices (Broxton et al., 2014a, 2014b). The variable *veg_change_10yr* captures the 10-year change in

https://doi.org/10.1017/eds.2026.10060 Published online by Cambridge University Press

vegetation from 2001 to 2010 by subtracting the vegetation index in 2001 from the index in 2010, and the variable *veg_change_5yr* captures the 5-year change in vegetation from 2006 to 2010 by subtracting the index in 2006 from 2010. All other predictors are from official government statistics offices in Brazil and organized at the municipal level (Ingram and Marchesini da Costa, 2019a, 2019b; Mendez and Sachdeva, 2025). Demographic factors include the proportion of the population composed of young men (*pop_young_men*), population density (*pop_density*), and the proportion of single-parent households (*single_parent*). Socio-economic factors include the Gini measure of income inequality (*inequality*), employment rate (*employment_rate*), the human development index, which is a composite measure of income, education, and life expectancy (*human_development*), and the proportion of households participating in a prominent conditional cash transfer program, Bolsa Familia (*Bolsa_Familia*). Public safety factors include the annual homicide rate in each of several years (2006, 2007, 2008, 2009, and 2010; e.g., *homicide_rate*_2010), and the 5-year change in homicide rate from 2005 to 2010, captured by subtracting the rate in 2006 from the rate in 2010 (*change_in_homicide_rate*_5*yr*).

Notably, vegetation data are in raster format (500-m cells), and the embargo and all other data from municipal units are in vector format, though the embargo locations are much smaller areas than municipal areas. All data were processed to the size of the 500 m cells of the vegetation rasters, yielding a grid across the Amazon of nearly 6 million cells (5,949,013). In all, 2,474 of these cells contain embargoes for illegal deforestation.

Figure 1 shows the geographic boundary of the study region, the Legal Amazon. Figure 2 shows the distribution of embargoes across the study region. Notably, the embargo locations are not points, but rather polygons that identify the boundary of the embargoed property. The red boundary identifies the state of Pará, and the inset focuses on the southeast corner of this state. The close-up in the inset shows the different geometric shapes of the polygons for various embargoes.

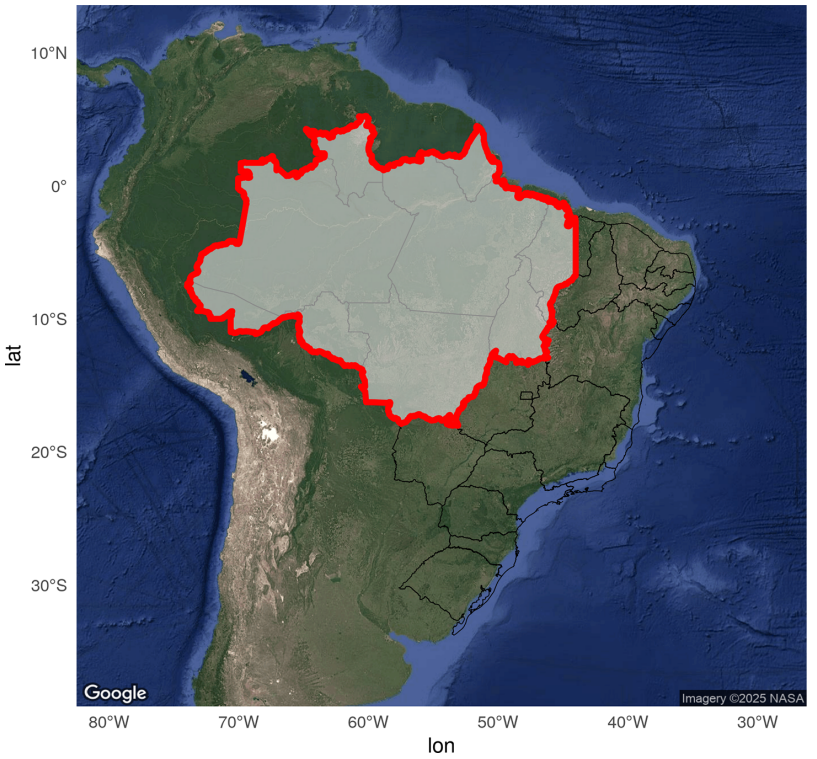



***Figure 1.*** *Map of study area, the Legal Amazon, shown in red within Brazil. The outline of Brazil and its first-level subnational units (states) are shown in gray against the backdrop of South America.*

https://doi.org/10.1017/eds.2026.10060 Published online by Cambridge University Press

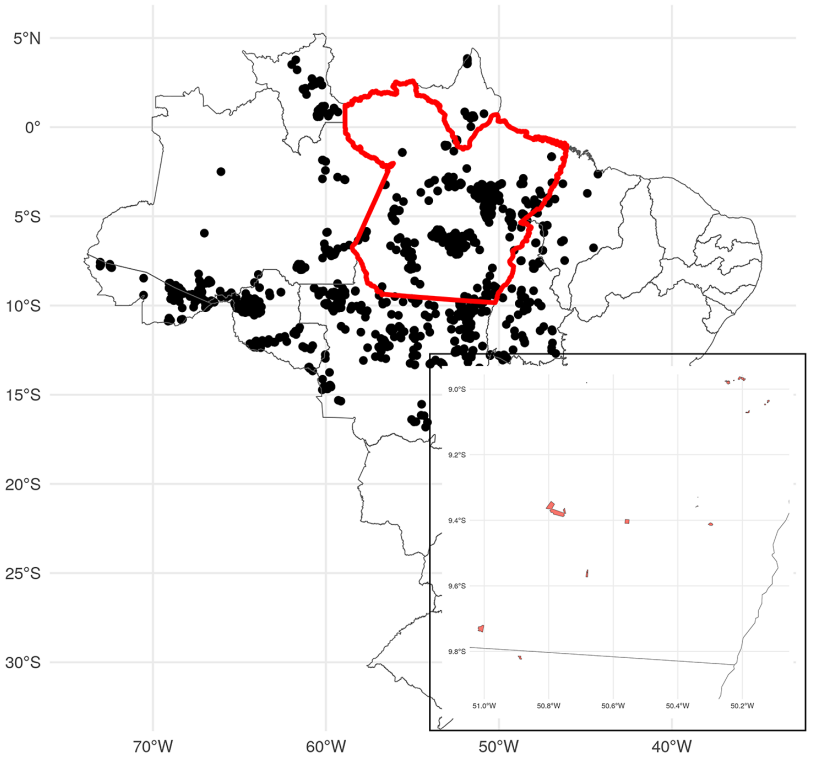



***Figure 2.*** *Distribution of embargoes across the study area. Boundaries of Brazilian states shown in gray. To illustrate that the embargo locations are polygons, not points, the state of Pará is outlined in red, and the inset image zooms in to the southeast corner of this state. The inset shows that embargo locations are polygons of varying sizes and shapes.*

### *3.2. Model selection and rationale*

In the predictive approach here, the core analytic goal is not to assess the statistical significance of any one predictor or set of predictors. Rather, the key goal is to reduce overall error in order to increase the accuracy of predictions, and then interpret the best model.

The process of selecting the best model covers six modeling options: the common least-squares model (ordinary least-squares, or OLS), ridge regression (Hoerl and Kennard, 1970), lasso regression (Tibshirani, 1996), elastic net (enet) regression (Zou and Hastie, 2005), adaptive best-subset selection (ABESS) (Zhu et al., 2022), random forest (RF) (Ho, 1995), and extreme gradient boosting (XGB) (Chen and Guestrin, 2016). Ridge and lasso regression are both regression-based approaches to large datasets and machine learning that minimize variance and reduce error. These models also regularize (penalize) coefficients, pushing coefficients toward zero, which can help identify the most relevant predictors. Elastic net (enet) models combine elements of both ridge and lasso techniques, pushing coefficients toward zero to highlight the most relevant predictors, but tend to generate more complex models than ridge or lasso. ABESS operates similarly to identify the most relevant explanatory factors, and has been described as a kind of "oracle" technique (Zhu et al., 2020) that identifies the best predictors of an outcome in the data as if it knew the true data-generating process. RF and XGB are both ensemble approaches based on combining classification and regression trees (CARTs). The difference is that RF builds a large number of trees (a forest) independent of each other and then engages in an averaging process to identify meaningful variables and interactions. Meanwhile, XGB builds individual trees sequentially, with each new tree learning (boosting) from the errors of the previous tree, and users can calibrate this learning.

https://doi.org/10.1017/eds.2026.10060 Published online by Cambridge University Press

***Table 1.*** *Summary of model performance*

| Model | MSE | MAE |
|---|---|---|
| lm | 0.0008337 | 0.001731 |
| ridge | 0.0008345 | 0.0016218 |
| lasso | 0.0008347 | 0.0016219 |
| elasticnet | 0.0008347 | 0.0016219 |
| abess | 0.0008337 | 0.0017319 |
| rf | 0.000828 | 0.0016075 |
| xgb | 0.0008845 | 0.0016869 |

*Note:* Models were evaluated based on two measures of predictive accuracy: MSE (mean squared error) and MAE (mean absolute error). Included models are: lm (basic linear model, ordinary least squares), ridge regression, lasso regression, elasticnet regression, abess (adaptive best-subset selection), rf (random forest), and xgb (extreme gradient boosting).

