Received: 30 April 2019 | Revised: 21 November 2019 | Accepted: 11 December 2019
DOI: 10.1002/ldr.3526

RESEARCH ARTICLE

WILEY

Landscape-scale spatial modelling of deforestation, land degradation, and regeneration using machine learning tools

Clovis Grinand[1,2] | Ghislain Vieilledent[3,4,5] | Tantely Razafimbelo[6] | Jean-Roger Rakotoarijaona[7] | Marie Nourtier[1] | Martial Bernoux[2,8]

[1]Research and Development, NITIDAE, Maison de la Télédétection, 34000 Montpellier, France
[2]Eco & Sols, IRD, 34000 Montpellier, France
[3]UMR AMAP, CIRAD, 34000 Montpellier, France
[4]AMAP, CIRAD, CNRS, INRA, IRD, University of Montpellier, 34000 Montpellier, France
[5]Bio-economy unit (JRC.D.1), Joint Research Centre of the European Commission, 21027 ISPRA, Italy
[6]Agronomy, Laboratoire des Radio-Isotopes, 101 Antananarivo, Madagascar
[7]Direction des Informations Environnementales, Office National pour l'Environnement, 101 Antananarivo, Madagascar
[8]MICCA, FAO, Food and Agriculture Organization of the United Nations, 00153 Rome, Italy

Correspondence
Clovis Grinand, Research and Development, NITIDAE, Maison de la Télédétection, 500 avenue Jean François Breton, 34000 Montpellier, France.
Email: c.grinand@nitidae.org

Funding information
ANRT PhD scholarship, Grant/Award Numbers: CIFRE N°2012-1153, CIFRE N 2012-1153; Centre de Coopération Internationale en Recherche Agronomique pour le Développement; Institut de Recherche pour le Développement; European Commission; Fondation pour la Recherche sur la Biodiversité, Grant/Award Number: AAP-SCEN-2013 I; Research Foundation

Abstract

Land degradation and regeneration are complex processes that greatly impact climate regulation, ecosystem service provision, and population well-being and require an urgent and appropriate response through land use planning and interventions. Spatially explicit land change models can greatly help decision makers, but traditional regression approaches fail to capture the nonlinearity and complex interactions of the underlying drivers. Our objective was to use a machine learning algorithm combined with high-resolution data sets to provide simultaneous and spatial forecasts of deforestation, land degradation, and regeneration for the next two decades. A 17,000-km^2 region in the south of Madagascar was taken as the study area. First, an empirical analysis of drivers of change was conducted, and then, an ensemble model was calibrated to predict and map potential changes based on 12 potential explanatory variables. These potential change maps were used to draw three scenarios of land change while considering past trends in intensity of change and expert knowledge. Historical observations displayed clear patterns of land degradation and relatively low regeneration. Amongst the 12 potential explanatory variables, distance to forest edge and elevation were the most important for the three land transitions studied. Random forest showed slightly better prediction ability compared with maximum entropy and generalized linear model. Business-as-usual scenarios highlighted the large areas under deforestation and degradation threat, and an alternative scenario enabled the location of suitable areas for regeneration. The approach developed herein and the spatial outputs provided can help stakeholders target their interventions or develop large-scale sustainable land management strategies.

KEYWORDS
ensemble method, land use change modelling, Madagascar, REDD+, scenarios

1 | INTRODUCTION

1.1 | International context: Targeting the drivers of change

The agriculture, forestry, and other land use sector, which is responsible for a quarter of global anthropogenic greenhouse gas emissions (Intergovernmental Panel on Climate Change, 2014), is under pressure to find pathways to mitigate climate change and improve population livelihood through sustainable land management. Mitigation initiatives under the scope of the environmental United Nations conventions, such as Reduction of Emissions due to Deforestation and Forest Degradation and enhancement of forest carbon stocks (REDD+) under the United Nations Framework on

Climate Change or Land Degradation Neutrality under the United Nations Convention to Combat Desertification, require the identification of drivers of land use change to (a) quantify the impact on ecosystem goods and services and (b) design appropriate strategies for conservation and sustainable development. However, a growing number of scientific assessments of the drivers of deforestation are reaching diverging conclusions (Ferretti-Gallon & Bush, 2014) and may explain why current REDD+ policies are struggling to demonstrate their effectiveness as a benefit-sharing solution (Weatherley-Singh & Gupta, 2015).

1.2 | Drivers of change analysis: No common framework

Assessing the driving forces behind land use change is the key for understanding changes in our global environment (Bax et al., 2016) and for building realistic models of land use change (Veldkamp & Lambin, 2001). However, they are difficult to quantify and assess because they have long underlying causal chains—also referred to as biophysical feedback (Verburg, 2006) or socioeconomical retroactions—and take different shapes depending on the perspective that is chosen (Wehkamp et al., 2015). For instance, the perspective described by Geist and Lambin (2001) is often used to distinguish direct drivers (or proximal) and indirect (or underlying causes) drivers. The former is defined as human activities or actions at the local level that directly lead to the conversion of land into another land use such as forest clearing due to agricultural expansion or mining. The latter implies complex social processes at various scales, which ‘underpin or sustain the direct drivers,’ such as the demographic expansion or the price of commodities. They are then analysed either from a process-driven or data-driven modelling framework. Nonetheless, all the models fail to capture all the complexity (Veldkamp & Lambin, 2001). Currently, no accepted framework exists to assess the driving forces of land change process because the availability of the input data set (quantity and quality) and assumptions used (correlation or causality) greatly influence the results.

1.3 | Land change modelling: Limitations and the way forward

Spatially explicit land use change models show a great advantage for the prediction of potential land change locations in a transparent and verifiable manner. The three most important and common criteria of land change models for policymakers are (a) compliance with Intergovernmental Panel on Climate Change good practice guidelines, (b) clarity, and (c) dynamic baseline updating (Huettner et al., 2009). However, land change models heavily rely on two key parameters: the input data set and model assumptions. The former usually refers to land use maps, used as the main input data, and the accuracy of these maps is affected by biases in operator and satellite image classification techniques. The latter refers to the digital relationship between land change observations and explanatory variables, either linear or nonlinear. Veldkamp and Lambin (2001) argue that linear models are prone to numerical instability as “small measurement errors in input data can propagate and lead to spurious results, given the intrinsic nonlinear behaviour of the modelled system.” In contrast, nonlinear algorithms, such as machine learning algorithms (neural network, support vector machines, decision tress, etc.), can capture nonlinear observation–variable relationships, but these have not been tested yet for land degradation and regeneration spatial modelling to our knowledge.

Two decades of high-resolution remote-sensing images allow the detection of land use change in an unprecedent manner. Notably, Hansen et al. (2013) published a globally consistent and locally relevant data set of vegetation cover gain and loss over a long historical period, from 2000 to 2018. This data set provides a means for assessing key ecosystem dynamics such as deforestation, land degradation, and regeneration while assuming that tree cover is a proxy for numerous ecosystem services. In this study, we explore the application of machine learning algorithms with an easy-to-access and globally available vegetation change data set. The overall objective of this research is to test a new, low bias, and adaptive land change modelling framework.

1.4 | Madagascar: A need for spatially explicit, sound, and comprehensive information

Madagascar is recognized as a major biodiversity reservoir in the world, and this reservoir is mainly located within Madagascar's intact or natural forest. Recent studies have highlighted a dramatic increase in deforestation in this country. On a national scale, a study revealed a shift from 0.5% of deforestation (21,710 ha yr^{-1}) for the 2005–2010 period to 0.92% by year (34,567 ha yr^{-1}) for the period 2010–2013 within the tropical humid ecoregion (Rakotomala et al., 2015), with dramatic values in the dry and spiny forest area (ONE et al., 2015). Today, the total remaining intact forest is less than 8,485,509 ha (ONE et al., 2015) relative to the 10,605,700 ha remaining in 1990 (Harper et al., 2007), which corresponds to a loss of 20% in 25 years. Madagascar has participated in both the REDD+ and land degradation neutrality schemes since 2008, and with the help of the Forest Carbon Partnership Facility (Readiness Plan Idea Note, 2008), Madagascar has recently validated its REDD+ Readiness Preparation Proposal as described in the national REDD+ strategy (Readiness Preparation Proposal, 2014) and has proposed an emission reduction programme in a rainforest pilot region (Emission Reductions Program Idea Note, 2015). In these documents for national and subnational scale REDD+, broad information is provided on the factors of deforestation and the driving forces that underlie these changes, but quantifiable and spatially explicit data are still missing. Land use change spatial assessment in REDD+ countries such as Madagascar is urgently required (a) to precisely estimate the impact of those deforestation programmes that have been avoided and the effectiveness of conservation efforts and

(b) to build comprehensive possible future scenarios with sound economic and environmental assessment.

1.5 | Objectives

The main aim of this paper was to develop, test, and validate a new tool with high-resolution, spatially explicit, potential change maps of deforestation, degradation, and regeneration. Then, we proposed land change scenarios at a regional scale. The approach was tested in southeastern Madagascar, which displays a high level of biodiversity and a high rate of deforestation.

