Conservation Letters

WILEY

Conservation Letters
A journal of the Society for Conservation Biology
Open Access

LETTER OPEN ACCESS

Practical and Ethical Issues in Big Data and Machine Learning Forecasts of Zambian Community Forestry Engagement

Thomas Pienkowski[1,2] | Morena Mills[1] | Matt Clark[1] | Kaala Moombe[3] | Henry Chilufya[4] | Alexandros Sfyridis[1] | Jocelyne Shimin Sze[5] | Erik Olsson[6] | Andreas Christ Sølvsten Jørgensen[7,8]

[1]Centre for Environmental Policy, Imperial College London, London, UK | [2]Durrell Institute of Conservation and Ecology, University of Kent, Canterbury, UK | [3]The Center for International Forestry Research, Lusaka, Zambia | [4]Independent, Lusaka, Zambia | [5]Institute of Environmental Science and Technology, Universitat Autònoma de Barcelona, Barcelona, Spain | [6]Betty and Gordon Moore Center for Science, Conservation International, Arlington, Virginia, USA | [7]Department of Mathematics, Imperial College London, London, UK | [8]I-X Centre for AI In Science, Imperial College London, London, UK

Correspondence: Thomas Pienkowski (t.pienkowski@kent.ac.uk)

Received: 15 May 2025 | **Revised:** 11 December 2025 | **Accepted:** 2 January 2026

Keywords: area-based conservation | artificial intelligence | community-based conservation | community forestry | forecasting | machine learning | predictive conservations science | scaling | socioecological systems

ABSTRACT

Approaches integrating geospatial "big data" and machine learning will likely be increasingly used to predict conservation-related human behavior, such as patterns of local engagement, in socioecological systems. Yet, few studies evaluate both the technical and ethical aspects of such applications. Here, we provide a nation-scale worked example that combines machine learning and publicly available data to predict spatial patterns of Community Forestry establishment among 539,221 settlements across Zambia. Our model accurately predicted out-of-sample spatial establishment patterns three-quarters of the time (balanced accuracy = 76.5%, sensitivity = 64.0%, specificity = 89.1%), though it had a high false positive rate (precision = 24.3%). Accurately forecasting conservation establishment patterns for effective resource allocation requires better data on local preferences and programmatic decision-making, among other factors. Furthermore, such artificial intelligence applications risk making decision-making more technocratic, top-down, and opaque; therefore, they should only inform deliberation over possible future scenarios within wider, multistakeholder governance processes.

1 | Introduction

Although the success of community-based conservation has been mixed (Blaikie 2006; Dressler et al. 2010; Brooks et al. 2013; Galvin et al. 2018), more equitable and devolved conservation governance systems tend to be associated with better social and ecological outcomes (Schreckenberg et al. 2016; Dawson et al. 2024; Clark et al. 2025a). Scaling out well-designed community-based conservation models may help achieve conservation goals (e.g., Target 3 of the Kunming–Montreal Global Biodiversity Framework) in more socially equitable and beneficial ways (Dudley et al. 2018; UN 2022). Here, scaling out means expanding an initiative to reach more people or locations (see *Supporting Information: Defining scaling*) (Moore et al. 2015).

Growing research explores the characteristics that facilitate adoption and scaling in conservation and restoration efforts (e.g., Abernethy et al. 2014; Battista et al. 2017; Romero-de-Diego et al. 2021; Clark, Andrews and Hillis 2022; Clark et al. 2024b; Jørgensen et al. 2024; Pienkowski et al. 2024b, 2025; Tavares et al. 2024; Jagadish et al. 2024; Hounkpati et al. 2024; Mills et al. 2025). Much of this research has been retrospective (but see Clark et al.

2024b and Jørgensen et al. 2024 for exceptions), but predictive conservation science approaches can help forecast changing socioecological systems and with designing, implementing, and evaluating initiatives in ways that anticipate and adapt to those changes (Travers et al. 2019).

A promising application of predictive science is forecasting geographic locations where conservation efforts are likely to be established. If sufficiently accurate, these approaches could enable forecasting of initiative emergence under plausible future scenarios, helping anticipate their context-dependent social and ecological impacts. For example, landowners in areas with more roads, higher populations, and greater wealth were more likely to engage in restoration of the Brazilian Atlantic Forest (Pienkowski et al. 2024b). Forecasting engagement patterns could, therefore, help predict the additionality of future conservation and restoration efforts.

Existing attempts at this (e.g., Jørgensen et al. 2024, who used an agent-based modeling approach) have relied on data collected through intensive fieldwork and collaboration with local practitioners, which provides valuable insights but can be prohibitively costly and time-consuming. In contrast, there is a growing abundance of geospatial "big data" on social, economic, and ecological features (often derived from remote sensing) that might be associated with conservation establishment. Moreover, data-led modeling approaches designed specifically for prediction may help forecast spatially explicit establishment patterns and explore underlying mechanisms driving involvement. In particular, nonparametric, nonstatistical machine learning algorithms (henceforth "machine learning methods," see *Supporting Information: Defining machine learning*) often outperform inferential methods in predictive accuracy (van Cranenburgh et al. 2022; Pichler & Hartig 2023). Machine learning methods are increasingly popular in ecological research (Tuia et al. 2022; Pichler & Hartig 2023; Han et al. 2023). However, their application to understanding patterns in the social dimensions of natural resource management remains comparatively uncommon (but see Rocha et al. 2020, Epstein et al. 2021, and Reynolds et al. 2025).

Processes shaping conservation establishment are inherently complex and context-dependent, influenced by local relationships, politics, power dynamics, and chance occurrences (Sterling et al. 2017). These complexities, embedded in adaptive socioecological systems, may pose fundamental barriers to predicting patterns of conservation establishment (Hofman et al. 2017; Preiser et al. 2018). Furthermore, the potential ethical consequences of using novel data sources and artificial intelligence have been extensively discussed in other fields (e.g., Hanna et al. 2025) and are increasingly addressed within conservation (e.g., Sandbrook et al. 2018, 2021; Wearn et al. 2019, 2019; Adams 2019; Pritchard et al. 2022; Sandbrook 2025). However, few studies explicitly integrate ethical considerations when combining big data (e.g., from remote sensing [York et al. 2023]) and machine learning to examine social dimensions of natural resource management, especially in the context of scaling. Here, we explore the use of machine learning and publicly available data to predict spatial patterns of Community Forestry establishment in Zambia (see *Community Forestry in Zambia*). We aim to assess the technical performance of this predictive approach, including its suitability for forecasting future establishment through a hypothetical illustrative scenario. We also aim to evaluate the associated ethical risks and uncertainties inherent to predictive approaches.

2 | Methods

2.1 | Community Forestry in Zambia

Our study focused on Community Forestry in Zambia, a model designed to enhance local involvement in managing forests under customary authority or within Local and National Forests (Government of Zambia 2018) (see *Supporting Information: Legal foundations of Community Forestry*). Each Community Forest has been formalized through agreements between a Community Forest Management Group (CFMG) and the Forestry Department. These agreements transfer authority to control access, use, and management of designated forest areas to CFMGs.

As of August 2024, the CFMG Database—a national repository of registered CFMGs—recorded 146 CFMGs, of which 103 signed Community Forestry Agreements with the government between October 2017 and November 2023 (Government of Zambia & USAID 2024). The CFMG Database is, to our knowledge, among the most comprehensive geospatial datasets on community-based conservation in Africa (though anecdotal evidence suggests significantly more CFMGs have been reported since 2024), with Community Forestry adopted at scale in Zambia, making it an illustrative case study.

