Article

Global potential for natural regeneration in deforested tropical regions

https://doi.org/10.1038/s41586-024-08106-4

Received: 4 August 2023

Accepted: 24 September 2024

Published online: 30 October 2024

Open access

Check for updates

Brooke A. Williams[1,2,3,4,15] ✉, Hawthorne L. Beyer[1,5,15], Matthew E. Fagan[6], Robin L. Chazdon[1,7], Marina Schmoeller[1], Starry Sprenkle-Hyppolite[8], Bronson W. Griscom[8], James E. M. Watson[3,4], Anazélia M. Tedesco[1,3,4], Mariano Gonzalez-Roglich[9,10], Gabriel A. Daldegan[9], Blaise Bodin[11], Danielle Celentano[12], Sarah Jane Wilson[13], Jonathan R. Rhodes[2,3,4], Nikola S. Alexandre[8,14], Do-Hyung Kim[14], Diego Bastos[1] & Renato Crouzeilles[1,5]

Extensive forest restoration is a key strategy to meet nature-based sustainable development goals and provide multiple social and environmental benefits[1]. Yet achieving forest restoration at scale requires cost-effective methods[2]. Tree planting in degraded landscapes is a popular but costly forest restoration method that often results in less biodiverse forests when compared to natural regeneration techniques under similar conditions[3]. Here we assess the current spatial distribution of pantropical natural forest (from 2000 to 2016) and use this to present a model of the potential for natural regeneration across tropical forested countries and biomes at a spatial resolution of 30 m. We estimate that an area of 215 million hectares—an area greater than the entire country of Mexico—has potential for natural forest regeneration, representing an above-ground carbon sequestration potential of 23.4 Gt C (range, 21.1–25.7 Gt) over 30 years. Five countries (Brazil, Indonesia, China, Mexico and Colombia) account for 52% of this estimated potential, showcasing the need for targeting restoration initiatives that leverage natural regeneration potential. Our results facilitate broader equitable decision-making processes that capitalize on the widespread opportunity for natural regeneration to help achieve national and global environmental agendas.

Extensive forest restoration is critical for mitigating climate change[4]. Natural climate solutions remain the most under-invested climate mitigation opportunity among all climate mitigation sectors[5] and, specifically, forest restoration is one of the five largest cost-effective climate mitigation options across all sectors[6]. Global climate mitigation policy platforms have set ambitious forest restoration targets to substantially increase the area of natural ecosystems by 2050, such as the Bonn Challenge, which aims to restore 350 million hectares (Mha) by 2030[7], and target 2 of the recently adopted Global Biodiversity Framework, which calls for 30% of the area of degraded ecosystems to be brought under restoration by 2030[8]. Forest restoration at this scale could sequester hundreds of gigatonnes of carbon and reduce the risk of extinction for thousands of species, particularly if strategic planning is used to prioritize areas for restoration[4]. Tropical forested regions are particularly important owing to their unparalleled biodiversity[9], the magnitude of economic, cultural and recreational services that they provide to people and their rapid growth rates relative to other forest types[10], and because large areas have already been cleared and degraded[11]. The salience of reforestation is recognized internationally, with 51 out of 55 nations' nationally determined contributions to the Paris Agreement containing targets related to forests as a nature-based solution to climate change mitigation and adaptation[12].

Identifying areas where forests can recover effectively with minimal intervention is critical for achieving forest restoration at scale. Tree planting and extensive site preparation are popular strategies and can be effective, particularly when using locally adapted native tree species in mixtures[13]. However, implementing tree planting at large scales is prohibitively expensive, especially for developing nations, and only sometimes effectively helps native biodiversity to recover[14]. In areas where ecological conditions are such that forests can grow back on their own or with low-cost assistance, natural regeneration methods are less costly (for example, US$12–3,880 per ha compared with US$105–25,830 for forest restoration projects in the tropics and subtropics[15]) and often are more effective than full tree planting in terms of their long-term success rates and biodiversity outcomes[3]. Yet natural regeneration has been underused as a restoration strategy, in part because planners and implementers lack knowledge of where the

[1]Institute for Capacity Exchange in Environmental Decisions, Canberra, Australian Capital Territory, Australia. [2]School of Biology and Environmental Science, Queensland University of Technology, Brisbane, Queensland, Australia. [3]School of the Environment, University of Queensland, St Lucia, Queensland, Australia. [4]Centre for Biodiversity and Conservation Science, University of Queensland, St Lucia, Queensland, Australia. [5]Mombak, São Paulo, Brazil. [6]Department of Geography and Environmental Systems, University of Maryland Baltimore County, Baltimore, MD, USA. [7]Forest Research Institute, University of the Sunshine Coast, Sippy Downs, Queensland, Australia. [8]Center for Natural Climate Solutions, Conservation International, Arlington, VA, USA. [9]Moore Center for Science, Conservation International, Arlington, VA, USA. [10]Wildlife Conservation Society, Buenos Aires, Argentina. [11]Amazon Conservation Association, Washington, DC, USA. [12]Instituto Socioambiental, São Paulo, Brazil. [13]School of Environmental Studies, University of Victoria, Victoria, British Columbia, Canada. [14]United Nations Children's Fund, New York, NY, USA. [15]These authors contributed equally: Brooke A. Williams, Hawthorne L. Beyer. ✉e-mail: brooke.a.williams@qut.edu.au