We first compiled a historical change data set from the global forest change data set, which recorded gain and loss at 30-m pixel for the 2000–2014 period (Hansen et al., 2013). Presenting a benchmark of the intact forest cover in 2000 (Grinand et al., 2013), this raw data set was used to derive a data set for three land change transitions: deforestation, land degradation, and land regeneration. In addition, we collected and prepared 12 potential land change explanatory variables that were constructed and statistically assessed for their contributions to the three land change processes. Validation of the model was performed using several commonly used land change accuracy metrics. Three land change scenarios were established and used to assess the potential impacts and opportunities in natural protected areas and areas with currently no protected status.

2 | MATERIAL AND METHODS

2.1 | Study area

The study area is located in the southern part of the tropical humid forest corridor of Madagascar (Figure 1), approximately 70 km wide along its east–west axis and 200 km long along its north–south axis (1,676,000 ha). The region is marked by a large east–west gradient of precipitation, from 2000 to 700 mm (WorldClim database, Hijmans et al., 2005). Four principal landscapes can be distinguished: the flat sandy coastline, the humid rough montane terrain, the downhill mosaic crop-savannah system, and the semi-arid gently sloping western corridor area. Two national parks are located in the study area. One to the south, the Andohahela National Park (82,000 ha), was created first as a national reserve in 1934, and one to the North, the Midongy du Sud National Park, was created more recently (1997) and covers 188,000 ha. In total, these two parks cover 16% of the study area and 46.8% (191,970 ha) of its forested area. Biodiversity is mainly located in the forested areas (Vieilledent et al., 2018). The soils are dominated with ferralitic soils developed from igneous rock, more or less truncated by erosion processes, leading to local deposits of soil particles in the valleys (Grinand et al., 2017). The agricultural system is dominated by irrigated rice cropping systems and shifting agriculture of food crops such as rice associated with cassava and maize in more or less long crop-fallow rotations. Other activities include cattle ranching and cash crop production, mainly coffee. The population is rural, with only 11 towns with more than 10,000 inhabitants and that hold more than one food market a week and with around 1,417 villages (Figure 1).

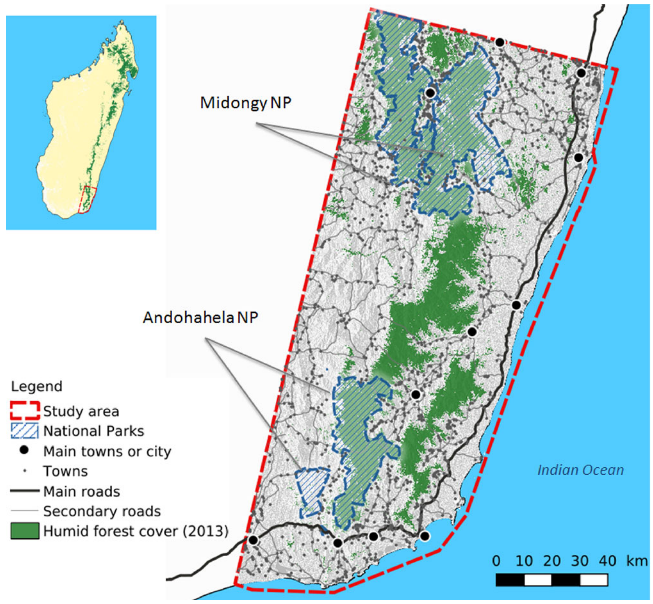

FIGURE 1 Location of the study area in the southeast tropical humid corridor. Sources: Système des Aires Protégées de Madagascar, 2010; BD200 Foiben-Taosarintanin'i Madagasikara; Rakotomalala et al, 2015. NP, National Park [Colour figure can be viewed at wileyonlinelibrary.com]

2.2 | Land use change data set

In this study, we combined two existing data sets. The vegetation change data set produced by Hansen et al. (2013) for the 2000–2014 period available globally and the intact forest map in 2000 produced by Grinand et al. (2013) in Madagascar. First, we collected the vegetation loss and gain information (Hansen et al., 2013) that was derived from vegetation reflectance change analysis. Vegetation index are correlated to biomass productivity and commonly used as an indicator of land health status to assess land degradation as a whole (Bai et al., 2013; United Nations Environment Programme, 2012; Yengoh et al., 2015). In Hansen et al. (2013), vegetation loss was defined as “a stand-replacement disturbance or complete removal or a change from a forest to nonforest state” for the 2000–2014 period, omitting selective removal of trees that do not lead to a nonforested state (forest degradation). Vegetation gain was defined as “the inverse of loss, or a nonforest to forest change entirely within the 2000–2012 period”, omitting areas that might have been considered as forest cover in 2000 (land regeneration that started before 2000). Second, we applied a mask of natural forest extent from another study that used intensive photointerpretation and the national forest definition (Grinand et al., 2013) in order to separate pixels representing vegetation loss or gain within and outside intact forest at the initial date (2000).

Finally, we defined three different land change processes and calculated the corresponding data set: 'deforestation' as vegetation loss inside intact forest, 'land degradation' as vegetation loss outside intact forest and with no vegetation gain observed at the same location, and 'land regeneration' as vegetation gain outside intact forest without any vegetation loss. Areas with loss and gain observed at the same location, which are likely to represent agricultural land that has been cleared and left fallow, were not included in this study.

2.3 | Potential land change explanatory variables

Twelve potential explanatory variables of land use change were converted into spatially explicit layers and included in our analysis (Table 1). They represent three types of variables usually used in spatially explicit land change studies (Aguilar-Amuchastegui et al., 2014; Bax et al., 2016; Ferretti-Gallon & Bush, 2014) and already tested in Madagascar (Thomas, 2007, Vieilledent et al., 2013). The first type represents variables related to the amount of time required to access the land and transport goods to market: elevation, slope, proximity to towns or villages, proximity to main roads or secondary roads, and proximity to the forest edge. The second type represents potential productivity factors of land under agriculture: orientation of the slope (aspect), proximity to water course, and the number of dry months, which is defined as the number of months with potential evapotranspiration higher than monthly rainfall (http://madaclim.cirad.fr). The last predictor type expresses land tenure and land regulation: the two national park delimitations collected from the Protected Areas system of Madagascar ("Système des Aires Protégées de Madagascar") and the population density aggregated at the county level ('communes'), which was taken from a 2006 to 2009 census collected by "Institut National de la Statistique à Madagascar".

TABLE 1 The 12 explanatory variables derived

Name	Source	Range (min–max)	Unit
Elevation	SRTM	1–1946	Metre
Slope	SRTM	0–69	Degree
Aspect	SRTM,	0–360	Degree
Proximity to forest edge	This study	30–12,041	Metre
Number of dry month	MadaClim	0–12	Month
Proximity to rivers	SRTM	0–9,1	km
Protected areas	SAPM	0,1,2	Category
Population density	INSTAT	8–1,623	People/km^2
Proximity to main roads	FTM	0–65	km
Proximity to main towns	FTM	2,3–65	km
Proximity to villages	FTM	60–18,541	Metre
Proximity to tracks	FTM	30–13,441	Metre

Note: SRTM (https://www2.jpl.nasa.gov/srtm); MadaClim, (http://madaclim.cirad.fr); SAPM, version 2010; Protected Area Network System, version 2010; INSTAT, census survey from 2006 to 2009; FTM from National Institute of Geography.
Abbreviations: FTM, Foiben-Taosarintanin'i Madagasikara; INSTAT, Institut National de la Statistique à Madagascar; MadaClim, Climate Data on Madagascar; SAPM, Système des Aires Protégées de Madagascar; SRTM, Shuttle Radar Topography Mission.

2.4 | Importance of drivers

Prior to modelling, spatial drivers of deforestation, land degradation, and regeneration processes were assessed using extractions of spatial predictor values for each land change process. We used a stratified random sampling scheme by randomly sampling 10,000 observations areas without changes and 10,000 observations in the land change category. Three data sets of 20,000 observations representing the three processes were thus compiled. Observed probabilities (ratio of change observation divided by the total observations) were computed for quantiles on the predictor range values. This approach allows a quick overview of the influence of each factor, with values above 0.5 having a positive effect on change, with values below 0.5 having a negative effect, and a straight line at the 0.5 value indicating no influence. This empirical analysis was complemented with a linear regression model to assess the direction (positive or negative) and correlation significance of each predictor using the same matrices.

2.5 | Model building

Leading spatially explicit land use change modelling software such as LAND CHANGE MODEL (Eastman, 2012), GEOMOD (Pontius, Jr., Cornell, & Hall., 2001), or DINAMICA EGO (Soares-Filho et al., 2009) uses statistical models or modelling chains that usually require fine tuning with numerous key parameters, which may greatly impact the results. This study considers the random forest algorithm (RF; Breiman, 2001), which is increasingly being used in many spatial applications that deal with nonlinear and complex nature–human interactions, appreciated for its good predictive ability and low parameterization requirements. Examples include global and high-resolution biomass mapping (e.g., Baccini et al., 2012; Vieilledent et al., 2016), land use and land cover (Gislason et al., 2006), and soil organic carbon change mapping (e.g., Grinand et al., 2017). Recently, this tool has been tested in land use change modelling applications (Gounaridis et al., 2019). RF combines the advantage of using bagging (random selection of individual and variable) and a simple decision tree (recursive binary split in the explanatory variable data set) that can be used to solve both regression and classification problems.

