



Trustworthy Predictive Algorithms for Complex Forest System Decision-Making

Pushpendra Rana^{1,2*} and Lav R. Varshney³

¹ Indian Forest Service, Shimla, India, ² Affiliated Researcher, University of Illinois at Urbana-Champaign, Urbana, IL, United States, ³ Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, 314 Coordinated Science Laboratory, Urbana, IL, United States

OPEN ACCESS

Edited by:

Amy E. Duchelle,
Center for International Forestry
Research (CIFOR), Indonesia

Reviewed by:

Natalia Vasilievna Lukina,
Russian Academy of Sciences
(RAS), Russia
Swetha Peteru,
Center for International Forestry
Research (CIFOR), Indonesia

*Correspondence:

Pushpendra Rana
pranaifs27@gmail
orcid.org/0000-0001-8626-3351

Specialty section:

This article was submitted to
People and Forests,
a section of the journal
Frontiers in Forests and Global
Change

Received: 25 July 2020

Accepted: 10 December 2020

Published: 11 January 2021

Citation:

Rana P and Varshney LR (2021)
Trustworthy Predictive Algorithms for
Complex Forest System
Decision-Making.
Front. For. Glob. Change 3:587178.
doi: 10.3389/ffgc.2020.587178

Advances in predictive algorithms are revolutionizing how we understand and design effective decision support systems in many sectors. The expanding role of predictive algorithms is part of a broader movement toward using data-driven machine learning (ML) for modalities including images, natural language, speech. This article reviews whether and to what extent predictive algorithms can assist decision-making in forest conservation and management. Although state-of-the-art ML algorithms provide new opportunities, adoption has been slow in forest decision-making. This review shows how domain-specific characteristics, such as system complexity, impose limits on using predictive algorithms in forest conservation and management. We conclude with possible directions for developing new predictive tools and approaches to support meaningful forest decisions through easily interpretable and explainable recommendations.

Keywords: forest system complexity, limits to prediction, trustworthy algorithms, forestry, machine learning

INTRODUCTION

Algorithmic decision-making is becoming ubiquitous. Predictive machine learning (ML) assists humans in natural language processing, speech, and image recognition, as applied in many sectors including healthcare and law (Jordan and Mitchell, 2015; Gómez et al., 2018; Mueller et al., 2019). Predictive algorithms help humans take effective and consistent decisions, assist them in prioritizing led attention (Varshney, 2016) through useful insights (Appel et al., 2014), process large-scale data into usable form (Simon, 1996), and monitor and control events in real-time. Algorithmic decisions are sometimes cast as an alternative to vague, noisy, and biased human decision-making (Appel et al., 2014). By augmenting human ability, algorithmic decisions may save time, energy, and resources in public delivery of services (Mehta et al., 2013; Appel et al., 2014; Varshney, 2016). When we use the term *ML* in this paper, we always refer to predictive algorithms, even though they are only a subset of ML. See **Table 1** for a glossary of ML terms.

Prediction, however, is not a new idea in forestry and is considered a critical component of forestry science along with knowledge and understanding (Kimmins et al., 2005). Forestry scholars have used forecasting and scenarios-based analyses (Kimmins et al., 2005; Heinonen et al., 2017), growth and yield models (Amaro et al., 2003; Burkhart and Tomé, 2012), and individual-based forest gap models to predict forest succession, composition, and effects of changes in the environment on forests (Botkin et al., 1972; Purves et al., 2008). They have also used Bayesian network models to predict forest fires (Sevinc et al., 2020), Markov chain models to predict forest dynamics (Feldman et al., 2005), and multilevel nonlinear mixed models to predict forest growth

variables (Hall and Bailey, 2001). ML is a promising data-driven approach to prediction in forestry science, enabling better predictions about future forest states and assisting in forest decisions.

Inspired by success in other sectors, scholars have started exploring ML for forest conservation and management. Recent examples include predicting wildlife poaching (Gurumurthy et al., 2018), classifying drivers of global forest loss (Curtis et al., 2018), and predicting deciduous tree species composition using unmanned aerial vehicle multispectral data (Franklin and Ahmed, 2018). Other applications include identifying fire risk zones (Sakr et al., 2010; Rodrigues and de la Riva, 2014; Dutta et al., 2016), producing spatially explicit carbon stock maps to monitor forest-based climate change mitigation mechanisms such as reducing emissions from deforestation and forest degradation (REDD+) (Baccini et al., 2012; Mascaro et al., 2014), and detecting subtle changes in forests, forest types, and land use from hyperspectral imagery (Li et al., 2013; Curtis et al., 2018; Holloway and Mengersen, 2018). A variety of algorithms are used, including random forests and deep neural networks. Despite recent progress, however, there has been limited research on whether and to what extent predictive algorithms can assist decision-making processes in forestry.

This article reviews relevant ML studies in forestry to identify trends and patterns of existing literature in suggesting solutions to forest decision-making. Forest decision-making, in this review, includes all decisions related to management and conservation of forests, wildlife, and biodiversity. ML-based applications may assist forest managers, policymakers, and frontline forestry staff to make better decisions to protect forests and wildlife by providing subtle and deeper insights into the various dimensions of particular forest management decisions. We find that any meaningful endeavor to design ML applications may require algorithms appropriate to the domain-specific characteristics of forestry including scale-dependence of complex human-forest relationships (Moran and Ostrom, 2005), as well as system-level dynamics, interactions, feedback loops, nonlinearities, surprises, and unintended consequences (Liu et al., 2007; Ostrom, 2009; Hofman et al., 2017).

Forest decision-making is context-specific and influenced by power and regulatory structures, incentives, and professional norms. Many factors lead to an observed forestry outcome, and the importance of these factors and the interactions among them vary across different contexts (Fleischman, 2014). Competing land uses such as carbon storage, livelihoods, biodiversity conservation, and timber harvesting, as well as heterogeneous stakeholders further complicates forest decision-making in the face of difficult tradeoffs among these multiple uses (Chhatre and Agrawal, 2009; Persha et al., 2011) and user groups. Moreover, there are considerable limits to predictability (Liverman and Cuesta, 2008). We argue that such complexity in forestry systems and decision-making has limited the use of prediction in forestry and necessitates new tools and approaches to support meaningful forest decisions through easily interpretable and explainable recommendations (Hofman et al., 2017; Mueller et al., 2019; Selbst et al., 2019; Salganik et al., 2020) that engender trust.

Drawing on a traditional review method (Jesson et al., 2011), we identify relevant studies that use ML approaches in forest decision-making. These studies include primary studies and reviews. We searched in Google Scholar and arXiv for scholarly articles using keywords including combinations of “forests,” “forest decision making,” “forest management,” “wildlife,” “biodiversity,” and “forest conservation” with “machine learning” or “artificial intelligence” or “predictive algorithms.” We also searched for “accountability” or “fairness” or “interpretability” with “machine learning” or “artificial intelligence” or “predictive algorithms.” The review was open to all years of publication and not delimited to a select time period. The latest search was done in April-May, 2020. We further examined the citation list of published reviews covering various aspects of machine learning and artificial intelligence in various fields and selected relevant articles to expand our list of articles for this review. We found 81 articles as a result of our searches and all of them were reviewed.

Based on broader literature in forest management, we identified three critical dimensions related to forest decision-making: (a) forest system complexity, (b) interpretability, and (c) fairness and justice, which require adequate consideration in expanding the use of ML in forest management decision support. We categorized the reviewed papers into these three classes and synthesized their findings within each category through narrative synthesis to examine evidence on the limitations of predictive algorithms in forest conservation and management.

After characterizing the crosscutting challenges inherent in the forestry sector relating to the use of predictive algorithms in forest decision-making, the review concludes by highlighting some especially promising research frontiers for ML in forest decision-making. Given our focus on synthesizing available evidence on ML in forest conservation and management, we do not claim that this review comprehensively covers the literature on ML applications in forestry sciences, though we believe it does provide an accurate depiction of current trends.

PREDICTIVE ALGORITHMS AND FOREST DECISION-MAKING: CRITICAL DIMENSIONS

ML scholars have started developing a range of applications based on predictive algorithms to assist forest decision-making. Our review showed a recent increase in the number of these applications with work focused on supervised learning rather than other forms of machine learning such as unsupervised, semi-supervised, or reinforcement learning. ML scholars have used a range of approaches to assist forest decision-making (please refer to **Table 1** for definitions of some of these ML approaches and terms).

Based on broader trends in these studies, we identified three major dimensions related to forest systems that impose limitations on the use of predictive algorithms in forest decision-making. **Table 2** provides a comprehensive view of critical dimensions to support forest decision-making, possible strategies, and limits of predictive algorithms to support these strategies due to cross-cutting challenges. Papers listed in the last

TABLE 1 | Terms and approaches in machine learning (definitions and references)*.

