International Journal of the Commons

The Problem of Institutional Fit: Uncovering Patterns with Boosted Decision Trees

RESEARCH ARTICLE

ubiquity press

GRAHAM EPSTEIN
CRISTINA I. APETREI
JACOPO BAGGIO
SIVEE CHAWLA
GRAEME CUMMING
GEORGINA GURNEY
TIFFANY MORRISON
HITA UNNIKRISHNAN
SERGIO VILLAMAYOR TOMAS

*Author affiliations can be found in the back matter of this article

ABSTRACT

Complex social-ecological contexts play an important role in shaping the types of institutions that groups use to manage resources, and the effectiveness of those institutions in achieving social and environmental objectives. However, despite widespread acknowledgment that "context matters", progress in generalising how complex contexts shape institutions and outcomes has been slow. This is partly because large numbers of potentially influential variables and non-linearities confound traditional statistical methods. Here we use boosted decision trees – one of a growing portfolio of machine learning tools – to examine relationships between contexts, institutions, and their performance. More specifically we draw upon data from the International Forest Resources and Institutions (IFRI) program to analyze (i) the contexts in which groups successfully self-organize to develop rules for the use of forest resources (local rulemaking), and (ii) the contexts in which local rulemaking is associated with successful ecological outcomes. The results reveal an unfortunate divergence between the contexts in which local rulemaking tends to be found and the contexts in which it contributes to successful outcomes. These findings and our overall approach present a potentially fruitful opportunity to further advance theories of institutional fit and inform the development of policies and practices tailored to different contexts and desired outcomes.

CORRESPONDING AUTHOR:
Graham Epstein
School of Environment,
Resources and Sustainability,
University of Waterloo, CA
graham.epstein@uwaterloo.ca

KEYWORDS:
Institutional fit; Collective action; Context; Community-based management; Environmental governance; Machine learning

TO CITE THIS ARTICLE:
Epstein, G., Apetrei, C. I., Baggio, J., Chawla, S., Cumming, G., Gurney, G., Morrison, T., Unnikrishnan, H., & Tomas, S. V. (2024). The Problem of Institutional Fit: Uncovering Patterns with Boosted Decision Trees. *International Journal of the Commons*, 18(1), pp. 1–16. DOI: https://doi.org/10.5334/ijc.1226

INTRODUCTION

The concept of contingency is a defining feature of the Bloomington School of institutional analysis and related strands of research in environmental governance and social-ecological systems (SESs) (Ostrom and Ostrom 1999, Young 2002, Sproule-Jones 2005, Acheson 2006, Ostrom et al. 2007). Indeed, there is a growing consensus that institutions and the outcomes they generate are heavily influenced by the complex social-ecological contexts in which they are found (Galaz et al. 2008, Young 2008, Cox 2012, Epstein et al. 2015) and that there are no institutional panaceas for addressing environmental problems (Brock and Carpenter 2007, Ostrom et al. 2007, Morrison 2017). However, despite many years of research and growing interest in middle-range theories and archetypes (Magliocca et al. 2018, Meyfroidt et al. 2018, Oberlack et al. 2019), progress in understanding the contexts in which different institutions are likely to be found and contribute to successful outcomes has been slow (Cumming et al. 2020).

Institutions, which are the formal and informal rules, norms and conventions that structure the incentives, opportunities, and constraints that actors face when making decisions (Crawford and Ostrom 1995), play a critical role in influencing how people interact with their environment and each other. However, institutions are not random features of the environment, but rather vary systematically across a gradient of social, and ecological characteristics (Leslie et al. 2015, Cumming and Epstein 2020). Protected areas, for example, tend to be situated far from major population centres (Joppa and Pfaff 2009) where there is less intense competition with alternative uses such as agriculture and resource extraction and associated political conflicts (Devillers et al. 2015). Communities exploiting mobile resources such as fish, wildlife, and water, meanwhile, tend to avoid quotas in favour of timing, location or technology rules owing to the relatively high costs associated with collecting and processing information about those resources (Schlager et al. 1994, Acheson and Wilson 1996, Cifdaloz et al. 2010). These and other patterns emerge over time as groups experiment with different institutions and retain those that deliver upon the social and ecological objectives of decision-makers. Indeed, research on the topic of institutional fit asserts that outcomes depend greatly upon the match between institutions and the contexts in which they are used (Cumming et al. 2006, Galaz et al. 2008, Young 2008, Epstein et al. 2015).

The literature on institutional fit has developed over the last few decades in the face of mounting evidence of the failure of theoretically optimal solutions to address environmental problems. This literature, which examines the fit between institutions and the context in which they are embedded (Epstein et al. 2015), offers a cautionary tale about the dangers of undermining local and traditional governance systems based on theories or models (Johannes 1978, Nalau et al. 2018) and transplanting institutions that work well in one context to others (Ban et al. 2009). On the other hand, it offers practical guidance about how institutions can be designed or adapted to different contextual conditions. Ultimately, the goal is to develop theory about the socio-ecological context (i.e., conditions) under which different types of institutions are most effective in steering behavior and solving environmental problems. For example, several studies have suggested that participatory environmental governance is more likely to contribute to successful outcomes when communities have strong local leaders and high levels of social capital (Singleton and Taylor 1992, Gutierrez et al. 2011; Frey 2017). Resources that cross jurisdictional boundaries (i.e., transboundary rivers, migratory species) and those that are closely connected through biophysical linkages (i.e., wildlife/habitat, predator/prey), meanwhile, tend to benefit from institutional mechanisms that facilitate coordination (Bodin 2017, Garrick 2018). These mechanisms can range from informal governance networks and arrangements that have been developed and adapted by communities over generations, (Bodin and Tengö 2012, Bodin et al. 2014) to large international organizations, such as Regional Fisheries Management Organizations and River Basin Organizations, designed to facilitate coordination of policies and sharing of information and resources (Rayfuse 2015; Milman and Gerlak 2020).

Recent research on the topic of institutional fit has used qualitative comparative analysis to better understand the efficacy of Ostrom's (1990) design principles across different social-ecological contexts and outcomes. Vallury et al. (2022), for instance, found that the design principles were more likely to contribute to successful irrigation outcomes in South India when communities were small and the reservoirs they used were small in relation to the size of the population they served. Whittaker et al. (2021), on the other hand, uncovered patterns between several design principles and outcomes from lakes in Wisconsin to generate several useful insights for theory and practice. These include highlighting the importance of graduated sanctioning in contexts that are more ecologically conducive to the establishment of an invasive species, and where the density of buildings on the surrounding shoreline is higher. Although these and other insights from the literature on institutional fit have proved useful for policy and practice, the literature has struggled to move beyond a few core principles and advance a more general understanding of how complex social-ecological contexts affect the performance of institutions.

A generalised understanding of the relationship between contexts, institutions and their performance has been hindered by several factors. First, there are relatively few large-n social-ecological datasets that consistently capture variation in the complex social-ecological contexts in which institutions are found and outcomes are realized. Second, there is a general tendency among institutional scholars to view large-n datasets as a source of information for hypothesis testing, rather than an opportunity for applying exploratory methods for hypothesis development. To be clear, we acknowledge the importance of developing large-n datasets for hypothesis testing, but also believe that there are certain types of datasets or moments in the life cycle of those datasets where they hold greater value with respect to hypothesis development. Third and finally, most conventional quantitative research methods are best suited to address questions about the average "effects" of institutions within a population of cases and tend to struggle with problems involving large numbers of potentially influential and oftentimes correlated variables, interactions and non-linear responses (Agrawal 2003, Rana and Miller 2018). Fortunately, recent advances in machine learning have contributed to the development of a growing portfolio of tools that can help to respond to some of these challenges (Frey 2020a). These allow researchers to sift through large numbers of contextual variables to identify those that are more or less important for predicting outcomes and developing insights about how institutions and their performance vary across a gradient of social and ecological conditions.

This paper applies one such machine learning tool – boosted decision trees – to move beyond high-level claims that "context matters" and to develop insights about the specific contexts in which institutions are likely to be found and contribute to successful outcomes. It does so by drawing upon data collected as part of the International Forest Resources and Institutions (IFRI) program (Huntington et al. 2016, Ostrom et al. 2016) to examine the social and ecological contexts in which forest user groups successfully self-organize to develop and maintain rules for the use of forest resources (local rulemaking), and the contexts in which local rulemaking is associated with better forest conditions.

In what follows, we provide a brief review of research on the social, ecological and institutional factors that influence prospects for collective action and sustainable community-based natural resource management (CBNRM). We then describe the data and methods used in this study and summarize the results of two boosted decision tree models examining the conditions in which local rulemaking is likely to be found and contribute to better forest conditions. We find noteworthy differences between the contexts that appear conducive to local rulemaking and those in which local rulemaking is likely to contribute to successful ecological outcomes. Finally, the paper concludes by discussing the implications of these results for future research on CBNRM, and opportunities to advance contextually explicit theories of environmental sustainability.