RF was then tested and compared with the generalized linear model, which is a commonly used regression algorithm for land use change modelling and maximum entropy; the latter of which is a famous 'two-class' species distribution model that has recently been applied with success in a deforestation modelling study (Aguilar-Amuchastegui, Riveros, & Forrest, 2014). RF was used in the classification mode using a two-class (change and no change) mode, and class membership was further processed. The three algorithms were calibrated using the same point data set presented above (20,000 observations). The calibrated model was applied to the spatial predictor layer stack to predict the probability of the land change category over the study area at a 30-m resolution. We referred to the three transition probability maps as the deforestation risk map, land degradation risk map, and land regeneration suitability map.

2.6 | Model assessment

Model accuracy assessment is a key step in land use change modelling because it involves providing sound information to stakeholders about potential future land use dynamics. In this study, we randomly sampled 20,000 points within the initial land cover, that is, the extent of forest in 2000 for an accurate assessment of deforestation and the extent of the nonforested area in 2000 for an accurate assessment of land degradation and regeneration. For these point locations, we predicted the land transition probabilities by using the above-mentioned calibrated models. We predicted the 2014 land allocation by using the historical 2000–2014 amount of change (Table 2) and assigning the highest probability values to 'change' value (value of 1) and assigning the remaining pixels to 'no change' (value of 0). We then calculated commonly used accuracy metrics: the area under the curve (AUC) 'receiver operating characteristic' (referred to as AUC in this following text), the figure of merit (FOM), and user's accuracy indexes. The AUC is the most commonly used metric for species distribution models (Elith et al., 2006). This statistic was computed using the pROC package available in R (Robin et al., 2011). It allows the predictive power of the land change model to be assessed, with a value of 1 indicating perfect predictive power, 0.5 meaning that the model is no better than random, and values below 0.5 indicating systematically incorrect predictions (Pontius et al., 2001). We also computed the FOM because it is a required indicator in the REDD+ methodologies (Shoch et al., 2013), although this indicator is correlated with the net area of change (Pontius et al., 2008), which underpins study-to-study comparisons. REDD+ methodologies usually require an FOM value greater or equal to the net change ratio. The formulas used to derive each accuracy metrics are summarized in Table 2.

TABLE 2 Illustration of the change matrix used for validation and to derive the accuracy indexes

		Reference	
		No change (0)	**Change (1)**
Predicted	No change (0)	**A**	D
	Change (1)	C	**B**

Note: Overall accuracy: OA = (A + B)/(A + B + C + D). User accuracy of change: UA_c = B/(B + C). User Accuracy of no change: UA_{NC} = A/(A + D). Balanced User Accuracy: UA = (UA_C + UA_{NC})/2. Figure of Merit: FOM = B/(B + C + D), where A indicates the correctly predicted (true negative); B is the correctly predicted presence of land change (true positive); C is the no change pixel predicted as change (false positive); and D is the change observations predicted as no change (false negative).

2.7 | Predicting future land use transitions areas

Land use change modelling outcomes are twofold: future rate (or quantity) of change and potential location of changes to come. These two overarching goals imply different data requirements and validation strategies (Veldkamp and Lambin, 2001). Several authors have suggested the need to clearly separate these processes to obtain a comprehensive validation framework (Geist & Lambin, 2001, Pontius et al., 2001). This study focuses on the spatial distribution of changes combined with simple expert decisions on the quantity of change.

The land change probability maps were used to derive the land change allocation maps by assigning the highest probability pixels to the change value until the expected quantity of change is reached. The remaining pixels were assigned as having no change value. Based on the observed land change, we developed three usual and easy-to-test scenarios: two business-as-usual (BAU) and one alternative scenario. The first two are considering either a historical average rate of change or the past trend, 'BAU average' and 'BAU trend,' respectively. These scenarios reflect two commonly used baseline scenarios, with and without accounting for historical trend, with the former being seen as the worst case scenario (i.e., steady increase for the next 20 years), whereas the latter is more conservative. The third scenario named hereafter "alternative scenario" represents policy targets that were discussed during meetings with local stakeholders (protected area managers and local authorities). This scenario reflects national policy of deforestation reduction (REDD+ commitments presented in the introduction) and landscape restoration. Regarding the "alternative scenario," an ambitious restoration plan was launched in March 2019 by the government, with a target to restore 40,000 ha of land each year. In this study, the alternative scenario depicts an optimistic view considering a 50% decrease of both deforestation and land degradation and considering an important effort of 10,000 ha converted to sustainable land management over the next 20 years (500 ha by year). The term sustainable land management here includes activities in the field that increase the vegetation response over years compared with the initial situation. Thus, sustainable land management covers a range of human activities or practices that span abandoned agricultural land, long crop-fallow rotation, tree plantation, and agroforestry.

Predicted land change allocation maps were constructed independently for each transition and finally combined into one unique land use change map. We assumed that, in nonforested areas where land regeneration and land degradation process were predicted in the same location, priority was being given to land regeneration.

All data processing steps (Figure 2) were carried out via free and open source software: GRASS GIS (GRASS Development Team 2015), QGIS (QGIS Development Team 2009), and R (R Core Team 2015).

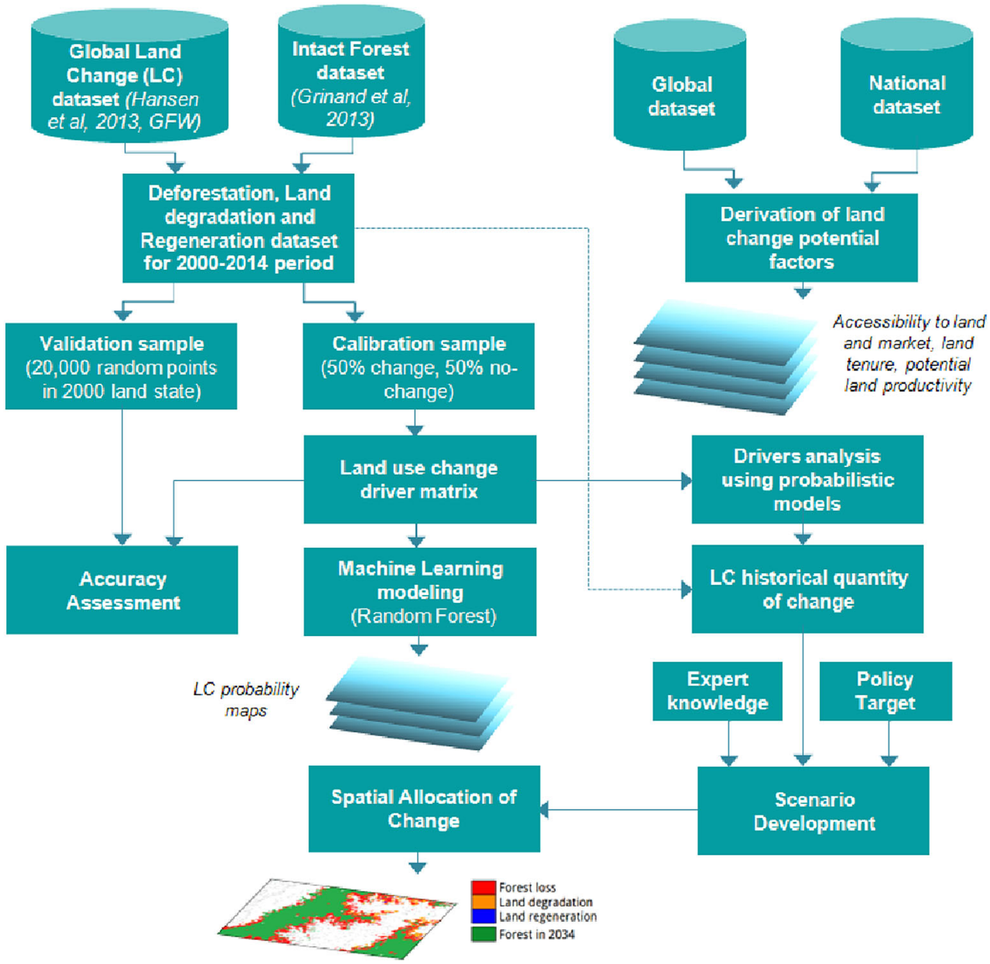

FIGURE 2 Work flow diagram of the different steps carried out in this study [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 3 Historical land change summary statistics for 2000–2014 period

Land change category	Area of change 2000–2014 (ha)	Percentage of initial state of forest or land (%)	Annual rate of change (ha yr^{-1})
Deforestation	24,834	5.76	1,774
Land degradation	38,320	3.41	2,737
Land regeneration	4,221	0.38	302

Note: These statistics were extracted using a combination of data sets from Hansen et al. (2013) and Grinand et al. (2013).

3 | RESULTS

3.1 | Observed historical land use changes

During the 2000 to 2014 period, 24,834 ha (5.76%) of forest were lost; 38,320 ha (3.41% of the nonforested initial state) of land were degraded; and 4,221 ha (0.38% of the nonforested initial state) of land underwent regeneration (Table 3). We observed an acceleration of forest loss and land degradation at a pace of +195 ha yr^{-1} and +113 ha yr^{-1}, respectively, in the last 14 years (Figure 3). Land degradation outside the intact forest was found to be quite important (2,737 ha yr^{-1}), and we observed only a few regeneration areas (302 ha yr^{-1}).