	Terms and approaches	Definition
Machine learning Terms	Machine learning	"Machine learning is the science (and art) of programming computers so that they learn from data" (Géron, 2019)
	Predictive algorithms	Predictive algorithms are subset of machine learning
	Regression problem	ML problems with real-valued outcome/response
	Classification problem	ML problems with categorical outcome/response
	Bias	Systematic error in machine learning models
	Variance	The amount by which a predicted estimate would change if a different training dataset is used (James et al., 2013)
	Overfitting	More complex models can closely fit a given dataset and therefore, overfit the data. It means these models follow the training data too closely and may not generalize in the field (James et al., 2013)
	Predictive accuracy	Measures to "quantify the extent to which the predicted response value for a given observation is close to the true response value for the observation." (James et al., 2013). E.g., mean squared error (MSE) in the regression setting
Problem classes	Supervised learning	Involves fitting a model that relates the outcome/response to the predictors with the objective to accurately predict the outcome/response for future observations (prediction) or to understand the relationship between the predictors and the response (inference) (James et al., 2013)
	Unsupervised learning	We have measurements on a set of variables but no associated outcome/response variables. Under such conditions, we aim to learn relationships between the variables or between the observations to understand the structure of the data. E.g., includes market segmentation study where customers are clustered as per their spending patterns (James et al., 2013)
	Semi-supervised learning	We only have outcome/response variables for a subset of total number of observations, though we have measurements for predictor variables in all the cases. Semi-supervised learning aims to use both the observations for which we have outcome/response variables and also, for which we do not have outcome/response variables (James et al., 2013)
	Reinforcement learning	"An agent is placed in an environment and must learn to behave successfully therein" by sequentially interacting and optimizing a reward function through an action policy (Russel and Norvig, 2013)
Algorithms	Supervised learning	
	Decision trees	Decision trees involve stratifying or segmenting the predictor space into number of simple regions in a hierarchical manner according to an appropriate split criterion (James et al., 2013)
	Random Forests	The method involves considering several decision trees simultaneously. Bootstrapped training samples are used to build a number of decision trees, and each time a split is considered, we select a random sample of m predictors as split candidates out of a complete set of p predictors (James et al., 2013)
	Naïve Bayes Classifier	A simple probabilistic classifier based on Bayes theorem that is well suited to situations with a large number of features (Lewis, 2017)
	Support Vector Machines	"Support Vector Machine produces nonlinear boundaries by constructing a linear boundary in a large, transformed version of the feature space" (Hastie et al., 2009)
	Neural networks	The method aims at extracting layers of "linear combinations of the inputs as derived features, and then model the target as a nonlinear function of these features" (Hastie et al., 2009). Examples architectures include multilayer perceptrons, Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN)
	Deep neural networks	Neural networks with many layers, also referred to as deep learning
	Ensemble learning	Ensemble learning involves two steps: "developing a population of base learners from the training data, and then combining them to form the composite predictor." (Hastie et al., 2009)
	Unsupervised learning	
	Clustering methods	Broad set of techniques used to finding clusters or subgroups in a given data set (James et al., 2013)
	K-means clustering	Here, we aim to partition a particular dataset into pre-specified number of clusters by minimizing mean squared error (James et al., 2013)
	Hierarchical clustering	Clustering method where we do not know how many clusters we need in the beginning; we get a dendrogram allowing us to visualize clusterings obtained for each possible number of clusters (James et al., 2013)

(Continued)

TABLE 1 | Continued

Terms and approaches	Definition
Semi-supervised learning	
Semi-supervised learning	Some of the algorithms used in semi-supervised learning include generative methods, graph-based models, semi-supervised support vector machines, mixture models, self-training and co-training algorithms (Chapelle et al., 2010).
Reinforcement learning	
Q-Learning	"A Q-learning agent learns an action-utility function, or Q-function, giving the expected utility of taking a given action in a given state" (Russel and Norvig, 2013)

*This is not an exhaustive list of all the machine learning terms or approaches. Please refer to cited references and other ML literature for further information.

column of **Table 2** represent only selected papers to highlight key limits of predictive algorithms in supporting forest decision-making. Below we detail each of these three dimensions: (a) forest system complexity, (b) interpretability, and (c) fairness and justice in algorithms, and synthesize relevant studies under each dimension through narrative syntheses to generate evidence on the limits of prediction in forest decision-making.

Forest System Complexity and Algorithmic Decision-Making

Forests are complex social-ecological systems with system-level dynamics, interactions, and feedback loops (Liu et al., 2007; Ostrom, 2009; Hofman et al., 2017). Social and ecological variables in forestry systems show dynamic, nonlinear growth or relationships with other variables, and have thresholds where their value and plausible impact on the outcome of interest change direction. Interventions in forest ecosystems involve synergies and tradeoffs among multiple objectives that unfold at multiple scales (Persha et al., 2011), experience regular surprises and unintended consequences (Liu et al., 2007), and often have social and ecological processes operating at different scales (Baylis et al., 2016). The presence of time lags in social-ecological impacts of conservation interventions (Miller et al., 2017), varying degrees of resilience in social-ecological systems such as forests (Liu et al., 2007), and difficulty in modeling human behavior (Hofman et al., 2017; Salganik et al., 2020) further leads to poor understanding of these systems.

Forest managers can make sound decisions only if they understand problems, have access to relevant information, and know how to use this information, which is challenging in the face of complex forest dynamics. Any decision support in forest management requires understanding the type, scale, and depth of available information and knowledge about forest systems (Stock and Rauscher, 1996). Moreover, given large uncertainty in human behavior (Nishant et al., 2020), it is important to understand how people affected by a particular forest management application will react. In the absence of cognitive support, forest managers largely make decisions relying on subjective values, individual preferences, perceptions, and expectations. This results in incoherent decisions that do not meet mutually-agreed-upon standards (Stock and Rauscher, 1996). Developing algorithms to support such multifarious decision-making is challenging.

Forestry decisions may be especially poor for problems that are not uniformly perceived, prioritized, or processed (Ordóñez

et al., 2020). For example, government officials perceive forestry threats differently than others (Yousefpour et al., 2017), and therefore collect forestry data in a particular way that reflects their biases. This may result in incomplete, inaccurate, and biased data, which may limit its utility in decision support systems. Moreover, there are a range of drivers and processes that determine forestry decisions (Fleischman, 2014). For example, external donors and environmental NGOs influence forestry decision-making around the world at the community and national levels (Ayana et al., 2018). ML-based predictions of forest growth in a changing climate may not be useful, as they rely on forest growth and yield models that abstract highly complex and nonlinear forest systems into simplistic forms (Ashraf et al., 2015). Similarly, ML-driven decisions to support wildlife habitat management may not be useful in areas with high faunal and floral diversity, since there is substantial forest complexity and dynamic relationships among species (Gonzalez et al., 2016). We argue that the complexity of forest systems and the factors that influence decisions require any predictive algorithm to incorporate domain-specific characteristics of forests. However, this may be daunting.

Limits to Prediction in Forest Systems

Due to dynamic, nonlinear relationships and thresholds, there may be uncertainty and limits to predictability, leading to ineffectual ML models (Gonzalez et al., 2016; Rey et al., 2017; Gholami et al., 2019). For example, Rana and Miller (2019) show that vegetation growth as measured by NDVI (Normalized Difference Vegetation Index) follows a nonlinear trend in Kangra district in northern India. Moreover, in the same district, different forest management units show different NDVI outcome trajectories, driven by varying social-ecological attributes and pathways, emphasizing that couplings between people and nature vary across space, time, and organizational units (Liu et al., 2007; Rana and Miller, 2019). As another example, increased variability in the biomass of big trees caused loss of accuracy and bias, under ML models that estimate biomass of tropical forest trees (Montano et al., 2017).

Even large datasets fail to capture the full range of outcomes due to geographic complexities, which can affect the class accuracies of a predictive algorithm. For example, despite large sample sizes comprising millions of 10 km × 10 km grid cells around the world and training sample cells ($n = 5,000$), Curtis et al. (2018) noticed considerable regional variation in accurately

TABLE 2 | Critical dimensions of forest decision-making, possible strategies and limits of predictive algorithms to support these strategies.

Critical dimensions of forest decision-making	Possible strategies to support forest decision-making	Short description of the limits of predictive algorithms identified in the review	Studies describing limits of predictive algorithms
System complexity	Model system-level dynamics, interactions and feedback loops	Predictive algorithms fail to capture system-level dynamics, scale-level dependence of human-forest interactions, social-ecological interactions, and feedback loops inherent in forest systems	Thompson et al., 2012; Varshney, 2016; Selbst et al., 2019; Gonzalez et al., 2016; Kroll et al., 2016; Hofman et al., 2017 Struss, 2004; Ashraf et al., 2015; Norouzzadeh et al., 2018; Debeljak et al., 2001; Ye et al., 2019; Rao et al., 2020
	Include social actors, institutions and broader context in decision-support systems	Algorithms often miss or simplify complex social-ecological contexts and diverse set of social actors and institutions found in forestry contexts	Rodrigues and de la Riva, 2014; Dutta et al., 2016; Hofman et al., 2017; Holloway and Mengersen, 2018; Mueller et al., 2019; Selbst et al., 2019; Salganik et al., 2020
	Model synergies and tradeoffs, surprises and unintended consequences, non-linear relationships, time lags	Failure of predictive algorithms to model domain characteristics of forest systems including dynamic and non-linear growths, thresholds, surprises, time lags, unintended characteristics and prevalence of synergies and tradeoffs among multiple objectives	Thompson et al., 2012; Varshney, 2016; Hofman et al., 2017; Selbst et al., 2019
	Understand human perceptions, behavior and attitudes	Failure of predictive algorithms to model human behavior, perceptions and attitudes	Dutta et al., 2016; Hofman et al., 2017; Nguyen et al., 2016; Fang et al., 2017
	System factors and outcome complexity due to regional and ecological variations	Predictive algorithms fail to model inherent complexity and variability in forest system factors and outcomes due to regional and ecological variations	Curtis et al., 2018; Franklin and Ahmed, 2018; Hethcoat et al., 2019
Interpretability	Usable and explainable recommendations	Data experts design algorithms purely as technical problems resulting in unusable and unexplainable recommendations in forest decision-making	Wagstaff, 2012; Padarian et al., 2020
	Inclusion of social and technical context while designing algorithms	Predictive algorithms fail to capture social and technical contexts and make simplistic assumptions about social actors, institutions, and their interactions	Wagstaff, 2012; Dutta et al., 2016; Mueller et al., 2019; Selbst et al., 2019
	Interpretation of ML results in specific contexts to support decision-making	Little scholarly tradition within ML community to interpret results in their specific socio-economic and political contexts narrows model interpretability	Aertsen et al., 2010; Wagstaff, 2012; Mueller et al., 2019
	Uniform model-based predictions to support a given decision	Predictive models lack uniformity in their predictions. For the same set of input features and prediction tasks, complex models can generate multiple accurate models with varying details of explanations	Adadi and Berrada, 2018; Hall and Gill, 2018
	Robust and verified unique causal solutions to a given problem	Predictive algorithms are only evaluated by their predictive success and are not optimized to answer causal questions	Drake et al., 2006; Aertsen et al., 2010; Nunes and Görgens, 2016; Pearl and Mackenzie, 2018
	Full understanding of how predictive algorithm is making decisions	Black-box nature of many ML algorithms make it difficult for humans to understand their decisions	Naidoo et al., 2012; Mascaro et al., 2014; Kar et al., 2017; Mueller et al., 2019
	Big, accurate and appropriate data to support interpretable decisions	Lack of data, class imbalance, data sparsity, noise in data quality and presence of spatial and temporal correlation further limits the development of interpretable ML models in forest management	Lippitt et al., 2008; Ali et al., 2015; Curtis et al., 2018; Franklin and Ahmed, 2018; Gurumurthy et al., 2018; Gholami et al., 2019; Hethcoat et al., 2019; Molnar, 2019; Bland et al., 2015; Kar et al., 2017; Debeljak et al., 2001; Ashraf et al., 2013; Kuiper et al., 2020
Fairness and justice	Fairer and just predictive algorithms that do no potential harm or perpetuate past injustices	Predictive ML models can be unfair and unjust and can perpetuate past injustices in forest management	O'neil, 2016; Corbett-Davies and Goel, 2018; Angwin et al., 2019; Selbst et al., 2019
	Reliable and unbiased recommendations	ML based predictions can be highly unreliable, biased and inaccurate	Kroll et al., 2016; Angwin et al., 2019; Gholami et al., 2019; Gillingham, 2019
	Accurate predictions to support decision-support	Large-scale uncertainty and variation in predictive accuracy of ML models	Lippitt et al., 2008; Li et al., 2013; Mascaro et al., 2014; Kroll et al., 2016; O'Connor et al., 2017; Dressel and Farid, 2018; Gholami et al., 2019; Gillingham, 2019; Rey et al., 2017