COMMUNITY-BASED MANAGEMENT, CONTEXT AND INSTITUTIONAL FIT

The CBNRM literature developed quickly following publication of Elinor Ostrom's (1990) *Governing the Commons*, which demonstrated that despite the fatalistic predictions of the 'tragedy of the commons', communities were capable of self-organizing to sustainably govern the use of common-pool resources. Further research built upon these findings to examine the conditions that facilitated sustainable CBNRM. Initially, group characteristics such as size and levels of social and economic heterogeneity received considerable attention as scholars examined the validity of theoretical predictions that voluntary collective action would be limited to small, relatively homogenous, groups (Olson 1965, Fearon and Laitin 1996). Empirical research has clearly refuted this hypothesis, suggesting instead that the effects of group characteristics vary across different public goods (Ostrom and Ostrom 1999, Ostrom 2003), and are mediated by factors such as leadership and social capital (Singleton and Taylor 1992, Glowacki and von Rueden 2015), government policies and property rights regimes, (Ostrom 2003, Coleman and Steed 2009, Coleman and Fleischman 2012) and markets (Meinzen-Dick et al. 2002, Araral Jr 2009).

The relationship between group size, collective action and sustainable natural resource management has been a subject of considerable debate over the last fifty years. In general, standard economic theories suggest that increases in group size are likely to undermine prospects for collective action through a range of mechanisms that collectively contribute to higher costs and lower levels of trust (Olson 1965). Empirical research, on the other hand, has revealed a more complex relationship. The likelihood of social monitoring in forests, for instance, appears to rise with increases in group size (Agrawal and Yadama 1997, Epstein et al. 2021), while contributions to third-party monitoring appear to be highest when user groups are between 40 and 100 households (Agrawal and Goyal 2001). Social, human, and economic capital can also influence prospects for collective action. Social capital and local leadership generally enable collective action by providing a foundation of trust and mechanisms for resolving disputes and securing agreements (Henrich et al. 2001, Agrawal 2003, Gibson et al. 2005, Gutierrez et al.

2011, Henrich et al. 2015, Epstein et al. 2021). Similarly, higher levels of education and other types of training may provide groups with the knowledge, skills and capacity they need to initiate and sustain collective action (Meinzen-Dick et al. 2002). By contrast, the effects of economic wealth and assets on collective action appear to be somewhat mixed. While high levels of poverty may create challenges for initiating and sustaining collective action by reducing the capacity of groups to defer flows of resource benefits and invest time and resources in management (Acheson 2006), it may also generate strong incentives for communities to work together to address shared problems and ensure the long-term sustainability of the resources upon which they depend (Poteete and Ostrom 2004, Andersson and Gabrielsson 2012). Endowments such as land, meanwhile, may increase the capacity of actors to make contributions to local governance and defer benefits from natural resources (Maskey et al. 2006), but may also erode incentives to invest in managing those resources (Doss and Meinzen-Dick 2015).

Variability in the social, cultural and economic characteristics of groups may also influence prospects for collective action. While conventional wisdom holds that group homogeneity is conducive to collective action (Fearon and Laitin 1996), theoretical models and empirical research have revealed a far more complex relationship. First, groups can differ along a number of different dimensions including their identities (i.e., race, ethnicity, religion, gender), values and interests (i.e., consumptive vs. non-consumptive use of resource), wealth, assets and the rights they hold with respect to the use and management of resources (Vedeld 2000). Second, measures of group heterogeneity are often highly correlated with each other and with group size which can make it difficult to isolate their respective effects. Third, mixed empirical evidence suggests that the effects of group heterogeneity are likely mediated by institutions and other factors (Poteete and Ostrom 2004). For example, while some suggest that economic heterogeneity might enable collective action when wealthy group members choose to bear all or most of the costs of public goods from which they will also benefit (Hardin 1982, Baland and Platteau 1999, Adhikari and Lovett 2006), others have found that inequality may undermine collective action by contributing to conflicts and mistrust (Varughese and Ostrom 2001, Bardhan and Dayton-Johnson 2002, Cardenas 2003, Janssen et al. 2011, Baggio et al. 2015).

Although research has tended to give priority to the impacts of group characteristics on collective action, a number of other factors have been shown to be influential. First, and perhaps foremost, laws and policies that provide local resource users with secure and enforceable rights to use and manage resources may act as a powerful catalyst for cooperation (Cardenas et al. 2000, Frey and Jegen 2001, Baragwanath and Bayi 2020, Romero and Saavedra 2021). Second, resource characteristics can have a significant influence on the costs and complexity of managing resources. For example, groups managing large forests may face significantly higher long-term monitoring costs undermining efforts towards raising funds to hire local guards (Agrawal and Goyal 2001). However, there may also be fewer constraints in trying to develop rules that achieve a balance between local livelihoods and long-term conservation (Chhatre and Agrawal 2009, Persha et al. 2011). Finally, access to markets for natural resources can be an important driver of collective action and sustainability outcomes, although evidence about the nature of their effects are somewhat mixed. On the one hand, several studies have found that external market incentives can rapidly overwhelm the capacity of local systems (Berkes et al. 2006, Cinner et al. 2021); while others have shown they may act as a powerful catalyst for collective action (Kaganzi et al. 2009, Epstein et al. 2021). Collectively, the literature presents a complicated picture of collective action and sustainable CBNRM, and the need for new approaches for advancing theory that can explicitly consider multiple factors within a single empirical analysis. Moreover, as policymakers and NGOs continue to roll out policies and programs to support CBNRM (Gurney et al. 2021) there is a need for practical guidance about the contexts in which groups are likely to successfully self-organize and achieve more sustainable outcomes.

DATA AND METHODS

This paper draws upon data from the International Forest Resources and Institutions (IFRI) database (Wertime et al. 2007). IFRI is a multidisciplinary program of collaborative research that has collected data from over 500 forests and 800 forest user groups across fifteen countries. The database includes information about forests, the people that use them, and the rules that govern their interactions with the forest and each other. The IFRI database is one of the few large-*n* datasets available to researchers on the commons to develop and test hypotheses concerning the role of communities and informal institutions for the governance of common-pool resources. Although the IFRI database suffers from a number of well-known limitations related to sampling and measurement of some concepts, it remains one of the largest available datasets that can be used to advance understanding of collective action in natural resource management and we believe that the use of legacy cases (i.e., prior to 2011) offers an important opportunity to apply exploratory methods to develop hypotheses, and avoid "burning" through newer cases and datasets that may undermine their utility for hypothesis testing.

The sample includes 176 unique observations of forest commons from six countries (Guatemala, Bolivia, India, Kenya, Madagascar, Nepal and Uganda) between 1993 and 2011. The sample was selected by first dropping observations of privately-owned forests and those with missing values for any of the variables included in the analysis. Next, in cases where forests had repeated observations over time, the most recent observation was retained for analysis. Finally, countries with fewer than five remaining observations were excluded. Further details on the variables used in this analysis can be found in Table 1.

This analysis includes two dichotomous dependent variables that record details about the source of rules for the use of forest products and forest conditions, respectively. Local rulemaking records whether all rules for the use of forest resources were developed by current or former members of local user groups (a value of 1), or whether rules for at least some resources were developed by external actors (a value of 0). Success, meanwhile, is based upon two subjective measures of the vegetation density and species diversity of the forest relative to other forests in the same ecological zone. A forest is defined as successful (a value of 1) if both vegetation density and species diversity are assessed as typical or better when compared to forests in the same ecological zone.

The explanatory variables were selected based upon theory and research on the drivers of collective action in natural resource management (Baland and Platteau 1999, Agrawal 2003, Poteete and Ostrom 2004, Epstein et al. 2021), and filtered by the availability of corresponding variables and data in the IFRI database. It is important to note that two different samples are used in this analysis to develop insights about the factors influencing local rulemaking and local rulemaking success, respectively. The local rulemaking model includes the full sample of 176 observations as summarized in Table 1. The successful local rulemaking model, meanwhile, analyzes a subsample of this dataset which includes only those observations in which local rulemaking is equal to one, resulting in a sample of 112 observations. Summary statistics for variables included in the successful local rulemaking subsample can be found in Table S1, while Table S2 compares means between the full sample and local rulemaking sample and Table S3 reports country-level means for outcomes (i.e., local rulemaking and successful local rulemaking).

BOOSTED DECISION TREES

Factors influencing the likelihood of local rulemaking and successful local rulemaking were analyzed using boosted decision trees. Boosted decision trees are an approach in machine learning that is used to predict outcomes using a set of features or variables by sequentially fitting decision trees to improve predictions and then combining these to develop a single predictive model. Decision trees are increasingly being applied in social-ecological systems research (Gutierrez et al. 2011, Jouffray et al. 2015, Di Franco et al. 2016, Rana and Miller 2018, Frey 2020a, Epstein et al. 2021) to advance understanding of the complex drivers of sustainability outcomes. They have several advantages over more traditional statistical techniques, specifically with respect to their ability to include large numbers of variables and model non-linear responses (Elith et al. 2008).