Here, the learning rate was set at 0.1. All models are evaluated based on their ability to minimize prediction error.

All analyses were done using R (R Core Team, 2024). To address large-memory demands based on the size of the data, the analysis ran on a high-performance computing cluster at the author's home institution, the University at Albany. Some variations remained computationally prohibitive. For instance, RF modeling did not include proximity matrices due to the computational challenges of processing an $N \times N$ matrix where $N$ is nearly six million. The conclusion returns to this point in addressing limitations and future research.

Least-squares regression is part of the base R software using the *lm*() syntax. Ridge, lasso, and enet regressions were implemented using the package *glmnet* (Simon et al., 2011; Tay et al., 2023). ABESS was implemented with the *abess* package (Zhu et al., 2020; Zhu et al., 2022; Zhang et al., 2023). Random forests were implemented with the *randomForest* package (Liaw and Wiener, 2002), with interpretation facilitated by the *randomForestExplainer* package (Paluszynska et al., 2020). XGB was implemented with the package *xgboost* (Chen et al., 2024). Replication materials are available from the author's Harvard Dataverse site (https://doi.org/10.7910/DVN/723MZQ). See also Data Availability Statement.

Table 1 reports mean squared error (MSE) and mean absolute error (MAE) for all models tested. The random forest (RF) model has the lowest MSE (0.000828) and MAE (0.0016075), so it outperforms all other options. The RF model also has an AUC of 0.92, shown in Figure 3, with a 95% CI of 0.908–0.922) (DeLong et al., 1988; Robin et al., 2011). Based on this performance, the interpretation of results focuses on the RF model.

## 4. Results and discussion

Figure 4 reports two variable importance plots from the random forest model. In both plots, the variables toward the top are the more important predictors. The plot on the left ranks predictors according to the proportion by which MSE increases if the predictor is left out (%IncMSE), and the plot on the right ranks predictors according to the increase in node purity (IncNodePurity). Changes in vegetation rank highly in both plots. In the first plot, the 5-year change in vegetation is the top feature and the 10-year change is the fourth feature; in the second plot, 10-year change and 5-year change in vegetation are the top two predictors. After that, there is a range of socioeconomic, demographic, public safety, and political features. In the first graph, employment rate, single-parent households, the homicide rate in 2009, prevalence of young men, abstention from voting, margin of victory, and the human development index are all in the top 10 predictors. In the second graph, Bolsa Familia (the financial assistance program for low-income families), the homicide rate in 2007, the prevalence of young men, effective number of candidates (ENC), abstention, employment rate, and the 5-year change in homicide rate are in the top 10.

https://doi.org/10.1017/eds.2026.10060 Published online by Cambridge University Press

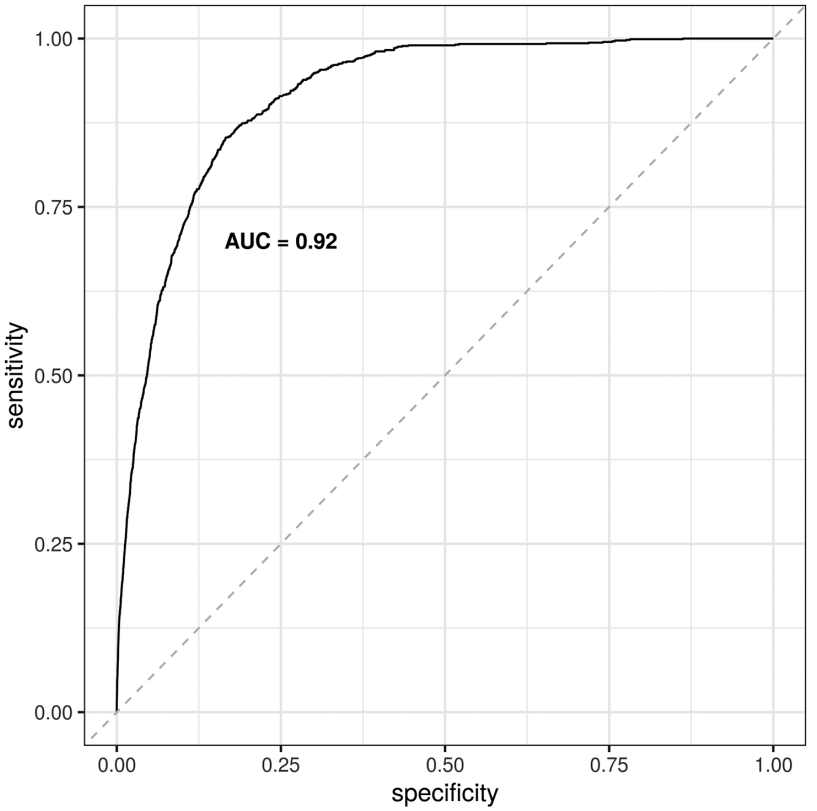



***Figure 3.*** *ROC-AUC curve for random forest (RF) model, showing an AUC of 0.92.*

There is substantial overlap across the two plots, with the two measures of vegetation change, employment rate, prevalence of young men, abstention, and measures of violence appearing in the top 10 on both.

Figure 5 complements the previous plots, reporting the distribution of minimal depth among the trees. Again, variables toward the top are most important. Minimal depth is the "depth of the node which splits on that variable and is the closest to the root of the tree. If it is low, then a lot of observations are divided into groups on the basis of this variable" (Paluszynska et al., 2020). The label identifies the mean minimal depth for each variable across all trees, and the *x*-axis identifies the number of trees any one variable was used for splitting. The following features appear in the top half of the graph: prevalence of young men, abstention, employment rate, Bolsa Familia, and human development, along with several measures of violence—homicide rates for 2006, 2007, 2008, and 2010, as well as the 5-year change in homicide rate—and the 10-year change in vegetation.

The next three graphs examine feature importance based on a combination of different measures of importance, that is, multi-way importance. Figure 6 captures multi-way importance based on the three least-correlated measures of importance: number of nodes, MSE increase, and times a root. Figure 7 captures multi-way importance based on measures of tree structure: number of nodes, times a root, and mean minimal depth. And Figure 8 captures multi-way importance based on the three measures of accuracy: MSE increase, node purity increase, and *p*-value. In Figures 6–8, variables that rank highly across all three measures are labeled to help identify the most important ones, and also indicate which combinations of variables are most important.

The features identified as most important across all three multi-way plots include: homicide rate in 2010, 5-year change in homicide rate, 10-year change in vegetation, Bolsa Familia, prevalence of young men, employment rate, ENC, and abstention. The 5-year change in vegetation, human

https://doi.org/10.1017/eds.2026.10060 Published online by Cambridge University Press

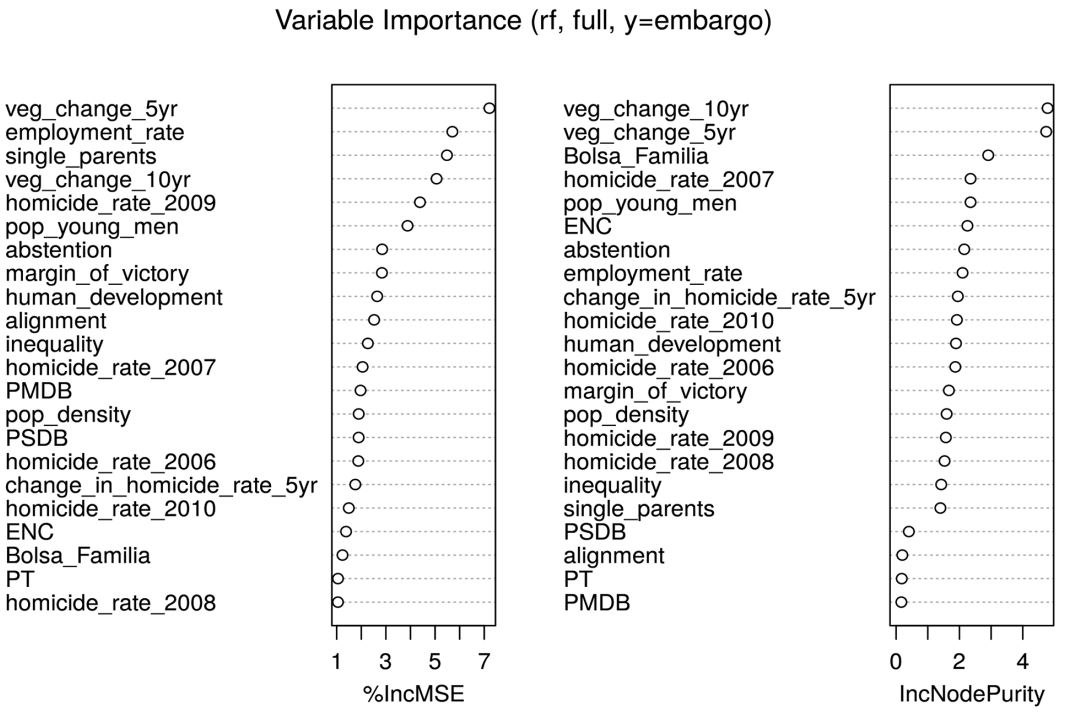



***Figure 4.*** *Variable importance plots based on the percent increase in MSE (%IncMSE) and increase in node purity (IncNodePurity). Variables included are: 5-year change in vegetation index (veg_change_5yr), 10-year change in vegetation index (veg_change_10yr), population density (pop_density), prevalence of young men (pop_young_male), homicide rate in various years from 2006 to 2010 (e.g., homicide_rate_2006), employment rate (employment_rate), proportion of single-parent households (single_parents), prevalence of the Bolsa Familia poverty-reduction program (Bolsa_Familia), economic inequality (inequality), human development index (human_development), margin of victory in last mayoral election (margin_of_victory), effective number of candidates from last mayoral election (ENC), partisan alignment of mayor and governor (alignment), abstention rate (abstention), and identifiers for the three major political parties of local mayors (PMDB, PSDB, and PT).*

development, and the homicide rate in 2008 also appear in two of the graphs. Measures that appear only in one graph are population density (Figure 6) and the homicide rate in 2007 (Figure 8).