(Continued)

TABLE 2 | Continued

Critical dimensions of forest decision-making	Possible strategies to support forest decision-making	Short description of the limits of predictive algorithms identified in the review	Studies describing limits of predictive algorithms
	Incorporate broader system dynamics, complete data and notions of fairness	ML experts often narrowly focus on readily available limited-use and uncertain data, and simplistic notions of fairness	Kroll et al., 2016; Selbst et al., 2019
	No discrimination on grounds of gender, health and ethnicity	Predictive algorithms can discriminate on grounds of gender, health or ethnicity	Hajian et al., 2016
	Do not use proxies to assess social-economic status	Understanding socio-economic position of an individual based on census zip code or other proxies and then judge his or her suitability for job, loan or conservation program can be discriminatory	O'neil, 2016
	Standardized predictive algorithms with common reporting and meaningful quantitative metrics to avoid subjectivity and bias	Lack of common reporting and meaningful quantitative metrics to evaluate ML models introduce subjective and bias in ML-based recommendations	Wagstaff, 2012; Hofman et al., 2017
	Inability of elites to manipulate decisions using their power	Local powerful elites may control the development and deployment of algorithm to serve their vested interests at the cost of the larger public	Selbst et al., 2019

classifying drivers of global forest loss due to insufficient distinction of one land use and management category from another, as well as sparse training data in certain classes. There can also be extremely imbalanced data that makes accurate prediction difficult, e.g., in developing an ML pipeline to identify areas at high risk of poaching in the protected areas of Uganda (Gholami et al., 2019). The size of raw data can also be problematic, e.g., in recent attempts to monitor audio signals of African elephants with real-time ML methods to offer protection against poachers, network bandwidth limitations required efficient audio compression (Bjorck et al., 2019).

Quantifying uncertainty is also difficult. For example, in using deep learning to project Australia's forest cover dynamics, it was difficult to make uncertainty projections due to the large number of model parameters (Ye et al., 2019). Transfer of models trained with particular sets of conditions in a given forest system to a new system with different kinds of conditions is difficult (Hart et al., 2019). Similarly, transferring computer vision models for classifying animal species in camera trap images trained in one region to another is difficult due to the presence of previously-unseen species (Beery et al., 2019).

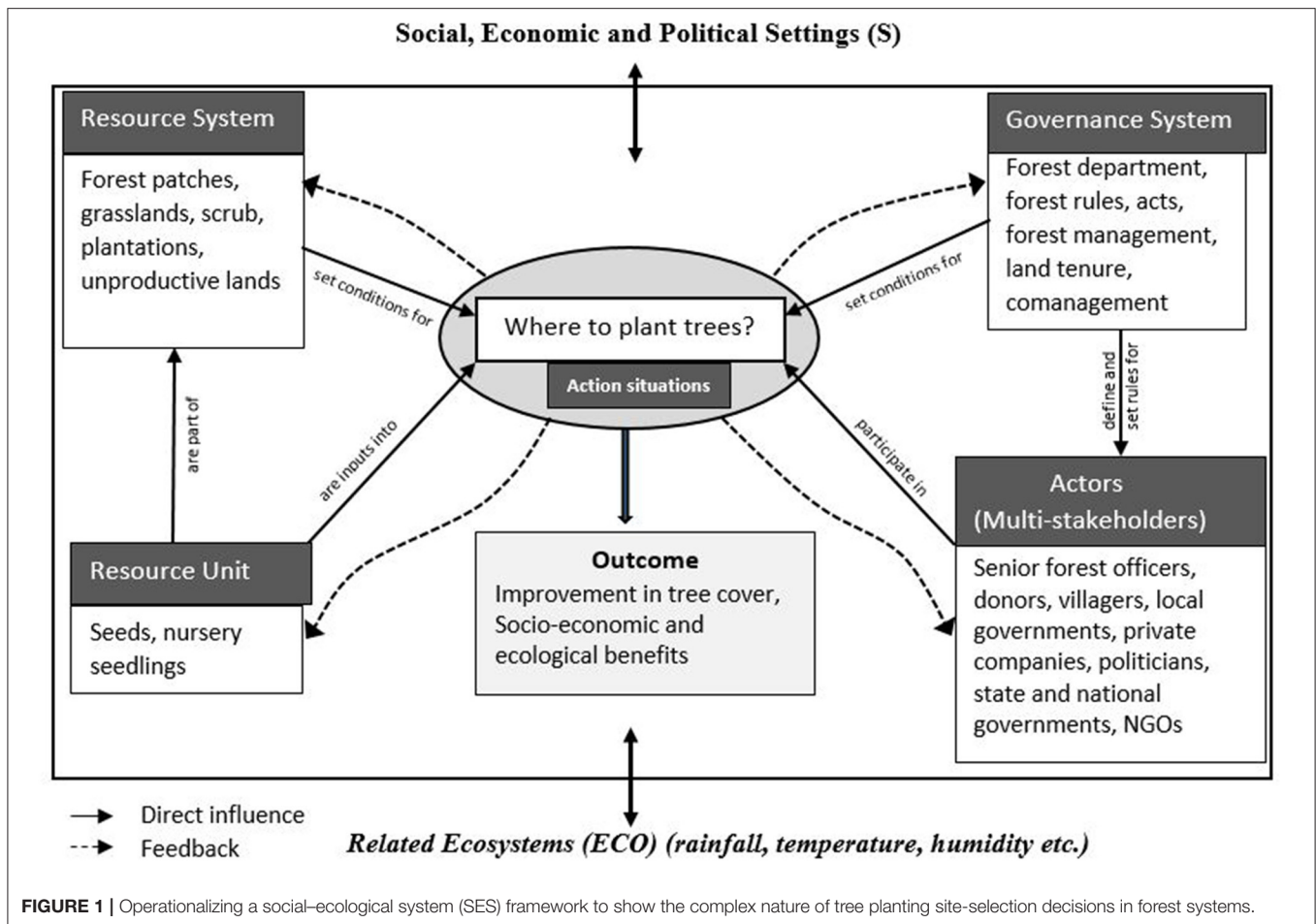
Simplification of Broader Contexts Limits ML Effectiveness in Forest Decision-Making

Predictive algorithms may face difficulty in encompassing all system-level processes such as spatial and temporal dynamics, interactions, and feedbacks (Struss, 2004; Liu et al., 2007; Gonzalez et al., 2016; Kroll et al., 2016; Rey et al., 2017; Curtis et al., 2018; Hethcoat et al., 2019). System-level information should include all factors that operate at the subsystem level (actor, governance systems, resource system, resource unit, and external influences) to influence a particular outcome (Ostrom, 2009; Bland et al., 2015). For example, the decision of private forest landowners to harvest trees from their farms is based on several factors including actor-level (e.g., education, age,

and income), resource system (farm size), location (distance, elevation, and slope), and market (timber prices) (Silver et al., 2015; Snyder and Kilgore, 2018). Usually, predictive algorithms model a component of social-ecological systems while ignoring the social actors, institutions, and interactions within these systems, resulting in the elimination of larger context (Rodrigues and de la Riva, 2014; Dutta et al., 2016; Holloway and Mengersen, 2018; Selbst et al., 2019). Missing such factors, interactions, and feedback loops may abstract the larger context into a simple model that provides inadequate decision-support in forest and wildlife management (Holloway and Mengersen, 2018; Selbst et al., 2019). We present below two decision-making contexts, tree-planting and wildlife management, to emphasize this point further.

First, let us consider how ignoring broader contexts in tree planting can restrict its effectiveness as a natural climate solution (Figure 1) (Bastin et al., 2019). We have operationalized a social-ecological system (SES) framework to show the complex nature of tree planting site-selection decisions. The decision problem is where to plant trees in a landscape. This is complicated by the presence of multiple stakeholders, diverse governance and resource system contexts, along with interactions and feedbacks involved in tree-planting site selection. Various rules, acts, and cultural norms of forest department, scheme-specific planting and budgetary guidelines, land tenure rights, and participatory provisions govern these planting decisions. Selected enclosures for growing trees are part of the resource system (forests, grasslands, plantations, unproductive lands). The availability of blank patches, site quality constraints, and socio-economic factors set conditions for tree planting site selection decisions (Rana and Varshney, 2020).

The final tree planting decisions on forests and or other types of lands determine the future social and ecological outcomes, which then influence each of the sub-systems (governance systems, actors, resource system and resource unit) through



positive or negative feedback. For example, the local community may be using the planting site for grazing animals, organizing village events, or cultivating crops. Such interactions are highly dynamic, unpredictable, and influenced by seasonality and the changing livelihood needs of local communities. Also, there is a high likelihood that overgrazing in that planting site or planting a community-unfriendly tree species may lead to extensive forest degradation. Moreover, growing a tree plantation into a secondary forest usually takes 15 or more years. Due to this inherent temporal uncertainty in forestry outcomes, it is difficult for any data scientist to appropriately model such time lags and capture all interactions, feedback loops, and dynamics over the longer-term (Thompson et al., 2012).