Several key stakeholders are typically involved in the Community Forest establishment process (see *Supporting Information: Community Forestry establishment process*), including the Forestry Department, traditional authorities, community members, and civil society and private sector actors (Government of Zambia 2018). Interactions between these groups, partly shaped by power dynamics, are expected to influence where Community Forests are established (e.g., Siangulube et al. 2023). Moreover, the Community Forestry establishment process is often top-down, led by civil society, private sector, and government actors who typically approach communities in areas of greater conservation value (see *Supporting Information: Community Forestry establishment process*). Technical and bureaucratic requirements can deter bottom-up establishment, though some communities do initiate the process themselves by seeking external support.

2.2 | Variable Description

The units of analysis were rural settlement boundaries identified in the Geo-Referenced Infrastructure and Demographic Data for Development (GRID3) database (version 2.0, CIESIN 2023). The original GRID3 dataset included 635,170 settlements. A series of inclusion criteria were used to define the final sample size of 404,005, excluding 231,165 settlements mostly far from forests (see *Supporting Information: Settlement exclusion*).

The response variable was the presence of a Community Forest within 2215 m of settlement boundaries (see *Supporting Information: Response variable* for a justification of this distance). This response variable was a coarse indicator of involvement in Community Forest establishment. All Community Forests registered within the CFMG Database were included in the analysis, representing areas that had reached at least the second (i.e., negotiation and signing of maps) of the seven broad steps involved in establishing Community Forests (see *Supporting Information: Community Forestry establishment process*) (Government of Zambia 2018).

Drawing on Diffusion of Innovations theory, we searched for publicly available geospatial datasets of features with theoretically plausible associations with the response variable (Table 1, see *Discussion* and *Supporting Information: Theoretical framework*). These data were aggregated at the scale of individual settlements (Table 1). We also expected that landscape and regional-scale socioecological characteristics might influence establishment patterns. Therefore, we took the mean of each feature across all settlements within (1) a hexagonal grid with 115 km x 115 km cells (see *Gradient boosting analysis* and *Supporting Information: Spatial block cross-validation*) and (2) administrative districts (of which there were 116 of highly variable sizes) and included these as additional features (see *Supporting Information: Excluded features*).

3 | Gradient Boosting Analysis

Our analysis utilized eXtreme Gradient Boosting (XGBoost) through the "xgboost" package (version 1.7.8.1, Chen et al. 2024) in R (version 4.4.2, R Core Team 2020) (see *Supporting Information: eXtreme Gradient Boosting*). We adjusted for class imbalance (see *Supporting Information: Class imbalance*), and all features were scaled to units of 1 standard deviation and centered at 0. We partitioned the data into training and test sets using spatial blocking (*Supporting Information: Spatial block cross-validation*). Approximately 70% of the blocks (with all associated observations) were randomly selected for the training set, and the remaining blocks (and their observations) were allocated to the test set. Within the training set, we used seven-fold block cross-validation and a grid search approach to tune the XGBoost hyperparameters (see *Supporting Information: Spatial block cross-validation*). We then tested the model's predictive performance using the remaining test dataset (i.e., all observations within the remaining ∼30% of spatial blocks). We illustrated feature importance and accumulated local effects plots (and their associated uncertainty obtained via 1000 bootstrap iterations, sampling spatial blocks of observations with replacement), which are informative for interpreting machine-learning outputs (Molnar 2019) (see *Supporting Information: Interpretable machine learning*).

A hallmark of predictive conservation science approaches is that they attempt to make forecasts about the state or outcomes of future socioecological systems (Travers et al. 2019). We constructed a hypothetical scenario where cattle densities increased at double the historical rate (from 2000 to 2022) among settlements with the lowest current cattle densities, projected to 2030 (see *Supporting Information: Scenario analysis*). This could occur if, for example, a new control program were introduced that effectively eradicates populations of Tsetse flies (trypanosome vectors that cause human and animal African trypanosomiasis [Muyobela et al. 2023]).

4 | Results

Our total sample included 539,221 settlements (Table 2). Of these, 20,675 (5.1%) were considered involved in Community Forest establishment, and 383,330 (94.9%) were considered not involved (but see *Variable description* for why this is a coarse indicator).

For our test dataset, our gradient-boosting analysis yielded a balanced accuracy of 76.5%, a sensitivity of 64.0%, and a specificity of 89.1%. In other words, our model accurately predicted whether settlements were considered involved in establishing Community Forests (or not) over three-quarters of the time. It correctly identified involved settlements in around two-thirds of cases and was highly accurate in identifying settlements that were considered not involved. However, it also had a precision of only 24.3%, suggesting that only around one in four of the settlements predicted to be involved actually were, indicating a high false positive rate (Figure 1). Settlements predicted as involved were characterized by higher district-level drought risk (at district level), closer proximity to protected areas (local, grid, district), lower cattle densities (local, grid), lower road densities (local), lower distances to hotels (local, though the association is complex), lower women's literacy (district), and higher topographic ruggedness (district) and megafauna densities (district) (Figure 2).

In the baseline scenario, predicted settlement involvement with Community Forest establishment (55,447) was approximately 2.7 times the actual observations. Under the increased cattle density scenario, allowing abandonment of Community Forests by settlements modestly reduced the forecasted involvement to 51,252, primarily in Northwestern Province (Figure 1). Without abandonment, the number of forecasted settlements involved was the same as in the baseline (i.e., 55,457).

5 | Discussion

5.1 | Forecasts Under Changing Socioecological Conditions

Our model's overprediction and poor differentiation of fine-scale patterns may reflect the omission of features known to be important in influencing establishment (see *Supporting Information: Interpreting performance metrics*) or big dataset limitations (see *Supporting Information: Feature dataset limitations*). Multiple studies suggest that external support and facilitation can play key roles in driving conservation engagement (e.g., Romero-de-Diego et al. 2021; Mills et al. 2025; Pienkowski et al. 2025). Within our study area, the World Bank and Global Environment Facility have funded projects targeted at specific provinces, so external interests likely drive Community Forest establishment. Similarly, prior research shows that local engagement is partly determined by perceptions of an initiative's relative benefits and costs (e.g., shaped by resource dependence [Gatiso 2019]), equity, and flexibility and alignment with local needs (e.g., Romero-de-Diego et al. 2021; Pienkowski et al. 2024b, 2025; Jagadish et al. 2024; Mills et al. 2025). Other potentially important omitted features include the strength of local leadership and governance systems (e.g., Gutiérrez et al. 2011), secure resource and tenure rights (Joglekar et al. 2025), and social cohesion, cooperation,

TABLE 1 | The features initially identified for inclusion in the statistical analysis drawing on the Diffusion of Innovations theory (see *Supporting Information: Theoretical framework*).