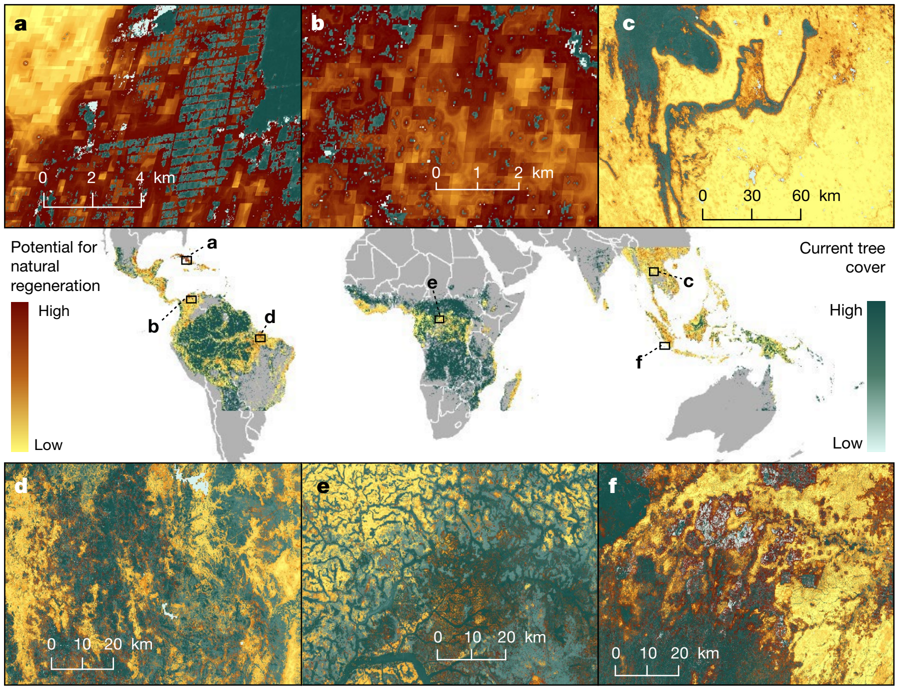


Fig. 1 | The potential for natural regeneration. a–f, The potential for natural regeneration across the global tropics and subtropics, with examples across Cuba (**a**), Colombia (**b**), Thailand (**c**), Brazil (**d**), the Democratic Republic of Congo (**e**) and Indonesia (**f**). The areas in grey were excluded from this analysis and were not considered to be available for forest regeneration. Tree cover is represented according to ref. 11 for the year 2018.

process can occur and the time it will take to deliver socioeconomic and environmental benefits[3].

The restoration community currently lacks an effective tool for predicting where natural regeneration is most likely to occur and therefore offer multiple benefits of native forest recovery with higher certainty. This gap is a key reason why approaches based on natural regeneration are not applied at larger scales[16]. Previous approaches to mapping forest restoration potential[17,18] have relied on coarse-scale data, expert opinion and assumptions about potential forest cover[19–21], rather than actual data on natural forest regrowth after deforestation, all of which leads to underestimates of natural regeneration potential. By using robust pantropical data on where forests have grown back to quantify the potential for natural regeneration at a high spatial resolution, our study provides essential guidance to enable forest restoration planning and policy in countries with tropical forest.

Tropical natural regeneration potential

We build on the methods used for a recently published pantropical remote sensing analysis[22] that identified 31.6 ± 11.9 Mha of natural regrowth in 4.78 million patches globally between 2000 and 2012 that persisted to 2016 (hereafter, 2000–2016), with an average patch area of 1.2 ha, to assess the scale and locations where natural regeneration has potential in tropical forest biomes (±25° latitude). The analysis[22] distinguished areas of natural regrowth from plantations on the basis of extensive training data and ground-truth information; patches of natural regrowth were defined as at least 0.45 ha in area with vegetation taller than 5 m in height.

Based on a sample of 5.4 Mha of natural regrowth detected previously[22], we used machine learning methods to distinguish areas in which natural regeneration did or did not occur across the global tropics as a function of a suite of geospatial, biophysical and socioeconomic variables. To predict potential tropical forest regeneration across Earth, we selected predictor variables spanning local (site) and landscape scales (from 30 m to country level) that are known to influence the potential for tropical forest regrowth. These biophysical and socioeconomic variables include distance to nearest tree cover; local forest density in a 1 km^2 area; land cover; 12 soil metrics reflecting conditions in the top 30 cm of soil; 19 bioclimatic variables reduced to 5 principal component axes; slope; net primary productivity; monthly average of burned area due to wildfires over the period of 2001–2017; distance to water; population density; gross domestic product; human development index; road density; distance to urban areas; and protected area status (Supplementary Table 2). Spatially explicit values of predictor variables were then used to model the relative potential for natural regeneration in the present (2015) and in the near future (2030), assuming that overall conditions from 2000 to 2016 apply to future scenarios (Fig. 1). For reporting purposes, we translate the resulting continuous potential for natural regeneration value (0–1) to area-based values by multiplying the relative potential for natural regeneration by the area of each pixel, resulting in a weighted-area value. We provide a sensitivity analysis to this calculation where area is regarded as cells that have greater than 50% potential (Supplementary Table 4). The continuous potential for natural regeneration can be interpreted as the probability of natural regeneration per pixel; the weighted-area value is therefore the expected area of natural regeneration per pixel.

We estimate that biophysical conditions can support natural regeneration in tropical forests over 215 Mha (confidence interval (CI) = 214.78–215.22 Mha) globally (Fig. 2; neotropics: 98 Mha (CI = 97.80–98.20 Mha); Indomalayan tropics: 90 Mha (CI = 89.82–90.18 Mha); and Afrotropics: 25.5 Mha (CI = 25.47–25.53 Mha)) until 2030. This is a smaller estimate, unsurprisingly given our smaller study region (the pan tropics), compared with other studies such as ref. 10, in which it was estimated that natural forest regrowth could be biophysically possible across 349 Mha (constrained by existing restoration commitments) and 678 Mha (maximum potential) globally. Five countries—Brazil, Indonesia, China, Mexico and Colombia—account for 52% of the global potential for natural regeneration (Fig. 2; 20.3, 13.6, 7.2, 5.6 and 5.2%, respectively). This result is unsurprising given the large areas of land that these countries cover (24.3% of the study region), but presents important implications for restoration action. We find that 29 other countries have natural regeneration potential areas of over 1 Mha

Fig. 2 | Land area that can support natural regeneration. Summary of the area of land that can support natural forest regeneration (has potential for natural regeneration) shown in green, with the total land area available for forest restoration shown in yellow (in Mha) for countries located at least in part within humid tropical and subtropical forest biomes (tropical and subtropical dry broadleaf forests, tropical and subtropical moist broadleaf forests, and tropical and subtropical coniferous forests). The full table is provided in Supplementary Table 3.

each (Supplementary Table 3). Although some of this area could regenerate without direct human intervention, assisted natural regeneration methods are likely to improve the rate of regeneration and carbon and biodiversity outcomes[2]. This means that some management actions to remove obstacles to forest regeneration (for example, excluding grazers, fire protection, removal of invasive species) would be required[2,13]. Importantly, our analysis does not specify the type of policy approach or management action, if any, that may be required to facilitate or enhance natural regeneration. Such management decisions should be made locally on the basis of local conditions and restoration objectives defined by local stakeholders.