The data was analyzed through several steps in R version 4.0.3 (R Core Team 2020). First, model hyperparameters were tuned using 10-fold cross-validation with five repeats in the caret package (Kuhn 2008) based on cross-validated accuracy. Cross-validation involves randomly dividing a sample into k folds (in this case 10), estimating a model using k-1 folds and testing predictions against the withheld data and then repeating for each fold. Repeated cross-validation simply involves repeating this process a set number of times (in this case 5). The tuning grid was adapted from Epstein et al. (2021) and included the number of trees (100–5,000 in increments of 100), interaction depth which defines the number of splits for each tree (1–5), shrinkage or learning rate which defines the contribution of each tree to the final model (0.001, 0.005, 0.01) and minimum number of observations in a terminal node (5, 10) with a fixed bag-fraction of 0.7.

We then used the set of parameters that maximized cross-validated accuracy to fit a boosted decision tree and characterize the relative influence of variables using the gbm package (Greenwell et al. 2019). Accuracy was chosen on the basis of its intuitiveness when compared to alternative performance metrics and to maintain consistency with similar research (Epstein et al. 2021). Relative influence provides a measure of the importance of a variable in classifying observations and is based upon the number of times a variable is used for splitting and is weighted by the squared improvement in model error from each split (Friedman and Meulman 2003). As a result high levels of correlation among variables will tend to reduce estimates of the relative influence of those variables. Finally, the relationships between variables and outcomes were visualized using partial dependence plots in the pdp package (Greenwell 2017), while trimming outliers to avoid extrapolating to points with limited data**.** Partial dependence plots calculate predictions for each value of a given variable while permuting other variables in the model, and then averaging these to show the average "effects" of a variable across a given range. While partial dependence plots are useful for interpreting the results of machine learning

VARIABLE	DESCRIPTION	MEAN	STD. DEV.	MIN	MAX
Local rulemaking	All rules for most important forest resources were designed by local forest user groups	0.64	0.48	0	1
Success	Forester's assessment of the vegetation density and species diversity of the forest	0.57	0.50	0	1
Group size	Number of households in user group	246.03	516.44	5	5116
Leadership	Presence of individuals investing time and resources to support local collective action	0.45	0.50	0	1
Land ownership	Percentage of households that own land	0.79	0.33	0	1
Poverty rate	Fraction of user group members considered poor	0.18	0.22	0	1
College education	Fraction of user group with college education	0.03	0.08	0	0.77
Literate	Literacy rate	0.46	0.27	0	1
Subsistence	Fraction of user group that derive subsistence benefits from forest resources	0.74	0.36	0	1
Interest heterogeneity	Fractionalization index based on households deriving subsistence, commercial, both or neither benefits from the forest	0.22	0.23	0	0.75
Economic heterogeneity	Fractionalization index based on households considered poor, wealthy and neither	0.26	0.23	0	0.66
Social capital	Composite index of user group cooperation on harvesting, processing and marketing of non-forest products	2.03	2.29	0	9
Distance to market	Travel time to markets (minutes)	179.84	636.34	1	7200
State forest	Forest is owned by the state	0.81	0.39	0	1
Forest size	Size of forest in hectares	1272.21	4299.10	0.9	40000
Year	Year of observation	2000.74	5.07	1993	2011
Guatemala	Forest is located in Guatemala	0.05	0.21	0	1
Bolivia	Forest is located in Bolivia	0.10	0.30	0	1
India	Forest is located in India	0.19	0.39	0	1
Kenya	Forest is located in Kenya	0.09	0.28	0	1
Madagascar	Forest is located in Madagascar	0.04	0.20	0	1
Nepal	Forest is located in Nepal	0.43	0.50	0	1
Uganda	Forest is located in Uganda	0.11	0.32	0	1

Table 1 Summary Statistics (n = 176).

models, they also assume independence among variables (Molnar 2022).

RESULTS

The results of this analysis are summarized in the following sections, providing details about the relative influence of variables and partial dependence plots for the top three most influential variables for local rulemaking (3.1) and successful local rulemaking (3.2), respectively.

LOCAL RULEMAKING

The results of the local rulemaking model, which correctly classifies approximately 79% of the observations in the sample (Table S4), with a cross-validated accuracy of 71%, are summarized in Figure 1. As can be seen, three features play a more influential role in determining the likelihood that forest user groups have developed rules for the use of forest resources: group size, forest size and Kenya. In general, it appears that forest user groups are more likely to develop and maintain rules when those groups and the forests they use are smaller in size (i.e., <100 households and 250 ha); while user groups located in Kenya are less likely to have developed rules than groups in other countries. Further details about the relationships between the remaining features (with non-zero relative influence) are supplied in Figure S1.

SUCCESSFUL LOCAL RULEMAKING

The successful local rulemaking model, meanwhile, correctly classifies approximately 81% of observations in

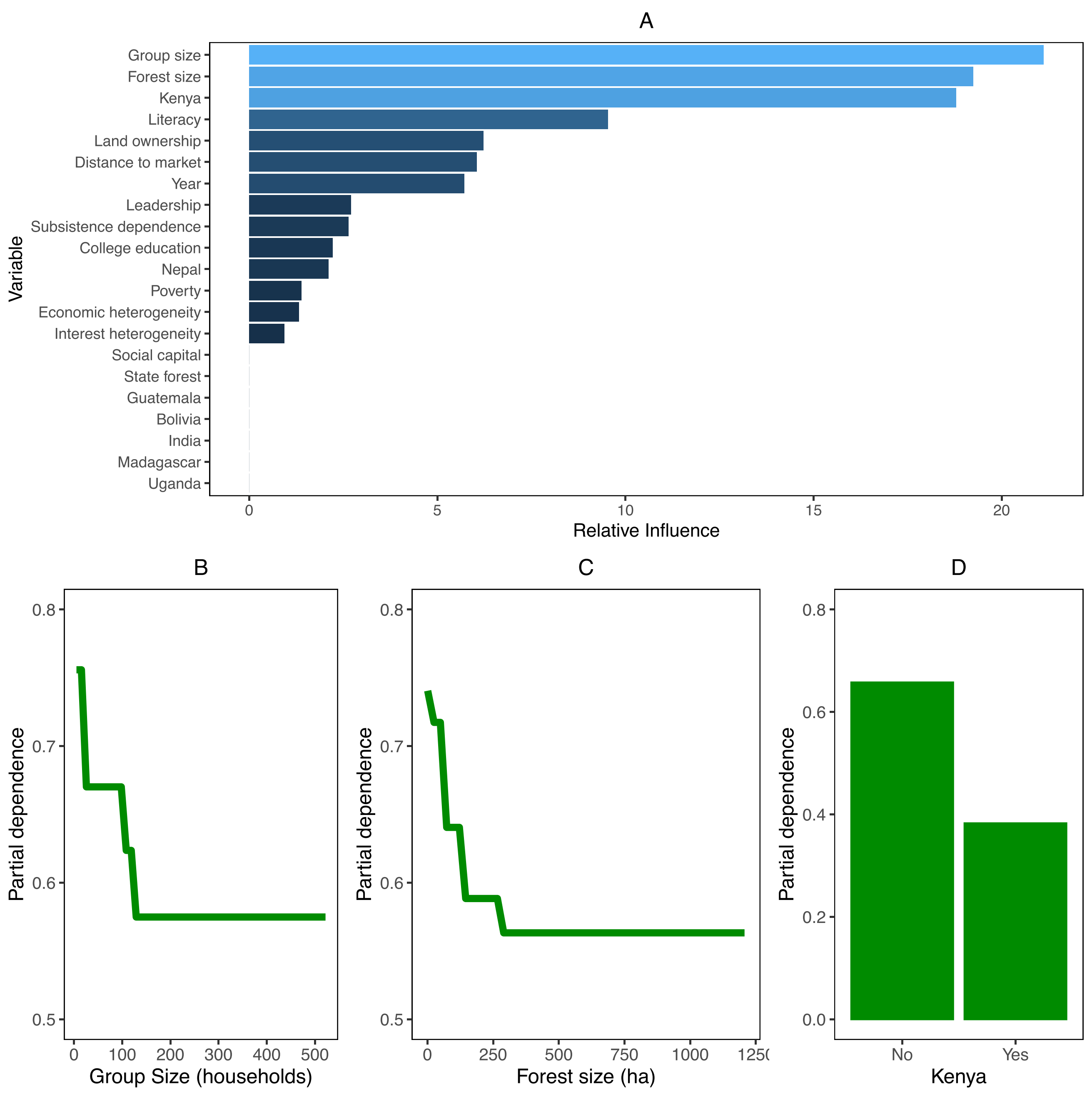

Figure 1 Results of the local rulemaking model.