Second, let us consider how it is difficult for data scientists to capture complex and dynamic contexts inherent in wildlife management while designing any ML-based applications. As an example, wildlife managers use habitat suitability models for individual wild animals to decide how to protect them. ML models can help create more general suitability models. However, developing such models to scale up to the population- or a landscape-level comprising multiple faunal and floral species is difficult due to their complex interactions (Debeljak et al., 2001). Modeling such complex wildlife systems requires

information about forest structure, tree species composition, herbal and shrub layers, plant species distribution, habitat management activities, and feeding places of animals. Moreover, animal tracking must identify sex, seasonality, and diurnal characteristics to improve modeling outcomes (Debeljak et al., 2001). A recent ML application to support ranger patrol strategies in Zimbabwe to protect elephants from poachers required participatory modeling processes and accounted for observer bias in modeling through robust and regular data collection of patrol efforts at a relevant scale (Kuiper et al., 2020). However, obtaining such complex and dynamic data to build useful ML models to support wildlife decisions is difficult and may not be prioritized by field staff with other pressing responsibilities.

Interpretability and Algorithmic Decision-Making

Clear and easily explainable information is a must for forest decision-making: it may reduce the trust deficit between stakeholders and forest officials for particular forest conservation and management tasks, especially in the context of prevailing mistrust between foresters and local communities in several parts of the world (Springate-Baginski and Blaikie, 2013). On

the other hand, uninterpretable decision support systems restrict the ability of forestry officials to persuade local communities and other stakeholders to support suitable forestry decisions. Forest managers and forest users are likely to feel alienated from decision-making processes without interpretability. Indeed, diffuse, inscrutable, and non-intuitive information can result in poor forestry outcomes. For example, if an ML tree planting support system fails to provide easily understood information to a forest official on why a particular site is preferred for growing trees, that forest guard will either ignore it or follow it unconvincingly, resulting in poor planting decisions. This would not only lead to poor survivorship of tree plantations, but also wasteful expenditure. There are several reasons predictive algorithms fail to produce interpretable explanations as detailed below.

Experts Design Algorithms as a Purely Technical Exercise

Experts design ML applications in isolation from the local social context, casting them as purely technical problems that end up yielding unusable and inexplicable recommendations (Wagstaff, 2012). As per a recent estimate, only 1% of ML papers interpret results in their specific contexts, as these interpretations are hard to make and further, there is little scholarly tradition within this field for reporting such interpretations (Wagstaff, 2012).

Limits to Designing Interpretable Algorithms Due to System Complexity

The complexity in social-ecological systems, such as forests, further restricts the ability of developers to produce standardized and interpretable algorithms (Hofman et al., 2017; Norouzzadeh et al., 2018; Ferraro et al., 2019; Mueller et al., 2019; Selbst et al., 2019). Many algorithms are based on simplistic assumptions about social actors, institutions, and their interactions, and may not serve forest officials due to their model choices (Rodrigues and de la Riva, 2014; Dutta et al., 2016). For example, an early fire detection model relied only on weekly climatic data and assumed humans have little influence on fire occurrences in Australia (Dutta et al., 2016). Often, algorithm developers, let alone policymakers, do not understand the mechanistic reasoning an algorithm has come up with, reasons behind certain assumptions about social-ecological systems, or choices of tuning and regularization parameters. For example, while using deep learning for wildlife species identification and counting from camera trap images, further explanation into choices made by data scientists in picking certain *hyperparameters* may improve modeling outcomes and their better understanding (Norouzzadeh et al., 2018). All of these phenomena make it difficult for people directly affected by implementing decisions from such algorithms to trust them or even to take appropriate decisions to support forest conservation and management (Mueller et al., 2019).

ML Models Often Lack Transparency, Restricting Their Deployment in Forest Decision-Making

ML models are often considered as black-box models that have highly entangled input features, which make their disaggregation into human understandable form difficult (Naidoo et al., 2012). For the same set of input features and prediction tasks, complex ML models can generate multiple accurate models with varying details of explanations (Adadi and Berrada, 2018). Simpler models on the other hand may find some variables as important predictors (Rodrigues and de la Riva, 2014) rather than incorporating the broader social-ecological context, which is necessary for meaningful ML-based decisions. Moreover, the absence of causal pathways from inputs to outputs in ML applications restricts their ecological interpretability and therefore limits their adoption in forestry decision-support systems (Drake et al., 2006; Aertsen et al., 2010; Nunes and Görgens, 2016). More importantly, causal relationships between predictors and outcomes may elude such algorithms that are only evaluated by predictive success and not optimized to answer causal questions. The problems of data overfitting further narrow the interpretability of such models, making their use in forestry decision-making difficult (Aertsen et al., 2010). Mascaro et al. (2014) find overfitting and spatial correlation of model errors as limitations of their model to map tropical forest carbon by upscaling LiDAR-based carbon estimates (Mascaro et al., 2014).

The lack of accurate and adequate data in forestry further limits developing interpretable models (Lippitt et al., 2008; Kar et al., 2017; O'Connor et al., 2017; Curtis et al., 2018; Franklin and Ahmed, 2018; Gurumurthy et al., 2018; Gholami et al., 2019; Hethcoat et al., 2019). Scholars have noticed significant class imbalance, sparsity, and noise in the patrolling datasets they use in predicting wildlife poaching (Bland et al., 2015; Kar et al., 2017; Gurumurthy et al., 2018; Gholami et al., 2019). They also identified geographic and language barriers in collecting and synthesizing data for forest conservation decisions (Gurumurthy et al., 2018). The absence of high-quality data (and lack of computing power and black-box nature of deep learning) is problematic in modeling the physical properties of forest ecosystems, as noticed during forest damage assessment in Bavaria, Germany (Hamdi et al., 2019). Moreover, while exploring convolutional neural networks to analyze biodiversity, lack of adequate training examples in existing datasets was a critical challenge that reduced model performance (Rodner et al., 2015).

Spatial and temporal correlation in data can limit the performance of ML prediction models as observed by Ashraf et al. (2013) while using NN-based growth model to predict volume increment of individual trees (Ashraf et al., 2013). Using thousands of fuel moisture content measurements, a state-of-the-art physics-assisted recurrent neural network model for Live Fuel Moisture Content (LFMC) failed to capture spatial and temporal variability of the outcome (Rao et al., 2020). Many parametric models work well with small datasets and yield interpretable results but are hard to automate and are not flexible. On the other hand, complex ML models may require a lot of data to distill knowledge in the form of interpretable suggestions. This suggests

that in the absence of large-scale forestry data, algorithms may be limited in producing insights for effective forest decision-making (Gholami et al., 2019; Hethcoat et al., 2019; Molnar, 2019) without new algorithmic developments (Yu et al., 2019).

Fairness and Justice and Algorithmic Decision-Making

Scholars have found that predictive ML algorithms often deliver inaccurate, unfair, or unjustified results (Li et al., 2013; Mascaro et al., 2014; Kroll et al., 2016; Franklin and Ahmed, 2018), discriminate on the grounds of gender, health, or ethnicity (Hajian et al., 2016), and fail to learn and adapt to changing circumstances (Mueller et al., 2019). Results from other fields indicate that predictive ML models may be unfair and unjust. For example, studies of a widely used criminal risk assessment tool showed that its predictions were racially biased (Angwin et al., 2019) and not more accurate than predictions made by a person with little or no criminal justice expertise (Dressel and Farid, 2018). Predictive algorithm-based decision support systems for child and social work also indicate many problems especially related to the accuracy of the data, algorithms, and proposed decisions (Gillingham, 2019). ML experts fail to deeply consider fairness in forest decision-making owing to their narrow focus on readily available limited-use data, neglect of broader system dynamics, and incorporation of only simplistic notions of fairness (Selbst et al., 2019).

Lack of Data Restricts Development of Fair and Just ML Applications

Datasets used in ML applications for forest conservation and management lack critical socio-economic, political, and biophysical dimensions related to forest decisions. In many cases, due to lack of detailed data on human behavior, proxies such as zip code or language patterns are used to approximate socioeconomic position of an individual and then judge her suitability for a job, loan, or conservation program. Understanding relationships with such simplistic correlations may be discriminatory (O'neil, 2016). The available data products used in ML research are inherently uncertain due to error propagation when combining multiple sources of data, modeling relationships, extrapolating to new locations, or making educated guesses about variables of interest (Kugler et al., 2019). Moreover, there is little guidance on how knowledge related to historically disadvantaged social groups such as indigenous forest peoples and women can be integrated in these data systems given the abstraction traps inherent in ML-based applications, potentially leading to unfair outcomes in forest decision-making (Selbst et al., 2019).

Data scientists developing algorithms are often completely unaware of the importance and interplay of various social-economic and political factors that influence forest decision-making. Hidden power dynamics and structures, vested economic interests, and social biases are widespread in the forestry sector, with powerful elites controlling decision-making processes to serve their objectives at the cost of forests and communities (Persha and Andersson, 2014; Rana, 2014). Without extra care in developing and deploying algorithmic

support, elites may alter algorithms to serve their objectives at the cost of the larger public (Selbst et al., 2019). This negatively impacts international goals aimed at ending poverty, hunger, and other forms of social and economic discrimination. Biased algorithms may fail to support poor forest-dependent communities if algorithms do not selectively include a wider range of concerns from these groups or incorporate elements of fairness and justice based on a system-level understanding of forestry contexts especially in developing countries (Selbst et al., 2019). Without considering these biases in any ML effort, it is not possible to achieve fair and just decision support in the forestry sector.

Neglect of Broader System-Level Contexts May Lead to Unfair and Unjust Algorithms

ML scholars often treat model development as an independent activity wherein only model parameters, inputs, and outputs matter. Such an approach omits broader system-level contexts from modeling efforts, and fails to produce fair and just machine learning algorithms (Rodrigues and de la Riva, 2014; Dutta et al., 2016; Selbst et al., 2019). Justice and fairness are properties of social systems and so measuring such concepts through simple metrics at the level of technical subsystems (ML algorithms) may lead to unethical and erroneous algorithms devoid of any meaningful insights into forest decision-making. Narrowing down broader concepts of justice and fairness to narrow technological tools leads to five major abstraction traps in modeling efforts (Selbst et al., 2019). These include failure to model the entire system where a fair concept is intended to be applied (framing trap), transferring one algorithmic solution developed in one social context to a different one (portability trap), simplifying fairness concepts (formalism trap), poor understanding of how an algorithm changes human behavior (ripple effect trap), and believing that algorithms provide solutions to all problems (solutionism trap) (Selbst et al., 2019). These traps potentially limit the use of ML algorithms in solving problems in forestry where concepts of fairness and justice are complex and multi-dimensional.