The accuracy of the predictions as assessed through validation was 87.9% based on autocorrelation effects (further details are provided in the Methods and Supplementary Information 1). Spatial variation in accuracy was considerable, with the lowest accuracies occurring in portions of Southeast Asia. By comparison, the out-of-bag accuracy estimated during model-fitting was 87.8%, indicating good model fit and minimal overfitting. At the biome level, validation accuracies were similarly high, with 87.9%, 87.9% and 87.8% in tropical and subtropical moist broadleaf forests, tropical and subtropical dry broadleaf forests, and tropical and subtropical coniferous forests, respectively. The validation accuracy estimate of 87.9% was based on an independent set of 4.87 million random points (equally stratified with respect to the two levels of the dependent variable; further details are provided in Supplementary Information 1 and Extended Data Fig. 4). We contrasted machine learning models of the presence or absence of natural regeneration that included both socioeconomic and biophysical variables to models that only included biophysical variables. As both models showed similar accuracy (0.886 and 0.880, respectively), we based the spatial predictions of natural regeneration potential on the model that contained only biophysical variables, as these factors are less subject to change and have a higher spatial resolution than socioeconomic variables (extended rationale is provided in Supplementary Information 1).

The potential for natural regeneration was positively associated with local forest density within a 1 km radius and negatively associated with distance to existing forest (Fig. 3a,b). For example, out of a random sample of 62,493 grid cells (of all 30 × 30 m grid cells) across the study region, 98.1% of cells with a potential of >0.5 (for illustrative purposes) occur within 300 m of a forest edge. We find that this association is especially pronounced in the neotropics (Fig. 3c,d),

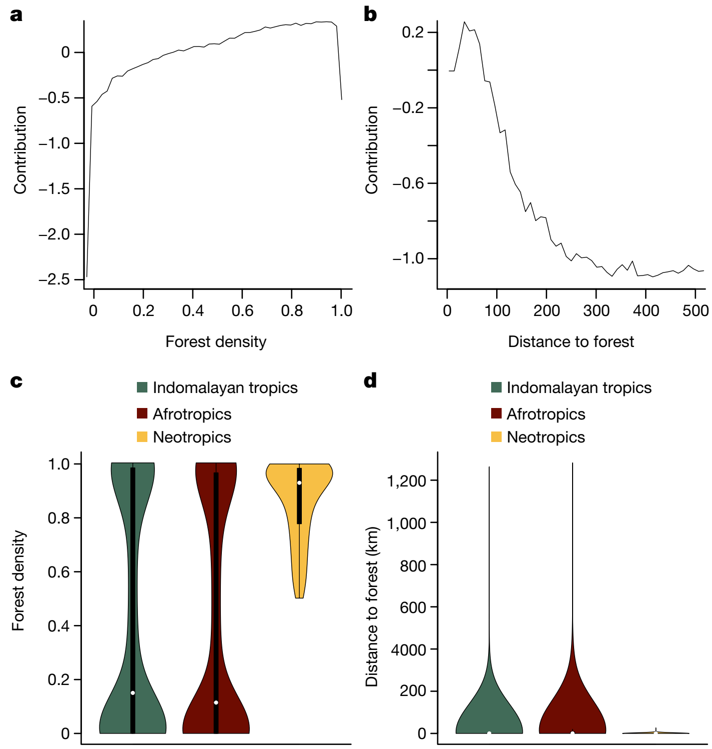


Fig. 3 | The relationship between forest density, distance to forest and potential for natural regeneration. a–d, Partial plots showing the relationship between forest density (frequency of forest within a 1 km circular buffer) (**a**) and distance to forest (**b**) and the potential for natural regeneration, and the frequency at which training points with a potential for natural regeneration value of over 0.5 (n = 3,533,732) fell within each forest density value (**c**) and distance to forest value (**d**). The distance to nearest forest partial plot (**b**) is plotted over a limited x axis range to improve legibility. The y axis of the partial plots (**a** and **b**) shows the predicted values (probability scores) associated with the forest density and distance to forest values (the x axes). The y axes of partial dependence plots do not typically represent the true value of the prediction because they visualize the marginal effect of a single feature on model predictions, while keeping other features fixed at certain values. The y axes represent the model's predicted values (not the actual true values) for the target variable.

probably owing to the higher density of persisting intact forest[23]. Our pantropical remote-sensing-based finding is consistent with extensive previous field-based studies that have shown that nearby local forests are essential for forest regeneration as they moderate climate, serve as seed sources and provide core habitat to seed-dispersing species[24]. Similarly, soil organic carbon content was a positive predictor of the potential for natural regeneration, perhaps because soil organic carbon content is higher within and near forests[25], and low levels of soil carbon are indicators of higher intensification of land use because of tillage and cultivation practices that destroy soil aggregates[26] and remove rootstocks that can promote woody regeneration through resprouts[27].

On the basis of recent spatially explicit assessments of above-ground carbon accumulation potential from natural regeneration[10], we estimate that forest regeneration over this 215 Mha could sequester 23.4 Gt of C (range, 21.1–25.7) in aboveground biomass alone over a 30 year period if these potential regrowing forests have the opportunity to establish and persist (neotropics: 11.1 Gt (CI = 10.0–12.2 Gt); Indomalayan tropics: 5.42 Gt (CI = 4.87–5.96 Gt); Afrotropics: 3.1 Gt (CI = 2.83–3.37 Gt)). This is more than three years' worth of gross carbon removals by primary and secondary tropical and subtropical forests globally at current C sequestration rates (7.01 Gt of C per year), considering that the values from our model are additional (occurring on currently deforested lands) to natural regeneration removal factors accounted for in this global estimate[28]. Our estimates do not account for any existing above-ground biomass that may be overtaken by regenerating forests, although such losses are expected to be negligible on converted lands such as pasture or croplands with relatively low initial aboveground biomass[29]. Furthermore, these estimates do not account for belowground biomass, which is difficult to estimate spatially but could account for around an additional 22–28%[30] of the aboveground biomass, implying total carbon sequestration of 28.6–30.0 Gt of C, and up to 1 Gt of potential C removals per year. New carbon sinks in naturally regenerating forests could increase current global carbon sequestration potential in primary and secondary tropical and subtropical forests by 14.3% per year[28].