(A) Relative influence of variables provides an estimate of the contributions of each variable to predictions that groups successfully self-organize to develop rules for the use of forest resources. **(B-D)** Partial dependence plots of top three variables, including **(B)** group size, **(C)** forest size, and **(D)** Kenya show the relationship between a variable while accounting for the average effects of other variable and trimming outliers to avoid over-extrapolation to values with limited data.

the sample (Table S5) with a cross-validated accuracy of 75%. The results of this model are summarized in Figure 2, showing that the likelihood of successful local rulemaking is heavily influenced by the size of the forests that are managed by local forest user groups and to a lesser extent, their level of subsistence dependence on the forest and the size of associated forest user groups. In general, it appears that groups that have developed and maintained rules for the use of forest resources are more likely to be successful in terms of maintaining a forest with higher levels of vegetation density and species diversity when forests are larger than 200 ha, and forest user groups have high levels of subsistence dependence and include more than 100 households. Further details about the relationships

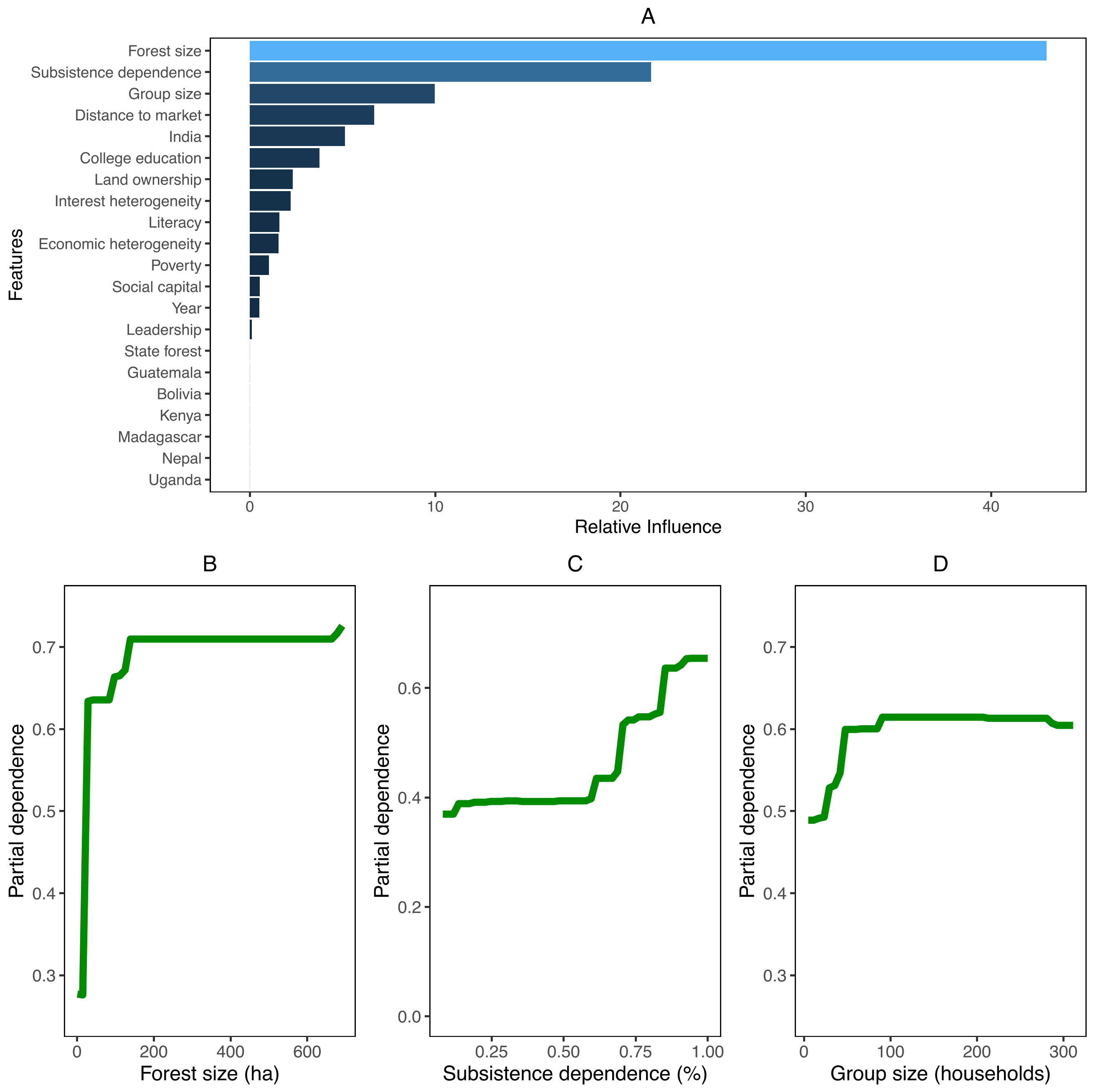

Figure 2 Results of the successful local rulemaking model.

(A) Relative influence of variables provides an estimate of the contributions of each variable to predictions that groups successfully self-organize to develop rules for the use of forest resources. **(B-D)** Partial dependence plots of top three variables, including **(B)** forest size, **(C)** subsistence dependence, and **(D)** group size show the relationship between a variable while accounting for the average effects of other variables and trimming outliers to avoid over-extrapolation.

between the remaining features (with non-zero relative influence) are supplied in Figure S2.

DISCUSSION

CBNRM continues to hold considerable promise as a strategy for protecting forest resources and supporting a transition toward a more sustainable, equitable and just relationship between people and the environment (Erbaugh et al. 2020). Indeed, there is a large and growing body of evidence linking community engagement and de facto management rights to better forest conditions (Chattre and Agrawal 2009; Persha et al. 2011; Blackman et al. 2017; Hajjar et al. 2021). At the same time, policies designed to support community forest management have often failed to adequately empower communities with rights to use and manage local forest resources to deliver upon their potential (Ribot and Agrawal 2006; Sapkota et al. 2020; Hajjar et al. 2021).

The results of this analysis, for instance, show that the contexts in which groups are able to successfully self-organize to develop and maintain rules for the use of forest products are systematically different from those in which local rulemaking is likely to contribute to successful ecological outcomes. More specifically, prospects for local rulemaking appear to be highest when user groups and the forests they use are smaller in size (<250 hectares; <100 households); while prospects for successful local rulemaking rise when they are larger in size (> 200 hectares; > 100 households). Whereas small group size and forests may ease the costs and complexity of developing rules (Olson 1965), larger groups with high levels of subsistence dependence may face lower per capita costs for governance activities such as enforcement and have greater incentives and lower discount rates to ensure the long-term sustainability of resource flows (Ostrom 2000; Epstein et al. 2021).

In addition to the core findings of this research, partial dependence plots (Figure S1 and S2) suggest several potentially useful directions for future research. These include suggestions that groups are more likely to successfully self-organize to develop rules for the use of forest resources when literacy and land ownership rates are higher, and settlements are located in close proximity to markets for forest products. They also provide some evidence to suggest that local rulemaking might be more likely to contribute to better forest conditions when settlements are located further from markets and user groups include at least some individuals with a college education. Collectively these and other insights from the two boosted decision tree models have potentially valuable insights concerning the fit between local rulemaking and complex social-ecological contexts and may inform further empirical research. However, before considering these insights, it is important to note the limitations of this study.

Boosted decision trees and related approaches in machine learning offer a powerful approach for sifting through large numbers of potentially influential variables to characterize their relative importance and develop insights about the nature of their relationships with outcomes. However, boosted decisions and the IFRI dataset we used may have some influence on the results. First, the IFRI database is not a random sample of forest commons and thus important questions remain concerning the population of cases to which the results of this analysis might apply. In general, similar research has suggested that findings are more likely to apply to cases that fall within the range of independent variables in the sample (Chhatre and Agrawal 2009, Coleman 2009). Although this assumption is generally accepted by many audiences (King et al. 2001), future research on the relationship between contexts, institutions and their performance would surely benefit from a more structured approach to sampling. Second, despite including over twenty predictors in each model they do not exhaust all factors that have been proposed to influence collective action and sustainable environmental governance (Ostrom 2009). Cultural heterogeneity, for example, which has been posited to affect collective action by influencing transaction costs (Fearon and Laitin 1996), was not included due to differences in how the cultural and religious backgrounds of group members were recorded across cases.

Third, decision trees are known to be biased towards the selection of continuous variables (Strobl et al. 2009), potentially resulting in higher estimates of the relative influence of variables such as forest size and group size. Correlation among variables (Figure S3 and S4), meanwhile, can have the opposite effect on the relative influence of variables, potentially contributing to lower scores for variables such as economic heterogeneity and poverty. Fourth and finally, although the models provide insights about the contexts in which local rulemaking is likely to be found and contribute to successful ecological outcomes, it does not provide any evidence to suggest that local rulemaking might be preferred to alternative institutional arrangements in these contexts. In fact, it seems plausible that most institutional arrangements would achieve better ecological performance in larger forests given well known ecological relationships between species diversity and area, edge effects and practical constraints on harvesting rates (Bowker et al. 2017). It is worth noting, however, that one recent study found that protected area deforestation rates in protected areas (many of which are managed by state agencies) tended to be higher when protected areas were larger in size (Wolf et al. 2021). Nevertheless, the findings of this research suggest several potentially fruitful directions for future research.