Feature	Plausible justification	Data aggregation notes	Source
Forest loss (2000–2015, %)	(+) Areas with greater recent historic forest cover loss are likely to be prioritized by the Forest Department, carbon credit partners, and nongovernmental organizations.	Percentage of 1179 m[a] buffer of settlements that experienced forest loss	Hansen et al. 2013
Distance to the nearest airport (km)	(?) Expected to relate to socioeconomic characteristics that might influence the relative costs and benefits of involvement with Community Forestry. For example, airport proximity might be associated with opportunity costs from tourism and other activities.	Euclidean distance from settlement centroid	OpenStreetMap 2023
Distance to the nearest hotel (km)	(?) See “Distance to the nearest airport (km)”	Euclidean distance from settlement centroid	OpenStreetMap 2023
Distance from the nearest protected area (km)	(−) Promotion of Community Forestry by government, civil society, and private sector actors may be higher within and around protected areas (especially Game Management Areas).	Euclidean distance from settlement centroid	UNEP-WCMC 2025
Drought severity score	(?) Settlements in areas with greater drought severity may prioritize non-crop-related income sources, thus being more likely to be involved. Alternatively, drought risk might motivate extractive forest use as a coping strategy, reducing incentives to be involved in sustainable Community Forestry.	Nearest Standardized Precipitation-Evapotranspiration Index score (between 1981 and 2016 at 5 km^2) to settlement centroids	Peng et al. 2020
Food insecurity score	(−) Settlements in areas with lower food security might have less capacity to invest in involvement.	Mean food stress score within 1179 m[a] buffer of settlements (frequency of passing food stress threshold, calculated at ~250 m^2 resolution, based on 2009–2019 data)	CIESIN 2020
Improved water sources (%)	(?) The same rationale as above for the wealth score, assuming that improved water sources are a proxy for wealth.	Mean percentage of households with improved water sources within 1179 m[a] buffer of settlements (derived from data at 5 km^2 resolution)	USAID 2024
Mean cattle density (per km^2)	(−) Areas with higher cattle densities may have coupled social (e.g., greater pastoralism and access to cattle markets) and ecological (e.g., lower Tsetse fly populations) conditions misaligned with Community Forestry.	Mean cattle density within 1179 m[a] buffer of settlements (density is the mean count of sheep per km^2 in 2010)	Gilbert et al. 2018
Mean elevation (m)	(+) Areas of higher elevation tend to experience less anthropogenic risk, thus potentially being more suitable for Community Forestry.	Mean elevation (30 m^2 resolution) within 1179 m[a] buffer of settlements	JAXA 2022

(Continues)

TABLE 1 | (Continued)

Feature	Plausible justification	Data aggregation notes	Source
Mean goat density (per km^2)	(−) See "Mean cattle density (per km^2)"	Mean goat density within 1179 m[a] buffer of settlements (density is the mean count of sheep per km^2 in 2010)	Gilbert et al. 2018
Mean sheep density (per km^2)	(−) See "Mean cattle density (per km^2)"	Mean sheep density within 1179 m[a] buffer of settlements (density is the mean count of sheep per km^2 in 2010)	Gilbert et al. 2018
Percentage of cropland cover (%)	(?) Areas with greater cropland cover may have coupled social (e.g., greater dependence on crops to meet needs) and ecological (e.g., less intact forests) conditions misaligned with Community Forestry.	Percentage cover within 1179 m[a] buffer of settlements	Karra et al. 2021
Percentage of forest cover (%)	(+) Areas with greater forest cover may offer greater opportunities for extractive use and income from carbon credits. Thus, settlements in these areas might be more likely to be involved.	Percentage cover within 1179 m[a] buffer of settlements	Karra et al. 2021
Population density (per 100 m^2)	(?) Potential mechanisms are unclear, but population density can influence socioecological context in multiple ways (e.g., potential land competition), so it should be included.	Mean population density within 1179 m[a] buffer of settlements (number of people per 100 m^2 in 2015)	Carioli et al. 2023
Road density score	(−) Settlements in poorly connected areas may have fewer alternative income sources and lower opportunity costs, making them more likely to be involved.	Mean road density within 1179 m[a] buffer of settlements (density is the total length of roads per km^2, within a 10-km search radius around 250 m^2 pixels)	OpenStreetMap 2023
Settlement area (km^2)	(?) Settlement size may interact with other features to influence the probability of involvement. For example, small settlements with disproportionately high forest cover loss might be prioritized by the Forest Department, carbon credit partners, and nongovernmental organizations.	Settlement area in km^2	CIESIN 2023
Density of African megafauna species ranges (count per 250 m^2), see *Supporting Information: African megafauna species*	(+) Promotion of Community Forestry by government, civil society, and private sector actors may be higher in areas where megafauna are more prevalent.	Mean within 1179 m[a] buffer of count (250 m^2 resolution)	IUCN 2025
Topographic ruggedness (SD)	(+) More topographically rugged areas tend to experience fewer anthropogenic risks, thus potentially being more suitable for Community Forestry.	Standard deviation of elevation (30 m^2 resolution) within 1179 m[a] buffer of settlements	JAXA 2022

(Continues)

TABLE 1 | (Continued)

Feature	Plausible justification	Data aggregation notes	Source
Travel time to population centers (min)	(−) Settlements in these areas may have fewer alternative income sources and lower opportunity costs, making them more likely to be involved in Community Forestry.	Mean travel time to population centers (with <50,000 inhabitants) within 1179 m[a] buffer of settlements (~1 km^2 resolution)	Weiss et al. 2018
Wealth score	(?) Wealthier settlements might have a greater capacity—thus being more likely—to invest in establishing Community Forests than poorer ones. Alternatively, poorer settlements might have fewer alternative income sources and lower opportunity costs, making them more likely to engage.	Mean estimated wealth index score within 1179 m[a] buffer of settlements (derived from data at 1.6 km^2 resolution)	Lee 2022
Women's literacy (%)	(+) Settlements in areas where there are higher literacy rates among women might have more technical capacity to engage.	Mean percentage of women who are literate within 1179 m[a] buffer of settlements (derived from data at 5 km^2 resolution)	USAID 2024

Key: (+) = expected positive association; (−) expected negative association; (?) unclear expected direction of association.
Abbreviation: SD, standard deviation.
[a]This value (1179 m) represents a plausible distance over which residents may travel when interacting with their local social and ecological environment. Bosina and Weidmann (2017) conducted an extensive review of studies on walking speeds worldwide, finding that walking speeds generally ranged between 1.0 and 1.6 m/s, with little variability between continents. In one of the few studies of its kind, Cook (2022) evaluated walking patterns among women in South Africa, finding that 75% of steps were accumulated in bouts >15 min. Therefore, using an average walking speed of 1.31 m/s, residents may walk approximately 1179 m from settlements during a typical 15-min walking bout. This rough approximation was confirmed as reasonable during discussions with experts.

and participation (Auer et al. 2020) (see *Supporting Information: Theoretical framework*). For instance, settlements with local champions and well-established decision-making structures may have been more likely to engage. Although features in our model might be partial proxies for these factors (e.g., populations in areas with less crop land cover and cattle densities might be more dependent on forest resources), none directly capture them. Consequently, our modeling approach might have yielded more reliable, fine-scale predictions if data on a wider diversity of relevant factors were available. Collecting such data at scale might be costly, indicating trade-offs between predictive performance and study feasibility. This could be overcome through nested approaches, where coarse, large-scale patterns are detected using inexpensive geospatial data, followed by intensive fieldwork to differentiate fine-scale patterns.

Additionally, we were unable to definitively determine which settlements were directly engaged with Community Forestry, instead having to use a coarse, distance-based rule to approximate involvement (see *Variable description*). Thus, our precision estimates may be contested. Generally, individual or community engagement with initiatives is often an ambiguous, dynamic process, which can fluctuate in intensity over time (Montes de Oca Munguia et al. 2021). Forecasting these dynamics is feasible, but requires corresponding data on the nature, degree, and variability of engagement.