Discussion

Assisted natural regeneration of forests after deforestation and land use has been successfully applied in many local contexts and has the potential to contribute to considerable carbon sinks and biodiversity conservation areas if upscaled. However, forest restoration is still largely based on tree planting approaches[2]. Here we show that there are substantial opportunities for tropical forests to regenerate with little or moderate assistance (including protection from reclearance and fires), particularly in the Neotropics and Indomalayan tropics. Moreover, if assisted natural regeneration is undertaken over the 30-year period, we show that these efforts will sequester substantial amounts of carbon, mitigating current pantropical forest carbon losses by approximately 90.5% per year[31] or accounting for 26.9% of total potential carbon across global deforested areas[32]. Our results also demonstrate fine-scale spatial variation in the potential for natural tropical forest regeneration (for example, near existing forests), highlighting the importance of using spatially explicit methods to target priority regions and set ambitious and feasible restoration targets. Five nations—Brazil, Indonesia, China, Mexico and Colombia—stand out, accounting for 52% of this potential for natural regeneration (20.3, 13.6, 7.2, 5.6 and 5.2%, respectively) showcasing the need for targeted forest restoration initiatives such as payment for ecosystem service schemes[33] and improving current structures for compensation for environmental damage[34]. However, beyond national opportunities, our maps could facilitate accounting for natural regeneration explicitly in multiobjective planning analyses and decision support tools that are increasingly being used by nations[4]. Our analysis can also inform priority locations for restoration to support incentives for nature-based solutions and to identify locations for carbon offsets where additionality and permanence may be greatest, therefore improving the effectiveness of climate mitigation strategies. We suggest using our predictive model as a first level of assessment, as local site factors and indicators of long-term land use also strongly determine the success and quality of natural regeneration[35].

Our analysis may underestimate regeneration potential in two ways. First, natural regeneration is a positive-feedback process (Fig. 3). As forests regrow better closer to existing forest cover, areas that regenerate naturally will extend the area that supports natural regeneration in the future by serving as seed sources, habitat for seed dispersers, making conditions such as soil quality and microclimates amenable to tree regeneration[36] and potentially enhancing local and regional climate conditions[37]. Our area estimates are based on contemporary conditions, but natural regeneration may be feasible over even larger areas over the long term. Second, our analysis estimates natural regeneration potential on the basis of observed occurrences of recent historic natural regeneration only in areas that support dense natural forest ecosystems, excluding non-forested or sparsely forested ecosystems such as savannas[20]. These two factors reinforce our conclusion that there is substantial and widespread but overlooked potential for natural regeneration across the global tropics. Note that the effects of climate change, such as increased drought and fire risk, increased precipitation and CO_2 fertilization, may alter the successional trajectories of tropical forests, impeding or encouraging natural regeneration processes[38].

The benefits and the benefactors of forest restoration depend heavily on the approaches used[3] and the locations and socioeconomic contexts in which restoration occurs[4]. Tree planting enables implementers to influence the precise species composition of restored forest, which could include planting commercially or culturally valuable species. However, in contexts in which forests can regenerate naturally or with some assistance, reforestation based on tree planting can actually reduce biodiversity conservation benefits if selected tree species or management practices do not provide suitable habitat for species of conservation concern or for keystone species[39]. This is especially true when planting and managing non-native tree species as monocultures[40]. Conversely, initial stages of natural regeneration may be delayed due to limited seed dispersal or poor germination and survival of native tree species[2]. This delay could translate to reduced rates of carbon sequestration and biodiversity value that could be mitigated by assisted natural regeneration interventions[41]. These limitations to establishment are offset by the potential to implement natural regeneration over much larger areas because of lower establishment costs[16].

Forests undergoing post-agricultural natural regeneration are highly vulnerable to reclearance[40]. Young forests fail to persist for many reasons, including market shifts, perverse policies, vulnerability to fire and a lack of awareness of the benefits they provide[42]. Where forests are recovering on unused agricultural land, global forces, such as market shifts in the prices of forest-risk commodities (for example, cattle, coffee, cacao or palm oil) can push landholders to reclear land if commodities become more valuable. In many countries, national policies aimed at protecting forests by imposing use or clearing restrictions on forests of a certain age, height or area can create perverse incentives to prevent forests from regrowing if landholders seek to retain control of their land[43]. At local levels, young regenerating forests are often perceived as unused or wasteland, and may be recleared to claim or retain control of land[44]. In all cases, taking appropriate measures to protect young regenerating forest (for example, firebreaks, fencing and/or patrolling at the local level, along with adequate incentives or compensation, appropriate legislation and effective resources for enforcement) is important for them to persist. Changing perceptions at the local and national levels around the value of these forests through education and/or market opportunities is a key intervention in many places[45]. Despite the ephemeral nature of naturally regenerating areas, they create additionality when actively protected, and could be used to leverage funds from carbon markets and other environmental service payment programs.

A number of pathways exist for achieving large-scale forest regeneration, but each comes with specific challenges that have limited implementation up to now. Compensating farmers and communities for engaging in natural regeneration, for example, through payments for ecosystem services or offset markets, is a possible solution, but should be combined with integrating longer-term benefits from forests to help increase persistence. In particular, carbon offsetting projects through the voluntary carbon market could be a large source of funding for natural regeneration activities (as it is for REDD+ (reducing emissions from deforestation and forest degradation in developing countries) methods and reforestation), but suitable methodologies to certify natural regeneration projects are lacking. Naturally regenerating forests do not fit into many current categories for carbon offsetting[46]. For example, currently, REDD+ methods require project areas to be forested for 10 years before a project starts, whereas afforestation, reforestation and revegetation methods require project areas to be unforested for 10 years before a project starts, which would invalidate recently cleared areas or areas already under regeneration[46,47]. Improved forest management project areas typically have to be within a logging concession. Robust, natural regeneration certification schemes may also be needed—without them, issuing carbon credits will be challenging, limiting the ability of projects to be financially viable and to provide benefits and incentives to landholders and communities. A major challenge for certification will be accurately quantifying the additionality of natural regenerating forests used as offsets against carefully constructed counterfactuals to avoid overestimating carbon sequestration[46,48].

At regional and local scales, providing local people with training to harvest and market products from naturally regenerating forests could be another pathway to finance and provide an incentive to keep young naturally regenerating forests standing[16]. However, additional interventions (and costs) would be required to ensure that the appropriate species return (for example, through enrichment planting[49]), and ongoing support to access and build market connections would be needed to reduce risks to farmers. Implementers could also provide economic or technical support for engaging in alternative, less-land-extensive livelihoods (such as micro-lending schemes, capacity building) that would free up land for forest recovery, provided that setting aside this 'extra' land for forest recovery is made explicit. Devolving control of forests and securing land rights to local communities through community forest arrangements is another way to practice restoration that meets local needs[50], provided that appropriate technical support and monitoring are provided. Creating multifunctional landscapes with local peoples and communities involved in the planning, implementation and monitoring phases[16] can also help to ensure that local needs and values are met. At national scales, legislation that prescribes forest management for landowners, such as Brazil's Forest Code, may be effective at scaling up restoration efforts but can be politically and socially unpopular and vulnerable to political change[51].