Lack of Standardization May Increase Bias and Uncertainty in Algorithmic Decision-Making

The absence of common reporting and lack of meaningful quantitative metrics to evaluate ML models are some of the critical factors that may lead to their limited adoption in forest decision-making. Individual researchers' decisions on the selection of questions, data, model, and evaluation metrics may result in a high level of subjectivity and bias, and therefore, a failure to replicate results (Drake et al., 2006; Hofman et al., 2017). Moreover, abstract metrics used in ML such as classification accuracy, R^2 (coefficient of determination), RMSE (root mean squared error), and AUC (areas under receiver operating characteristic) may not correspond to impact of forest conservation and management interventions (Wagstaff, 2012; Hofman et al., 2017; Gholami et al., 2019). Moreover, not only do many users fail to decipher decisions made by algorithms but even the developers often fail to understand how their system works (Mueller et al., 2019). These findings suggest that ML

applications can have high model uncertainties (Gholami et al., 2019), and may lead to biased and unfair outcomes in the forestry sector.

DISCUSSION AND FUTURE DIRECTIONS

This review examines current trends in the use of ML applications in forest conservation and management. As evident from this review, ML can assist forest decision-making by characterizing numerous aspects of the contexts that shape forest decision-making or other social phenomenon by bringing forth new plausible hypotheses, patterns, and relationships, which are not readily apparent to social scientists or practitioners. ML algorithms are also quite valuable in exploring complex and composite patterns, identifying new features to model human-environment interactions, which are not easily discernible (National Research Council, 1998).

To realize the full potential of ML applications in forest management, this review calls for addressing three critical challenges that restrict the widespread adoption of ML in forest conservation and management: complexity, justice, and interpretability. First, any meaningful forest decision support system based on ML must characterize limits on prediction in complex, uncertain, and dynamic social-ecological systems such as forests (Hofman et al., 2017; Mueller et al., 2019; Salganik et al., 2020). Second, any ML application must maximize the chance of reducing potential social harm and achieving fairness rather than perpetuating past injustices associated with forest conservation and management practices (Corbett-Davies and Goel, 2018). Third, in the future, the adoption of predictive algorithms in forestry will depend on how interpretable and explainable such algorithms are to local forest officials and the general public (Holloway and Mengersen, 2018). The review further provides promising future directions for ML-based predictive algorithms to support forest decision-making.

Characterize Limits on Prediction in Complex, Uncertain, and Dynamic SES Systems

There should be research on incorporating system-level attributes, interactions, and feedback loops in any prediction endeavor in forest decision-making (Struss, 2004; Holloway and Mengersen, 2018; Selbst et al., 2019). In socio-ecological systems such as forests, there could be system level dynamics, regular surprises, unintended consequences, interactions, and feedback loops, which may lower the theoretical best performance of a given model (Liu et al., 2007; Hofman et al., 2017). In this scenario, we must devise interventions that do not require accurate predictions. Under conditions where ML models perform much below the theoretical limits, it is advisable to lower the expectations about the success of the proposed algorithmic decisions in terms of predictive accuracy accordingly (Hofman et al., 2017).

Although, how to define a theoretical limit to predictive accuracy in a given complex system, such as forests, is still under debate, scholars have advocated use of better data and model

classes with more informative features to construct models. For example, if a hypothesized mechanism driving a particular outcome in a forest decision-support system explains less observed variance than the theoretical limit, it is apparent that other likely mechanisms must be identified. On the other hand, if the outcomes in forest systems are intrinsically unpredictable (theoretical limit is low), our expectations about the utility of the suggested ML-algorithm should be reduced accordingly (Hofman et al., 2017). Moreover, metrics that evaluate whether the model is capable enough to explain the complexities of social-ecological systems to suggest appropriate forest decisions might be justified (Hoffman et al., 2018; Kim, 2018). To get meaningful insights, ML scholars can use simple abstraction and stochastic analysis or exploit known scientific theories in forestry science to provide useful decisions using metrics that public officials or other stakeholders care about (Struss, 2004; Varshney, 2016; Karpatne et al., 2017).

There has been recent interest in capturing system dynamics in coupled social-ecological systems. As an example, in modeling lake temperature, without using the key physical relationships between the temperature, density, and depth of water in a physics-based loss function used by neural networks, scientifically-consistent physics-based solutions cannot be obtained (Karpatne et al., 2017). These findings suggest the importance of using physics-based equations in modeling complex social-ecological systems. In addition, some scholars have suggested combining traditional forestry science knowledge, whether from professional foresters or from indigenous peoples, with an ML classifier in the form of algorithm fusion to reduce epistemic uncertainty and maintain AI decision safety in forest decision-making (Kshetry and Varshney, 2019; Rana and Varshney, 2020).

ML Applications Must Maximize the Chance of Reducing Social Harm

Making algorithms in forestry transparent to allow more scrutiny and establishing clear governance frameworks including elements of regulatory oversight, awareness-raising, and accountability in the public sector may improve algorithmic decision-making processes. This may reduce the chances of human rights violations through unfair decision-making (Koene et al., 2019; Mueller et al., 2019). Any model developer should enable algorithmic verification, validation, security, and human control over ML systems to maximize the social benefits (Russell et al., 2015) including pre-registering their models with a designated agency, as well as disclosing all choices and assumptions. They should provide a detailed account of origins and use of training and test data, choice of models and other components used in their research so that users keep these facts in mind when judging the suitability of these algorithms for forest decision-making (Whittaker et al., 2018; Mueller et al., 2019). Moreover, they should consider the data-generating process and should increasingly use theories to guide the choice of variables and other regularization parameters to enhance user confidence in algorithmic decision-making (Rana and Miller, 2018). Even organizations that create algorithms should

bear some responsibility for algorithmic decision-making and associated risks (Martin, 2019).

Forest decision-making can specifically be enhanced if algorithm developers follow specific standards, rules, and best practices to ensure fairness and nondiscrimination (Kroll et al., 2016; Corbett-Davies et al., 2017; Kehl and Kessler, 2017). While crafting any fair ML algorithm to support decision-making, it is important to include social scientists, indigenous peoples, forest management committees and other institutions, and their interactions to get a holistic idea of local decision-making cultures, regulatory norms, and incentive structures in a particular forest decision-making context (Gurumurthy et al., 2018; Selbst et al., 2019). Bringing different stakeholders on one platform may, however, be tedious. For example, data scientists participating in developing ML algorithms have little overlap with social scientists in their theoretical frameworks, terminology, or empirical and epistemological approaches.

New technical solutions can enable algorithms to avoid biased data, produce equitable outcomes under various contexts, and ensure procedural regularity such that a consistent set of decision rules are used in each case (Kroll et al., 2016). Some technical tools to ensure procedural regularity may include software verification, zero-knowledge proofs, cryptographic commitments, and fair random choices. We may need to measure the impact of ML algorithms and devise algorithmic audits to understand the assumptions embedded in these models and then score them for fairness to promote their use in forest decision-making (O'neil, 2016). ML algorithms may also benefit from improving validation, conducting uncertainty analysis, incorporating qualitative data at appropriate scales, and including interactions and feedbacks (Liverman and Cuesta, 2008).

Our review shows that ML algorithms are not as “objective” as one might outwardly think and are produced within power laden systems with negligible involvement of stakeholders managing the forests. ML algorithms may be unjust and unfair to local communities if data scientists design them only using data and expert input provided by national forestry agencies. As an example, providing predictive algorithms to foresters—who hold decision making power and ultimately interpret algorithm outputs—can centralize decision-making to national forestry agencies and further widen the power gap between state agencies and local communities with negative impacts on rural livelihoods. Moreover, as algorithms by data scientists have components that they decide to include or not to include based on their own subjective judgements, it is important that social safeguards are in place where such algorithms are to be implemented, and due legal process is carried out to evaluate the social and environmental impacts of such algorithms before they are tried on ground. Any ML-research involving any plausible threat to disadvantaged groups including women, smallholder landowners, or indigenous communities should involve strict adherence to confidential norms as prescribed by various universities or research institutions through institutional review boards.

Promote Interpretable ML Models to Improve Their Adoption for Forest Decision-Making

Researchers should proactively address explainability by promoting easily interpretable ML models to improve their adoption for forest decision-making (Hoffman and Klein, 2017; Herweijer and Waughray, 2018; Padarian et al., 2020). Explanations in the form of easy to understand “coherent stories” may also improve the performance of human-ML systems (Mueller et al., 2019). Others have emphasized responsible and accountable AI, interdisciplinary approach, and adequate funding to minimize environmental harms (Herweijer and Waughray, 2018), and to develop causal models to support explanations (Lake et al., 2017) in algorithmic decision-making in forestry. Efforts should also be made to develop techniques to audit black-box predictive models to have a deeper understanding of model behavior and to identify features important in model prediction (Ribeiro et al., 2016). Some have suggested model-agnostic interpretability tools as they scale much better and are easier to automate in terms of interpretability (Molnar, 2019). Others have noted that there will always be a tradeoff between interpretability and performance of models. But, there can be cases, where such tradeoff may not exist and an interpretable model may have the best performance (Kar et al., 2017).

Theory and System-Analysis Based ML Approaches and Use of Fine-Resolution Datasets

Forest decision-making can benefit if ML-based algorithms include underlying theory and prediction of human behavior (Nguyen et al., 2016; Fang et al., 2017; Gómez et al., 2018), complex system analysis approaches to manage and conserve forest resources (Coulson et al., 1987), spatial information through satellite or unmanned aerial vehicles (Mascaro et al., 2014; Ali et al., 2015; Gonzalez et al., 2016; Rey et al., 2017; Gewali et al., 2018) and ground truth data (O'Connor et al., 2017). Our review suggests that most applications prioritize supervised learning over other forms of machine learning such as unsupervised, semi-supervised, or reinforcement learning mainly because of the focus on labeled training data. We noted that the choice of best machine learning algorithm to assist forest decision-making depends upon the nature of available data, problem, and the solution sought.