3.2 | Drivers of the location of deforestation, degradation, and regeneration

The importance of factors was analysed using empirical (Figure 4) and regression methods (Table 4) to visualize and quantify the correlation between observed change and the selected explanatory variables. We first observed a major influence of elevation and distance to the forest edge for the three land change processes under study. Areas below 700 m of elevation display a high risk of deforestation and land degradation. We observed two elevation suitability peaks at 110 and 570 m for land regeneration. Areas above 700 m are much less threatened by degradation or are much more suitable for regeneration. The proximity to the forest edge effect displayed a clear decreasing trend, with a high risk of deforestation within 200 m inside the forest, a high risk of land degradation within the 500-m buffer around the forest, and up to 800 m for land regeneration.

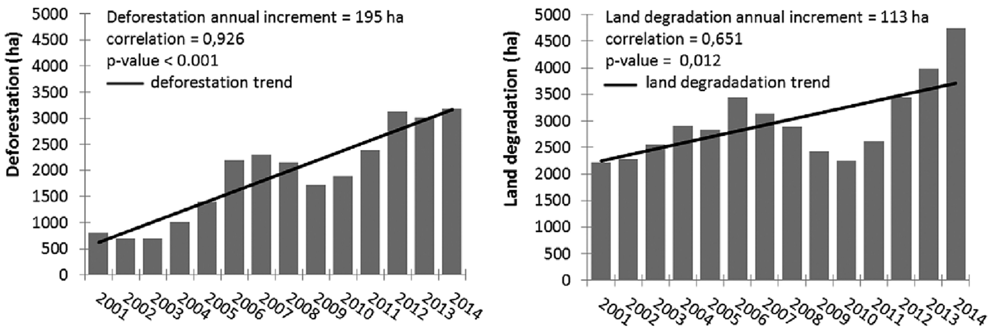

FIGURE 3 Annual deforestation and land degradation in hectares for the historical period. Values were extracted from forest loss year data product (Hansen et al., 2013) and intact forest extent (Grinand et al., 2013); deforestation is the forest loss within intact forest, and land degradation the tree loss outside intact forest. The data were smoothed with a moving window of 3 years

Slope has no effect on deforestation, in contrast to degradation and regeneration, which are more likely to occur in steep areas (>8°). Slope orientation (aspect) had the same influence for the three process, with higher suitability for sun-facing slope (north), and conversely. The number of dry months seems moderately important, with less suitability values on all transitions over areas with more than 4 dry months. This finding applies to the western part of the study area and approximately one third of the study area. Distance to the rivers does not seem to influence any land change transitions.

Proximity to main roads and towns shows a broad decreasing trend for deforestation risk but with sometimes irregular patterns. Proximity to villages and tracks is, however, clearly affecting the probability of deforestation, with high values up to 4 km. Regarding land degradation and regeneration, both transitions are affected by the main roads and towns, in a large spatial fringe, from 7 to 30 km. Proximity to villages and tracks has no importance for regeneration; however, we observed a slight increase of land degradation in areas at more than 2 km away. According to the regression analysis (Table 4), population density is significant despite a low *z* value. The relationship between the three processes and population density is low (Table 4), with no clear pattern (Figure 4). Finally, the two national parks showed contrasted responses regarding land transition (Figure 5). This will be further addressed in a subsequent section (Section 3.5).

3.3 | Land use change model accuracies

The three models were applied on an independent sampling validation data set in order to calculate accuracy measurements (Table 5). The three models showed overall accuracy above 75% for the three land use changes modelled. The RF model performed systematically better compared with the two others regarding the AUC and FOM metrics. AUC was above 0.87 for the three transitions, indicating that the three models are much better than a random model. FOM was 0.19 for deforestation, 0.11 for land degradation model, and 0.02 for land regeneration model using RF. Maximum entropy and generalized linear model were slightly better compared with RF regarding the user accuracy of change (UA_c) and the balanced user accuracy (UA).

3.4 | Land use change scenarios on the horizon 2034

Land use change maps under BAU scenarios (Table 6) revealed three land change hot spots. The average and trend BAU scenarios did not show great differences. First, the forested land that displays the highest risk of deforestation is located between the two national parks. A second change area displays land degradation around the southeast forested area and the remaining northern forested patches. Finally, the land regeneration area is essentially located in the northern area, close to the town of Midongy and adjoining the national Park (Figure 6). The alternative scenario displays reduced patches of deforestation and degradation and further highlights the northern area as being the best suited location for sustainable land management.

3.5 | Conservation threats and restoration opportunities

As we saw in Section 3.2, the two parks display a contrasted pattern regarding historical land use change. We further analysed these differences by extracting the estimated area of change for the three scenarios (Figure 7). We observed that the Andohahela National Park is weakly affected by land changes, with less than 3,000 ha of cumulative land use change estimated for the next two decades regardless of the scenario. On the other hand, land use change in the Midongy National Park can represent up to 18% of its overall area for both the BAU 'trend' scenario (more than 34,321 ha of change for the next two decades). The alternative scenario in this park shows high potential for reduced deforestation, degradation, and a clear pattern of potential regeneration (8,253 ha compared with 2,180 ha under the BAU 'average' scenario).

The remaining unprotected area shows the great extent of both deforestation and land degradation. Deforestation may affect more than 59,613 ha of forested areas in 20 years under the two BAU scenarios. The alternative scenario offers a relatively high amount of potential land regeneration (8,485 ha), but this regeneration represents only a small share (0.6%) of the total unprotected area and is located mainly in the northern part of the study area (Figure 8).

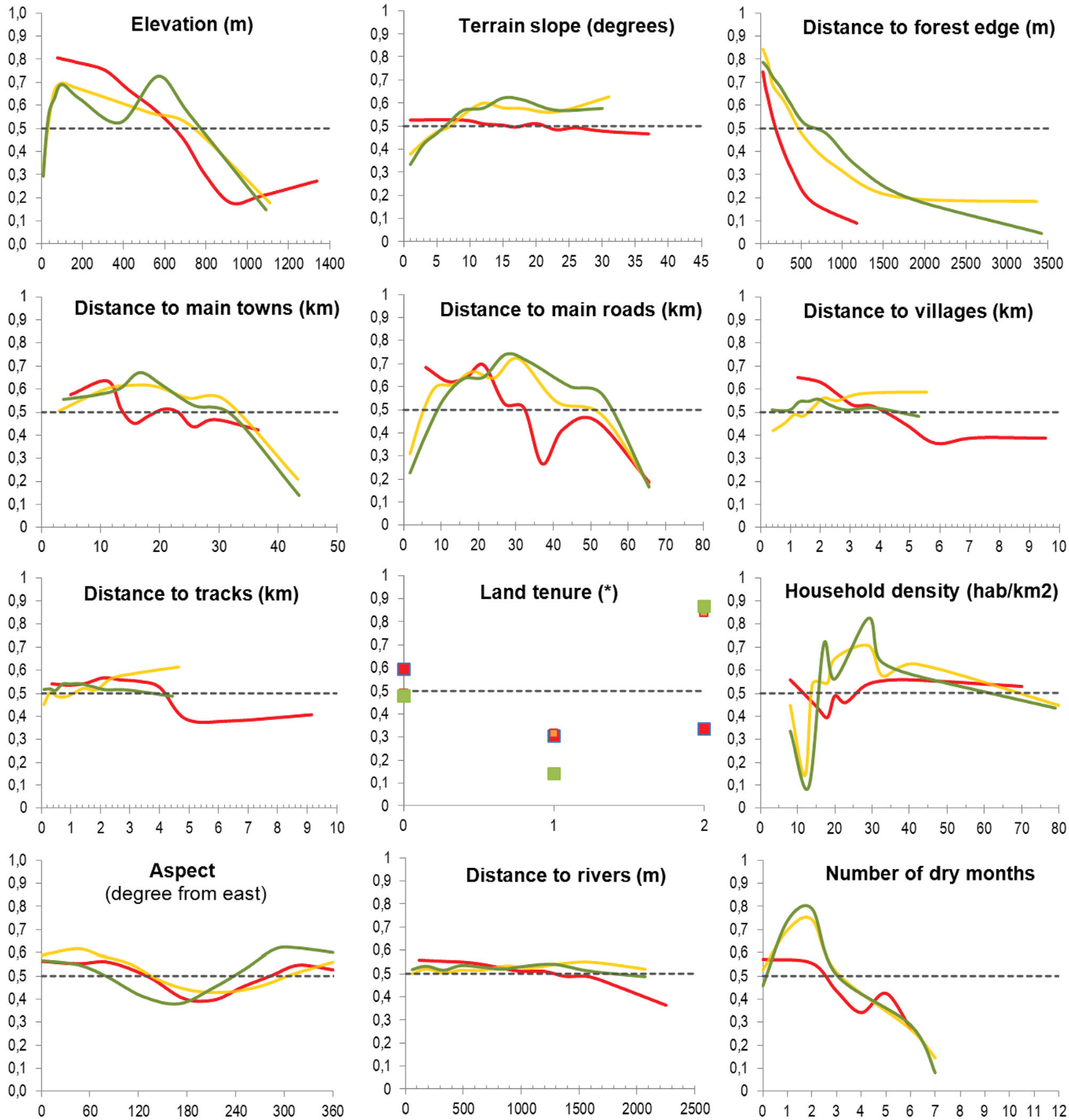

FIGURE 4 Probability distribution of deforestation (red line), land degradation (orange line), and regeneration (green line) observations. The dashed line represents the 50% probability; values above indicates high probability of land change; values below indicates low probability of land change. * indicates land tenure factor. 0 = no protected areas 1 = National park of Andohahela, 2 = National Park of Midongy [Colour figure can be viewed at wileyonlinelibrary.com]