Previous studies[10,52,53] used general potential areas for reforestation (that is, Forest Landscape Restoration Opportunities[54]) and assumed that natural regeneration was possible and would occur across the entire potential areas (excluding a small fraction (7%) set aside for plantation forestry[52,53]). A previous study[52] estimated the maximum additional global reforestation mitigation potential to be 1.03 Gt C per year over 215 Mha (scaled to our area with potential for natural regeneration based on supplementary table 9 of ref. 52). To determine this area, Griscom et al.[52] assumed that natural regeneration would occur on 93% of potential reforestation areas, with the remainder going to plantations to maintain the status quo fraction of plantations to forest (7%). The Griscom et al.[53] update provided an estimated cost-effective potential of reforestation for tropical regions of 0.28 Gt C yr^{-1} (based on values from ref. 55). Similar to refs. 52,53, our spatially explicit estimates equate to 0.8 Gt C per year. However, we note that each of these analyses is quite distinct, and key differences in our analyses make direct comparison to previously published estimates convoluted. For example, cost-effective reforestation from ref. 53 summarized country-level mitigation potential values from ref. 55, which was based on an economic modelling process to simulate the aggregate response of abatement suppliers to a variable carbon price. To our knowledge, there are no studies that exist for the global tropics similar enough to ours (that is, informed by data on actual forest regeneration) to provide directly comparable estimates of natural regeneration or C sequestration potential. Given that our model (and therefore estimates of C sequestration potential) accounts for the heterogeneity of potential for natural regeneration across nations, our estimates are an advance compared with total estimates derived from flat sequestration rates.

Importantly, our analysis does not define where restoration activities should or should not occur, nor does it take into account opportunities such as feasibility, socioeconomic constraints or benefits (beyond potential C sequestration) such as specific biodiversity conservation objectives, as other studies do[56]. Our high-resolution maps of natural regeneration (and therefore climate change mitigation) potential are intended to be used as one input into broader decision-making processes for conservation and land-use planning. Our analysis focuses on tropical forest natural regeneration, specifically in response to the demand for cost-effective restoration solutions for nations that are under-resourced (compared with those at temperate latitudes) and where natural regeneration has a proven potential for positive outcomes. A recent assessment of a prioritization analysis sought to identify places where restoration of agricultural land might provide the greatest biodiversity and carbon sequestration benefits at the lowest cost. The findings showed that restoration priorities tended to be concentrated in areas that have emerging economies, and are more populated and more economically unequal, have less food security and employ more people in agriculture[57]. We echo calls to ensure that all tropical forest restoration planning and implementation activities should be undertaken with procedural, distributional, recognitional and contextual equity taken into account[57]. The biophysical potential for natural regeneration is only one of many factors to be considered in restoration planning decisions and target setting.

Note that, although our spatially explicit model is produced at a 30-m resolution, and our subsequent analyses are carried out using this produced dataset, the input covariate and predictor datasets used were a mix of resolutions—subject to the best data available for the necessary time periods (a list of data and sources used in the analysis is provided in Supplementary Table 2). Thus, users should be aware that, in some places, predictions of the potential for natural regeneration are driven by data at resolutions coarser than 30 m (visually evident in Fig. 1). This emphasizes the need for our product to be improved over time, and for implementation activities that take into account the local context.

The total pantropical potential carbon sequestration values that we present here reflect the maximum potential of natural regeneration if it takes place over the entire 215 Mha area for 30 years. The realized potential is likely to be much lower, as reforestation may include commercial or exotic species, not persist or cause leakage. Leakage is sometimes considered at the project level (although not necessarily mitigated against), but rarely at broad scales, and is likely to lead to reduced overall additionality of restoration projects[58]. In the context of restoration, leakage occurs when a restoration intervention has an effect outside the accounting boundary used to track mitigation effects (for example, an action causing emissions reductions in one place may also cause increases elsewhere) and frequently results from restoration activities[58]. That said, regional forest expansions are commonly observed across the tropics when deforestation pressures decline[59], indicating that societal and policy changes to increase reforestation are possible.

Validation of our modelling approach indicates reasonable predictive performance of our model (approximately 87.9% accuracy), similar to the 74–79% accuracy reported in an analysis of the potential for natural regeneration within Brazil's Atlantic Forest[60]. The high level of accuracy and the 30-m mapping resolution of the predictions make our model predictions useful for informing a range of forest restoration planning activities at international to subnational scales. For example, high-resolution maps of natural regeneration potential could be used to identify regions for testing national forest restoration incentive programs, or to estimate implementation costs in global-scale restoration planning activities[4], thereby improving the estimates of return-on-investment that such approaches attempt to maximize through formal optimization. Our analysis provides a tool to highlight and leverage natural regeneration potential as an integral component of local, national and global restoration planning and implementation.

Conclusion

Natural forest regeneration presents an opportunity to achieve cost-effective forest restoration at scale, delivering major climate mitigation benefits through carbon sequestration and contributing to net-zero emission targets by mid-century. It can also produce critical co-benefits including conserving biodiversity, regulating water resources, reducing erosion and increasing resilience; thus, the maps that we present here can inform multiple and interlinked environmental agendas at national and international scales. As regenerating forests can promote further natural regeneration in the future through a positive-feedback loop, there should be a strong incentive to begin supporting these processes immediately. Our pantropical analysis clarifies the extent of this immediately available opportunity, showing where costs of tree planting can be avoided or minimized while launching ambitious global forest restoration objectives for this decade. Recognizing the massive regeneration capacity of tropical forests is key to stabilizing climate change and reducing biodiversity loss alongside protecting intact forests, achieving regenerative land use and reducing deforestation, emphasizing the need to enhance recovery of degraded forests and target conservation efforts and land-use planning in human-modified landscapes where natural regeneration is most likely to succeed.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-024-08106-4.

1. Holl, K. D. Restoring tropical forests from the bottom up. *Science* **355**, 455–456 (2017).
2. Chazdon, R. L. & Guariguata, M. R. Natural regeneration as a tool for large-scale forest restoration in the tropics: prospects and challenges. *Biotropica* **48**, 716–730 (2016).
3. Crouzeilles, R. et al. Ecological restoration success is higher for natural regeneration than for active restoration in tropical forests. *Sci. Adv.* **3**, e1701345 (2017).
4. Strassburg, B. B. N. et al. Global priority areas for ecosystem restoration. *Nature* **586**, 724–729 (2020).
5. Wolosin, M. et al. *Exponential Roadmap for Natural Climate Solutions* (Conservation International, 2022).
6. IPCC Working Group. *Climate Change 2022 Mitigation of Climate Change Working Group III Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge Univ. Press, 2014).
7. *The Bonn Challenge* www.bonnchallenge.org/about (IUCN, 2020).
8. *First Draft of the Post-2020 Global Biodiversity Framework* (CBD, 2021).