While exploring e.g. hyperspectral remote sensing data for ML applications, researchers should focus on addressing the problem of high dimensionality of data, build models invariant under different conditions, promote the use of unsupervised classification in the absence of ground truth data, and create and use new public standardized datasets (O'Connor et al., 2017; Gewali et al., 2018; Bjorck et al., 2019; Rolnick et al., 2019). Given the requirements of data needed to capture the forest complexity, we argue that predictive ML applications in forest management must be developed at a very fine scale as used by scholars (Kelling et al., 2013; Curtis et al., 2018; Norouzzadeh et al., 2018). Therefore, to benefit from ML, scholars should explore finer

spatial and temporal resolution datasets for ML (Kelling et al., 2013; Norouzzadeh et al., 2018) to improve the size of their datasets to identify and solve forest conservation problems and their drivers (Lippitt et al., 2008; Ali et al., 2015; Czimmer and Gálos, 2016; Curtis et al., 2018; Gewali et al., 2018; Holloway and Mengersen, 2018; Rolnick et al., 2019). Moreover, scholars may need to employ GIS and spatial analytical approaches to integrate disparate data sources and should be careful of scale of analysis (Moran and Ostrom, 2005).

However, use of fine-resolution data may instill fear among some landowners, who may worry that ML combined with fine-resolution data may reveal secrets to government officials about their land-use practices, which may violate existing regulations. Such algorithms can therefore negatively affect rural livelihoods (National Research Council, 1998). Under such cases, governments must develop norms of accountability in using fine-resolution data from remote sensing satellites to ensure transparency and fairness in any ML-based decision support system to avoid any potential decline in rural income, loss of community rights over forest resources, unjust and inequitable forest decision-making, or civic unrest and legal complications (Molnar, 2019). A standardized system for data storage, geocoding, and processing data for algorithm development can be developed and it should be open to public scrutiny.

CONCLUSION

In conclusion, the performance of algorithms in supporting forestry decisions should be judged on metrics that directly affect human life such as time and money saved, effort reduced, and effectiveness of conservation interventions increased (Wagstaff, 2012). To explore how much these algorithms can contribute to forest conservation and management given theoretical

and practical limits of prediction in the forestry sector, an interdisciplinary partnership between foresters, ecologists, data scientists, and local communities is a must (Struss, 2004; Czimmer and Gálos, 2016; Kroll et al., 2016; Salganik et al., 2020). We must also establish verifiable and safe ML systems, create adaptable and flexible algorithms suitable to different social-ecological contexts, and improve the transportability or external validity of ML models. Progress on these research themes may help build robust, credible, and productive ML systems to support forest decision-making to effectively conserve and manage forests and wildlife.

Finally, acknowledging fundamental limits to predicting human decisions and activities and maintaining awareness about the multifaceted uncertainties in data can provide alternative and innovative ML algorithms in supporting useful and meaningful forest decisions (Liverman and Cuesta, 2008; Hofman et al., 2017; Kugler et al., 2019). Drawing lessons from this review, we argue that developing effective ML algorithms to support forest decision-making requires fusion of quite different scientific traditions of ML community and forest social science, which necessitate cross-fertilization of discipline-specific theories, and empirical/epistemological cultures.

AUTHOR CONTRIBUTIONS

PR conceptualized the study and methodology, analyzed and interpreted the data, and wrote the original draft. LV helped in interpretation of the results, provided analytical insights, and in critical revision and editing of the manuscript.

FUNDING

This work was supported in part by the National Science Foundation under grant CCF-1717530.

REFERENCES

- Adadi, A., and Berrada, M. (2018). Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). *IEEE Access*. 6, 52138–52160. doi: 10.1109/ACCESS.2018.2870052
- Aertsen, W., Kint, V., Van Orshoven, J., Özkan, K., and Muys, B. (2010). Comparison and ranking of different modelling techniques for prediction of site index in Mediterranean mountain forests. *Ecol. Modell.* 221, 1119–1130. doi: 10.1016/j.ecolmodel.2010.01
- Ali, I., Greifeneder, F., Stamenkovic, J., Neumann, M., and Notarnicola, C. (2015). Review of machine learning approaches for biomass and soil moisture retrievals from remote sensing data. *Remote Sens.* 7, 16398–16421. doi: 10.3390/rs71215841
- Amaro, A., Reed, D., and Soares, P. (2003). *Modelling Forest Systems*. Wallingford: CABI.
- Angwin, J., Larson, J., Mattu, S., and Kirchner, L. (2019). Machine bias: there's software used across the country to predict future criminals. and it's biased against blacks. 2016. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> (accessed May 23, 2016).
- Appel, S. U., Botti, D., Jamison, J., Plant, L., Shyr, J. Y., and Varshney, L. R. (2014). Predictive analytics can facilitate proactive property vacancy policies for cities. *Technol. Forecast. Soc. Change* 89, 161–173. doi: 10.1016/j.techfore.2013.08.028
- Ashraf, M. I., Meng, F.-R., Bourque, C. P.-A., and MacLean, D. A. (2015). A novel modelling approach for predicting forest growth and yield under climate change. *PLoS ONE* 10:e0132066. doi: 10.1371/journal.pone.0132066
- Ashraf, M. I., Zhao, Z., Bourque, C. P.-A., MacLean, D. A., and Meng, F.-R. (2013). Integrating biophysical controls in forest growth and yield predictions with artificial intelligence technology. *Can. J. For. Res.* 43, 1162–1171. doi: 10.1139/cjfr-2013-0090
- Ayana, A. N., Arts, B., and Wiersum, K. F. (2018). How environmental NGOs have influenced decision making in a semi-authoritarian state: the case of forest policy in Ethiopia. *World Dev.* 109, 313–322. doi: 10.1016/j.worlddev.2018.05.010
- Baccini, A., Goetz, S. J., Walker, W. S., Laporte, N. T., Sun, M., Sulla-Menashe, D., et al. (2012). Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nat. Clim. Change* 2:182. doi: 10.1038/nclimate1354
- Bastin, J.-F., Finegold, Y., Garcia, C., Mollicone, D., Rezende, M., Routh, D., et al. (2019). The global tree restoration potential. *Science* 365, 76–79. doi: 10.1126/science.aax0848
- Baylis, K., Honey-Rosés, J., Börner, J., Corbera, E., Ezzine-de-Blas, D., Ferraro, P. J., et al. (2016). Mainstreaming impact evaluation in nature conservation. *Conserv. Lett.* 9, 58–64. doi: 10.1111/conl.12180
- Beery, S., Morris, D., and Yang, S. (2019). Efficient pipeline for camera trap image review. *arXiv [preprint] arXiv:1907.06772*.
- Bjorck, J., Rappazzo, B. H., Chen, D., Bernstein, R., Wrege, P. H., and Gomes, C. P. (2019). Automatic detection and compression for passive acoustic monitoring of the African forest elephant. *arXiv [preprint] arXiv:1902.09069*. doi: 10.1609/aaai.v33i01.3301476