Furthermore, much of the promise of predictive conservation science comes from its potential to forecast outcomes under future socioecological conditions (Travers et al. 2019). Yet, the reliability and utility of this scenario analysis approach rests on several core considerations (see *Supporting Information: Scenario analysis reliability*). Crucially, it is well established that many socioecological systems are complex, adaptive, and historically dependent (Preiser et al. 2018), which may make some outcomes inherently unpredictable (Hofman et al. 2017). For example, cattle densities are driven by and subsequently affect numerous coupled socioeconomic (e.g., cattle market distances), cultural (e.g., pastoralist identities and traditions), historical (e.g., colonial veterinary fences to prevent wildlife-livestock transmission and land allocations for white settlers [Cumming et al. 2015]), and ecological factors (e.g., forest habitat suitability), which likely collectively influence establishment (Clark et al. 2025b). Developing plausible scenarios that center on features endogenously coupled with others—that is, where changes in one part of a system have strong, adaptive, nonlinear cascading effects on others—is extremely challenging. However, other features might be comparatively more exogenous, allowing for more credible scenarios. For example, some programmatic decisions—such as international investment levels or the spatial allocation of resources by external organizations—may be comparatively less closely coupled with other features influencing establishment. Even in these cases, developing plausible scenarios requires sophisticated domain knowledge generated through, for instance, qualitative (e.g., Brittain et al. 2022; Pienkowski et al. 2022; Booth et al. 2023), participatory (e.g., Oteros-Rozas et al. 2015), or expert-oriented (e.g., Jørgensen et al. 2024) exercises.

Although our analysis uses cross-sectional data, the availability of time-stamped engagement and feature data would allow

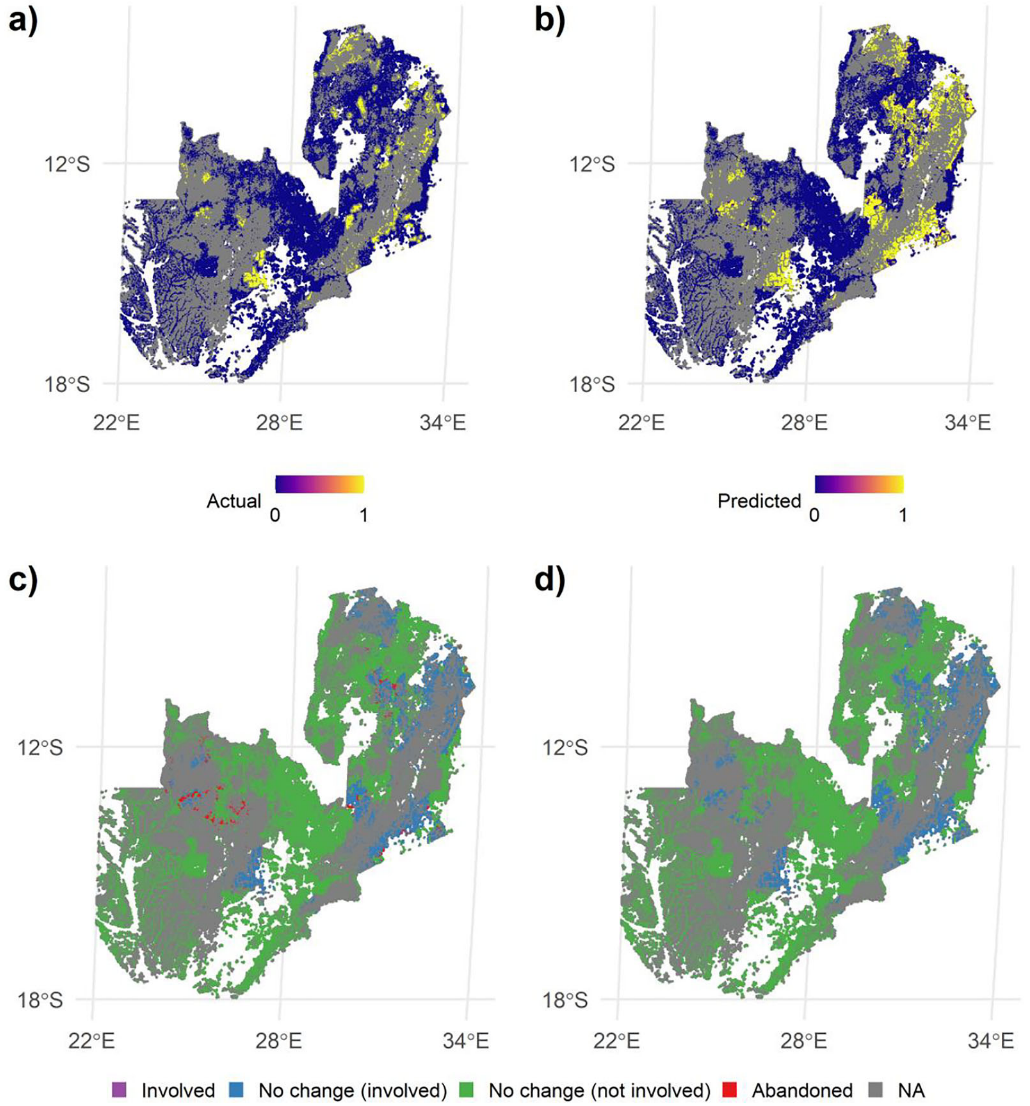

FIGURE 1 | Spatial patterns of settlement involvement with Community Forest establishment in Zambia (mean aggregated at 2.5 km^2 resolution), showing (a) actual and (b) predicted, and forecast scenarios under increased cattle density (c) with abandonment (where settlements currently classified as "involved" may change to "not involved" if predicted to do so) and (d) without abandonment (where currently "involved" settlements retain that status regardless of predictions). Predictions were generated from the full dataset.

for spatio-temporal forecasting. Such forecasting would enable simultaneous prediction of where and over what time horizons adoption might occur under alternative scenarios, as illustrated by Jørgensen et al. (2024).

5.2 | Ethical Dimensions of Scaling Prediction

Recent technological and infrastructural advances have catalyzed enthusiasm and funding for artificial intelligence, machine learning, and big data in conservation (Reynolds et al. 2025). For example, the Bezos Earth Fund recently launched a multiyear USD $100 million AI for Climate and Nature Grand Challenge, while Google's AI for Social Good and Microsoft's AI for Earth aim to utilize private-sector machine learning resources for conservation (Joppa 2017; Wearn et al. 2019). To date, these techniques have been mostly applied to wildlife monitoring (Tuia et al. 2022; Pichler & Hartig 2023; Han et al. 2023), alongside growing interest in their potential for evidence synthesis and communication (Tyler et al. 2023; Berger-Tal et al. 2024). Correspondingly, much of the emerging scholarship on ethical and social justice risks focuses on wildlife monitoring and related applications (e.g., Sandbrook et al. 2018, 2021; Wearn et al. 2019; Adams 2019; York et al. 2023), with important exceptions (e.g., Sandbrook 2025).