First, given the limitations discussed above there is a clear need for further empirical testing of the findings. More specifically, although cross-validation provides an estimate of model performance against withheld observations, it does not provide an independent test to validate model predictions. Instead, new data developed from case studies or other large-*n* datasets are needed to validate the hypotheses that local rulemaking is more likely to be found in contexts where user groups and forests are smaller in size and more likely to contribute to successful ecological outcomes when forests and user groups are larger in size and include a larger fraction of users deriving subsistence benefits from forests. If these and other hypotheses derived from this study are confirmed, they may provide insights to inform the development of policies and programs to support community-based forest management tailored to different communities and ecological contexts. For example, external actors might consider investing additional time and resources in building capacity and supporting local self-organization when communities and the forests they use are larger in size.

Second, the unfortunate mismatch between the contexts in which local rulemaking is likely to be found and those in which it is likely to be ecologically successful has several implications for theory and future research. On the one hand, this finding connects with a longstanding theoretical debate in the literature on community-based natural resource management questioning whether the conditions that explain institutional development are the same as those that explain the robustness of those institutions and the environmental outcomes they produce (Singleton and Taylor 1992; Morrison 2017). While the literature has tended to assume that the factors contributing to the establishment of CBNRM and its robustness are broadly similar, others have questioned the validity of this assumption (Agrawal 2003, Basurto 2013). Our findings present additional empirical evidence in favour of the latter hypothesis. On the other hand, differences in the conditions that support local rulemaking and those in which it is likely to be successful highlight a clear need for research examining the barriers that groups face with respect to self-organization and how they vary with changes in the size of communities and the forests they use. Indeed, while many countries have taken steps to create space for communities in natural resource management through decentralization, co-management and other participatory models of environmental governance (Njaya 2007, Agrawal et al. 2008), the results have been decidedly mixed (Blaikie 2006, Ribot et al. 2006; Sapkota et al. 2020; Hajjar et al. 2021).

In some cases, barriers to collective action are embedded within the government policies that enable community participation themselves. These can include placing limits on where, when and how communities can manage resources and/or establishing onerous conditions for the transfer of powers (Ribot et al. 2006, Larson and Soto 2008). Central governments in many countries including Nepal and Uganda have, for instance, retained significant control over forest resources by establishing spatial limitations on where local communities and governments are permitted to manage their forests. This allows governments to transfer powers and avoid costs of managing smaller or less valuable forests, while retaining control over the use and management of larger, or more valuable forests. Central governments can also exercise *de facto* control by directly restricting the scope of local rulemaking, or alternatively by requiring them to seek approval for their management plans (Ribot et al. 2006, Epstein et al. 2020). Although these and other external factors may create barriers for local rulemaking, internal factors may also influence prospects for collective action. Transaction costs may, for instance, be higher in larger groups (Olson 1965), exacerbating challenges associated with developing and securing agreements about rules.

Third and finally, this research highlights an important opportunity to advance understanding of the relationship between complex social-ecological contexts, institutions, and outcomes by leveraging a growing portfolio of tools to explore and uncover patterns in social-ecological datasets. Methodological limitations of conventional statistical techniques have played an important role in limiting progress. Boosted decision trees and many other machine learning techniques are not subject to the same limitations (Elith et al. 2008), and can therefore be used to gain traction on increasingly important questions about tailoring institutions to address diverse and mounting sustainability challenges. However, computational tools are not a panacea for advancing theory and improving environmental policies, but there are steps that can be taken to enhance their contributions. These include further efforts to develop large-n datasets that allow researchers to examine the distribution and performance of institutions across a gradient of social and ecological contexts, tools that enable consistent collection and coding of data (Lam 1998, Poteete et al. 2010, Gurney et al. 2019, Cox et al. 2021) and a more structured approach to sampling such as a census of well-defined populations (Rana and Miller 2018; Brewer et al. 2022).

As environmental problems continue to mount in a wide range of different social, ecological and institutional contexts, the need for a diverse portfolio of potential strategies for addressing them and knowledge concerning the contexts in which these strategies are more (or less) likely to give rise to sustainable outcomes has never been greater. Collectively the insights presented here are a useful starting point for future research on the relationship between contexts, institutions, and their performance, particularly with respect to the contexts that enable or constrain local rulemaking and contribute to its success.

ADDITIONAL FILE

The additional file for this article can be found as follows:

- **Supplementary Information.** Additional details about the data and results. DOI: https://doi.org/10.5334/ijc.1226.s1

FUNDING INFORMATION

This work was supported by the National Socio-Environmental Synthesis Center (SESYNC) through funding received from the National Science Foundation (DBI-1639145). GE would to acknowledge support from the Canada First Research Excellence Fund under the Global Water Futures Programme.

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR AFFILIATIONS

Graham Epstein orcid.org/0000-0003-3431-8511
School of Environment, Resources and Sustainability, University of Waterloo, CA

Cristina I. Apetrei orcid.org/0000-0003-2109-6470
Faculty of Sustainability, Leuphana University, Universitätsallee 1, 21335 Lüneburg, Germany

Jacopo Baggio orcid.org/0000-0002-9616-4143
School of Politics, Security and International Affairs, University of Central Florida, US

Sivee Chawla orcid.org/0000-0001-8565-1037
ARC Centre of Excellence for Coral Reef Studies, James Cook University, AU

Graeme Cumming orcid.org/0000-0002-3678-1326
ARC Centre of Excellence for Coral Reef Studies, James Cook University, AU

Georgina Gurney orcid.org/0000-0002-4884-7468
ARC Centre of Excellence for Coral Reef Studies, James Cook University, AU

Tiffany Morrison orcid.org/0000-0001-5433-037X
ARC Centre of Excellence for Coral Reef Studies, James Cook University, AU

Hita Unnikrishnan orcid.org/0000-0002-4298-4939
Urban Institute, Sheffield University, GB

Sergio Villamayor Tomas orcid.org/0000-0002-5170-1718
Institute of Environmental Science and Technology, Autonomous University of Barcelona, Barcelona, Spain

REFERENCES

Acheson, J. (2006). Institutional Failure in Resource Management. *Annual Review of Anthropology*, *35*, 117–134. DOI: https://doi.org/10.1146/annurev.anthro.35.081705.12323

Acheson, J. M., & **Wilson, J. A.** (1996). Order out of Chaos: The Case for Parametric Fisheries Management. *American Anthropologist*, *98*, 579–594. DOI: https://doi.org/10.1525/aa.1996.98.3.02a0011

Adhikari, B., & **Lovett, J. C.** (2006). Transaction costs and community-based natural resource management in Nepal. *Journal of Environmental Management*, *78*, 5–15. DOI: https://doi.org/10.1016/j.jenvman.2005.04.00

Agrawal, A. (2003). Sustainable Governance of Common-Pool Resources: Context, Methods, and Politics. *Annual Review of Anthropology*, *32*, 243–262. DOI: https://doi.org/10.1146/annurev.anthro.32.061002.09311

Agrawal, A., Chhatre, A., & **Hardin, R.** (2008). Changing Governance of the World's Forests. *Science*, *320*, 1460–1462. DOI: https://doi.org/10.1126/science.115536

Agrawal, A., & **Goyal, S.** (2001). Group Size and Collective Action. *Comparative Political Studies*, *34*, 63–93. DOI: https://doi.org/10.1177/0010414001034001003

Agrawal, A., & **Yadama, G.** (1997). How do Local Institutions Mediate Market and Population Pressures on Resources? *Forest Panchayats in Kumaon, India. Development and Change*, *28*, 435–465. DOI: https://doi.org/10.1111/1467-7660.00050

Andersson, E., & **Gabrielsson, S.** (2012). 'Because of poverty, we had to come together': collective action for improved food security in rural Kenya and Uganda. *International journal of agricultural sustainability*, *10*, 245–262. DOI: https://doi.org/10.1080/14735903.2012.666029

Araral, E., Jr. (2009). What Explains Collective Action in the Commons? *Theory and Evidence from the Philippines. World Development*, *37*, 687–697. DOI: https://doi.org/10.1016/j.worlddev.2008.08.002

Baggio, J. A., Rollins, N. D., Pérez, I., & **Janssen, M. A.** (2015). Irrigation experiments in the lab: trust, environmental variability, and collective action. *Ecology and Society*, *20*. DOI: https://doi.org/10.5751/ES-07772-200412

Baland, J.-M., & **Platteau, J.-P.** (1999). The Ambiguous Impact of Inequality on Local Resource Management. *World Development*, *27*, 773–788. DOI: https://doi.org/10.1016/S0305-750X(99)00026-1

Ban, N. C., Caldwell, I. R., Green, T. L., Morgan, S., O'Donnell, K., & **Selgrath, J. C.** (2009). Diverse fisheries require diverse solutions. *Science*, *323*, 338–338. DOI: https://doi.org/10.1126/science.323.5912.338

Baragwanath, K., & **Bayi, E.** (2020). Collective property rights reduce deforestation in the Brazilian Amazon. *Proceedings of the National Academy of Sciences*, *117*, 20495–20502. DOI: https://doi.org/10.1073/pnas.1917874117

Bardhan, P., & **Dayton-Johnson, J.** (2002). Unequal irrigators: heterogeneity and commons management in large-scale multivariate research. Pages 87–112. In E. Ostrom, T. Dietz, N. Dolsak, Stern P. C., S. Stonich & E. U. Weber (Eds.), *The drama of the commons*.