- Bland, L. M., Collen, B. E. N., Orme, C. D. L., and Bielby, J. O. N. (2015). Predicting the conservation status of data-deficient species. *Conserv. Biol.* 29, 250–259. doi: 10.1111/cobi.12372
- Botkin, D. B., Janak, J. F., and Wallis, J. R. (1972). Some ecological consequences of a computer model of forest growth. 60, *J. Ecol.* 849–872. doi: 10.2307/2258570
- Burkhardt, H. E., and Tomé, M. (2012). *Modeling Forest Trees and Stands*. Berlin/Heidelberg: Springer Science and Business Media. doi: 10.1007/978-90-481-3170-9
- Chapelle, O., Schölkopf, B., and Zien, A. (eds.). (2010). *Semi-Supervised Learning*. Cambridge, MA: MIT Press. 542–542. doi: 10.1109/TNN.2009.2015974
- Chhatre, A., and Agrawal, A. (2009). Trade-offs and synergies between carbon storage and livelihood benefits from forest commons. *Proc. Natl. Acad. Sci.* 106, 17667–17670. doi: 10.1073/pnas.0905308106
- Corbett-Davies, S., and Goel, S. (2018). The measure and mismeasure of fairness: a critical review of fair machine learning. *arXiv [preprint] arXiv:1808.00023*.
- Corbett-Davies, S., Pierson, E., Feller, A., Goel, S., and Huq, A. (2017). “Algorithmic decision making and the cost of fairness,” in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Halifax, NS), 797–806. doi: 10.1145/3097983.3098095
- Coulson, R. N., Folse, L. J., and Loh, D. K. (1987). Artificial intelligence and natural resource management. *Science* 237, 262–267. doi: 10.1126/science.237.4812.262
- Curtis, P. G., Slay, C. M., Harris, N. L., Tyukavina, A., and Hansen, M. C. (2018). Classifying drivers of global forest loss. *Science* 361, 1108–1111. doi: 10.1126/science.aau3445
- Czimer, K., and Gálos, B. (2016). A new decision support system to analyse the impacts of climate change on the Hungarian forestry and agricultural sectors. *Scan. J. For. Res.* 31, 664–673. doi: 10.1080/02827581.2016.1212088
- Debeljak, M., Džeroski, S., Jerina, K., Kobler, A., and Adamič, M. (2001). Habitat suitability modelling for red deer (*Cervus elaphus* L.) in South-central Slovenia with classification trees. *Ecol. Modell.* 138, 321–330. doi: 10.1016/S0304-3800(00)00411-7
- Drake, J. M., Randin, C., and Guisan, A. (2006). Modelling ecological niches with support vector machines. *J. Appl. Ecol.* 43, 424–432. doi: 10.1111/j.1365-2664.2006.01141.x
- Dressel, J., and Farid, H. (2018). The accuracy, fairness, and limits of predicting recidivism. *Sci. Adv.* 4:eaa05580. doi: 10.1126/sciadv.aao5580
- Dutta, R., Das, A., and Aryal, J. (2016). Big data integration shows Australian bush-fire frequency is increasing significantly. *R. Soc. Open Sci.* 3:150241. doi: 10.1098/rsos.150241
- Fang, F., Nguyen, T. H., Pickles, R., Lam, W. Y., Clements, G. R., An, B., et al. (2017). PAWS-A deployed game-theoretic application to combat poaching. *AI Mag.* 38, 23–36. doi: 10.1609/aimag.v38i1.2710
- Feldman, O., Korotkov, V. N., and Logofet, D. O. (2005). The monoculture vs. rotation strategies in forestry: formalization and prediction by means of Markov-chain modelling. *J. Environ. Manag.* 77, 111–121. doi: 10.1016/j.jenvman.2005.03.005
- Ferraro, P. J., Sanchirico, J. N., and Smith, M. D. (2019). Causal inference in coupled human and natural systems. *Proc. Natl. Acad. Sci.* 116, 5311–5318. doi: 10.1073/pnas.1805563115
- Fleischman, F. D. (2014). Why do foresters plant trees? Testing theories of bureaucratic decision-making in central India. *World Dev.* 62, 62–74. doi: 10.1016/j.worlddev.2014.05.008
- Franklin, S. E., and Ahmed, O. S. (2018). Deciduous tree species classification using object-based analysis and machine learning with unmanned aerial vehicle multispectral data. *Int. J. Remote Sens.* 39, 5236–5245. doi: 10.1080/01431161.2017.1363442
- Géron, A. (2019). *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. Sebastopol, CA: O'Reilly Media.
- Gewali, U. B., Monteiro, S. T., and Saber, E. (2018). Machine learning based hyperspectral image analysis: a survey. *arXiv [preprint] arXiv:1802.08701*.
- Gholami, S., Xu, L., Carthy, S. M., Dilkina, B., Plumptre, A., Tambe, M., et al. (2019). Stay ahead of poachers: illegal wildlife poaching prediction and patrol planning under uncertainty with field test evaluations. *arXiv [preprint] arXiv:1903.06669*.
- Gillingham, P. (2019). Can predictive algorithms assist decision-making in social work with children and families? *Child Abuse Rev.* 28, 114–126. doi: 10.1002/car.25477
- Gómez, E., Castillo, C., Charisi, V., Dahl, V., Deco, G., Delipetrev, B., et al. (2018). Assessing the impact of machine intelligence on human behaviour: an interdisciplinary endeavour. *arXiv [preprint] arXiv:1806.03192*.
- Gonzalez, L., Montes, G., Puig, E., Johnson, S., Mengersen, K., and Gaston, K. (2016). Unmanned aerial vehicles (UAVs) and artificial intelligence revolutionizing wildlife monitoring and conservation. *Sensors* 16:97. doi: 10.3390/s16010097
- Gurumurthy, S., Yu, L., Zhang, C., Jin, Y., Li, W., Zhang, X., et al. (2018). “Exploiting data and human knowledge for predicting wildlife poaching,” in *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies* (New York, NY), 29. doi: 10.1145/3209811.3209879
- Hajian, S., Bonchi, F., and Castillo, C. (2016). “Algorithmic bias: from discrimination discovery to fairness-aware data mining,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (New York, NY), 2125–2126. doi: 10.1145/2939672.2945386
- Hall, D. B., and Bailey, R. L. (2001). Modeling and prediction of forest growth variables based on multilevel nonlinear mixed models. *For. Sci.* 47, 311–321. doi: 10.1093/forestscience/47.3.311
- Hall, P., and Gill, N. (2018). “An introduction to machine learning interpretability,” in *An Applied Perspective on Fairness, Accountability, Transparency, and Explainable AI*. Boston, MA: O'Reilly.
- Hamdi, Z. M., Brandmeier, M., and Straub, C. (2019). Forest damage assessment using deep learning on high resolution remote sensing data. *Remote Sens.* 11:1976. doi: 10.3390/rs11171976
- Hart, E., Sim, K., Kamimura, K., Meredieu, C., Guyon, D., and Gardiner, B. (2019). Use of machine learning techniques to model wind damage to forests. *Agric. For. Meteorol.* 265, 16–29. doi: 10.1016/j.agrformet.2018.10.022
- Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Berlin/Heidelberg: Springer Science and Business Media.
- Heinonen, T., Pukkala, T., Mehtätalo, L., Asikainen, A., Kangas, J., and Peltola, H. (2017). Scenario analyses for the effects of harvesting intensity on development of forest resources, timber supply, carbon balance and biodiversity of Finnish forestry. *For. Policy Econ.* 80, 80–98. doi: 10.1016/j.forpol.2017.03.011
- Herweijer, C., and Waughray, D. (2018). *Fourth Industrial Revolution for the Earth. Harnessing Artificial Intelligence for the Earth*. A Report of PricewaterhouseCoopers (PwC). New York, NY.
- Hethcoat, M. G., Edwards, D. P., Carreiras, J. M., Bryant, R. G., Franca, F. M., and Quegan, S. (2019). A machine learning approach to map tropical selective logging. *Remote Sens. Environ.* 221, 569–582. doi: 10.1016/j.rse.2018.11.044
- Hoffman, R. R., and Klein, G. (2017). Explaining explanation, part 1: theoretical foundations. *IEEE Intell. Syst.* 32, 68–73. doi: 10.1109/MIS.2017.54
- Hoffman, R. R., Mueller, S. T., Klein, G., and Litman, J. (2018). Metrics for explainable AI: challenges and prospects. *arXiv [preprint] arXiv:1812.04608*.
- Hofman, J. M., Sharma, A., and Watts, D. J. (2017). Prediction and explanation in social systems. *Science* 355, 486–488. doi: 10.1126/science.aal3856
- Holloway, J., and Mengersen, K. (2018). Statistical machine learning methods and remote sensing for sustainable development goals: a review. *Remote Sens.* 10:1365. doi: 10.3390/rs10091365
- James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An Introduction to Statistical Learning*. New York, NY: Springer. doi: 10.1007/978-1-4614-7138-7
- Jesson, J., Matheson, L., and Lacey, F. M. (2011). *Doing Your Literature Review: Traditional and Systematic Techniques*. London: Sage.
- Jordan, M. I., and Mitchell, T. M. (2015). Machine learning: trends, perspectives, and prospects. *Science* 349, 255–260. doi: 10.1126/science.aaa8415
- Kar, D., Ford, B., Gholami, S., Fang, F., Plumptre, A., Tambe, M., et al. (2017). “Cloudy with a chance of poaching: adversary behavior modeling and forecasting with real-world poaching data,” in *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems* (Richland, SC), 159–167.
- Karpatne, A., Watkins, W., Read, J., and Kumar, V. (2017). Physics-guided neural networks (PGNN): an application in lake temperature modeling. *arXiv preprint arXiv:1710.11431*.
- Kehl, D. L., and Kessler, S. A. (2017). *Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing*. Responsive Communities.