Beyond predictive performance, omitting data on local preferences means overlooking which communities might actively want or oppose Community Forestry (see *Supporting Information: Data justice framework*). Examples from elsewhere show that pressures to scale can combine with state control, rent-seeking, and territoriality to drive the spread of conservation, even in the face of local opposition (Benjaminsen et al. 2013; Bluwstein & Lund 2018; Sungusia et al. 2020; Pienkowski et al. 2024a). Failing to center local preferences and priorities within this modeling may shift influence away from local people and reinforce top-

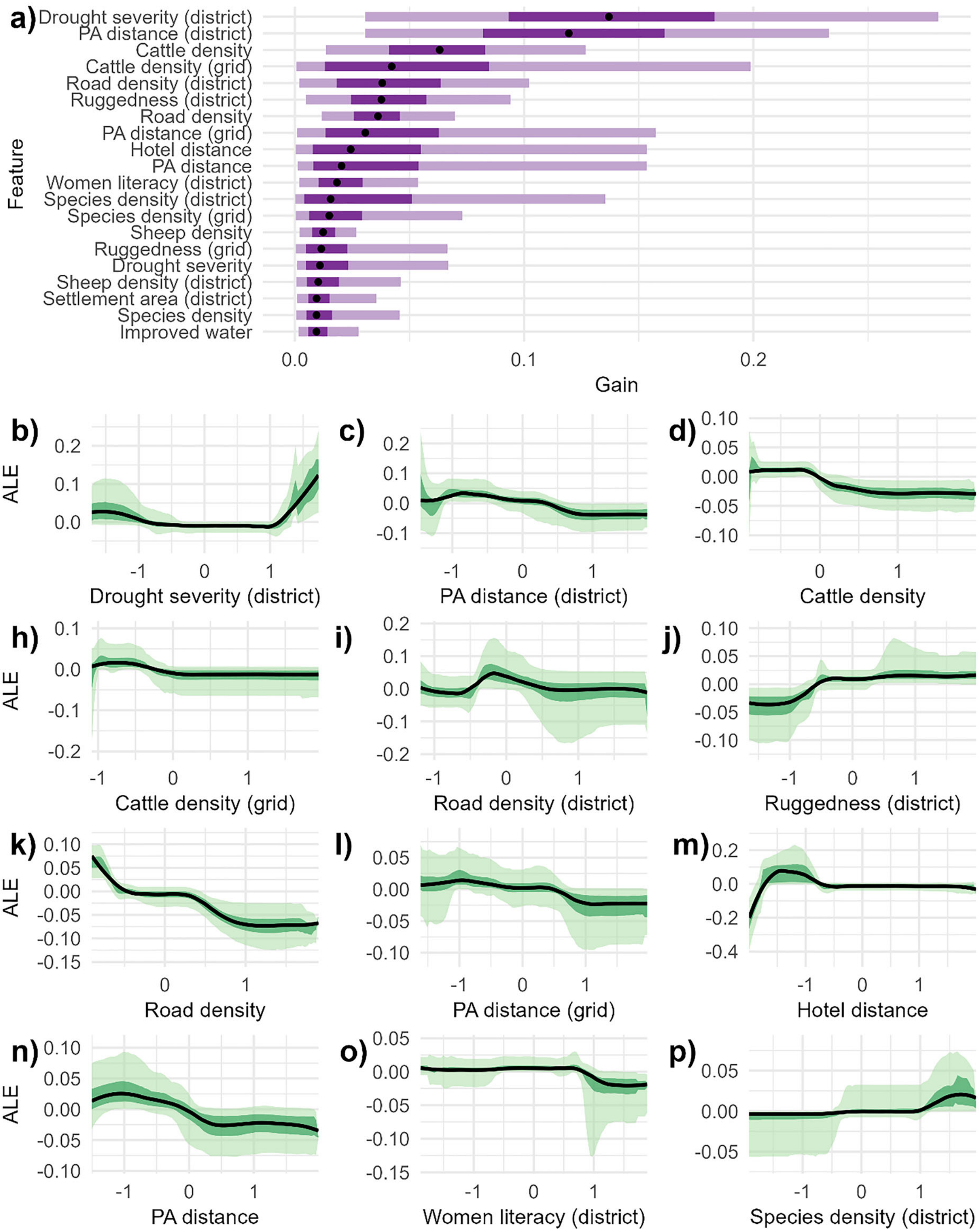

FIGURE 2 | Bootstrap estimated (B = 1000) (a) feature importance scores and (b–p) accumulated local effects (ALE) plots for features from an eXtreme Gradient Boosting model trained to predict settlement involvement in Community Forestry in Zambia. (a) Higher-ranked features contribute more to the model's predictive performance. Median score (points), central 50% interval of scores (dark purple), and central 95% interval of scores (light purple) are shown for the 20 most important features. (b–p) ALE values (y-axis, centered at zero, zero = baseline) show the isolated, average effect of a feature value (x-axis) on the predicted probability of involvement: upward slopes indicate that higher feature values increase that probability, flat regions indicate no change, and downward slopes indicate reduced probability. Loess-smoothed (span = 0.3) median of ALE curves (black line), central 50% interval of ALE curves (dark green), and central 95% interval of ALE curves (light green) are shown for the top 12 features. PA, protected area.

TABLE 2 | Descriptive statistics illustrating differences in features between settlements considered involved and not involved in Community Forest establishment.

Variable	Not involved in establishment (N = 383,330)	Involved in establishment (N = 20,675)
Forest loss (2000–2015, %)	6.1 (6.8)	4.7 (4.4)
Distance to the nearest hotel (km)	609 (248)	501 (250)
Distance to the nearest protected area (km)	94 (61)	44 (33)
Drought severity score	−0.57 (0.77)	−0.76 (1.11)
Food insecurity score	0.46 (1.11)	0.23 (0.67)
Improved water sources (%)	51 (25)	61 (26)
Mean cattle density (per km^2)	4.9 (5.3)	2.3 (2.1)
Mean elevation (m)	1183 (176)	1088 (245)
Mean goat density (per km^2)	5.4 (6.3)	4.1 (3.7)
Mean sheep density (per km^2)	0.27 (0.60)	0.50 (0.84)
Percentage of cropland cover (%)	6 (15)	6 (13)
Population density (per 100 m^2)	0.22 (0.45)	0.20 (0.38)
Road density score	0.02 (1.02)	−0.44 (0.39)
Settlement area (km^2)	0.024 (0.077)	0.025 (0.091)
Density of African megafauna species ranges (count per 250 m^2)	11.96 (1.95)	13.63 (1.95)
Topographic ruggedness (SD in m)	12 (11)	15 (15)

(Continues)

TABLE 2 | (Continued)

Variable	Not involved in establishment (N = 383,330)	Involved in establishment (N = 20,675)
Travel time to population centers (min)	119 (94)	146 (91)
Wealth score	18 (7)	17 (5)
Women's literacy	53 (11)	50 (12)

Note: Values in brackets are in standard deviations.

down decision-making processes. This risk is compounded by the fact that the methods employed here are highly specialized, resulting in an asymmetric understanding that might conflict with data access (Pritchard et al. 2022) and model interpretability (Yousefzadeh & Cao 2022; Cao & Yousefzadeh 2023) principles. Gathering rich contextual data at scale is challenging and costly, and privacy concerns may deter local communities from disclosing their views. These challenges are common across centrally coordinated conservation planning, where ensuring meaningful local participation (Adams et al. 2019; Lécuyer et al. 2024) and appropriately integrating social well-being data (Stephanson & Mascia 2014) are cross-cutting issues.

Furthermore, the absence of data on programmatic decisions could lead to misguided choices based on modeling results. Rather than approximating factors influencing local establishment decisions, our results may instead reflect where conservation actors have focused their efforts so far. This likely explains why proximity to protected areas emerged as an important feature as conservation actors promoted Community Forests to buffer these areas. Consequently, decisions derived from this analysis may merely mirror historical conservation decisions rather than indicate where "organic" engagement is most likely, undermining the decentralization of natural resource governance. This is loosely analogous to machine learning models trained on biased crime data that produce predictive policing approaches, perpetuating discriminatory law enforcement (Hung & Yen 2023). Valid concerns have been raised about the emergence of novel big data sources in conservation (e.g., Pritchard et al. 2022; York et al. 2023). Yet, we believe more comprehensive data is needed on the actions of conservation organizations for transparency, collaboration, and research, especially actions carried out on public land and with public funding (Pienkowski et al. 2024a).