Basurto, X. (2013). Linking multi-level governance to local common-pool resource theory using fuzzy-set qualitative comparative analysis: Insights from twenty years of biodiversity conservation in Costa Rica. *Global Environmental Change*, *23*, 573–587. DOI: https://doi.org/10.1016/j.gloenvcha.2013.02.011

Berkes, F., Hughes, T. P., Steneck, R. S., Wilson, J. A., Bellwood, D. R., Crona, B., Folke, C., Gunderson, L. H., Leslie, H. M., Norberg, J., Nyström, M., Olsson, P., Österblom, H., Scheffer, M., & **Worm, B.** (2006). Globalization, Roving Bandits, and Marine Resources. *Science*, *311*, 1557–1558. DOI: https://doi.org/10.1126/science.1122804

Blackman, A., Corral, L., Lima, E. S., & **Asner, G. P.** Titling indigenous communities protects forests in the Peruvian Amazon. *Proc. Natl. Acad. Sci. USA*, *114*, 4123–4128 (2017). DOI: https://doi.org/10.1073/pnas.1603290114

Blaikie, P. (2006). Is Small Really Beautiful? *Community-based Natural Resource Management in Malawi and Botswana. World Development*, *34*, 1942–1957. DOI: https://doi.org/10.1016/j.worlddev.2005.11.023

Bodin, Ö. (2017). Collaborative environmental governance: Achieving collective action in social-ecological systems. *Science*, *357*, eaan1114. DOI: https://doi.org/10.1126/science.aan1114

Bodin, Ö., Crona, B., Thyresson, M., Golz, A.-L., & **TengÖ, M.** (2014). Conservation Success as a Function of Good Alignment of Social and Ecological Structures and Processes. *Conservation Biology*, *28*, 1371–1379. DOI: https://doi.org/10.1111/cobi.12306

Bodin, Ö., & **Tengö, M.** (2012). Disentangling intangible social–ecological systems. *Global Environmental Change*, *22*, 430–439. DOI: https://doi.org/10.1016/j.gloenvcha.2012.01.005

Bowker, J., De Vos, A., Ament, J., & **Cumming, G.** (2017). Effectiveness of Africa's tropical protected areas for maintaining forest cover. *Conservation Biology*, *31*, 559–569. DOI: https://doi.org/10.1111/cobi.12851

Brewer, T. D., Andrew, N., Gruber, B., & **Kool, J.** (2022). Large-sample-size assessment of socioeconomic predictors of community-level resource management occurrence. *Conservation Biology*, *36*, e13800. DOI: https://doi.org/10.1111/cobi.13800

Brock, W. A., & **Carpenter, S. R.** (2007). Panaceas and diversification of environmental policy. *Proceedings of the National Academy of Sciences*, *104*, 15206–15211. DOI: https://doi.org/10.1073/pnas.0702096104

Cardenas, J.-C. (2003). Real wealth and experimental cooperation: experiments in the field lab. *Journal of Development Economics*, *70*, 263–289. DOI: https://doi.org/10.1016/S0304-3878(02)00098-6

Cardenas, J. C., Stranlund, J., & **Willis, C.** (2000). Local Environmental Control and Institutional Crowding-Out. *World Development*, *28*, 1719–1733. DOI: https://doi.org/10.1016/S0305-750X(00)00055-3

Chhatre, A., & **Agrawal, A.** (2009). Trade-offs and synergies between carbon storage and livelihood benefits from forest commons. *Proceedings of the National Academy of Sciences*, *106*, 17667–17670. DOI: https://doi.org/10.1073/pnas.0905308106

Cifdaloz, O., Regmi, A., Anderies, J. M., & **Rodriguez, A. A.** (2010). Robustness, vulnerability, and adaptive capacity in small-scale social-ecological systems: The Pumpa Irrigation System in Nepal. *Ecology and Society*, *15*, 39. DOI: https://doi.org/10.5751/ES-03462-150339

Cinner, J. E., Barnes, M. L., Gurney, G. G., Lockie, S., & **Rojas, C.** (2021). Markets and the crowding out of conservation-relevant behavior. *Conservation Biology*, *35*, 816–823. DOI: https://doi.org/10.1111/cobi.13606

Coleman, E. A. (2009). Institutional factors affecting biophysical outcomes in forest management. *Journal of Policy Analysis and Management*, *28*, 122–146. DOI: https://doi.org/10.1002/pam.20405

Coleman, E. A., & **Fleischman, F. D.** (2012). Comparing Forest Decentralization and Local Institutional Change in Bolivia, Kenya, Mexico, and Uganda. *World Development*, *40*, 836–849. DOI: https://doi.org/10.1016/j.worlddev.2011.09.008

Coleman, E. A., & **Steed, B. C.** (2009). Monitoring and sanctioning in the commons: an application to forestry. *Ecological Economics*, *68*, 2106–2113. DOI: https://doi.org/10.1016/j.ecolecon.2009.02.006

Cox, M. (2012). Diagnosing Institutional Fit: a Formal Perspective. *Ecology and Society*, *17*, 54. DOI: https://doi.org/10.5751/ES-05173-170454

Cox, M., Gurney, G., Anderies, J., Coleman, E., Darling, E., Epstein, G., Frey, U., Nenadovic, M., Schlager, E., & **Villamayor-Tomas, S.** (2021). Lessons learned from synthetic research projects based on the Ostrom Workshop frameworks. *Ecology and Society, 26*. DOI: https://doi.org/10.5751/ES-12092-260117

Crawford, S. E. S., & **Ostrom, E.** (1995). A Grammar of Institutions. *The American Political Science Review*, *89*, 582–600. DOI: https://doi.org/10.2307/2082975

Cumming, G. S., Cumming, D. H. M., & **Redman, C. L.** (2006). Scale mismatches in social-ecological systems: causes, consequences and solutions. *Ecology and Society*, *11*(1). DOI: https://doi.org/10.5751/ES-01569-110114

Cumming, G. S., & **Epstein, G.** (2020). Landscape sustainability and the landscape ecology of institutions. *Landscape Ecology*. DOI: https://doi.org/10.1007/s10980-020-00989-8

Cumming, G. S., Epstein, G., Anderies, J. M., Apetrei, C. I., Baggio, J., Bodin Ö., Chawla, S., Clements, H. S., Cox, M., Egli, L., Gurney, G. G., Lubell, M., Magliocca, N., Morrison, T. H., Müller, B., Seppelt, R., Schlüter, M., Unnikrishnan, H., Villamayor-Tomas, S., & **Weible, C. M.** (2020). Advancing understanding of natural resource governance: a post-Ostrom research agenda.

Current Opinion in Environmental Sustainability, 44, 26–34. DOI: https://doi.org/10.1016/j.cosust.2020.02.005

Devillers, R., Pressey, R. L., Grech, A., Kittinger, J. N., Edgar, G. J., Ward, T., & **Watson, R.** (2015). Reinventing residual reserves in the sea: are we favouring ease of establishment over need for protection? *Aquatic Conservation: Marine and Freshwater Ecosystems*, *25*, 480–504. DOI: https://doi.org/10.1002/aqc.2445

Di Franco, A., Thiriet, P., Di Carlo, G., Dimitriadis, C., Francour, P., Gutiérrez, N. L., de Grissac, A. J., Koutsoubas, D., Milazzo, M., & **del Mar Otero, M.** (2016). Five key attributes can increase marine protected areas performance for small-scale fisheries management. *Scientific Reports, 6*. DOI: https://doi.org/10.1038/srep38135

Doss, C. R., & **Meinzen-Dick, R.** (2015). Collective action within the household: Insights from natural resource management. *World Development*, *74*, 171–183. DOI: https://doi.org/10.1016/j.worlddev.2015.05.001

Elith, J., Leathwick, J. R., & **Hastie, T.** (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, *77*, 802–813. DOI: https://doi.org/10.1111/j.1365-2656.2008.01390.x

Epstein, G., Gurney, G., Chawla, S., Anderies, J. M., Baggio, J., Unnikrishnan, H., Villamayor Tomas, S., & **Cumming, G. S.** (2021). Drivers of compliance monitoring in forest commons. *Nature Sustainability*. DOI: https://doi.org/10.1038/s41893-020-00673-4

Epstein, G., Morrison, T. H., Lien, A., Gurney, G. G., Cole, D. H., Delaroche, M., Villamayor Tomas, S., Ban, N., & **Cox, M.** (2020). Advances in understanding the evolution of institutions in complex social-ecological systems. *Current Opinion in Environmental Sustainability*, 44, 58–66. DOI: https://doi.org/10.1016/j.cosust.2020.06.002