4 | DISCUSSION

4.1 | On the drivers of the location of land use change

Elevation and proximity to the forest edge were the two first drivers explaining land use transitions. Those two biophysical and proximity local drivers were also reported to largely influence deforestation in many countries (Green et al., 2013; Armenteras et al., 2019; Bax et al., 2016; Aguilar-Amuchastegui, Riveros, & Forrest, 2014). Elevation in Madagascar is a physical barrier to human presence; the highlands above 800 m are not suitable for human settlement because of their steep slopes, dense forest, and distance from the current villages. As expected, the proximity to forest edge is positively correlated to deforestation because it is easier to clear-cut the forest at the edge than inside the forest. Land degradation also occurs at the forest edge

TABLE 4 Results of linear logistic regression

	Deforestation		Land degradation		Land regeneration	
Factors	z value	Significance	z value	Significance	z value	Significance
Intercept	*25.484*	***	*13.931*	***	*17.390*	***
Elevation	**−24.960**	***	**−23.625**	***	**−24.199**	***
Slope	−1.930	*	**11.450**	***	**13.624**	***
Aspect	−8.099	***	−6.161	***	7.031	***
Proximity forest edge	**−36.923**	***	**−33.613**	***	**−38.835**	***
Number of dry months	−5.782	***	−5.190	***	**−23.562**	***
Proximity rivers	−0.866		3.933	***	4.839	***
Parks—Andohahela	−0.229		−3.479	***	−3.475	***
Parks—Midongy	**−12.394**	***	**18.210**	***	**16.967**	***
Population Density	−3.647	***	−4.018	***	−3.650	***
Proximity main roads	3.577	***	8.932	***	**21.505**	***
Proximity main towns	5.161	***	−9.044	***	**−14.830**	***
Proximity villages	**−12.675**	***	3.592	***	−4.696	***
Proximity tracks	**11.944**	***	6.136	***	4.171	***

Note: Bold values indicate factors with *z* value above 10.
***0.001.
**0.01.
*0.05.

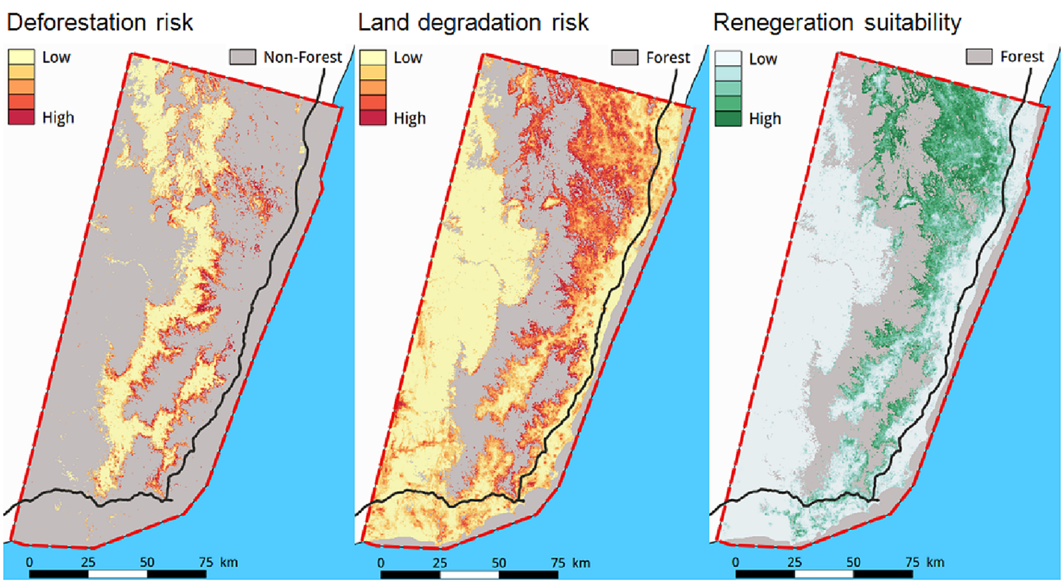

FIGURE 5 Illustration of the three land transitions maps [Colour figure can be viewed at wileyonlinelibrary.com]

area and is related to shifting cultivation practices combined with the intense rainfall that triggers soil erosion (Grinand et al., 2017). Proximity to forest edge is also a factor facilitating forest regeneration. Trees of the native forest can regenerate at a higher rate and with more diversity thanks to the presence of seed trees or frugivore seed dispersers that do not move far from the forest edge (Cubiña & Aide, 2001; McConkey et al., 2012; Wijdeven & Kuzee, 2000). Notably, most of the Malagasy tree species are adapted to dispersion by frugivorous vertebrates (Razafindratsima, 2014). The presence of small patches of forest near the intact forest edge (weakly fragmentated forest) can also contribute to the displacement of seeds dispersers on previously deforested areas (McConkey et al., 2012; Razafindratsima, 2014).

The other factors are less statistically influential but still provide important knowledge on the underlying processes. The slope had no influence on deforestation but did influence land degradation and regeneration, especially for areas of high slope. This indicates that the steep areas are subjected to deforestation (slash and burn practices or

TABLE 5 Accuracy assessment results

Land change	Model	AUC	OA	UA_C	UA_{NC}	UA	FOM
Deforestation	RF	0.90	0.78	0.19	0.99	0.59	0.19
	ME	0.84	0.91	0.26	0.95	0.61	0.15
	GLM	0.81	0.91	0.23	0.95	0.59	0.13
Land degradation	RF	0.88	0.75	0.11	0.99	0.55	0.11
	ME	0.84	0.94	0.18	0.97	0.57	0.10
	GLM	0.79	0.94	0.17	0.97	0.57	0.09
Land regeneration	RF	0.93	0.77	0.02	1.00	0.51	0.02
	ME	0.87	0.99	0.06	1.00	0.53	0.03
	GLM	0.86	0.99	0.07	1.00	0.53	0.04

Note: See Table 2 for accuracy metric definitions and formulas.
Abbreviations: AUC, area under the curve; FOM, figure of merit; GLM, generalized linear model; ME, maximum entropy; RF, random forest algorithm; OA, overall accuracy.

TABLE 6 Land change quantity scenarios for the 2014–2034 period

	Land change quantity scenario		
Land change transitions	BAU average (ha yr^{-1})	BAU trend (ha yr^{-1})	Alternative scenario
Deforestation	1,774	BAU average +195	50% decrease from 2013 level
Land degradation	2,737	BAU average +113	50% decrease from 2013 level
Land regeneration	302	302	BAU average + 10,000 ha of sustainable land management

Abbreviation: BAU, business-as-usual.

uncontrolled fire) but are also more likely to be rapidly abandoned. Abandonment could result in two contrasting phenomena in those areas, either severe and accelerated land degradation (bare soils are rapidly eroded) or soil regeneration when soils still have regenerative capacity (organic soil layer not yet eroded, well structured, and with a proximate seed 'bank'). The influence of the orientation of slope indicates that plots suitable for shifting cultivation or regeneration have a longer sun exposure, as expected. The results obtained for proximity to roads, towns, or villages suggest that the main roads and towns have different levels of attractiveness according to the city involved. To better account for these socioeconomic factors, one should go deeper into the type of the location or roads (not only two types), for instance, according to the number of food markets, density, or quality of the road. The influence of population density also displays an odd shape. This display was interpreted as being caused by specific local conditions, where population density is not the key factor but instead indicates the local governance or planning leadership, which can be different from one county (*fokontany*) to another.

Surprisingly, distance to the rivers did not appear to influence any land change transitions. This could be explained by two factors: First, rivers are not used as the main transportation means as in other countries, and second, irrigation systems are not well developed, so the agriculture relies essentially on rain-fed crop. Furthermore, numerous water courses exist over the studied area, yielding an explanatory variable with a limited range of values (from 0 to 2.5 km, Figure 4), which may hinder detection of its effect.

4.2 | On the contrasted effectiveness of contrasting efforts

We observed that the two national parks that lie in the study area have very distinct threats of and opportunities for land change, the former being only little affected in contrast to the latter, which exhibits a high rate of change. The reasons for such differences are the historical conservation activities and the socioeconomic conditions in the neighbouring communities. Indeed, Andohahela was created 60 years ago (in 1939) in contrast to Midongy, which was created more recently (in 1997). This underlines the effectiveness of long-term conservation activities. From a modelling perspective, this difference highlights the role of time or the time feedback involved in such land use explanatory variables. This understanding should be considered carefully when building scenarios based on change in land tenure or rights, as these factors imply a lag in the cause–effect relationship or elasticity.