Overall, our "worked example" illustrates the potential benefits and limitations of forecasting spatial patterns of conservation establishment. These approaches risk making conservation decision-making more technocratic, top-down, and opaque, potentially countering the devolution of governance and its associated social and ecological advantages (Schreckenberg et al. 2016; Dawson et al. 2024). Many of these concerns are generalizable to other attempts to predict social dimensions of natural resource management when using big data and quantitative modeling approaches. Yet, such forecasts might be more robust and ethically defensible if they integrate data on programmatic decision-making, local preferences, and the strength of governance institutions, among other factors. Moreover, rather than being treated as reliable predictions, the approaches used here might instead be integrated into broader multistakeholder landscape governance processes, where these maps are discussed and critiqued but do not drive decision-making. Communities in regions where engagement appears most likely might be invited by conservation actors to partner on Community Forestry. The need for meaningful local participation in conservation planning is widely recognized (Adams et al. 2019; Lécuyer et al. 2024), as is the value of using scenarios to help coproduce conservation science, policy, and practice (Bennett et al. 2016; Pereira et al. 2020; Cork et al. 2023). In this context, combining machine learning and big data might most usefully support multistakeholder processes exploring potential futures of conservation-related human behavior.

Author Contributions

Conceptualization, TP and ACSJ; Methodology, TP, ACSJ, MM, MC, and AS; Investigation, TP, ACSJ, and AS; Writing – original draft, TP, ACSJ, MM, and MC; Writing – review and editing, TP, MM, MC, KM, HC, AS, JSS, EO, and ACSJ; Funding acquisition, TP, ACSJ, and MM.

Acknowledgments

TP, ACSJ, and AS thank the Turner Kirk Trust (TKT24LFEPR); TP, MC, AJ, and MM thank the Leverhulme Trust for the research grant (RPG-2021-440); and TP thanks the Research England's Expanding Excellence in England Fund, UK Research and Innovation. JSS acknowledges funding by the European Union (ERC, CONDJUST, 101054259). Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. This work contributes to the ICTA-UAB "María de Maeztu" Programme for Units of Excellence of the Spanish Ministry of Science and Innovation (CEX2024-001506-M/funded by MICIU/AEI/ 10.13039/501100011033). We thank Dr Matt Sommerville for valuable comments on a draft of the article. This is contribution #21 from the Insights for Catalyzing Conservation at Scale initiative.

Conflicts of Interest

KM is an employee of the Center for International Forestry Research (a nonprofit scientific research organization), which conducts participatory action research on integrated landscape approaches that include Community Forest Management themes and practices.

Ethics Approval Statement

Our study used all publicly available secondary datasets, except for the Community Forest boundary data from the Government of Zambia. Therefore, no ethical approval was sought for this study (but see *Ethical dimensions of predictive conservation science*, where we discuss ethical considerations).

Data Availability Statement

Code and anonymized data are available at: https://doi.org/10.5281/zenodo.17201264.

References

Abernethy, K. E., Ö. Bodin, P. Olsson, Z. Hilly, and A. Schwarz. 2014. "Two Steps Forward, Two Steps Back: the Role of Innovation in Transforming Towards Community-Based Marine Resource Management in Solomon Islands." *Global Environmental Change* 28: 309–321.

Adams, V. M., M. Mills, R. Weeks, et al. 2019. "Implementation Strategies for Systematic Conservation Planning." *Ambio* 48: 139–152.

Adams, W. M 2019. "Geographies of Conservation II: Technology, Surveillance and Conservation by Algorithm." *Progress in Human Geography* 43: 337–350.

Auer, A., J. Von Below, L. Nahuelhual, et al. 2020. "The Role of Social Capital and Collective Actions in Natural Capital Conservation and Management." *Environmental Science & Policy* 107: 168–178.

Battista, W., A. Tourgee, C. Wu, and R. Fujita. 2017. "How to Achieve Conservation Outcomes at Scale: an Evaluation of Scaling Principles." *Frontiers in Marine Science* 3: 1–21.

Benjaminsen, T. A., M. J. Goldman, M. Y. Minwary, and F. P. Maganga. 2013. "Wildlife Management in Tanzania: State Control, Rent Seeking and Community Resistance." *Development and Change* 44: 1087–1109.

Bennett, E. M., M. Solan, R. Biggs, et al. 2016. "Bright Spots: Seeds of a Good Anthropocene." *Frontiers in Ecology and the Environment* 14: 441–448.

Berger-Tal, O., B. B. M. Wong, C. A. Adams, et al. 2024. "Leveraging AI to Improve Evidence Synthesis in Conservation." *Trends in Ecology & Evolution* 39: 548–557.

Blaikie, P. 2006. ""Is Small Really Beautiful?" Community-Based Natural Resource Management in Malawi and Botswana." *World Development* 34: 1942–1957.

Bluwstein, J., and J. F. Lund. 2018. "Territoriality by Conservation in the Selous–Niassa Corridor in Tanzania." *World Development* 101: 453–465.

Booth, H., M. S. Ramdlan, A. Hafizh, et al. 2023. "Designing Locally-Appropriate Conservation Incentives for Small-Scale Fishers." *Biological Conservation* 277: 109821.

Bosina, E., and U. Weidmann. 2017. "Estimating Pedestrian Speed Using Aggregated Literature Data." *Physica A: Statistical Mechanics and its Applications* 468: 1–29.

Brittain, S., C. T. Kamogne Tagne, F. Booker, et al. 2022. "Using Scenarios-Based Interviews to Predict Changes in Wild Meat Hunting and Consumption in Response to Different Alternatives Projects." *Conservation Science and Practice* 4: e12777.

Brooks, J., K. A. Waylen, and M. B. Mulder. 2013. "Assessing Community-Based Conservation Projects: a Systematic Review and Multilevel Analysis of Attitudinal, Behavioral, Ecological, and Economic Outcomes." *Environmental Evidence* 2: 1–34.

Cao, X., and R. Yousefzadeh. 2023. "Extrapolation and AI Transparency: Why Machine Learning Models Should Reveal When They Make Decisions Beyond Their Training." *Big Data & Society* 10: 1–5.

Carioli, A., M. Schiavina, S. Freire, and K. MacManus. 2023. GHS-POP R2023A—GHS Population Grid Multitemporal (1975–2030). European Commission, Joint Research Centre (JRC). http://data.europa.eu/89h/2ff68a52-5b5b-4a22-8f40-c41da8332cfe(accessed January 23, 2025).

Center for International Earth Science Information Network (CIESIN). 2020. Food Insecurity Hotspots Data Set.

Center for International Earth Science Information Network (CIESIN). 2023. GRID3 Data Hub. Flow Minder Foundation, United Nations Population Fund (UNFPA), WorldPop.

Chen, T., T. He, M. Benesty, et al. 2024. xgboost: Extreme Gradient Boosting. https://cran.r-project.org/web/packages/xgboost/index.html (accessed February 11, 2025).

Clark, M., A. Sykulski, S. Liu, et al. 2025a. Community Management and Wellbeing Increase the Resilience of Nepali Forests to Climate Change. Research Square. https://www.researchsquare.com/article/rs-6413705/v1 (accessed April 16, 2025).