Epstein, G., Pittman, J., Alexander, S. M., Berdej, S., Dyck, T., Kreitmair, U., Rathwell, K. J., Villamayor-Tomas, S., Vogt, J., & **Armitage, D.** (2015). Institutional fit and the sustainability of social–ecological systems. *Current Opinion in Environmental Sustainability*, *14*, 34–40. DOI: https://doi.org/10.1016/j.cosust.2015.03.005

Erbaugh, J. T., Pradhan, N., Adams, J., Oldekop, J. A., Agrawal, A., Brockington, D., Pritchard, R., & **Chhatre, A.** (2020). Global forest restoration and the importance of prioritizing local communities. *Nature Ecology & Evolution*, 4, 1472–1476. DOI: https://doi.org/10.1038/s41559-020-01282-2

Fearon, J. D., & **Laitin, D. D.** (1996). Explaining Interethnic Cooperation. *The American Political Science Review*, *90*, 715– 735. DOI: https://doi.org/10.2307/2945838

Frey, B. S., & **Jegen, R.** (2001). Motivation Crowding Theory. *Journal of Economic Surveys*, *15*, 589–611. DOI: https://doi.org/10.1111/1467-6419.00150

Frey, U. J. (2017). A synthesis of key factors for sustainability in social–ecological systems. *Sustainability Science*, *12*, 507– 519. DOI: https://doi.org/10.1007/s11625-016-0395-z

Frey, U. (2020a). Sustainable Governance of Natural Resources: Uncovering Success Patterns with Machine Learning. USA: Oxford University Press. DOI: https://doi.org/10.1093/oso/9780197502211.001.0001

Friedman, J. H., & **Meulman, J. J.** (2003). Multiple additive regression trees with application in epidemiology. *Statistics in Medicine*, *22*, 1365–1381. DOI: https://doi.org/10.1002/sim.1501

Galaz, V., Olsson, P., Hahn, T., Folke, C., & **Svedin, U.** (2008). The problem of fit among biophysical systems, environmental and resource regimes, and broader governance systems: insights and emerging challenges. In O. R. Young, (Ed.), *Institutions and Environmental Change*. Cambridge, MA: The MIT Press. pp. 147–186. DOI: https://doi.org/10.7551/mitpress/9780262240574.003.0005

Garrick, D. (2018). Decentralisation and drought adaptation: Applying the subsidiarity principle in transboundary river basins. *International Journal of the Commons, 12*. DOI: https://doi.org/10.18352/ijc.816

Gibson, C. C., Williams, J. T., & **Ostrom, E.** (2005). Local Enforcement and Better Forests. *World Development*, *33*, 273–284. DOI: https://doi.org/10.1016/j.worlddev.2004.07.013

Glowacki, L., & **von Rueden, C.** (2015). Leadership solves collective action problems in small-scale societies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*, 20150010. DOI: https://doi.org/10.1098/rstb.2015.0010

Greenwell, B. (2017). pdp: *An R Package for Constructing Partial Dependence Plots*. *9*, 421–436. DOI: https://doi.org/10.32614/RJ-2017-016

Greenwell, B., Boehmke, B., Cunningham, J., & **GBM Developers.** (2019). gbm: Generalized Boosted Regression Models. R package version 2.1.5.

Gurney, G. G., Darling, E. S., Ahmadia, G. N., Agostini, V. N., Ban, N. C., Blythe, J., Claudet, J., Epstein, G., Himes-Cornell, A., & **Jonas, H. D.** (2021). Biodiversity needs every tool in the box: use OECMs. Nature Publishing Group. DOI: https://doi.org/10.1038/d41586-021-02041-4

Gurney, G. G., Darling, E. S., Jupiter, S. D., Mangubhai, S., McClanahan, T. R., Lestari, P., Pardede, S., Campbell, S. J., Fox, M., & **Naisilisili, W.** (2019). Implementing a social-ecological systems framework for conservation monitoring: lessons from a multi-country coral reef program. *Biological Conservation*, *240*, 108298. DOI: https://doi.org/10.1016/j.biocon.2019.108298

Gutierrez, N. L., Hilborn, R., & **Defeo, O.** (2011). Leadership, social capital and incentives promote successful fisheries. *Nature*, *470*, 386–389. DOI: https://doi.org/10.1038/nature09689

Hajjar, R., Oldekop, J. A., Cronkleton, P., Newton, P., Russell, A. J., & **Zhou, W.** (2021). A global analysis of the social and environmental outcomes of community forests. *Nature Sustainability*, 4(3), 216–224. DOI: https://doi.org/10.1038/s41893-020-00633-y

Hardin, R. (1982). Collective Action. Baltimore, USA: The Johns Hopkins University Press.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & **McElreath, R.** (2001). In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies. *The American Economic Review*, *91*, 73–78. DOI: https://doi.org/10.1257/aer.91.2.73

Henrich, J., Chudek, M., & **Boyd, R.** (2015). The Big Man Mechanism: how prestige fosters cooperation and creates prosocial leaders. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*, 20150013. DOI: https://doi.org/10.1098/rstb.2015.0013

Huntington, H., Liang, W., Rice, M., Wilson, S. J., & **Agrawal, A.** (2016). IFRI Codebook.

Janssen, M. A., Anderies, J. M., & **Joshi, S. R.** (2011). Coordination and cooperation in asymmetric commons dilemmas. *Experimental Economics*, *14*, 547. DOI: https://doi.org/10.1007/s10683-011-9281-9

Johannes, R. E. (1978). Traditional Marine Conservation Methods in Oceania and their Demise. *Annual Review of Ecology and Systematics*, *9*, 349–364. DOI: https://doi.org/10.1146/annurev.es.09.110178.002025

Joppa, L. N., & **Pfaff, A.** (2009). High and far: biases in the location of protected areas. *PLoS ONE*, *4*, e8273. DOI: https://doi.org/10.1371/journal.pone.0008273

Jouffray, J.-B., Nyström, M., Norström, A. V., Williams, I. D., Wedding, L. M., Kittinger, J. N., & **Williams, G. J.** (2015). Identifying multiple coral reef regimes and their drivers across the Hawaiian archipelago. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*, 20130268. DOI: https://doi.org/10.1098/rstb.2013.0268

Kaganzi, E., Ferris, S., Barham, J., Abenakyo, A., Sanginga, P., & **Njuki, J.** (2009). Sustaining linkages to high value markets through collective action in Uganda. *Food policy*, *34*, 23–30. DOI: https://doi.org/10.1016/j.foodpol.2008.10.004

King, G., Keohane, R. O., & **Verba, S.** (2001). Designing social inquiry: Scientific inference in qualitative research. Princeton University Press.

Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of statistical software*, *28*, 1–26. DOI: https://doi.org/10.18637/jss.v028.i05

Lam, W. F. (1998). Governing irrigation systems in Nepal: institutions, infrastructure, and collective action. Oakland, CA: ICS Press Institute for Contemporary Studies.

Larson, A. M., & **Soto, F.** (2008). Decentralization of Natural Resource Governance Regimes. *Annual Review of Environment and Resources*, *33*, 213–239. DOI: https://doi.org/10.1146/annurev.environ.33.020607.095522

Leslie, H. M., Basurto, X., Nenadovic, M., Sievanen, L., Cavanaugh, K. C., Cota-Nieto, J. J., Erisman, B. E., Finkbeiner, E., Hinojosa-Arango, G., & **Moreno-Báez, M.** (2015). Operationalizing the social-ecological systems framework to assess sustainability. *Proceedings of the National Academy of Sciences*, *112*, 5979–5984. DOI: https://doi.org/10.1073/pnas.1414640112

Magliocca, N. R., Ellis, E. C., Allington, G. R., De Bremond, A., Dell'Angelo, J., Mertz, O., Messerli, P., Meyfroidt, P., Seppelt, R., & **Verburg, P. H.** (2018). Closing global knowledge gaps: producing generalized knowledge from case studies of social-ecological systems. *Global Environmental Change*, *50*, 1–14. DOI: https://doi.org/10.1016/j.gloenvcha.2018.03.003

Maskey, V., Gebremedhin, T. G., & **Dalton, T. J.** (2006). Social and cultural determinants of collective management of community forest in Nepal. *Journal of forest economics*, *11*, 261–274. DOI: https://doi.org/10.1016/j.jfe.2005.10.004

Meinzen-Dick, R., Raju, K. V., & **Gulati, A.** (2002). What Affects Organization and Collective Action for Managing Resources? *Evidence from Canal Irrigation Systems in India. World Development*, *30*, 649–666. DOI: https://doi.org/10.1016/S0305-750X(01)00130-9

Meyfroidt, P., Chowdhury, R. R., de Bremond, A., Ellis, E. C., Erb, K.-H., Filatova, T., Garrett, R., Grove, J. M., Heinimann, A., & **Kuemmerle, T.** (2018). Middle-range theories of land system change. *Global Environmental Change*, *53*, 52–67. DOI: https://doi.org/10.1016/j.gloenvcha.2018.08.006

Milman, A., & **Gerlak, A. K.** (2020). International river basin organizations, science, and hydrodiplomacy. *Environmental Science & Policy*, 107, pp.137–149. DOI: https://doi.org/10.1016/j.envsci.2020.02.023

Molnar, C. (2022). *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable* (2nd ed.). https://christophm.github.io/interpretable-ml-book/cite.html

Morrison, T. H. (2017). Evolving polycentric governance of the Great Barrier Reef. *Proceedings of the National Academy of Sciences*, *114*, E3013–E3021. DOI: https://doi.org/10.1073/pnas.1620830114

Nalau, J., Becken, S., Schliephack, J., Parsons, M., Brown, C., & **Mackey, B.** (2018). The role of indigenous and traditional knowledge in ecosystem-based adaptation: A review of the literature and case studies from the Pacific Islands.