Moreover, population density is more important around Midongy than around Andohahela (0.17 villages by square kilometre for Midongy versus 0.14 for Andohahela in the 5-km buffer around the National Parks), which increases the anthropogenic pressure on the forest. At the northeastern edge of Midongy, many people are settled, and the National Park is the only significant forest area accessible to the local population. In addition, several roads crossover the park, making it accessible. All these factors, which determine the pressure of the population that seeks access to forest for agriculture, wood fuel, and timber, can explain the higher rate of deforestation in

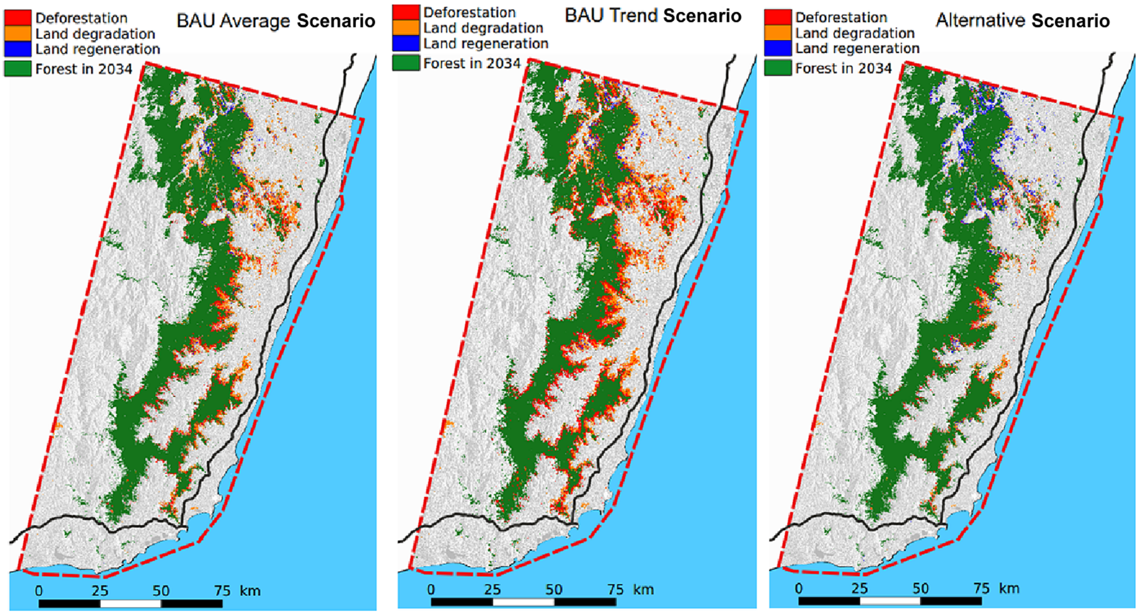

FIGURE 6 Land change maps for the 2014–2034 period according to the three scenarios described in Table 4. BAU, business-as-usual [Colour figure can be viewed at wileyonlinelibrary.com]

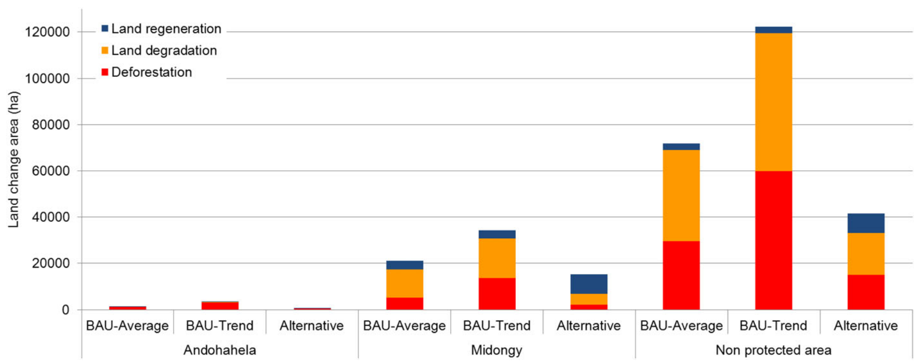

FIGURE 7 Land change allocation results according to the three scenarios for different extents: the Andohela National Park, the Midongy National Park, and outside those two perimeters. BAU, business-as-usual [Colour figure can be viewed at wileyonlinelibrary.com]

Midongy National Park than in Andohahela, even though those parks are managed by the same public entity.

4.3 | On the methodology

Spatially explicit land change models are legitimate for their scientific empirical soundness, reproducibility, and ability to be assessed by validation procedures (Castella & Verbug, 2007). In this study, the use of the RF machine learning algorithm provided satisfactory results, although it was not as robust for user accuracy of change as the other inference models tested. This model was recently applied in a deforestation modelling application (Dezécache et al., 2017) but was not compared with other models to our knowledge. No unique good model exists; however, the machine learning algorithm and model averaging may provide new solutions to increase our prediction ability. We observed, as others before have reported (Pontius et al., 2008, Sloan and Pelletier, 2012), that the accuracy of the predicted change relies on the amount of change observed. This was illustrated with the land regeneration models that provided very low FOM values. This shortage could be remediated by increasing the number of years of historical observations (Sloan and Pelletier, 2012). The use of a distinct calibration and validation period is often seen as a good practice for accuracy assessment (Shoch et al., 2013), but changes between the calibration and validation period in terms of quantity of change or relative importance of drivers can generate systematic errors (Camacho Olmedo et al., 2015). In addition, a distinct validation period reduces the number of observations required for calibrating the models and our ability to understand ongoing changes. The 14-year period used in this study is considered sufficient to capture such subtle land change processes as land degradation and regeneration. Indeed, if one takes the soil organic carbon as the biophysical indicator of both land degradation and land regeneration, as suggested by the United Nations Convention to Combat Desertification, the literature reports that significant changes could occur—and can be detected—at times scales of a few years and of decades for both processes, respectively, in the tropics (Don, Schumacher, & Freibauer, 2011).

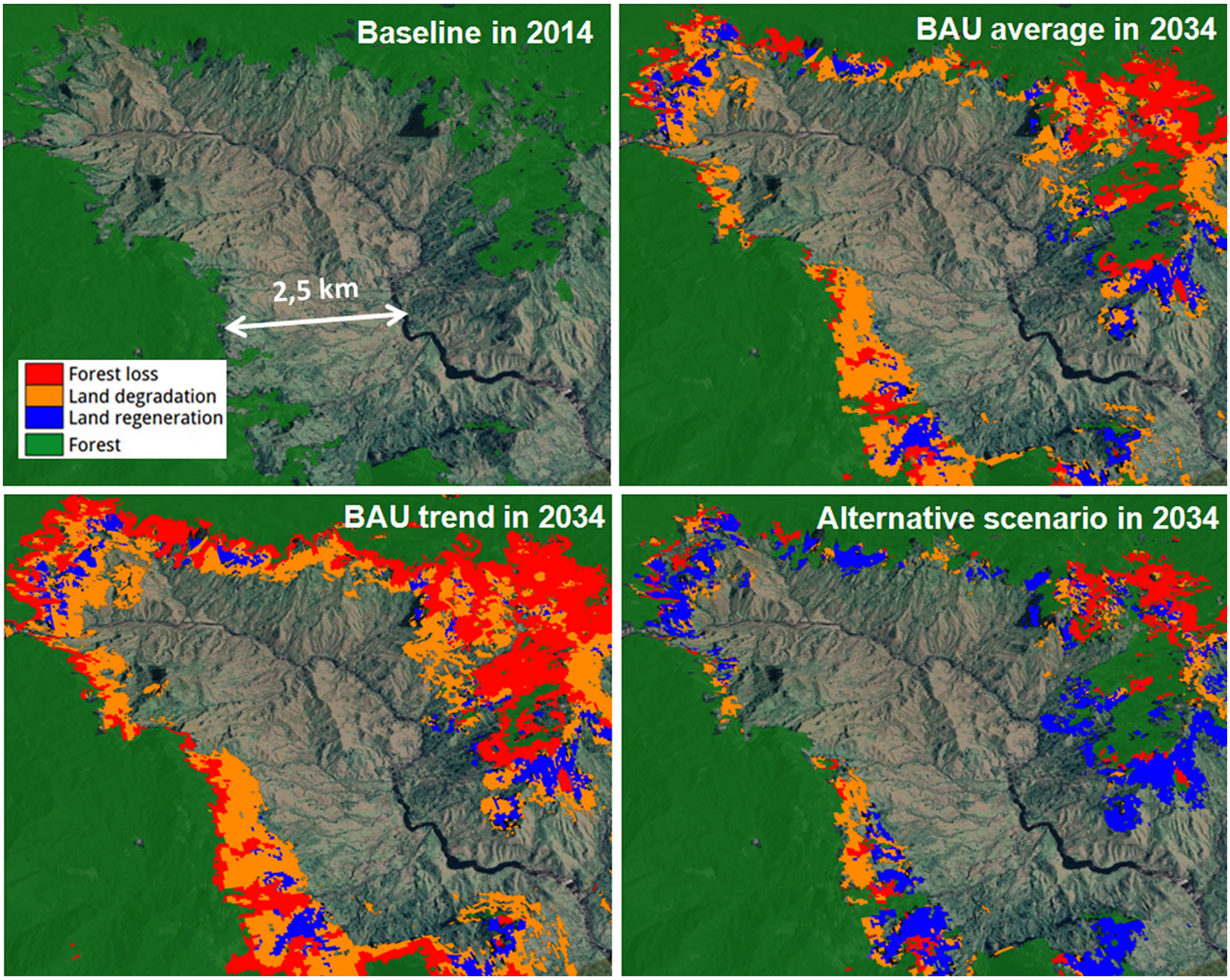

FIGURE 8 Landscape 3D view of the 2014 baseline and the land change outputs in 2034 obtained for the three scenarios and overlaid over Google Earth imagery. BAU, either using the historical average amount of change ("BAU average") or with consideration of the historical trend ('BAU trend'). BAU, business-as-usual [Colour figure can be viewed at wileyonlinelibrary.com]