Clark, M., J. Andrews, and V. Hillis. 2022. "A Quantitative Application of Diffusion of Innovations for Modeling the Spread of Conservation Behaviors." *Ecological Modelling* 473: 110145.

Clark, M., I. T. Gallizioli, O. Crowe, et al. 2025b. "Community Wellbeing Mediates Drought Adaptation in South African Rangelands". Journal of Environmental Management 393: 127066.

Clark, M., T. Pienkowski, A. Jagadish, et al. 2024. "Forecasting Adoption With Epidemiological Models Can Enable Adaptively Scaling Out Conservation." *One Earth* 7: 1820–1832.

Cook, I. 2022. "Objectively-Measured Step Cadence and Walking Patterns in a Rural African Setting: a Cross-Sectional Analysis." *BMC Research Notes* 15: 155.

Cork, S., C. Alexandra, J. G. Alvarez-Romero, et al. 2023. "Exploring Alternative Futures in the Anthropocene." *Annual Review of Environment and Resources* 48: 25–54.

Cumming, D. H. M., S. A. Osofsky, S. J. Atkinson, and M. W. Atkinson. 2015. Beyond Fences: Wildlife, Livestock and Land Use in Southern Africa. One Health: The Theory and Practice of Integrated Health Approaches. Available from. https://www.cabidigitallibrary.org/doi/10.1079/9781780643410.0243 (accessed April 16, 2025).

Dawson, N. M., B. Coolsaet, A. Bhardwaj, et al. 2024. "Is It Just Conservation? A Typology of Indigenous Peoples' and Local Communities' Roles in Conserving Biodiversity." *One Earth* 7: 1007–1021.

Dressler, W., B. Büscher, M. Schoon, et al. 2010. "From Hope to Crisis and Back Again? A Critical History of the Global CBNRM Narrative." *Environmental Conservation* 37: 5–15.

Dudley, N., H. Jonas, F. Nelson, et al. 2018. "The Essential Role of Other Effective Area-Based Conservation Measures in Achieving Big Bold Conservation Targets." *Global Ecology and Conservation* 15: e00424.

Epstein, G., G. Gurney, S. Chawla, et al. 2021. "Drivers of Compliance Monitoring in Forest Commons." *Nature Sustainability* 4: 450–456.

Galvin, K. A., T. A. Beeton, and M. W. Luizza. 2018. "African Community-Based Conservation: a Systematic Review of Social and Ecological Outcomes." *Ecology and Society* 23: 39.

Gatiso, T. T 2019. "Households' Dependence on Community Forest and Their Contribution to Participatory Forest Management: Evidence From Rural Ethiopia." *Environment, Development and Sustainability* 21: 181–197.

Gilbert, M., G. Nicolas, G. Cinardi, et al. 2018. "Global Distribution Data for Cattle, Buffaloes, Horses, Sheep, Goats, Pigs, Chickens and Ducks in 2010." *Scientific Data* 5: 180227.

Government of Zambia. 2018. National Guidelines for Community Forestry in Zambia.

Government of Zambia, USAID. 2024. CFMG Database. Available from. https://cfmg.mgee.gov.zm/cfmg/map (accessed August 1, 2024).

Gutiérrez, N. L., R. Hilborn, and O. Defeo. 2011. "Leadership, Social Capital and Incentives Promote Successful Fisheries." *Nature* 470: 386–389.

Han, B. A., K. R. Varshney, S. LaDeau, A. Subramaniam, K. C. Weathers, and J. Zwart. 2023. "A Synergistic Future for AI and Ecology." *Proceedings of the National Academy of Sciences* 120: e2220283120.

Hanna, M. G., L. Pantanowitz, B. Jackson, et al. 2025. "Ethical and Bias Considerations in Artificial Intelligence/Machine Learning." *Modern Pathology* 38: 100686.

Hansen, M. C., P. V. Potapov, R. Moore, et al. 2013. "High-Resolution Global Maps of 21st-Century Forest Cover Change." *Science* 342: 850–853.

Hofman, J. M., A. Sharma, and D. J. Watts. 2017. "Prediction and Explanation in Social Systems." *Science* 355: 486–488.

Hounkpati, K., H. Moluh Njoya, K. Adjonou, K. Kokou, S. Sieber, and K. Löhr. 2024. "Drivers Affecting Adoption of Forest Landscape Restoration Among Smallholder Farmers in Central Togo." *Frontiers in Sustainable Food Systems* 8: 1425797. https://www.frontiersin.org/journals/sustainable-food-systems/articles/10.3389/fsufs.2024.1425797/full.

Hung, T.-W., and C.-P. Yen. 2023. "Predictive Policing and Algorithmic Fairness." *Synthese* 201: 206.

International Union for Conservation of Nature (IUCN). 2025. IUCN Red List of Threatened Species.

Jagadish, A., A. Freni-Sterrantino, Y. He et al. 2024. "Scaling Indigenous-Led Natural Resource Management." *Global Environmental Change* 84: 102799.

Japan Aerospace Exploration Agency (JAXA). 2022. ALOS Global Digital Surface Model "ALOS World 3D - 30m (AW3D30)".

Joglekar, A., A. Jagadish, C. Jardim, et al. 2025. "Landholders' Engagement in Restoring Brazil's Atlantic Forest Is Linked With Livelihood Compatibility and Legal Compliance." *International Forestry Review* 27: 1–13.

Joppa, L. N 2017. "The Case for Technology Investments in the Environment." *Nature* 552: 325–328.

Jørgensen, A. C. S., T. Pienkowski, M. Clark, et al. 2024. "Forecasting the Adoption and Spread of a Community-Based Marine Management Initiative Using Agent-Based Models." *BioRxiv* 599026.

Karra, K., C. Kontgis, Z. Statman-Weil, J. C. Mazzariello, M. Mathis, and S. P. Brumby. 2021. "Global Land Use/Land Cover With Sentinel 2 and Deep Learning." 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. https://ieeexplore.ieee.org/document/9553499 (accessed April 17, 2025).

Lécuyer, L., E. Balian, J. R. A. Butler, et al. 2024. "The Importance of Understanding the Multiple Dimensions of Power in Stakeholder Participation for Effective Biodiversity Conservation." *People and Nature* 6: 1407–1420.

Lee, K. 2022. High-Resolution Poverty Maps in Sub-Saharan Africa. Harvard Dataverse.

Mills, M., M. V. Touchon, E. Denis, et al. 2025. "Scaling Out Community Conservation Initiatives: Experts Identify Economic and Social Benefits, Compatibility With Needs, and External Support as Key." *Conservation Letters* 18: e13100.

Molnar, C. 2019. Interpretable Machine Learning: A Guide for Making Black Box Models Explainable. https://christophm.github.io/interpretable-ml-book/(accessed February 11, 2025).

Montes de Oca Munguia, O., D. J. Pannell, R. Llewellyn, and P. Stahlmann-Brown. 2021. "Adoption Pathway Analysis: Representing the Dynamics and Diversity of Adoption for Agricultural Practices." *Agricultural Systems* 191: 103173.

Moore, M. L., D. Riddell, and D. Vocisano. 2015. "Scaling Out, Scaling Up, Scaling Deep: Strategies of Non-Profits in Advancing Systemic Social Innovation." *Journal of Corporate Citizenship* 58: 67–84.