Weather, Climate, and Society, *10*, 851–865. DOI: https://doi.org/10.1175/WCAS-D-18-0032.1

Njaya, F. (2007). Governance Challenges of the Implementation of Fisheries Co-Management: Experiences from Malawi. *International Journal of the Commons*, *1*, 137–153. DOI: https://doi.org/10.18352/ijc.21

Oberlack, C., Sietz, D., Bonanomi, E. B., De Bremond, A., Dell'Angelo, J., Eisenack, K., Ellis, E. C., Epstein, G., Giger, M., & **Heinimann, A.** (2019). Archetype analysis in sustainability research. *Ecology and Society*, *24*. DOI: https://doi.org/10.5751/ES-10747-240226

Olson, M. (1965). The logic of collective action: public goods and the theory of groups. Harvard University Press.

Ostrom, E. (1990). Governing the Commons. New York: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511807763

Ostrom, E. (2000). The danger of self-evident truths. *PS: Political Science & Politics*, *33*(1), 33–46. DOI: https://doi.org/10.2307/420774

Ostrom, E. (2003). How Types of Goods and Property Rights Jointly Affect Collective Action. *Journal of Theoretical Politics*, *15*, 239–270. DOI: https://doi.org/10.1177/0951692803015003002

Ostrom, E. (2009). A General Framework for Analyzing Sustainability of Social-Ecological Systems. *Science*, *325*, 419–422. DOI: https://doi.org/10.1126/science.1172133

Ostrom, E., Huntington, H., Andersson, K., Banana, A., Castallenos, E., Chhatre, A., England, J., Ghate, R., Gombya-Ssembajjwe, W., Karna, B., Leon, R., Liang, W., Marquez, L., Newton, P., Persha, L., Salk, C., Tucker, C., & **Agrawal, A.** (2016). IFRI Dataset.

Ostrom, E., Janssen, M. A., & **Anderies, J. M.** (2007). Going beyond Panaceas. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 15176–15178. DOI: https://doi.org/10.1073/pnas.0701886104

Ostrom, V., & **Ostrom, E.** (1999). Public Goods and Public Choices. Pages 75–130. In M. D. McGinnis (Ed.), *Polycentricity and Local Public Economies*. Ann Arbor: The University of Michigan Press.

Persha, L., Agrawal, A., & **Chhatre, A.** (2011). Social and Ecological Synergy: Local Rulemaking, Forest Livelihoods, and Biodiversity Conservation. *Science*, *331*, 1606–1608. DOI: https://doi.org/10.1126/science.1199343

Poteete, A. R., Janssen, M. A., & **Ostrom, E.** (2010). Working Together. Princeton, NJ: Princeton University Press. DOI: https://doi.org/10.1515/9781400835157

Poteete, A. R., & **Ostrom, E.** (2004). Heterogeneity, Group Size and Collective Action: The Role of Institutions in Forest Management. *Development and Change*, *35*, 435–461. DOI: https://doi.org/10.1111/j.1467-7660.2004.00360.x

R Core Team. (2020). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Rana, P., & **Miller, D. C.** (2018). Machine learning to analyze the social-ecological impacts of natural resource policy: insights from community forest management in the Indian Himalaya. *Environmental Research Letters*. DOI: https://doi.org/10.1088/1748-9326/aafa8f

Rayfuse, R. (2015). Regional fisheries management organizations. *The Oxford handbook of the law of the sea*. DOI: https://doi.org/10.1093/law/9780198715481.003.0020

Ribot, J. C., Agrawal, A., & **Larson, A. M.** (2006). Recentralizing While Decentralizing: How National Governments Reappropriate Forest Resources. *World Development*, *34*, 1864–1886. DOI: https://doi.org/10.1016/j.worlddev.2005.11.020

Romero, M., & **Saavedra, S.** (2021). Communal property rights and deforestation. *The Journal of Development Studies*, *57*, 1038–1052. DOI: https://doi.org/10.1080/00220388.2020.1817394

Sapkota, L. M., Dhungana, H., Poudyal, B. H., Chapagain, B., & **Gritten, D.** (2020). Understanding the barriers to community forestry delivering on its potential: An illustration from two heterogeneous districts in Nepal. *Environmental management*, *65*(4), pp. 463–477. DOI: https://doi.org/10.1007/s00267-019-01224-0

Schlager, E., Blomquist, W., & **Tang, S. Y.** (1994). Mobile Flows, Storage, and Self-Organized Institutions for Governing Common-Pool Resources. *Land Economics*, *70*, 294–317. DOI: https://doi.org/10.2307/3146531

Singleton, S., & **Taylor, M.** (1992). Common Property, Collective Action and Community. *Journal of Theoretical Politics*, *4*, 309–324. DOI: https://doi.org/10.1177/0951692892004003004

Sproule-Jones, M. (2005). The Concept of Contingency and the scholarship of Elinor Ostrom on the Commons: Commentary on Tom Dietz's "the Darwinian Trope in the drama of the Commons". *Journal of Economic Behavior & Organization*, *57*, 231–235. DOI: https://doi.org/10.1016/j.jebo.2004.06.020

Strobl, C., Hothorn, T., & **Zeileis, A.** (2009). Party on! A New, Conditional Variable Importance Measure for Random Forests Available in the party Package. DOI: https://doi.org/10.32614/RJ-2009-013

Vallury, S., Shin, H., Janssen, M., Meinzen-Dick, R., Kandikuppa, S., Rao, K., & **Chaturvedi, R.** (2022). Assessing the institutional foundations of adaptive water governance in South India. *Ecology and Society*, *27*(1). DOI: https://doi.org/10.5751/ES-12957-270118

Varughese, G., & **Ostrom, E.** (2001). The Contested Role of Heterogeneity in Collective Action: Some Evidence from Community Forestry in Nepal. *World Development*, *29*, 747–765. DOI: https://doi.org/10.1016/S0305-750X(01)00012-2

Vedeld, T. (2000). Village politics: Heterogeneity, leadership and collective action. *Journal of Development Studies*, *36*, 105–134. DOI: https://doi.org/10.1080/00220380008422648

Wertime, M., Ostrom, E., Gibson, C., & **Lehoucq, F.** (2007). International Forestry Resources and Institutions (IFRI) Research Program: Field Manual. Version 13. Center for the Study of Institutions, Population and Environmental Change, Indiana University, Bloomington.

Whittaker, D., Crippen, A., Johnson, C., & **Janssen, M. A.** (2021). Social-ecological institutional fit in volunteer-based organizations: a study of lake management organizations in Vilas County, Wisconsin, USA. *International Journal of the Commons*, *15*(1). DOI: https://doi.org/10.5334/ijc.1059

Wolf, C., Levi, T., Ripple, W. J., Zárrate-Charry, D. A., & **Betts, M. G.** (2021). A forest loss report card for the world's protected areas. *Nature Ecology & Evolution*, *5*, 520–529. DOI: https://doi.org/10.1038/s41559-021-01389-0

Young, O. R. (2002). The institutional dimensions of environmental change: fit, interplay, and scale. Cambridge, MA: MIT press. DOI: https://doi.org/10.7551/mitpress/3807.001.0001

Young, O. R. (2008). The Architecture of Global Environmental Governance: Bringing Science to Bear on Policy. *Global Environmental Politics*, *8*, 14–32. DOI: https://doi.org/10.1162/glep.2008.8.1.14

TO CITE THIS ARTICLE:
Epstein, G., Apetrei, C. I., Baggio, J., Chawla, S., Cumming, G., Gurney, G., Morrison, T., Unnikrishnan, H., & Tomas, S. V. (2024). The Problem of Institutional Fit: Uncovering Patterns with Boosted Decision Trees. *International Journal of the Commons*, 18(1), pp. 1–16. DOI: https://doi.org/10.5334/ijc.1226

Submitted: 29 October 2022 **Accepted:** 27 October 2023 **Published:** 10 January 2024

COPYRIGHT:
© 2024 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See http://creativecommons.org/licenses/by/4.0/.

International Journal of the Commons is a peer-reviewed open access journal published by Ubiquity Press.