One of the strengths of RF models like the one here is that they capture nonlinearities and interactions. There are 242 interactions in the model, and Figure 9 reports the 30 most frequent interactions. Table 2 complements the figure by showing these interactions in greater detail. In Figure 9, the *x*-axis identifies pairs of variables. The interactions are sorted in descending order of frequency, so the most frequent interactive pair is on the far left (and at the top of Table 2). The *y*-axis identifies the depth of each interaction. The top interactions involve the same variables highlighted earlier: homicide rates in several years, change in homicide rates, change in vegetation (5-year and 10-year), population density, the prevalence of young men, employment rate, Bolsa Familia, effective number of candidates (ENC), and abstention.

Summing up thus far, the variables that consistently emerge as most important include those that capture changes in the natural environment (vegetation change), violence and changes in violence (homicide rate and change in the homicide rate), demographic factors (population density and prevalence of young men), socioeconomic factors (Bolsa Familia, employment rate, and human development), and

https://doi.org/10.1017/eds.2026.10060 Published online by Cambridge University Press

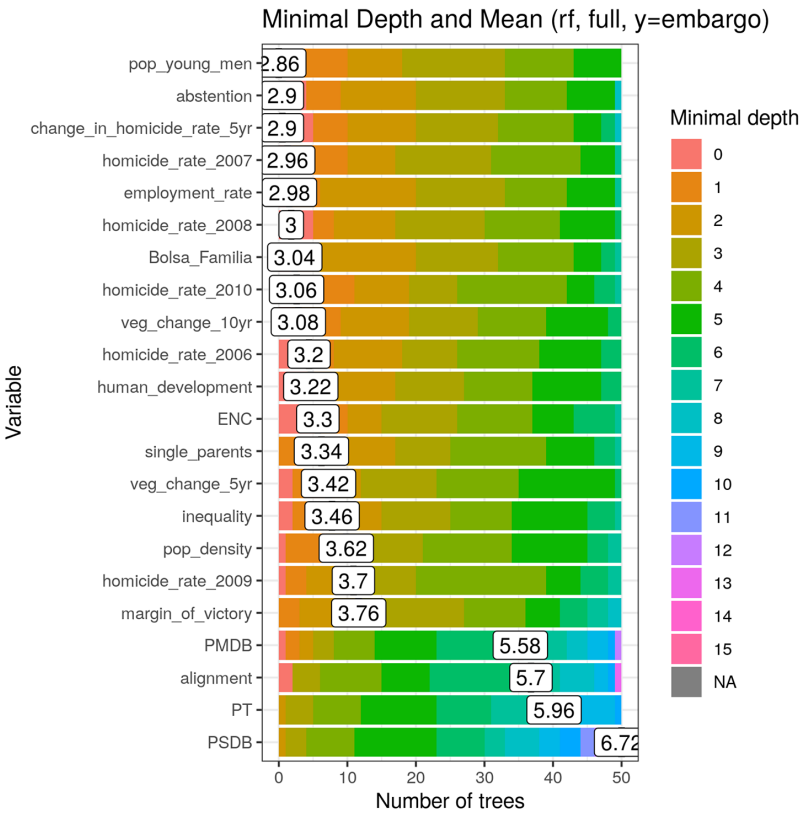



***Figure 5.*** *Variable importance plot based on minimal depth and mean. Variables included are: 5-year change in vegetation index (veg_change_5yr), 10-year change in vegetation index (veg_change_10yr), population density (pop_density), prevalence of young men (pop_young_male), homicide rate in various years from 2006 to 2010 (e.g., homicide_rate_2006), employment rate (employment_rate), proportion of single-parent households (single_parents), prevalence of the Bolsa Familia poverty-reduction program (Bolsa_Familia), economic inequality (inequality), human development index (human_development), margin of victory in last mayoral election (margin_of_victory), effective number of candidates from last mayoral election (ENC), partisan alignment of mayor and governor (alignment), abstention rate (abstention), and identifiers for the three major political parties of local mayors (PMDB, PSDB, and PT).*

political factors that capture participation and the competitiveness of local politics (abstention and ENC). Some of the features found to be important—on their own and in combination with others in the interactions above—might seem unsurprising to some audiences, while others may be more surprising and deserve more attention.

Among unsurprising features, violence and vegetation changes are typically linked to illegal deforestation. Recalling the opening sections, deforestation is often associated with areas of prior vegetation loss, and illegal deforestation often occurs at the boundaries of legal areas of land conversion, for example, at the edges of cattle ranches or soy farms (Skidmore et al., 2021). Other studies have used the registry of rural areas (Cadastro de Areas Rurais, CAR) (Carvalho et al., 2019) and similar administrative data on land use to try to approximate where these peripheral, rural areas at the edges of large landholdings might be. Violence is also a common feature of these rural areas where there is conflict over land grabbing and dispossession, and where the state has a weak presence (O'Donnell, 2001; Carvalho et al., 2019). For these reasons, it is also unsurprising that population density is linked to illegal deforestation. In Brazil, studies of local areas, including indigenous areas, find that roads are related to deforestation by facilitating

https://doi.org/10.1017/eds.2026.10060 Published online by Cambridge University Press

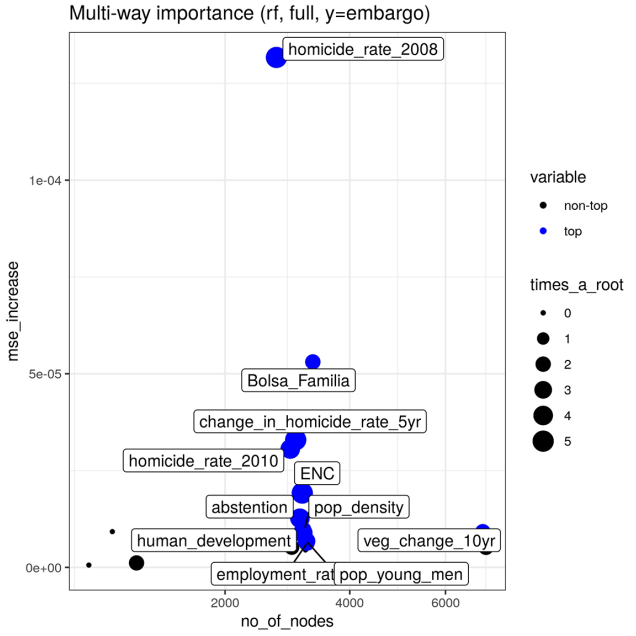



***Figure 6.*** *Multi-way importance based on three measures of importance: number of nodes (no_of_nodes), MSE increase, and times a root. The top 10 most important features are labeled, including: 10-year change in vegetation index (veg_change_10yr), population density (pop_density), prevalence of young men (pop_young_men), homicide rate in 2008 and 2010 (homicide_rate_2008 and homicide_rate_2010), 5-year change in homicide rates (change_in_homicide_rate_5yr), employment rate (employment_rate), prevalence of the Bolsa Familia poverty-reduction program (Bolsa_Familia), human development index (human_development), effective number of candidates from last mayoral election (ENC), and the abstention rate (abstention). More than 10 features are identified if there is a tie.*

land-grabbing and environmental crimes (Ferrante et al., 2021; da Silva et al., 2023b). This might be especially true of new or informal, unregulated roads, as well as other types of transportation networks (e.g., airstrips and waterways).

For these reasons, related work focuses on human–environment interactions, including how development goals can collide with environmental protection. For instance, building infrastructure (e.g., roads) can be beneficial for development, but infrastructure projects can also drive environmental degradation (Kaczan, 2020).

Other features may be more surprising and are not typically included in environmental monitoring or off-site approaches to detecting environmental crimes, so they merit greater attention. Indeed, these demographic, socioeconomic, and political factors highlight the interdisciplinary resonance and implications of the findings. These more surprising features include: (1) demographic factors like the prevalence of young men; (2) socioeconomic factors like the level of human development, the employment rate, and the poverty-reduction program Bolsa Familia; and (3) political factors like local margin of victory, effective number of candidates (ENC), and abstention.