Muyobela, J., C. W. W. Pirk, A. A. Yusuf, and C. L. Sole. 2023. "Spatial Distribution of *Glossina morsitans* (Diptera: Glossinidae) in Zambia: a Vehicle-Mounted Sticky Trap Survey and Maxent Species Distribution Model." *PLOS Neglected Tropical Diseases* 17: e0011512.

OpenStreetMap. 2023. OpenStreetMap. https://www.openstreetmap.org/export#map=5/51.50/-0.10 (accessed January 8, 2025).

Oteros-Rozas, E., B. Martín-López, T. M. Daw, et al. 2015. "Participatory Scenario Planning in Place-Based Social-Ecological Research: Insights and Experiences From 23 Case Studies." *Ecology and Society* 20: 1–66. https://www.jstor.org/stable/26270296.

Peng, J., S. Dadson, F. Hirpa, et al. 2020. "A Pan-African High-Resolution Drought Index Dataset." *Earth System Science Data* 12: 753–769.

Pereira, L. M., K. K. Davies, E. den Belder, et al. 2020. "Developing Multiscale and Integrative Nature–People Scenarios Using the Nature Futures Framework." *People and Nature* 2: 1172–1195.

Pichler, M., and F. Hartig. 2023. "Machine Learning and Deep Learning—A Review for Ecologists." *Methods in Ecology and Evolution* 14: 994–1016.

Pienkowski, T., M. Clark, A. Jagadish, et al. 2024a. "Five Lessons for Avoiding Failure When Scaling in Conservation." *Nature Ecology & Evolution* 8: 1804–1814.

Pienkowski, T., M. Clark, A. Jagadish, et al. 2025. "Diagnosing Scaling Bottlenecks in 10 Community Conservation Initiatives in Southern and Eastern Africa." *Conservation Biology* e70149.

Pienkowski, T., A. Freni Sterrantino, A. M. Tedesco, et al. 2024b. "Spatial Predictors of Landowners' Engagement in the Restoration of the Brazilian Atlantic Forest." *People and Nature* 7: 247–264. https://onlinelibrary.wiley.com/doi/abs/10.1002/pan3.10765.

Pienkowski, T., A. Keane, E. Kinyanda, C. Asiimwe, and E. J. Milner-Gulland. 2022. "Predicting the Impacts of Land Management for Sustainable Development on Depression Risk in a Ugandan Case Study." *Scientific Reports* 12: 11607.

Preiser, R., R. Biggs, A. De Vos, and C. Folke. 2018. "Social-Ecological Systems as Complex Adaptive Systems: Organizing Principles for Advancing Research Methods and Approaches." *Ecology and Society* 23: 1–15. https://www.jstor.org/stable/26796889.

Pritchard, R., L. A. Sauls, J. A. Oldekop, W. A. Kiwango, and D. Brockington. 2022. "Data Justice and Biodiversity Conservation." *Conservation Biology* 36: e13919.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.

Reynolds, S. A., S. Beery, N. Burgess, et al. 2025. "The Potential for AI to Revolutionize Conservation: a Horizon Scan." *Trends in Ecology & Evolution* 40: 191–207.

Rocha, J., K. Malmborg, L. Gordon, K. Brauman, and F. DeClerck. 2020. "Mapping Social-Ecological Systems Archetypes." *Environmental Research Letters* 15: 034017.

Romero-de-Diego, C., A. Dean, A. Jagadish, B. Witt, M. B. Mascia, and M. Mills. 2021. "Drivers of Adoption and Spread of Wildlife Management Initiatives in Mexico." *Conservation Science and Practice* 3: e438.

Sandbrook, C. 2025. "Beyond the Hype: Navigating the Conservation Implications of Artificial Intelligence." *Conservation Letters* 18: e13076.

Sandbrook, C., D. Clark, T. Toivonen, et al. 2021. "Principles for the Socially Responsible Use of Conservation Monitoring Technology and Data." *Conservation Science and Practice* 3: e374.

Sandbrook, C., R. Luque-Lora, and W. M. Adams. 2018. "Human Bycatch: Conservation Surveillance and the Social Implications of Camera Traps." *Conservation and Society* 16: 493–504.

Schreckenberg, K., P. Franks, A. Martin, and B. Lang. 2016. "Unpacking Equity for Protected Area Conservation." *PARKS* 22: 11–28.

Siangulube, F. S., R.-T. MAF, J. Reed, H. Djoudi, D. Gumbo, and T. Sunderland. 2023. "Navigating Power Imbalances in Landscape Governance: a Network and Influence Analysis in Southern Zambia." *Regional Environmental Change* 23: 41.

Stephanson, S. L., and M. B. Mascia. 2014. "Putting People on the Map Through an Approach That Integrates Social Data in Conservation Planning." *Conservation Biology* 28: 1236–1248.

Sterling, E. J., E. Betley, A. Sigouin, et al. 2017. "Assessing the Evidence for Stakeholder Engagement in Biodiversity Conservation." *Biological Conservation* 209: 159–171.

Sungusia, E., J. F. Lund, and Y. Ngaga. 2020. "Decolonizing Forestry: Overcoming the Symbolic Violence of Forestry Education in Tanzania." *Critical African Studies* 12: 354–371.

Tavares, M. F., P. Gallo, N. Nascimento, J. Bauhus, P. H. S. Brancalion, and M. Feurer. 2024. "Smallholders' Perspectives, Motivations, and Incentives for Restoring the Brazilian Atlantic Forest." *Restoration Ecology* 32: e14270.

Travers, H., M. Selinske, A. Nuno, et al. 2019. "A Manifesto for Predictive Conservation." *Biological Conservation* 237: 12–18.

Tuia, D., B. Kellenberger, S. Beery, et al. 2022. "Perspectives in Machine Learning for Wildlife Conservation." *Nature Communications* 13: 792.

Tyler, C., K. L. Akerlof, A. Allegra, et al. 2023. "AI Tools as Science Policy Advisers? The Potential and the Pitfalls." *Nature* 622: 27–30.

United Nations Environment World Conservation Monitoring Centre (UNEP-WCMC). International Union for Conservation of Nature (IUCN). 2025. Protected Planet: The World Database on Protected Areas (WDPA).

United Nations (UN). 2022. Kunming-Montreal Global Biodiversity Framework.

U.S. Agency for International Development (USAID). 2024. Spatial Data Repository—Modeled Surfaces. https://spatialdata.dhsprogram.com/modeled-surfaces (accessed January 23, 2025).

van Cranenburgh, S., S. Wang, A. Vij, F. Pereira, and J. Walker. 2022. "Choice Modelling in the Age of Machine Learning—Discussion Paper." *Journal of Choice Modelling* 42: 100340.

Wearn, O. R., R. Freeman, and D. M. P. Jacoby. 2019. "Responsible AI for Conservation." *Nature Machine Intelligence* 1: 72–73.

Weiss, D. J., A. Nelson, H. S. Gibson et al. 2018. "A Global Map of Travel Time to Cities to Assess Inequalities in Accessibility in 2015." *Nature* 553: 333–336.

York, N. D. L., R. Pritchard, L. A. Sauls, C. Enns, and T. Foster. 2023. "Justice and Ethics in Conservation Remote Sensing: Current Discourses and Research Needs." *Biological Conservation* 287: 110319.

Yousefzadeh, R., and X. Cao. 2022. To What Extent Should We Trust AI Models When They Extrapolate? arXiv. http://arxiv.org/abs/2201.11260 (accessed April 1, 2025).

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

Supplementary information: conl70022-sup-0001-SuppMat.docx