Other limitations exist in the application of spatial modelling techniques to forecast land degradation and regeneration that are related to the definition and to pattern recognition. First, no commonly agreed upon quantitative land degradation definition exists at the global or local scale. In this study, we considered land degradation as the removal of vegetation or tree cover at a 30-m resolution. This is a similar approach as the land productivity change indicator used as a proxy of land degradation worldwide (Brandt et al., 2018; Cherlet et al., 2018; United Nations Environment Programme, 2012; Yengoh et al., 2015). Second, the gain of vegetation is a slow process and currently available only for 2000–2012 in Hansen et al. (2013) and considers a no-tree cover in 2000. Other definitions or input data set of land regeneration or degradation would impact the results.

Finally, spatially explicit projections fail to capture change other than that of "frontier" change, that is, deforestation front along the forest edge (Sloan and Pelletier, 2012). In this study, deforestation and degradation were fairly accurately predicted as both processes relied highly on the forest edge variability. The location of regeneration was also partly explained by forest edge, which, in reality, may provide spurious results because small-scale regeneration may occur far from forest resources. Indeed, regeneration potential is steered by forest or agricultural management strategy, at a fine scale. Addressing the regeneration potential requires more than spatial factors and requires an understanding of sociocultural and economic drivers. For instance, an important land regeneration factor is the reduction of the rotation of the crop-fallow length system (Labrière et al., 2015), but this factor cannot be spatialized. However, we believe our results on land regeneration allocation maps may help policymakers and stakeholders to define appropriate interventions, even at a local scale (Figure 8).

5 | CONCLUSIONS

The objective of this paper was to test and evaluate a new spatially explicit land change modelling approach for the simultaneous forecast and assessment of three main environmental processes (deforestation, land degradation, and land regeneration) in one of the most biodiversity-rich areas of the world.

Empirical driver analysis allowed the identification of threshold values or tipping points regarding the potential land change drivers tested. Amongst them, two biophysical and socioeconomic factors (elevation and proximity to the forest edge) stand out in explaining the three processes studied. The results highlight the nonlinear relationship of the drivers with the processes, which argue for the use of nonlinear inference models such as machine learning algorithms.

The land change modelling approach developed in this study is a first attempt to explore the potential of machine learning tools combined with easy-to-access global land change data sets in an open source modelling framework. The land change allocation and suitability maps can be easily improved, as new or better quality input data sets are made available or replicated to other regions. We believe that such an approach can produce consistent and scalable information that can accompany land use planning processes and help target interventions for preventing land degradation and maximizing the chance of land regeneration. Land restoration planning can benefit from such information on appropriate areas, but stakeholder participation and ground surveys are needed to assess land rights, land use conflicts, and implications for food security.

ACKNOWLEDGMENTS

This research was funded by the French Biodiversity Research Foundation (FRB–FFEM (Fondation pour la Recherche sur la Biodiversité—Fond Français pour l'Environnement Mondial) through the BioSceneMada project (project agreement AAP-SCEN-2013 I) and European Commission through the Roadless Forests project. The first author was funded by an ANRT PhD scholarship (CIFRE N 2012-1153) provided jointly by the Nitidae/Etc Terra association, the Institut de Recherche pour le Développement (IRD), the Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD), and the Laboratoire des Radio Isotopes (LRI).

ORCID

Clovis Grinand https://orcid.org/0000-0003-2650-2829
Ghislain Vieilledent https://orcid.org/0000-0002-1685-4997

REFERENCES

Don, A., Schumacher, J., Freibauer, A. (2011). Impact of tropical land-use change on soil organic carbon stocks – A meta-analysis. *Global Change Biology*, *17*, 1658–1670. https://doi.org/10.1111/j.1365-2486.2010.02336.x

Aguilar-Amuchastegui, N., Riveros, J. C., & Forrest, J. L. (2014). Identifying areas of deforestation risk for REDD+ using a species modeling tool. *Carbon Balance and Management*, *9*, 10. https://doi.org/10.1186/s13021-014-0010-5

Armenteras, D., Murcia, U., González, T. M., Barón, O. J., & Arias, J. E. (2019). Scenarios of land use and land cover change for NW Amazonia: Impact on forest intactness. *Global Ecology and Conservation*, *17*, 2351–9894, e00567. https://doi.org/10.1016/j.gecco.2019.e00567

Baccini, A., Goetz, S. J., Walker, W. S., Laporte, N. T., Sun, M., Sulla-Menashe, D., … Friedl, M. A. (2012). Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nature Climate Change*, *2*, 182–185. https://doi.org/10.1038/nclimate1354

Bai, Z., Dent, D., Wu, Y., de Jong, R.. 2013. Land degradation and ecosystem services. Chapter 15. 357–381. In R. Lal et al. (eds.), Ecosystem Services and Carbon Sequestration in the Biosphere. https://doi.org/10.1007/978-94-007-6455-2_15

Bax, V., Francesconi, W., & Quintero, M. (2016). Spatial modeling of deforestation processes in the central Peruvian Amazon. *Journal for Nature Conservation*, *29*(2016), 79–88. http://dx.doi.org/10.1016/j.jnc.2015.12.002

Brandt, M., Rasmussen, K., Hiernaux, P., Herrmann, S., Tucker, C.J., Tong, X., Tian, F., Mertz, O., Kergoat, L., Mbow, C., David, J.L., Melocik, K., Dendoncker, M., Vincke, C., Fensholt, R. 2018. Reduction of tree cover in west African woodlands and promotion in semi-arid farmlands. *Nature Geoscience 11*, 328–333. https://doi.org/10.1038/s41561-018-0092-x

Camacho Olmedo, M. T., Pontius, R. G., Jr., Paegelow, M., & Mas, J. F. (2015). Comparison of simulation models in terms of quantity and allocation of land change. *Environmental Modeling & Software*, *69*, 214–221. https://doi.org/10.1016/j.envsoft.2015.03.003

Castella, J.-C., & Verbug, P. H. 2007. Combination of process-oriented and pattern-oriented models of land-use change in a mountain area of Vietnam. *Ecological Modelling*, *202*, 410–420. https://doi.org/10.1016/j.ecolmodel.2006.11.011

Cherlet, M., Hutchinson, C., Reynolds, J., Hill, J., Sommer, S., von Maltitz, G. (Eds.). (2018). *World atlas of desertification 3rd edition, rethinking land degradation and sustainable land managemen*, Luxembourg: Publication Office of the European Union.

Cubiña, A., & Aide, T. M. (2001). The effect of distance from forest edge on seed rain and soil seed bank in a tropical Pasture 1. *Biotropica*, *33*, 260–267. https://doi.org/10.1111/j.1744-7429.2001.tb00177.x

Dezécache, C., Faure, E., Gond, V., Jean-Michel Salles, J. M., Ghislain Vieilledent, G., & Hérault, B. (2017). Gold-rush in a forested El Dorado: Deforestation leakages and the need for regional cooperation. *Environmental Research Letters*, *12*(3). https://doi.org/10.1088/1748-9326/aa6082

Eastman, J.R. 2012. *The land change modeler for ecological sustainability (chapter 21) IDRISI Selva manual (version 17)*. Clark Labs.

Ferretti-Gallon, K., & Bush, J. (2014). What drives deforestation and what stops it? A meta-analysis of spatially explicit econometric studies. In *CGD working paper 361*. Washington, DC: Center for Global Development. http://www.cgdev.org/publication/what-drives-deforestation-and-what-stops-it-meta-analysis-spatially-explicit-econometric

Geist, HJ, Lambin, EF. 2001. What drives tropical deforestation? A meta-analysis of proximate and underlying causes of deforestation based on subnational case study evidence. Land-Use and Land-Cover Change International Project Office. Report Series 4. Louvain-la-Neuve, Belgium, 136 p.