Many of these more surprising features are familiar to scholars in criminology, as well as sociologists and demographers who study crime and violence. Social disorganization theory (Shaw and McKay, 1942; Sampson and Groves, 1989) provides a lens from those disciplines through which to understand these features. According to early versions of this theory, residential instability, resource deprivation, and ethnic heterogeneity were markers of unstable or disorganized communities, and this community disorganization was associated with higher rates of crime and violence (Shaw and McKay, 1942). Later work fleshed

https://doi.org/10.1017/eds.2026.10060 Published online by Cambridge University Press

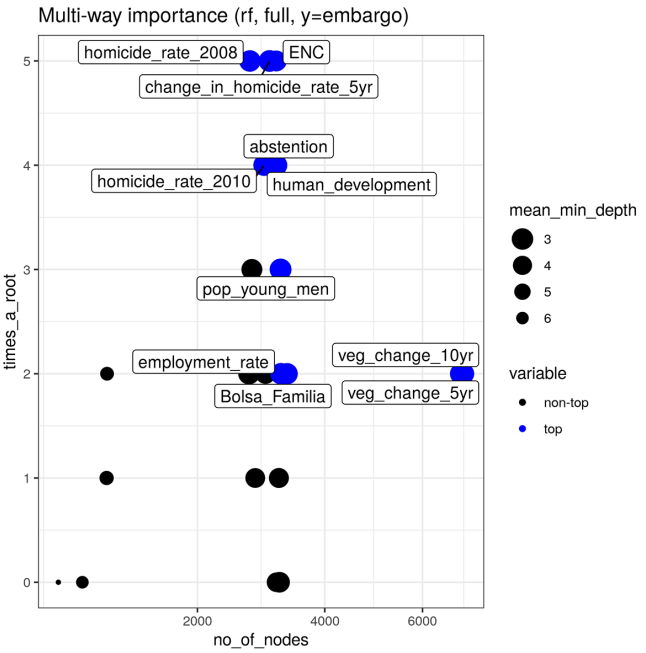



***Figure 7.*** *Multi-way importance based on measures related to the structure of trees: number of nodes (no_of_nodes), times a root, and mean minimal depth (mean_min_depth). The top 10 most important features are labeled, including: 5-year change in vegetation index (veg_change_5yr), 10-year change in vegetation index (veg_change_10yr), prevalence of young men (pop_young_male), homicide rate in 2008 and 2010 (homicide_rate_2008 and homicide_rate_2010), 5-year change in homicide rates (change_in_homicide_rate_5yr), employment rate (employment_rate), prevalence of the Bolsa Familia poverty-reduction program (Bolsa_Familia), human development index (human_development), effective number of candidates from last mayoral election (ENC), and abstention rate (abstention). More than 10 features are identified if there is a tie.*

out the role of related factors like unsupervised youth, low participation or engagement in civic organizations, weak or sparse friendships, and family disruption, linking these factors to low collective efficacy or social cohesion, or the absence of community resilience to violence (Sampson and Groves, 1989; Sampson et al., 1997). Generally, there is substantial support for social disorganization theory over time and across comparative contexts (Sampson and Groves, 1989; Sampson and Wilson, 2012; Bellair, 2017).

Social disorganization theory helps to understand the importance of the prevalence of young men, employment rates, human development, and Bolsa Familia. The prevalence of young men is one of the structural aspects of population pressures associated with interpersonal violence, including homicides (Land et al, 1990; Ingram and Marchesini da Costa, 2019a; Cullen et al., 2024). If violence is associated with illegal deforestation, and the prevalence of young men is associated with violence, it makes sense that the prevalence of young men would also be associated with illegal deforestation.

Considering that much illegal deforestation takes place at the peripheral boundaries of large cattle or soy farms, the kinds of labor required in these areas may attract young men. Alternatively, it may be supplied by criminal networks where young men are prevalent. In these remote, rural areas, the transient nature of this kind of labor mimics the residential instability, resource deprivation, and low supervision that are hallmarks of social disorganization accounts of crime and violence. Low or unstable employment rates also influence social disorganization and are correlated with crime (Land et al, 1990;

https://doi.org/10.1017/eds.2026.10060 Published online by Cambridge University Press

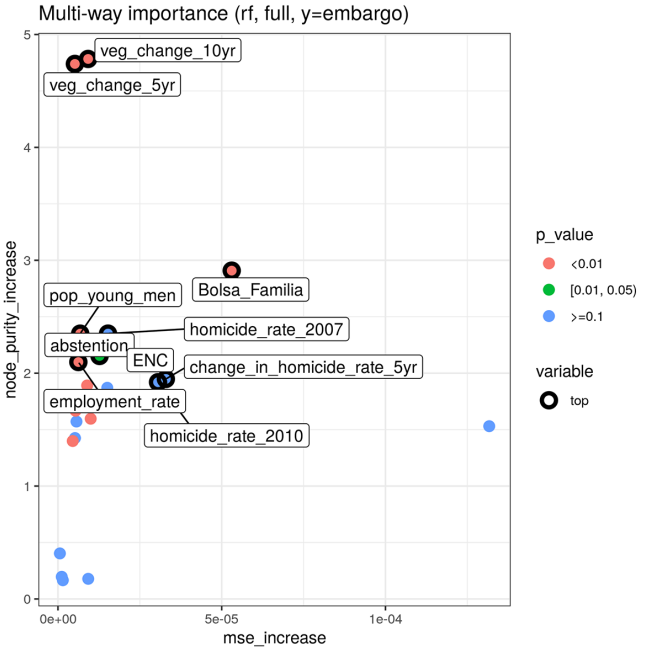



***Figure 8.*** *Multi-way importance based on measures related to predictive accuracy: MSE increase, node purity increase, and p-value. The top 10 most important features are labeled, including: 5-year change in vegetation index (veg_change_5yr), 10-year change in vegetation index (veg_change_10yr), prevalence of young men (pop_young_male), homicide rate in 2007 to 2010 (homicide_rate_2007 and homicide_rate_2010), 5-year change in homicide rate (change_in_homicide_rate_5yr), employment rate (employment_rate), prevalence of the Bolsa Familia poverty-reduction program (Bolsa_Familia), effective number of candidates from last mayoral election (ENC), and abstention rate (abstention). More than 10 features are identified if there is a tie.*

Baller et al., 2001). Another persistent finding in sociological and criminological studies is that adverse childhood experiences and exclusion—including maltreatment, family conflict, poverty, and low education—predict future gang involvement and entry into criminal networks (Carvalho and Soares, 2016; Bauer et al., 2024; Barker et al., 2026). This also helps explain the association of human development (a composite measure of income, life expectancy, and education) with illegal deforestation, since low values of human development would tend to associate with poverty, mortality, illiteracy, and other adverse conditions, which resonate with social disorganization and are linked to crime (Baller et al., 2001). The poverty-reduction program Bolsa Familia would generally be expected to ameliorate these adverse conditions. However, it is also targeted at these conditions, so it may be identifying these markers of social disorganization, including resource deprivation, clarifying its importance. Though the prevalence of single-parent households is not among the top predictors in this study, the criminology literature also helps understand the relevance of this feature, since family disruption undermines social cohesion and can be an adverse childhood experience that contributes to social disorganization and reduced efficacy (Land et al, 1990; Sampson et al., 1997), and is a predictor of crime and violence (Ingram and Marchesini da Costa, 2019a).

Lastly, among the political features, margin of victory and the effective number of candidates measure the level of electoral competition and capture local political incentives. Where competition is low, a single party dominates local politics, which is often associated with traditional, patron-client exchanges where

https://doi.org/10.1017/eds.2026.10060 Published online by Cambridge University Press

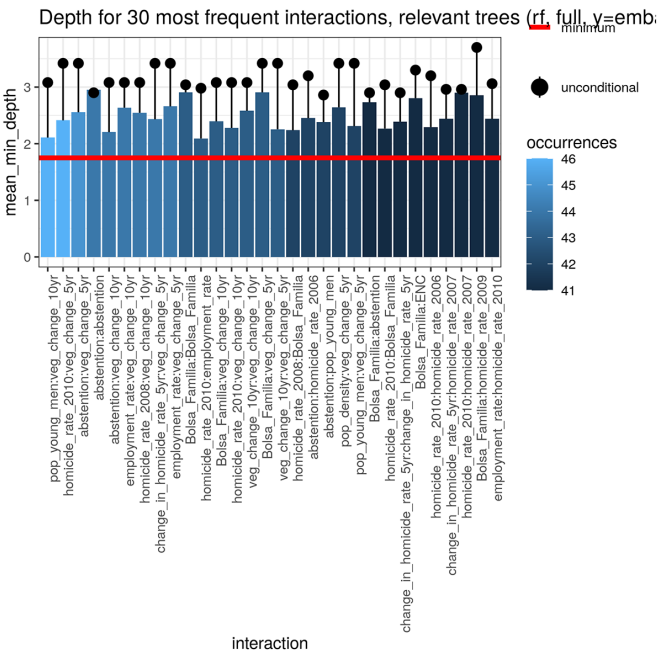



***Figure 9.*** *Depth of the 30 most frequent interactions, relevant trees only. Interactions appear left to right in descending order according to the number of occurrences. Variables included are: 5-year change in vegetation index (veg_change_5yr), 10-year change in vegetation index (veg_change_10yr), population density (pop_density), prevalence of young men (pop_young_male), homicide rate in each year from 2006 to 2010 (e.g., homicide_rate_2006), 5-year change in homicide rate (change_in_homicide_rate_5yr), employment rate (employment_rate), prevalence of the Bolsa Familia poverty-reduction program (Bolsa_Familia), effective number of candidates from last mayoral election (ENC), and the abstention rate (abstention).*