9. Pillay, R. et al. Tropical forests are home to over half of the world's vertebrate species. *Front. Ecol. Environ.* **20**, 10–15 (2022).
10. Cook-Patton, S. C. et al. Mapping carbon accumulation potential from global natural forest regrowth. *Nature* **585**, 545–550 (2020).
11. Hansen, M. C. et al. High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–853 (2013).
12. *NDC's—a Force for Nature?* (WWF, 2021).
13. Brancalion, P. H. S. et al. What makes ecosystem restoration expensive? A systematic cost assessment of projects in Brazil. *Biol. Conserv.* **240**, 108274 (2019).
14. Hua, F. et al. The biodiversity and ecosystem service contributions and trade-offs of forest restoration approaches. *Science* **376**, 839–844 (2022).
15. Bodin, B. et al. A standard framework for assessing the costs and benefits of restoration: introducing The Economics of Ecosystem Restoration. *Restor. Ecol.* **30**, e13515 (2022).
16. Chazdon, R. L. et al. Fostering natural forest regeneration on former agricultural land through economic and policy interventions. *Environ. Res. Lett.* **15**, 043002 (2020).
17. Bastin, J.-F. et al. The global tree restoration potential. *Science* **365**, 76–79 (2019).
18. Laestadius, L. et al. *Mapping Opportunities for Forest Landscape Restoration* (FAO, 2011).
19. Veldman, J. W. et al. Comment on "The global tree restoration potential". *Science* **366**, eaay7976 (2019).
20. Fagan, M. E. A lesson unlearned? Underestimating tree cover in drylands biases global restoration maps. *Glob. Change Biol.* **26**, 4679–4690 (2020).
21. Ling, P.-Y. et al. Mapping global forest regeneration—an untapped potential to mitigate climate change and biodiversity loss. *Environ. Res. Lett.* **18**, 054025 (2023).
22. Fagan, M. E. et al. The expansion of tree plantations across tropical biomes. *Nat. Sustain.* **5**, 681–688 (2022).
23. Grantham, H. S. et al. Anthropogenic modification of forests means only 40% of remaining forests have high ecosystem integrity. *Nat. Commun.* **11**, 5978 (2020).
24. Sloan, S., Goosem, M. & Laurance, S. G. Tropical forest regeneration following land abandonment is driven by primary rainforest distribution in an old pastoral region. *Landsc. Ecol.* **31**, 601–618 (2016).
25. Cerri, C. C., Volkoff, B. & Andreaux, F. Nature and behaviour of organic matter in soils under natural forest, and after deforestation, burning and cultivation, near Manaus. *For. Ecol. Manage.* **38**, 247–257 (1991).
26. Damian, J. M. et al. Deforestation and land use change mediate soil carbon changes in the eastern Brazilian Amazon. *Reg. Environ. Change* **21**, 64 (2021).
27. Shoo, L. P. & Catterall, C. P. Stimulating natural regeneration of tropical forest on degraded land: approaches, outcomes, and information gaps. *Restor. Ecol.* **21**, 670–677 (2013).
28. Harris, N. L. et al. Global maps of twenty-first century forest carbon fluxes. *Nat. Clim. Change* **11**, 234–240 (2021).
29. Chazdon, R. L. et al. Carbon sequestration potential of second-growth forest regeneration in the Latin American tropics. *Sci. Adv.* **2**, e1501639 (2016).
30. *Annex 3A.1 Biomass Default Tables for Section 3.2 Forest Land* (ICPP, 2003).
31. Baccini, A. et al. Tropical forests are a net carbon source based on aboveground measurements of gain and loss. *Science* **358**, 230–234 (2017).
32. Mo, L. et al. Integrated global assessment of the natural forest carbon potential. *Nature* https://doi.org/10.1038/s41586-023-06723-z (2023).
33. López-Cubillos, S. et al. Spatial prioritization to achieve the triple bottom line in payment for ecosystem services design. *Ecosyst. Serv.* **55**, 101424 (2022).
34. Bustamante, M. M. C. et al. Ecological restoration as a strategy for mitigating and adapting to climate change: lessons and challenges from Brazil. *Mitig. Adapt. Strateg. Glob. Change* **24**, 1249–1270 (2019).
35. César, R. G. et al. It is not just about time: agricultural practices and surrounding forest cover affect secondary forest recovery in agricultural landscapes. *Biotropica* https://doi.org/10.1111/btp.12893 (2021).
36. Chazdon, R. L. *Second Growth: The Promise of Tropical Forest Regeneration in an Age of Deforestation* (Univ. Chicago Press, 2014).
37. Lawrence, D., Coe, M., Walker, W., Verchot, L. & Vandecar, K. The unseen effects of deforestation: biophysical effects on climate. *Front. For. Glob. Change* **5**, 756115 (2022).
38. Uriarte, M. et al. Impacts of climate variability on tree demography in second growth tropical forests: the importance of regional context for predicting successional trajectories. *Biotropica* **48**, 780–797 (2016).
39. Chazdon, R. L. et al. The potential for species conservation in tropical secondary forests. *Conserv. Biol.* **23**, 1406–1417 (2009).
40. Brancalion, P. H. S. & Holl, K. D. Guidance for successful tree planting initiatives. *J. Appl. Ecol.* **57**, 2349–2361 (2020).
41. Shono, K., Chazdon, R., Bodin, B., Wilson, S. J. & Durst, P. Assisted natural regeneration: harnessing nature for restoration. *Unasylva* **252**, 71–81 (2020).
42. Holl, K. D., Loik, M. E., Lin, E. H. V. & Samuels, I. A. Tropical Montane forest restoration in Costa Rica: overcoming barriers to dispersal and establishment. *Restor. Ecol.* **8**, 339–349 (2000).
43. Chazdon, R. L. et al. When is a forest a forest? Forest concepts and definitions in the era of forest and landscape restoration. *Ambio* **45**, 538–550 (2016).
44. Zahawi, R. A., Reid, J. L. & Holl, K. D. Hidden costs of passive restoration. *Restor. Ecol.* **22**, 284–287 (2014).
45. Reid, J. L. et al. How long do restored ecosystems persist? *Ann. Mo. Bot. Gard.* **102**, 258–265 (2017).
46. Brancalion, P. H. S. et al. A call to develop carbon credits for second-growth forests. *Nat. Ecol. Evol.* **8**, 179–180 (2024).
47. *Warsaw Framework for REDD+* (UNFCCC, 2023).
48. West, T. A. P., Börner, J., Sills, E. O. & Kontoleon, A. Overstated carbon emission reductions from voluntary REDD+ projects in the Brazilian Amazon. *Proc. Natl Acad. Sci. USA* **117**, 24188–24194 (2020).
49. Yeong, K. L. et al. Enrichment planting to improve habitat quality and conservation value of tropical rainforest fragments. *Biodivers. Conserv.* **25**, 957–973 (2016).
50. Wilson, S. J. Communal management as a strategy for restoring cloud forest landscapes in Andean Ecuador. *World Dev. Perspect.* **3**, 47–49 (2016).
51. Soares-Filho, B. et al. Land use. Cracking Brazil's forest code. *Science* **344**, 363–364 (2014).
52. Griscom, B. W. et al. Natural climate solutions. *Proc. Natl Acad. Sci. USA* **114**, 11645–11650 (2017).
53. Griscom, B. W. et al. National mitigation potential from natural climate solutions in the tropics. *Philos. Trans. R. Soc. Lond. B* **375**, 20190126 (2020).
54. Potapov, P., Laestadius, L. & Minnemeyer, S. *Global Map of Potential Forest Cover* www.wri.org/resources/maps/atlas-forest-and-landscape-restoration-opportunities/data-info (2011).
55. Busch, J. et al. Potential for low-cost carbon dioxide removal through tropical reforestation. *Nat. Clim. Change* **9**, 463–466 (2019).
56. Brancalion, P. H. S. et al. Global restoration opportunities in tropical rainforest landscapes. *Sci. Adv.* **5**, eaav3223 (2019).
57. Schultz, B. et al. Recognizing the equity implications of restoration priority maps. *Environ. Res. Lett.* **17**, 114019 (2022).
58. Streck, C. REDD+ and leakage: debunking myths and promoting integrated solutions. *Clim. Policy* **21**, 843–852 (2021).
59. Meyfroidt, P. & Lambin, E. F. Global forest transition: prospects for an end to deforestation. *Annu. Rev. Environ. Resour.* **36**, 343–371 (2011).
60. Crouzeilles, R. et al. Achieving cost-effective landscape-scale forest restoration through targeted natural regeneration. *Conserv. Lett.* **13**, e12709 (2020).