Gislason, P. O., Benediktsson, J. A., & Sveinsson, J. R. (2006). Random forests for land cover classification. *Pattern Recognition Letters*, *27*, 294–300. https://doi.org/10.1016/j.patrec.2005.08.011

Gounaridis, D., Ioannis Chorianopoulos, I, Symeonakis, E., Sotirios Koukoulas, S., 2019. A random forest-cellular automata modelling approach to explore future land use/cover change in Attica (Greece), under different socio-economic realities and scales. *Science of the Total Environment. 646*, 320–335. https://doi.org/10.1016/j.scitotenv.2018.07.302

Green, J. M. H., Larrosa, C., Burgess, N. D., Balmford, A., Johnston, A., Mbilinyi, B. P., Platts, P. J., Coad, L. (2013). Deforestation in an African biodiversity hotspot: Extent, variation and the effectiveness of protected areas, *Biological Conservation*, *164*, 62–72, ISSN 0006–3207, https://doi.org/10.1016/j.biocon.2013.04.016

Grinand, C., Le Maire, G., Vieilledent, G., Razakamanarivo, H., Razafimbelo, T., & Bernoux, M. (2017). Estimating temporal changes in soil carbon stocks at ecoregional scale in Madagascar using remote-sensing. *Int. Journal of Applied Earth Observation and Geoinformation*, *54*, 1–14. https://doi.org/10.1016/j.jag.2016.09.002

Grinand, C., Rakotomalala, F., Gond, V., Vaudry, R., Bernoux, M., & Vieilledent, G. (2013). Estimating deforestation in tropical humid and dry forests in Madagascar from 2000 to 2010 using multi-date Landsat satellite images and the random forests classifier. *Remote Sensing of Environment*, *139*, 68–80. https://doi.org/10.1016/j.rse.2013.07.008

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., … Townshend, J. R. G. (2013). High-resolution global maps of 21st-century Forest cover change. In *Science 342 (15 November): 850–53. Data available on-line from* (Vol. 342, pp. 850–853). http://earthenginepartners.appspot.com/science-2013-global-forest

Harper, G., Steininger, M. K., Tucker, C. J., Juhn, D., & Hawkins, F. (2007). Fifty years of deforestation and forest fragmentation in Madagascar. *Environmental Conservation*, *34*, 1–9. https://doi.org/10.1017/S0376892907004262

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*(15), 1965–1978. https://doi.org/10.1002/joc.1276

Huettner, M., Leemans, R., Kok, K., & Ebeling, J. (2009). A comparison of baseline methodologies for 'Reducing Emissions from Deforestation and Degradation. *Carbon Balance and Management*, *4*, 4. https://doi.org/10.1186/1750-0680-4-4

IPCC. (2014). *Climate change 2014: Impacts, adaptation, and vulnerability. Part a: Global and sectoral aspects.Contribution of working group II to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge, United Kingdom and NewYork, NY, USA: Cambridge University Press, (p. 1132).

RPP, (2014). REDD+ Readiness Preparation Proposal – Madagascar. https://www.forestcarbonpartnership.org/sites/fcp/files/2014/August/R-PP%20June%202014.pdf

R-PIN, (2008). REDD+ Readiness Project Idea Note – Madagascar. https://www.forestcarbonpartnership.org/sites/forestcarbonpartnership.org/files/Madagascar_FCPF_R-PIN.pdf

ER-PIN, 2015. Emission reductions program idea note - testing emissions reductions in the rainforest ecoregion. https://www.forestcarbonpartnership.org/sites/fcp/files/2015/September/MDG_ERPIN_English%20with%20annexes.pdf

Labrière, N., Laumonier, Y., Locatelli, B., Vieilledent, G., & Comptour, M. (2015). Ecosystem services and biodiversity in a rapidly ransforming landscape in northern borneo. *PLoS ONE*, *10*(10), e0140423. https://doi.org/10.1371/journal.pone.0140423

McConkey, K. R., Prasad, S., Corlett, R. T., Campos-Arceiz, A., Brodie, J. F., Rogers, H., & Santamaria, L. (2012). Seed dispersal in changing landscapes. *Biological Conservation*, *146*, 1–13. https://doi.org/10.1016/j.biocon.2011.09.018

ONE, D. G. F., MNP, W. C. S., & Terra, E. (2015). *Changement de la couverture de forêts naturelles à Madagascar, 2005–2010-2013*. Madagascar: Antananarivo 21p.

Pontius, R., Boersma, W., Castella, J. C., Clarke, K., de Nijs, T., Dietzel, C., … Verburg, P. H. (2008). Comparing the input, output, and validation maps for several models of land change. *The Annals of Regional Science*, *42*, 11–13. https://doi.org/10.1007/s00168-007-0138-2

Pontius, R. G., Jr., Cornell, J., & Hall, C. (2001). Modeling the spatial pattern of land-use change with GEOMOD2: Application and validation for Costa Rica. *Agriculture, Ecosystems & Environment*, *85*(1–3), 191–203. https://doi.org/10.1016/S0167-8809(01)00183-9

Rakotomala, F.A, Rabenandrasana, J. C., Andriambahiny, J. E. 4, Rajaonson, R., Andriamalala, F., Burren, C., Rakotoarijaona, J.R., Parany, L., Vaudry, R., Rakotoniaina, S., Grinand, C. 2015. Estimation de la déforestation des forêts humides à Madagascar entre 2005, 2010 et 2013. *Revue Française de Télédétection et Photogrammétrie*, 211–212, 11–23

Razafindratsima, O. H. (2014). Seed dispersal by vertebrates in Madagascar's forests: Review and future directions. *Madagascar Conservation & Development*, *9*, 90–97. https://doi.org/10.4314/mcd.v9i2.5

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

Shoch, D, Eaton, J, Settelmyer, S. 2013. Project Developer's guidebook to Volontary carbon standard REDD methodologies. *Conservation International*, 97pp.

Sloan, S., & Pelletier, J. (2012). How accurately may we project tropical forest-cover change? A validation of a forward-looking baseline for REDD. *Global Environmental Change*, *22*(2012), 440–453. https://doi.org/10.1016/j.gloenvcha.2012.02.001

Soares-Filho, B.S., Rodrigues, H.O.and W.L.S. Costa. 2009. Modeling environmental dynamics with Dinámica EGO. 115 p. ISBN: 978-85-910119-0-2. Available at: www.csr.ufmg.br/dinamica/tutorial/Dinamica_EGO_guidebook.pdf.

UNEP. (2012). *Land health surveillance: An evidence-based approach to land ecosystem management* (p. 211). Nairobi: Illustrated with a Case Study in the West Africa Sahel. United Nations Environment Programme.

Veldkamp, A., & Lambin, E. F. (2001). Predicting land-use change. *Agriculture, Ecosystems and Environment*, *85*, 1–6. https://doi.org/10.1016/S0167-8809(01)00199-2

Verburg, P. H. (2006). 2006. Simulating feedbacks in land use and land cover models. *Landscape Ecology*, *21*, 1171–1183. https://doi.org/10.1007/s10980-006-0029-4

Vieilledent, G., Gardi, O., Grinand, C., Burren, C., Andriamanjato, M., Camara, C., … Rakotoarijaona, J. (2016). Bioclimatic envelope models predict a decrease in tropical forest carbon stocks with climate change in Madagascar. *Journal of Ecology*, *104*, 703–715. https://doi.org/10.1111/1365-2745.12548

Vieilledent, G., Grinand, C., Rakotomalala, F. A., Ranaivosoa, R., Rakotoarijaona, J.-R., Allnutt, T. F., & Achard, F. (2018). Combining global tree cover loss data with historical national forest-cover maps to look at six decades of deforestation and forest fragmentation in Madagascar. *Biological Conservation*, *222*, 189–197. https://doi.org/10.1016/j.biocon.2018.04.008

Vieilledent, G., Grinand, C., & Vaudry, R. (2013). Forecasting deforestation and carbon emissions in tropical developing countries facing demographic expansion: A case study in Madagascar. *Ecology and Evolution.*, *3*, 1702–1716. https://doi.org/10.1002/ece3.550

Weatherley-Singh, J., & Gupta, A. (2015). Drivers of deforestation and REDD + benefit-sharing: A meta-analysis of (missing) link. *Environmental Science & Policy*, *54*, 97–105. https://doi.org/10.1016/j.envsci.2015.06.017

Wehkamp, J., Aquino, A., Fuss, S., & Reed, E. W. (2015). Analyzing the perception of deforestation drivers by African policy makers in light of possible REDD+ policy responses. *Forest Policy and Economics*, *59*, 7–18. https://doi.org/10.1016/j.forpol.2015.05.005

Wijdeven, S. M. J., & Kuzee, M. E. (2000). Seed availability as a limiting factor in Forest recovery processes in Costa Rica. *Restoration Ecology*, *8*, 414–424. https://doi.org/10.1046/j.1526-100x.2000.80056.x

Yengoh, G. T., Dent, D., Olsson, L., Tengberg, A. E., & Tucker, C. J. (2015). *Use of the normalized difference vegetation index (NDVI) to assess land degradation at multiple scales current status, future trends, and practical considerations*. London, UK: Springer Briefs in Environmenal Science, 124 pages.

How to cite this article: Grinand C, Vieilledent G, Razafimbelo T, Rakotoarijaona J-R, Nourtier M, Bernoux M. Landscape-scale spatial modelling of deforestation, land degradation, and regeneration using machine learning tools. *Land Degrad Dev*. 2020;1–14. https://doi.org/10.1002/ldr.3526