one or two families dominate politics, and which make it easier to establish networks of patronage or corruption, and generate few incentives to change (Beer, 2006; Ríos-Figueroa, 2007; Ingram, 2013). Indeed, there is a literature in Brazil on local bosses or "colonels" (*coroneis*) (Leal, 1977; Pinto Gil, 2022) who wield substantial economic power and blur the lines between this private power and the public power of local government. Where a single political party is dominant, this private power and local government can become deeply enmeshed, and it would be easier to sidestep environmental controls. Conversely, where competition is higher, there is more turnover, traditional patronage networks tend to be disrupted, and politicians are generally more dynamic and responsive, including to environmental concerns. Regarding abstention, it is important to note that voting is mandatory in Brazil, so any abstention from voting is a strong signal of voter disengagement and lack of trust and confidence in the political process. In this regard, abstention is a signal of withdrawal from organizations and participation, a lack of efficacy, and a kind of anomie (Merton, 1938) that fits within the social disorganization theory discussed above. Indeed, research on social capital in political science and the risks of declining participation in civic organizations (Putnam, 1993, 1995) resonates with the literature on social cohesion, collective efficacy, and community resilience and the connection of these community characteristics and crime (Ingram, 2014) Thus, the competitiveness of local politics and abstention provide a bridge between political science and sociological approaches to crime and violence. Prior research establishes a relationship between these political factors and violence in Brazil (Ingram and Marchesini da Costa, 2019a), and this helps understand why these features are important predictors of illegal deforestation.

https://doi.org/10.1017/eds.2026.10060 Published online by Cambridge University Press

***Table 2.*** *Variables involved in the 30 most frequent interactions, with mean minimal depth and frequency*

| Variable | root_variable | mean_min_depth | Occurrences |
|---|---|---|---|
| veg_change_10yr | pop_young_men | 2.11 | 46 |
| veg_change_5yr | homicide_rate_2010 | 2.41 | 46 |
| veg_change_5yr | abstention | 2.56 | 45 |
| abstention | abstention | 2.95 | 44 |
| veg_change_10yr | abstention | 2.20 | 44 |
| veg_change_10yr | employment_rate | 2.64 | 44 |
| veg_change_10yr | homicide_rate_2008 | 2.55 | 44 |
| veg_change_5yr | change_in_homicide_rate_5yr | 2.43 | 44 |
| veg_change_5yr | employment_rate | 2.66 | 44 |
| Bolsa_Familia | Bolsa_Familia | 2.91 | 43 |
| employment_rate | homicide_rate_2010 | 2.09 | 43 |
| veg_change_10yr | Bolsa_Familia | 2.40 | 43 |
| veg_change_10yr | homicide_rate_2010 | 2.28 | 43 |
| veg_change_10yr | veg_change_10yr | 2.58 | 43 |
| veg_change_5yr | Bolsa_Familia | 2.91 | 43 |
| veg_change_5yr | veg_change_10yr | 2.26 | 43 |
| Bolsa_Familia | homicide_rate_2008 | 2.24 | 42 |
| homicide_rate_2006 | abstention | 2.45 | 42 |
| pop_young_men | abstention | 2.38 | 42 |
| veg_change_5yr | pop_density | 2.64 | 42 |
| veg_change_5yr | pop_young_men | 2.31 | 42 |
| abstention | Bolsa_Familia | 2.73 | 41 |
| Bolsa_Familia | homicide_rate_2010 | 2.27 | 41 |
| change_in_homicide_rate_5yr | change_in_homicide_rate_5yr | 2.39 | 41 |
| ENC | Bolsa_Familia | 2.80 | 41 |
| homicide_rate_2006 | homicide_rate_2010 | 2.29 | 41 |
| homicide_rate_2007 | change_in_homicide_rate_5yr | 2.44 | 41 |
| homicide_rate_2007 | homicide_rate_2010 | 2.90 | 41 |
| homicide_rate_2009 | Bolsa_Familia | 2.85 | 41 |
| homicide_rate_2010 | employment_rate | 2.44 | 41 |

*Note:* Interactions appear top-to-bottom in descending order of frequency. Variables included are: 5-year change in vegetation index (veg_change_5yr), 10-year change in vegetation index (veg_change_10yr), population density (pop_density), prevalence of young men (pop_young_male), homicide rate in each year from 2006 to 2010 (e.g., homicide_rate_2006), 5-year change in homicide rate (change_in_homicide_rate_5yr), employment rate (employment_rate), prevalence of the Bolsa Familia poverty-reduction program (Bolsa_Familia), effective number of candidates from last mayoral election (ENC), and the abstention rate (abstention).

## 5. Conclusion

### *5.1. General summary*

Detecting environmental crimes is important, especially in the Amazon biome. Existing approaches adopt either on-site or off-site approaches with various tradeoffs. Off-site approaches in Brazil tend to focus on the features of the physical environment derived from remote sensing, though a smaller set of approaches integrate administrative data from government registries of rural properties.

This study follows the off-site, machine-learning approach, includes features of the physical environment, and contributes to the advancement of environmental crime detection primarily in two ways: (1) leveraging a different kind of administrative data, namely embargoes—legal liens placed on properties for known environmental crimes—and (2) integrating a broad array of demographic, social, economic,

https://doi.org/10.1017/eds.2026.10060 Published online by Cambridge University Press

political, and public safety measures as predictors. The embargo data identify known sites of environmental crimes, specifically illegal deforestation, and the analysis trains models to detect yet-unknown sites of illegal deforestation. The inclusion of the sociodemographic, economic, political, and public safety data expands environmental crime detection in a more interdisciplinary direction consistent with multicausal understandings of environmental crime.

### *5.2. Methods*

The approach proceeded in three steps. First, using public administrative data on embargoes, we identified documented locations of illegal deforestation in the Legal Amazon. Second, we merged this information with data on a wide range of potential explanatory factors, including features of the physical environment from remote sensing, as well as demographic, socioeconomic, political, and public safety attributes at each location. Third, we divided the data into training and testing data, and trained a series of models. The random forest (RF) model performed best, and interpretation focused on this model.

### *5.3. Findings*

Key findings show that the combination of violence, change in violence, vegetation change, population density, the proportion of the population that is young and male, human development, employment rate, effective number of parties (ENC), and abstention are predictive of illegal deforestation. Some of the findings are unsurprising, including the relevance of vegetation change and violence. However, other findings are more surprising, including those that show the predictive importance of socioeconomic and political features.

### *5.4. Practical implications*

The practical implications of these findings relate primarily to (1) the value of interdisciplinary approaches to environmental crime detection, (2) collaborations with government, and (3) enhanced safety of environmental crime detection. The surprising importance of social, economic, and political variables highlights the value of interdisciplinary approaches in this field and supports multicausal or multifactorial approaches to illegal deforestation. The socioeconomic and political features in particular indicate that integrating a political-economy framework to environmental crime detection can yield valuable insights, and motivate future interdisciplinary work. The administrative data on embargoes were vital. It would be promising to coordinate with government agencies to see what other data are available, or to engage in a mutual feedback process to see how future analyses could be improved, or how administrative data collection and sharing could be improved. And the remote, off-site detection of environmental crimes has several benefits, including reducing costs and avoiding security risks, so the approach here—and improving it over time—helps build better and safer means of environmental crime detection.

### *5.5. Limitations*

The analysis has some limitations, including the use of somewhat older demographic, socioeconomic, and political data from 2000 to 2011. These data were used because a wide range of predictors were well-organized across 10 years during this time frame. The vegetation data from older satellite imagery also have reduced spatial resolution. Further, computational complexity limited some aspects of the analysis, including interpretation of a proximity matrix.

### *5.6. Future prospects*

Future research can address the limitations noted above. A natural next step is to update and expand the data. The quality and availability of data are increasing in Brazil, especially with satellite imagery with better spatial resolution since 2018. Other projects that systematize geo-spatial data of higher quality can

https://doi.org/10.1017/eds.2026.10060 Published online by Cambridge University Press

also be leveraged in the future. Further, Brazil's federal government is more supportive of environmental protection since 2022, which should make administrative data like that used here more accessible and transparent. Future work can employ the approach here with these more recent data, and access to a high-performance computing cluster can solve some of the computational challenges. Interdisciplinary collaboration is also promising, including combining efforts with both (a) researchers from other fields and (b) government. Doing so promises improved results and additional insights, and would strengthen forest enforcement. The approach can also be extended to other illicit behavior, in Brazil and elsewhere.

**Data availability statement.** Replication materials are available at Harvard Dataverse: https://doi.org/10.7910/DVN/723MZQ.