Methods

We model the potential for natural regeneration using a binary dependent variable representing point locations where natural regeneration occurred (1) or did not occur (0) between 2000 and 2016 and a suite of biophysical and socioeconomic independent variables. Locations where natural regeneration occurred from 2000 to 2016 were identified by Fagan et al.[22], who delineated patches of at least 0.45 ha of gain of vegetation taller than 5 m height from 2000 to 2016 using the 30-m-resolution Global Forest Watch time-series of tree cover[11]. Natural regeneration was differentiated from planted tree crop monocultures using a combination of geospatial datasets and a machine learning algorithm that was trained and validated using a large sample of known natural regeneration and planted patches[22]. The overall model accuracy for the three-class classification (plantation, open, natural regrowth) of regrowth patches is 90.6% (±0.7), and the mean global model accuracy estimate of the natural regrowth class is 88.9% (±1.2)[22]. Like the product it is derived from[11], the final predicted map omitted the majority of natural regrowth patches in the humid biome. The class producer's accuracy for natural regrowth patches in the humid biome was 78.8 (±5.6) but dropped to 18.7 (±5.4) when based on estimated area (Extended Data Fig. 1). The low producer's accuracy when based on estimated area is in part due to Fagan et al.[22] conservatively erring on the side of omission errors, and also because natural regeneration occupies a proportionately small area compared to the class it is separated from (all other classes), meaning that any natural regeneration patch erroneously mapped as other classes had a large influence on estimated regrowth area. Although we refer to both the mapped and estimated area, we emphasize the higher (mapped) estimate, given that this is a common and known issue with reporting on producer's accuracy[61,62]. Furthermore, regrowth patches were accurately classified when detected (class user's accuracy of 85.1 ± 5.6), permitting robust estimates of the overall area and distribution of regrowth (Supplementary Information 1 and Extended Data Figs. 1–4).

Study region

Our study area comprises all tropical and subtropical dry broadleaf forests, tropical and subtropical moist broadleaf forests, and tropical and subtropical coniferous forests[63] within ±25° latitude. The domain-defining areas where natural regeneration could have occurred but did not in the interval 2000–2016 were defined as areas within the aforementioned biomes excluding areas associated with: (1) land cover classes that precluded forest regeneration (European Space Agency (ESA) CCI 2000 land cover classes[64]: water bodies, bare areas, urban areas and sparse vegetation, codes 210, 200, 190 and 150, respectively); (2) areas that were already classed as tree cover in the Global Forest Watch dataset for the year 2000[11]; (3) areas that were deemed to have regenerated naturally between 2000 and 2016; and (4) areas associated with forestry activity[22]. We adopted this liberal definition to allow the subsequent modelling to identify empirically areas associated with low probabilities of natural regeneration, rather than imposing subjective assumptions in the form of more stringent criteria a priori. There are three reasons why we do not adopt stricter criteria for defining the domain of areas available for regeneration. First, there is substantial scope for the introductions of error when adopting constraints on a 30-m-resolution analysis with data mapped at coarser resolutions. For example, 51.8% of the random points generated within the regeneration patches delineated at a 30-m resolution fall within the forest-related ESA CCI land cover categories mapped at a 300-m resolution. This is consistent with our findings of the importance of distance to existing forest and local forest density being important predictors of the potential for natural regeneration, but highlights the problems with using coarser-resolution data to impose strict constraints on an analysis at a finer resolution. Conversely, only 0.3% of the random points generated within the regeneration patches fall within the excluded ESA CCI land cover categories (such as bare ground or urban areas), which we suggest is an acceptable rate of potential misidentification for those areas. Second, land cover classes are estimated with error implying that even classes that we would expect to be associated with a low probability of regeneration are likely to include some regeneration as a result of this misidentification. This could explain, for example, why a small rate of regeneration appears to occur in the excluded land cover classes. Finally, decisions regarding areas within which regeneration are deemed to be possible are subjective. Here, we prefer to limit such assumptions in the definition of the sampling domain and allow the model to identify empirical relationships about the probability of regeneration.

Modelling approach

The statistical modelling and validation is based on a set of six million random points generated within the regeneration and non-regeneration spatial domains defined above, stratified so that each level of the dependent variable was equally represented but with no other form of spatial stratification. For each of these random points, we used geographical information systems to extract values for a range of biophysical and socioeconomic variables (see the 'Predictor variables' section). We based the spatial predictions on the model containing biophysical variables only (the rationale is provided in Supplementary Information 1).