- Kelling, S., Gerbracht, J., Fink, D., Lagoze, C., Wong, W.-K., Yu, J., et al. (2013). A human/computer learning network to improve biodiversity conservation and research. *AI Mag.* 34, 10–10. doi: 10.1609/aimag.v34i1.2431
- Kim, T. W. (2018). Explainable artificial intelligence (XAI), the goodness criteria and the grasp-ability test. *arXiv [preprint] arXiv:1810.09598*.
- Kimmins, J. P., Welham, C., Seely, B., Meitner, M., Rempel, R., and Sullivan, T. (2005). Science in forestry: why does it sometimes disappoint or even fail us? *For. Chronicle* 81, 723–734. doi: 10.5558/tfc81723-5
- Koene, A., Clifton, C., Hatada, Y., Webb, H., and Richardson, R. (2019). *A Governance Framework for Algorithmic Accountability and Transparency*. European Parliamentary Research Service.
- Kroll, J. A., Barocas, S., Felten, E. W., Reidenberg, J. R., Robinson, D. G., and Yu, H. (2016). Accountable algorithms. *Univ. Penn. Law Rev.* 165, 633.
- Kshetry, N., and Varshney, L. R. (2019). “Safety in the face of unknown unknowns: algorithm fusion in data-driven engineering systems,” in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (Olomouc), 8162–8166. doi: 10.1109/ICASSP.2019.8683392
- Kugler, T. A., Grace, K., Wrathall, D. J., de Sherbinin, A., Van Riper, D., Aubrecht, C., et al. (2019). People and pixels 20 years later: the current data landscape and research trends blending population and environmental data. *Popul. Environ.* 41, 209–234. doi: 10.1007/s11111-019-00326-5
- Kuiper, T., Kavhu, B., Ms, N. A. N., Mandisodza-Chikerema, R., and Milner-Gulland, E. J. (2020). Rangers and modellers collaborate to build and evaluate spatial models of African elephant poaching. *Biol. Conserv.* 243:108486. doi: 10.1016/j.biocon.2020.108486
- Lake, B. M., Ullman, T. D., Tenenbaum, J. B., and Gershman, S. J. (2017). Building machines that learn and think like people. *Behav. Brain Sci.* 40:e253. doi: 10.1017/S0140525X16001837
- Lewis, N. D. (2017). *Machine Learning Made Easy With R*. Available online at: <https://www.amazon.com/Machine-Learning-Made-Easy-Intuitive/dp/1546483756>
- Li, M., Im, J., and Beier, C. (2013). Machine learning approaches for forest classification and change analysis using multi-temporal Landsat TM images over Huntington wildlife forest. *GISci. Remote Sens.* 50, 361–384. doi: 10.1080/15481603.2013.819161
- Lippitt, C. D., Rogan, J., Li, Z., Eastman, J. R., and Jones, T. G. (2008). Mapping selective logging in mixed deciduous forest. *Photogramm. Eng. Remote Sens.* 74, 1201–1211. doi: 10.14358/PERS.74.10.1201
- Liu, J., Dietz, T., Carpenter, S. R., Alberti, M., Folke, C., Moran, E., et al. (2007). Complexity of coupled human and natural systems. *Science* 317, 1513–1516. doi: 10.1126/science.1144004
- Liverman, D. M., and Cuesta, R. M. R. (2008). Human interactions with the Earth system: people and pixels revisited. *Earth Surf. Proc. Landforms* 33, 1458–1471. doi: 10.1002/esp.1715
- Martin, K. (2019). Ethical implications and accountability of algorithms. *J. Business Ethics* 160, 835–850. doi: 10.1007/s10551-018-3921-3
- Mascaro, J., Asner, G. P., Knapp, D. E., Kennedy-Bowdoin, T., Martin, R. E., Anderson, C., et al. (2014). A tale of two “forests”: random forest machine learning aids tropical forest carbon mapping. *PLoS ONE* 9:e85993. doi: 10.1371/journal.pone.0085993
- Mehta, S., Pimplikar, R., Singh, A., Varshney, L. R., and Visweswariah, K. (2013). “Efficient multifaceted screening of job applicants,” in *Proceedings of the 16th International Conference on Extending Database Technology* (New York, NY), 661–671. doi: 10.1145/2452376.2452453
- Miller, D. C., Rana, P., and Wahlén, C. B. (2017). A crystal ball for forests? Analyzing the social-ecological impacts of forest conservation and management over the long term. *Environ. Soc.* 8, 40–62. doi: 10.3167/ares.2017.080103
- Molnar, C. (2019). *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. <https://christophm.github.io/interpretable-ml-book> (accessed December 7, 2020).
- Montano, R. A. N. R., Sanquetta, C. R., Wojciechowski, J., Mattar, E., Dalla Corte, A. P., and Todt, E. (2017). Artificial intelligence models to estimate biomass of tropical forest trees. *Polibits* 56, 29–37. doi: 10.17562/PB-56-4
- Moran, E. F., and Ostrom, E. (2005). *Seeing the Forest and the Trees: Human-Environment Interactions in Forest Ecosystems*. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/6140.001.0001
- Mueller, S. T., Hoffman, R. R., Clancey, W., Emrey, A., and Klein, G. (2019). Explanation in human-AI systems: a literature meta-review, synopsis of key ideas and publications, and bibliography for explainable AI. *arXiv [preprint] arXiv:1902.01876*.
- Naidoo, L., Cho, M. A., Mathieu, R., and Asner, G. (2012). Classification of savanna tree species, in the Greater Kruger National Park region, by integrating hyperspectral and LiDAR data in a Random Forest data mining environment. *ISPRS J. Photogramm. Remote Sens.* 69, 167–179. doi: 10.1016/j.isprsjprs.2012.03.005
- National Research Council (1998). *People and Pixels: Linking Remote Sensing and Social Science*. Washington, DC: National Academies Press.
- Nguyen, T. H., Sinha, A., Gholami, S., Plumptre, A., Joppa, L., Tambe, M., et al. (2016). “Capture: a new predictive anti-poaching tool for wildlife protection.” in *Proceedings of the 2016 International Conference on Autonomous Agents and Multiagent Systems* (Singapore), 767–775.
- Nishant, R., Kennedy, M., and Corbett, J. (2020). Artificial intelligence for sustainability: challenges, opportunities, and a research agenda. *Int. J. Inf. Manag.* 53:102104. doi: 10.1016/j.ijinfomgt.2020.102104
- Norouzzadeh, M. S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M. S., Packer, C., et al. (2018). Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proc. Natl. Acad. Sci.* 115, E5716–E5725. doi: 10.1073/pnas.1719367115
- Nunes, M. H., and Görgens, E. B. (2016). Artificial intelligence procedures for tree taper estimation within a complex vegetation mosaic in Brazil. *PLoS ONE* 11:e0154738. doi: 10.1371/journal.pone.0154738
- O'Connor, C. D., Calkin, D. E., and Thompson, M. P. (2017). An empirical machine learning method for predicting potential fire control locations for pre-fire planning and operational fire management. *Int. J. Wildland Fire* 26, 587–597. doi: 10.1071/WF16135
- O'neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York, NY: Broadway Books.
- Ordóñez, C., Threlfall, C. G., Livesley, S. J., Kendal, D., Fuller, R. A., Davern, M., et al. (2020). Decision-making of municipal urban forest managers through the lens of governance. *Environ. Sci. Policy* 104, 136–147. doi: 10.1016/j.envsci.2019.11.008
- Ostrom, E. (2009). A general framework for analyzing sustainability of social-ecological systems. *Science* 325, 419–422. doi: 10.1126/science.1172133
- Padarian, J., Minasny, B., and McBratney, A. B. (2020). Machine learning and soil sciences: a review aided by machine learning tools. *SOIL Discuss.* 6, 35–52. doi: 10.5194/soil-6-35-2020
- Pearl, J., and Mackenzie, D. (2018). *The Book of Why: The New Science of Cause and Effect*. New York, NY: Basic Books.
- Persha, L., Agrawal, A., and Chhatre, A. (2011). Social and ecological synergy: local rulemaking, forest livelihoods, and biodiversity conservation. *Science* 331, 1606–1608. doi: 10.1126/science.1199343
- Persha, L., and Andersson, K. (2014). Elite capture risk and mitigation in decentralized forest governance regimes. *Global Environ. Change* 24, 265–276. doi: 10.1016/j.gloenvcha.2013.12.005
- Purves, D. W., Lichstein, J. W., Strigul, N., and Pacala, S. W. (2008). Predicting and understanding forest dynamics using a simple tractable model. *Proc. Natl. Acad. Sci.* 105, 17018–17022. doi: 10.1073/pnas.0807754105
- Rana, P. (2014). *Elite Capture and Forest Governance in India*. University of Illinois at Urbana-Champaign.
- Rana, P., and 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. *Environm. Res. Lett.* 14: 024008. doi: 10.1088/1748-9326/aafa8f
- Rana, P., and Miller, D. C. (2019). Explaining long-term outcome trajectories in social-ecological systems. *PLoS ONE* 14:e0215230. doi: 10.1371/journal.pone.0215230
- Rana, P., and Varshney, L. R. (2020). Planting trees at the right places: recommending suitable sites for growing trees using algorithm fusion. *arXiv [preprint] arXiv 2009.08002*.
- Rao, K., Williams, A. P., Flefil, J. F., and Konings, A. G. (2020). SAR-enhanced mapping of live fuel moisture content. *Remote Sens. Environ.* 245:111797. doi: 10.1016/j.rse.2020.111797

- Rey, N., Volpi, M., Joost, S., and Tuia, D. (2017). Detecting animals in African Savanna with UAVs and the crowds. *Remote Sens. Environ.* 200, 341–351. doi: 10.1016/j.rse.2017.08.026
- Ribeiro, M. T., Singh, S., and Guestrin, C. (2016). “Why should i trust you? Explaining the predictions of any classifier,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (New York, NY), 1135–1144. doi: 10.1145/2939672.2939778
- Rodner, E., Simon, M., Brehm, G., Pietsch, S., Wägele, J. W., and Denzler, J. (2015). Fine-grained recognition datasets for biodiversity analysis. *arXiv preprint arXiv:1507.00913*.
- Rodrigues, M., and de la Riva, J. (2014). An insight into machine-learning algorithms to model human-caused wildfire occurrence. *Environ. Modell. Softw.* 57, 192–201. doi: 10.1016/j.envsoft.2014.03.003
- Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., et al. (2019). Tackling climate change with machine learning. *arXiv [preprint] arXiv:1906.05433*.
- Russel, S., and Norvig, P. (2013). *Artificial Intelligence: A Modern Approach*. Essex: Pearson Education Limited.
- Russell, S., Dewey, D., and Tegmark, M. (2015). Research priorities for robust and beneficial artificial intelligence. *AI Mag.* 36, 105–114. doi: 10.1609/aimag.v36i4.2577
- Sakr, G. E., Elhajj, I. H., Mitri, G., and Wejinya, U. C. (2010). “Artificial intelligence for forest fire prediction,” in *2010 IEEE/ASME International Conference on Advanced Intelligent Mechatronics* (Budapest), 1311–1316. doi: 10.1109/AIM.2010.5695809
- Salganik, M. J., Lundberg, I., Kindel, A. T., Ahearn, C. E., Al-Ghoneim, K., Almaatouq, A., et al. (2020). Measuring the predictability of life outcomes with a scientific mass collaboration. *Proc. Natl. Acad. Sci.* 117, 8398–8403. doi: 10.1073/pnas.1915006117
- Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., and Vertesi, J. (2019). “Fairness and abstraction in sociotechnical systems,” in *Proceedings of the Conference on Fairness, Accountability, and Transparency* (New York, NY), 59–68. doi: 10.1145/3287560.3287598
- Sevinc, V., Kucuk, O., and Goltas, M. (2020). A Bayesian network model for prediction and analysis of possible forest fire causes. *For. Ecol. Manag.* 457:117723. doi: 10.1016/j.foreco.2019.117723
- Silver, E. J., Leahy, J. E., Weiskittel, A. R., Noblet, C. L., and Kittredge, D. B. (2015). An evidence-based review of timber harvesting behavior among private woodland owners. *J. For.* 113, 490–499. doi: 10.5849/jof.14-089
- Simon, H. A. (1996). Designing organizations for an information-rich world. *Int. Libr. Crit. Writings Econ.* 70, 187–202.
- Snyder, S. A., and Kilgore, M. A. (2018). The influence of multiple ownership interests and decision-making networks on the management of family forest lands: evidence from the United States. *Small-Scale For.* 17, 1–23. doi: 10.1007/s11842-017-9370-5
- Springate-Baginski, O., and Blaikie, P. (2013). *Forests People and Power: The Political Ecology of Reform in South Asia*. Abingdon: Routledge. doi: 10.4324/9781849771399
- Stock, M. W., and Rauscher, H. M. (1996). Artificial intelligence and decision support in natural resource management. *NZ J. For. Sci.* 26, 145–157.
- Struss, P. (2004). “Artificial intelligence methods for environmental decision support,” in *e-Environment: Progress and Challenge* (Research in Computing Science), 1–14.
- Thompson, J. R., Wiek, A., Swanson, F. J., Carpenter, S. R., Fresco, N., Hollingsworth, T., et al. (2012). Scenario studies as a synthetic and integrative research activity for long-term ecological research. *BioScience* 62, 367–376. doi: 10.1525/bio.2012.62.4.8
- Varshney, L. R. (2016). Fundamental limits of data analytics in sociotechnical systems. *Front. ICT* 3:2. doi: 10.3389/fict.2016.00002
- Wagstaff, K. (2012). Machine learning that matters. *arXiv [preprint] arXiv:1206.4656*.
- Whittaker, M., Crawford, K., Dobbe, R., Fried, G., Kaziunas, E., Mathur, V., et al. (2018). *AI Now Report 2018*. AI Now Institute at New York University.
- Ye, L., Gao, L., Marcos-Martinez, R., Mallants, D., and Bryan, B. A. (2019). Projecting Australia's forest cover dynamics and exploring influential factors using deep learning. *Environ. Modell. Softw.* 119, 407–417. doi: 10.1016/j.envsoft.2019.07.013
- Yousefpour, R., Temperli, C., Jacobsen, J. B., Thorsen, B. J., Meilby, H., Lexer, M. J., et al. (2017). A framework for modeling adaptive forest management and decision making under climate change. *Ecol. Soci.* 22:40. doi: 10.5751/ES-09614-220440
- Yu, H., Mineyev, I., and Varshney, L. R. (2019). “To abstract via algebraic innateness: hierarchical, interpretable, and task-free clustering,” in *Proceedings of the 2019 Information Theory and its Applications Workshop (ITA)* (San Diego, CA).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Rana and Varshney. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.