**Acknowledgments.** Earlier versions of the methods in this article were presented in 2024 at the University of California, Irvine, and in 2025 at the annual meeting of the New England Political Science Association (NEPSA), and at Conference on Applied Research Methods and Data Collection with Hard-to-Reach (H2R) and Hidden Populations, Nuffield College, Oxford University. I am grateful to all participants for their comments. I am also especially grateful to Bruce Desmarais for his instruction and continued correspondence on methods. All errors are my own.

**Author contribution.** The author made the following contributions. Conceptualization: M.I. Funding acquisition: M.I. Data curation: M.I. Formal analysis: M.I. Supervision: M.I. Validation: M.I. Methodology: M.I. Project administration: M.I. Resources: M.I. Software: M.I. Investigation: M.I. Writing—original draft preparation: M.I. Writing—review and editing: M.I. Visualization: M.I. Final approval: M.I.

**Funding statement.** The methods for this project were developed in part with funding from the Faculty Resource Allocation Program (FRAP-A) at the University at Albany, award no. 97969.

**Competing interests.** The authors declare none.

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

## References

**Alston LJ**, **Libecap GD and Mueller B** (2000) Land reform policies, the sources of violent conflict, and implications for deforestation in the Brazilian amazon. *Journal of Environmental Economics and Management 39*(2), 162–188. https://doi.org/10.1006/jeem.1999.1103.

**Anderson JL** (2024) The Brazilian special-forces unit fighting to save the amazon. Available at https://www.newyorker.com/magazine/2024/04/08/the-brazilian-special-forces-unit-fighting-to-save-the-amazon (last accessed 7 August 2026).

**Araujo R**, **Possebom V and Setti G** (2025) Deterring more than deforestation: Environmental enforcement reduces violence in the amazon. arXiv.org, Papers 2509.06076. Available at https://ideas.repec.org/p/arx/papers/2509.06076.html (last accessed 7 August 2026).

**Artiola JF**, **Pepper IL and Brusseau ML** 2004 Monitoring and characterization of the environment. In *Environmental Monitoring and Characterization*. San Diego, CA: Academic Press, pp. 1–9. https://doi.org/10.1016/B978-012064477-3/50003-5.

**Baller RD**, **Anselin L**, **Messner SF**, **Deane G and Hawkins DF** (2001) Structural covariates of U.S. county homicide rates: Incorporating spatial effects. *Criminology 39*(3), 561–588. https://doi.org/10.1111/j.1745-9125.2001.tb00933.x.

**Barker R**, **Bonell C and Melendez-Torres G** (2026) 'In it from the beginning': How do young men and boys explain their entry into criminal gangs? *Youth Justice 26*(1), 35–57.

**Bauer A**, **Martins RC**, **Hammerton G**, **Gomes H**, **Gonçalves H**, **Menezes AMB**, **Wehrmeister FC and Murray J** (2024) Prevalence and risk factors of gang membership in a Brazilian birth cohort. *JAMA Network Open 7*(10), e2. https://doi.org/10.1001/jamanetworkopen.2024.40393.

**Beer CC** (2006) Judicial performance and the rule of law in the Mexican states. *Latin American Politics and Society 48*(03), 33–61. https://doi.org/10.1111/j.1548-2456.2006.tb00355.x.

**Bellair P** (2017) Social disorganization theory. Available at https://doi.org/10.1093/acrefore/9780190264079.013.253.

**Berti L** (2025) Amazon deforestation spikes as brazil blames criminal fires. Available at https://news.mongabay.com/2025/07/amazon-deforestation-spikes-as-brazil-blames-criminal-fires/ (last accessed 7 August 2026).

**Boekhout van Solinge T** (2010) Deforestation crimes and conflicts in the amazon. *Critical Criminology 18*(4), 263–277. https://doi.org/10.1007/s10612-010-9120-x.

**Boekhout van Solinge T** (2014) Researching illegal logging and deforestation. *International Journal for Crime, Justice and Social Democracy 3*(2), 35–48. https://doi.org/10.5204/ijcjsd.v3i2.179.

**Bragagnolo L**, **da Silva R and Grzybowski J** (2021) Towards the automatic monitoring of deforestation in Brazilian rainforest. *Ecological Informatics 66*, 101454. https://doi.org/10.1016/j.ecoinf.2021.101454.

**Broxton PD**, **Zeng X**, **Sulla-Menashe D and Troch PA** (2014a) A global land cover climatology using modis data. *Journal of Applied Meteorology and Climatology 53*(6), 1593–1605. https://doi.org/10.1175/JAMC-D-13-0270.1.

**Broxton PD**, **Zeng X**, **Scheftic W and Troch PA** (2014b) A modis-based global 1-km maximum green vegetation fraction dataset. *Journal of Applied Meteorology and Climatology 53*(8), 1996–2004. https://doi.org/10.1175/JAMC-D-13-0356.1.

**Carpio-Domínguez JL** (2024) The harms and crimes of logging and deforestation. In Henry N. Pontell, ed. *Oxford Research Encyclopedia of Criminology and Criminal Justice*. Available at https://doi.org/10.1093/acrefore/9780190264079.013.769 (last accessed 7 August 2026).

**Carvalho LS and Soares RR** (2016) Living on the edge: Youth entry, career and exit in drug-selling gangs. *Journal of Economic Behavior & Organization 121*, 77–98. https://doi.org/10.1016/j.jebo.2015.10.018.

**Carvalho WD**, **Mustin K**, **Hilário RR**, **Vasconcelos IM**, **Eilers V and Fearnside PM** (2019) Deforestation control in the Brazilian amazon: A conservation struggle being lost as agreements and regulations are subverted and bypassed. *Perspectives in Ecology and Conservation 17*(3), 122–130. https://doi.org/10.1016/j.pecon.2019.06.002.

**Chen T and Guestrin C** (2016) Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Ser. KDD '16*. New York, NY: Association for Computing Machinery, pp. 785–794. https://doi.org/10.1145/2939672.2939785.

**Chen T**, **T. He**, **M. Benesty**, **V. Khotilovich**, **Y. Tang**, **H. Cho**, **K. Chen**, **R. Mitchell**, **I. Cano**, **T. Zhou**, **M. Li**, **J. Xie**, **M. Lin**, **Y. Geng**, **Y. Li, and J. Yuan**, *xgboost: Extreme Gradient Boosting, r package version 1.7.8.1*, 2024. Available at https://CRAN.R-project.org/package=xgboost.

**Chowdhury MS** (2024) Comparison of accuracy and reliability of random forest, support vector machine, artificial neural network and maximum likelihood method in land use/cover classification of urban setting. *Environmental Challenges 14*, 100800. https://doi.org/10.1016/j.envc.2023.100800.

**Comissão Pastoral da Terra** (2025) *Atlas dos Conflitos no Campo Brasileiro.* Comissão Pastoral da Terra (CPT). Available at https://cptnacional.org.br/2025/07/20/atlas-conflitos-campo-brasileiro/# (last accessed 8 August 2026).

**Cullen P**, **Peden AE**, **Francis KL**, **Cini KI**, **Azzopardi P**, **Möller H**, **Peden M**, **Sawyer SM**, **Nathan S**, **Joshi R**, **Patton GC and Ivers RQ** (2024) Interpersonal violence and gender inequality in adolescents: A systematic analysis of global burden of disease data from 1990 to 2019. *Journal of Adolescent Health 74*(2), 232–245. https://doi.org/10.1016/j.jadohealth.2023.08.044.

**Dal'Asta AP**, **Farias WAA**, **Monteiro AMV**, **Amaral S**, **Gontijo AB**, **Fernandes DA and Escada MIS** (2025) Violence, land conflicts and the dynamics of the agrarian economy in the Brazilian amazon. *Cadernos De Saude Publica 41*(suppl 1). https://doi.org/10.1590/0102-311xen064024.

**Damonte GH**, **Godfrid J**, **López Minchán AP and Alata Quispe E** (2024) The surge for local innovative institutions for transformation: Community-based monitoring in the Andes. *Global Sustainability 7*, e34. https://doi.org/10.1017/sus.2024.26.

**DeLong ER**, **DeLong DM and Clarke-Pearson DL** (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics 44*(3), 837–845. https://doi.org/10.2307/2531595.

**Doshi-Velez F and Kim B** (2017) Towards a rigorous science of interpretable machine learning. Available at https://arxiv.org/abs/1702.08608 (last accessed 7 August 2026).

**Ferrante L**, **Andrade MB and Fearnside PM** (2021) Land grabbing on Brazil's highway br-319 as a spearhead for amazonian deforestation. *Land Use Policy 108*, 105559. https://doi.org/10.1016/j.landusepol.2021.105559.

**Garnelo L**, **Fearnside PM and Ferrante L** (2023) Amazon: Between devastation, violence, and threads of hope. *Cadernos De Saude Publica 39*(12). https://doi.org/10.1590/0102-311xen152723.

**Gil GP** (2022) Os repasses do poder: Coronelismo, gestão pública e política no Amazonas (2002-2022). *Revista Novos Rumos Sociolóogicos 10*(18), 260–290. https://doi.org/10.15210/norus.v10i18.4851.

**Ho TK** (1995) Random decision forests. *Proceedings of 3rd international conference on document analysis and recognition 1*, 278–282.

**Hoerl AE and Kennard RW** (1970) Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics 12*(1), 55–67. 10.1080/00401706.1970.10488634.

**Hou B**, **Lin C**, **Chen M**, **Gouda MM**, **Zhao Y**, **Chen Y**, **Liu F and Feng X** (2025) Integrating UAV-Based RGB imagery with semi-supervised learning for tree species identification in heterogeneous forests. *Remote Sensing 17*(15), 2541. https://doi.org/10.3390/rs17152541.

**Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis (IBAMA)** (2026). Serviços - Consultas - Autuações e embargos. Available at https://www.gov.br/ibama/pt-br/servicos/consultas/autuacoes-e-embargos.

**Ingram MC** (2013) Elections, ideology, or opposition? Assessing competing explanations of judicial spending in the Mexican states. *Journal of Law, Economics, & Organization 29*(1), 178–209. https://doi.org/10.1093/jleo/ewr028.

**Ingram MC** (2014) Community resilience to violence: Local schools, regional economies, and homicide in Mexico's municipalities. In ShirkDA, WoodD and OlsonEL (eds), *Building Resilient Communities in Mexico: Civic Responses to Crime and Violence*. Washington, DC: Woodrow Wilson Center.

**Ingram MC and Marchesini da Costa M** (2019a) Political geography of violence: Municipal politics and homicide in Brazil. *World Development 124*, 104592. https://doi.org/10.1016/j.worlddev.2019.06.016.

**Ingram MC and Marchesini da Costa, M.**, "Replication Data for: Political Geography of Violence: Municipal Politics and Homicide in Brazil," 2019b. https://doi.org/10.7910/DVN/NX5QIU

**Instituto Nacional de Pesquisas Espaciais** (2025a) "Deter – monitoramento diário da supressão e degradação da vegetação nativa – alertas – amazoÌ‚nia/cerrado/pantanal," Instituto Nacional de Pesquisas Espaciais. Coordenação-Geral de Cieì‚Ncias da Terra. Programa de Monitoramento Dos Biomas Brasileiros, Tech. Rep. Available at https://terrabrasilis.dpi.inpe.br/downloads/ (accessed 22 July 2025).

https://doi.org/10.1017/eds.2026.10060 Published online by Cambridge University Press

**Instituto Nacional de Pesquisas Espaciais** (2025b) Prodes – monitoramento anual da supressão da vegetação nativa – amazoÌ‚nia/cerrado/pantanal/caatinga/mata atlaÌ‚ntica/pampa, INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. COORDENAÇÃO-GERAL DE CIEÌ‚NCIAS DA TERRA. PROGRAMA DE MONITORAMENTO DOS BIOMAS BRASILEIROS, Tech. Rep. Available at https://terrabrasilis.dpi.inpe.br/downloads/ (accessed 22 July 2025).

**Jessee E and Skinner M** (2005) A typology of mass grave and mass grave-related sites. *Forensic Science International 152*(1), 55–59. https://doi.org/10.1016/j.forsciint.2005.02.031.

**Jiménez-Muñoz JC**, **Mattar C**, **Sobrino JA and Malhi Y** (2016) Digital thermal monitoring of the amazon forest: An intercomparison of satellite and reanalysis products. *International Journal of Digital Earth 9*(5), 477–498. https://doi.org/10.1080/17538947.2015.1056559.

**Kaczan DJ** (2020) Can roads contribute to forest transitions? *World Development 129*, 104898. https://doi.org/10.1016/j.worlddev.2020.104898.

**Land KC**, **McCall PL and Cohen LE** (1990) Structural covariates of homicide rates: Are there any invariances across time and social space? *American Journal of Sociology 95*(4), 922–963. https://doi.org/10.1086/229381.

**Lappicy T and Lima CHR** (2023) Enhancing monthly streamflow forecasting for Brazilian hydropower plants through climate index integration with stochastic methods. *RBRH 28*, e48. https://doi.org/10.1590/2318-0331.282320230118.

**Leal VN** (1977) *Coronelismo: The Municipality and Representative Government in Brazil, Ser.* Cambridge University Press. https://doi.org/10.1017/CBO9780511759802.

**Liaw A and Wiener M** (2002) Classification and regression by randomForest. *R News 2*(3), 18–22.

**Lima CHR and AghaKouchak A** (2017) Droughts in Amazonia: Spatiotemporal variability, teleconnections, and seasonal predictions. *Water Resources Research 53*(12), 10824–10840. https://doi.org/10.1002/2016WR020086.

**Lima CH**, **Kwon H-H and Kim Y-T** (2021) A bayesian kriging model applied for spatial downscaling of daily rainfall from gcms. *Journal of Hydrology 597*, 126095. https://doi.org/10.1016/j.jhydrol.2021.126095.

**Mendes K** (2024) Revealed: Illegal cattle ranching booms in arariboia territory during deadly year for indigenous guajajara. Available at https://news.mongabay.com/2024/06/revealed-illegal-cattle-boom-in-arariboia-territory-in-deadliest-year-for-indigenous-guajajara/ (last accessed 7 August 2026).

**Mendes K** (2025)'we're getting back on track': Interview with ibama head rodrigo agostinho. Available at https://news.mongabay.com/2025/01/were-getting-back-on-track-interview-with-ibama-head-rodrigo-agostinho/ (last accessed 7 August 2026).

**Mendez C and Sachdeva M** (2025) On the political and socioeconomic geography of violence: Spatial heterogeneity and scale effects in Brazil. *Spatial Economic Analysis 20*(4), 768–780.

**Merton RK** (1938) Social structure and anomie. *American Sociological Review 3*(5), 672–682. https://doi.org/10.2307/2084686.

**Messias CG**, **de Almeida**, **Silva DE**, **Soler LS**, **Maurano LE**, **Camilotti VL**, **Alves FC**, **da Silva LJ**, **Reis MS**, **de Lima TC**, **Renó V**, **Lima DLC**, **Belluzzo AP**, **Quadros CB**, **Barradas DCM**, **de DRV**, **Bastos EFM**, **Cunha IP**, **de JJ**, **de Barros**, **Gusmão LHA**, **de R**, **de Moraes DRV**, **Silva DM**, **Chrispim EHS**, **Pinto JFSKC**, **Neto MRR**, **Matos MHH**, **Moreira NAP**, **Teixeira RC d S**, **Alves GMR**, **de ACS**, **Perez LP**, **Bento BMP**, **de HC**, **dos Santos IS**, **de LCL**, **Matias M**, **da MB**, **Pinheiro F d C**, **Carvalho A**, **Xaud H**, **Xaud M**, **Matos AP**, **Baumann L**, **da Silva EB and Adami LGFM** (2024) Unaccounted for nonforest vegetation loss in the Brazilian amazon. *Communications Earth & Environment 5*(1). https://doi.org/10.1038/s43247-024-01542-0.

**Miller T** (2019) Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence 267*, 1–38. https://doi.org/10.1016/j.artint.2018.07.007.

**Ministério do Meio Ambiente e Mudança do Clima** (2025) Proposta do brasil para combater crimes ambientais é aprovada pela onu. Available at https://www.gov.br/mma/pt-br/noticias/proposta-do-brasil-para-combater-crimes-ambientais-e-aprovada-pela-onu (last accessed 12 May 2026).

**Molnar C** (2025) *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*, 3rd Edn. Munich: Christoph Molnar.

**Muggah R** (2023) How drugs are destroying the amazon. *Foreign Policy.* Available at https://foreignpolicy.com/2023/08/06/amazon-drugs-coca-cocaine-deforestation-environment-biodiversity-climate-change-criminal-brazil-peru-colombia-bolivia-lula-logging/ (last accessed 7 August 2026).

**Muggah R and P. Smith** (2024) New technologies to map environmental crime in the amazon basin (commentary). Available at https://news.mongabay.com/2024/04/new-technologies-to-map-environmental-crime-in-the-amazon-basin-commentary/ (last accessed 7 August 2026).

**Nunes FSM**, **Soares-Filho BS**, **Oliveira AR**, **Veloso LVS**, **Schmitt J**, **Van der Hoff R**, **Assis DC**, **Costa RP**, **Börner J**, **Ribeiro SMC**, **Rajão RGL**, **de Oliveira U and Costa MA** (2024) Lessons from the historical dynamics of environmental law enforcement in the Brazilian amazon. *Scientific Reports 14*(1), 1828. https://doi.org/10.1038/s41598-024-52180-7.

**O'Donnell G** (2001) Democracy, law, and comparative politics. *Studies in Comparative and International Development 36*(1), 7–36. https://doi.org/10.1007/BF02687583.

**Olawade DB**, **Wada OZ**, **Ige AO**, **Egbewole BI**, **Olojo A and Oladapo BI** (2024) Artificial intelligence in environmental monitoring: Advancements, challenges, and future directions. *Hygiene and Environmental Health Advances 12*, 100114. https://doi.org/10.1016/j.heha.2024.100114.

**Oliveira GM d and Miranda BV** (2024) Environmental enforcement, property rights, and violence: Evidence from the Brazilian amazon. *Journal of Institutional Economics 20*, e27. https://doi.org/10.1017/S1744137424000122.

https://doi.org/10.1017/eds.2026.10060 Published online by Cambridge University Press

**Oliveira MC and Siqueira L** (2022) Digitalization between environmental activism and counter-activism: The case of satellite data on deforestation in the Brazilian amazon. *Earth System Governance 12*, 100135. https://doi.org/10.1016/j.esg.2022.100135.

**de Oliveira ES**, **Tuma LSR**, **Santos CAG**, **Mishra M and da Silva RM** (2026) Groundwater contamination risks and land use changes in a typical agreste/caatinga transition zone in northeastern Brazil. *Environmental Monitoring and Assessment 198*(2), 209. https://doi.org/10.1007/s10661-025-14961-z.

**Ouaknine A**, **Kattenborn T**, **Laliberté E and Rolnick D** (2025) Openforest: A data catalog for machine learning in forest monitoring. *Environmental Data Science 4*, e15. https://doi.org/10.1017/eds.2024.53.

**Paluszynska A**, **Biecek P and Jiang Y** (2020) *randomForestExplainer: Explaining and Visualizing Random Forests in Terms of Variable Importance*, r package version 0.10.1. Available at https://github.com/ModelOriented/randomForestExplainer.

**Putnam RD** (1993) *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton, NJ: Princeton University Press.

**Putnam RD** (1995) Bowling alone: America's declining social capital. *Journal of Democracy 6*, 65–78. https://doi.org/10.1353/jod.1995.0002.

**R Core Team**, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2024. Available at https://www.R-project.org/.

**Ramesh V**, **Ouaknine A and Rolnick D** (2025) Tree semantic segmentation from aerial image time series. *Environmental Data Science 4*, e37. https://doi.org/10.1017/eds.2025.10013.

**Rani S**, **Singh S and Purohit S** (2025) Evaluating daytime and nighttime land surface temperature pattern and trends in India: A comparative analysis of satellite and reanalysis data. *Earth Systems and Environment 9*(4), 3479–3500. https://doi.org/10.1007/s41748-024-00555-1.

**Ribeiro V**, **De Souza Mendes F**, **Singh C and Heilmayr R** (2024) Leveraging remote sensing for transparency and accountability in amazonian commodity supply chains. *One Earth 7*(11), 1935–1940. https://doi.org/10.1016/j.oneear.2024.10.011.

**Ríos-Figueroa J** (2007) Fragmentation of power and the emergence of an effective judiciary in Mexico, 1994–2002. *Latin American Politics and Society 49*(1), 31–57. https://doi.org/10.1111/j.1548-2456.2007.tb00373.x.

**Robin X**, **Turck N**, **Hainard A**, **Tiberti N**, **Lisacek F**, **Sanchez J-C and Müller M** (2011) Proc: An open-source package for r and s + to analyze and compare roc curves. *BMC Bioinformatics 12*(1), 77. https://doi.org/10.1186/1471-2105-12-77.

**Roy DP and Kumar SS** (2017) Multi-year modis active fire type classification over the Brazilian tropical moist forest biome. *International Journal of Digital Earth 10*(1), 54–84. https://doi.org/10.1080/17538947.2016.1208686.

**Saad-Diniz E and Gianecchini JV** (2021) Regulatory rollbacks in the amazon rainforest: A nuanced look into the effects of environmental victimization. *State Crime Journal 10*(2), 257–283. https://doi.org/10.13169/statecrime.10.2.0257.

**Sampson RJ and Groves WB** (1989) Community structure and crime: Testing social-disorganization theory. *American Journal of Sociology 94*(4), 774–802. https://doi.org/10.1086/229068.

**Sampson RJ and Wilson WJ** (2012) *Great American City: Chicago and the Enduring Neighborhood Effect*. Chicago: University of Chicago Press. https://doi.org/10.7208/chicago/9780226733883.001.0001.

**Sampson RJ**, **Raudenbush SW and Earls F** (1997) Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science 277*, 918–924. https://doi.org/10.1126/science.277.5328.918.

**Santos MS**, **Mediero L**, **Lima CHR and Moura LZ** (2018) Links between different classes of storm tracks and the flood trends in Spain. *Journal of Hydrology 567*, 71–85. https://doi.org/10.1016/j.jhydrol.2018.10.003.

**Shaw CR and McKay HD** (1942) *Juvenile Delinquency and Urban Areas*. Chicago, IL: University of Chicago Press.

**Silva VCS d**, **Vieira ICG**, **Galbraith D**, **Potapov P**, **M. Rivero SL d**, **Lima AMM d**, **S. Pimentel MA and Adami M** (2022) Marked non-compliance with deforestation embargoes in the Brazilian amazon. *Environmental Research Letters 17*(5), 054033. https://doi.org/10.1088/1748-9326/ac6398.

**da Silva RM**, **Lopes AG and Santos CAG** (2023a) Deforestation and fires in the Brazilian amazon from 2001 to 2020: Impacts on rainfall variability and land surface temperature. *Journal of Environmental Management 326*, 116664. https://doi.org/10.1016/j.jenvman.2022.116664.

**da Silva CFA**, **de Andrade MO**, **dos Santos AM and de Melo SN** (2023b) Road network and deforestation of indigenous lands in the Brazilian amazon. *Transportation Research Part D: Transport and Environment 119*, 103735. https://doi.org/10.1016/j.trd.2023.103735.

**Simon N**, **Friedman J**, **Tibshirani R and Hastie T** (2011) Regularization paths for cox's proportional hazards model via coordinate descent. *Journal of Statistical Software 39*(5), 1–13. https://doi.org/10.18637/jss.v039.i05.

**Skidmore ME**, **Moffette F**, **Rausch L**, **Christie M**, **Munger J and Gibbs HK** (2021) Cattle ranchers and deforestation in the Brazilian amazon: Production, location, and policies. *Global Environmental Change 68*, 102280. https://doi.org/10.1016/j.gloenvcha.2021.102280.

**Socrates P** (2025) GEF IBAMA 2025: Como funciona a gestão de fiscalização ambiental? Available at https://somaconcursos.com.br/gef-ibama-2025-como-funciona-a-gestao-de-fiscalizacao-ambiental/ (last accessed 8 August 2026).

**Suprayogi BS**, **Winardi RD and Maulana AO** (2025) Mapping the landscape of environmental crime: A literature review of emerging research themes and future directions. *Journal of Economic Criminology 8*, 100144. https://doi.org/10.1016/j.jeconc.2025.100144.

**Tay JK**, **Narasimhan B and Hastie T** (2023) Elastic net regularization paths for all generalized linear models. *Journal of Statistical Software 106*(1), 1–31. https://doi.org/10.18637/jss.v106.i01.

**Thompson ST and Magrath WB** (2021) Preventing illegal logging. *Forest Policy and Economics 128*, 102479. https://doi.org/10.1016/j.forpol.2021.102479.

https://doi.org/10.1017/eds.2026.10060 Published online by Cambridge University Press

**Tibshirani R** (1996) Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology 58*(1), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x.

**Vancine MH**, **Muylaert RL**, **Niebuhr BB**, **de Faria Oshima JE**, **Tonetti V**, **Bernardo R**, **De Angelo C**, **Rosa MR**, **Grohmann CH and Ribeiro MC** (2024) The Atlantic forest of south america: Spatiotemporal dynamics of the vegetation and implications for conservation. *Biological Conservation 291*, 110499. https://doi.org/10.1016/j.biocon.2024.110499.

**Vigh H and Sausdal D** (2025) Para-crime: Notes towards a 'grey criminology'. *The British Journal of Criminology*, azaf024. https://doi.org/10.1093/bjc/azaf024.

**Ye Z**, **Yang J**, **Zhong N**, **Tu X**, **Jia J and Wang J** (2020) Tackling environmental challenges in pollution controls using artificial intelligence: A review. *Science of the Total Environment 699*, 134279. https://doi.org/10.1016/j.scitotenv.2019.134279.

**Zhang Y**, **Zhu J**, **Zhu J and Wang X** (2023) A splicing approach to best subset of groups selection. *INFORMS Journal on Computing 35*(1), 104–119. https://doi.org/10.1287/ijoc.2022.1241.

**Zhou Z**, **Qiu C and Zhang Y** (2023) A comparative analysis of linear regression, neural networks and random forest regression for predicting air ozone employing soft sensor models. *Scientific Reports 13*(1), 22420. https://doi.org/10.1038/s41598-023-49899-0.

**Zhu J**, **Wen C**, **Zhu J**, **Zhang H and Wang X** (2020) A polynomial algorithm for best-subset selection problem. *Proceedings of the National Academy of Sciences of the United States of America 117*(52), 33117–33123. https://doi.org/10.1073/pnas.2014241117.

**Zhu J**, **Wang X**, **Hu L**, **Huang J**, **Jiang K**, **Zhang Y**, **Lin S and Zhu J** (2022) Abess: A fast best-subset selection library in python and r. *Journal of Machine Learning Research 23*(202), 1–7.

**Zou H and Hastie T** (2005) Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology 67*(2), 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x.

**Cite this article:** Ingram MC (2026). A machine learning approach to detecting environmental crimes in Brazil. *Environmental Data Science*, 5: e28. doi:10.1017/eds.2026.10060

https://doi.org/10.1017/eds.2026.10060 Published online by Cambridge University Press