We used randomForests[65] to fit models of natural regeneration occurrence and estimate spatially explicit predictions of the potential for natural regeneration in the near future (2030), assuming that overall conditions from 2000 to 2016 apply to future scenarios. We adopted a variable-selection procedure to improve parsimony and reduce the high computational burden of making predictions at 30-m resolution over large scales. RandomForests involves stochastic fitting processes, so two models fit to the same data may differ. We therefore fit ten models with all biophysical covariates by randomly selecting 500,000 training records (balanced with respect to the levels of the dependent variable) from a pool of six million records. The variables were ordered by the mean decrease in accuracy resulting from their omission[65], and the overall ordering of variable importance was based on summing the rank position of each variable among each of these ten models. Following a previous study[60], we then iteratively fit new models adding one additional variable in order of importance each time to identify the point at which the addition of new variables does not improve model accuracy. The final model was fit using only those variables with one million records. Model validation was based on an independent set of random points that was balanced with respect to the dependent variable. A concern with this approach to validation is that it may overestimate model performance, as it does not account for autocorrelation[66]. We evaluated this effect by quantifying the distance between the validation and training locations and assessing model performance in 1-km distance increments from the training locations.

Predictions of the potential for natural regeneration were based on this model with three covariates updated to reflect contemporary conditions rather than conditions in 2000. The revised forest density and distance to forest covariates were derived from the Global Forest Watch 2018 tree cover dataset[11] at a 30-m resolution, and the revised land user/cover dataset was derived from the combination of the 2015 ESA CCI dataset with the aforementioned tree cover data. Locations associated with open water, urban areas and rock or bare ground were deemed to be unavailable for forest regeneration and coded as NoData in the predictions. Further details on the modelling approach are provided in Supplementary Information 1. We used GADM (database of global administrative areas[67]) for all country-based summaries in this analysis.

Predictor variables

ESA CCI land cover (cropland density and distance to urban area). Land cover is an important consideration for predicting the potential for natural regeneration for various reasons. For example, land with higher cropland density is less likely to regenerate, and distance to urban areas is positively associated with natural regeneration (regenerating patches far from urban areas are less disturbed and less likely to be deforested)[68]. The Climate Change Initiative from the ESA makes available several products on land cover with global coverage. We obtained the 300-m-resolution land-cover rasters, available annually from 1992 to 2015[64]. The typology used by ESA/CCI follows the Land Cover Classification System (LCCS) developed by the United Nations Food and Agriculture Organization, being globally consistent, but also presenting a more detailed classification at regional scales.

The land-cover classification scheme was simplified from 31 to 11 classes based on ecological and land use similarity as follows (semi-colon delimited lists of the ESA CCI land cover classes):

(1) Cropland, rainfed; cropland, irrigated or post-flooding.
(2) Herbaceous cover.
(3) Mosaic cropland (>50%)/natural vegetation (tree, shrub, herbaceous cover) (<50%); mosaic natural vegetation (tree, shrub, herbaceous cover) (>50%)/cropland (<50%).
(4) Tree cover, broadleaved, evergreen, closed to open (>15%); shrubland evergreen.
(5) Tree cover, broadleaved, deciduous, closed to open (>15%); tree cover, broadleaved, deciduous, closed (>40%); tree cover, broadleaved, deciduous, open (15–40%).
(6) Tree cover, needle leaved, evergreen, closed to open (>15%); tree cover, needle leaved, deciduous, closed to open (>15%); tree cover, mixed leaf type (broadleaved and needle leaved).
(7) Mosaic tree and shrub (>50%)/herbaceous cover (<50%); mosaic herbaceous cover (>50%)/tree and shrub (<50%).
(8) Shrubland; shrubland deciduous.
(9) Grassland.
(10) Sparse vegetation (tree, shrub, herbaceous cover) (<15%); sparse tree (<15%); sparse shrub (<15%); sparse herbaceous cover (<15%); urban areas; bare areas; consolidated bare areas; unconsolidated bare areas; water bodies.
(11) Tree cover, flooded, fresh or brackish water; tree cover, flooded, saline water; shrub or herbaceous cover, flooded, fresh/saline/brackish water.

We derived a cropland density covariate from the ESA CCI dataset representing the mean cropland area within a 5-km circular buffer for each pixel within the study region, using the focal statistics tool in ArcMap (v.10.8)[69]. Some of the CCI classes represent composites of several land cover classes. When estimating density, we adopted the following reclassification for determining the proportion of each cell that was considered to represent cropland: 1: cropland; 1: cropland (irrigated or post-flooding); 0.75: mosaic cropland (>50%)/natural vegetation (tree); 0.25: mosaic natural vegetation, herbaceous cover (>50%)/cropland (<50%); 0: all other classes.

We also derived a distance to urban area covariate representing the Euclidean distance from each cell to the nearest cell classed as urban.

Tree cover (forest density and distance to forest patch). Previous studies have shown that distance to forest patches is negatively associated with natural regeneration potential (lower seedling density and species richness as the distance from the forest edge increases)[68]. The Global Forest Watch provides data on tree cover, gain and loss for the interval from 2000 to 2018 at a global scale, at 30-m resolution[11]. This dataset was used in three ways in our modelling processes. First, it was used to define the areas that were already forested in 2000, thereby precluding the possibility of natural regeneration within those areas. Random points for the non-regeneration (0) level of the dependent variable were excluded from these locations. Second, the 2000 tree cover data were used to estimate mean forest area within a 1-km circular buffer for each pixel within the study region using the focal statistics tool in ArcMap (v.10.8)[69], thereby resulting in two new derived 30-m-resolution datasets that were covariates in the models. Previous work[60] has demonstrated that forest density and distance to forest are strong predictors of the potential for natural regeneration. Finally, when generating the spatial predictions of the potential for natural regeneration, the 2018 tree cover data were used to define the areas that were already forested (and could therefore not regenerate naturally), and to update the forest density and distance to forest covariates.

We caution users that forest-based predictor (2000–2012 dataset) and explanatory variables (2011–2018 dataset) from the Global Forest Watch[11] used to generate the layers used in our analysis may differ in accuracy given that, for the 2011–2018 dataset, Landsat 8 OLI data were used, loss data were reprocessed, improved training data for calibrating the loss model were used, improved per sensor quality assessment models to filter input data were used, and improved input spectral features for building and applying the loss model were applied. The variety of mapping methods and attribution approaches means that the input datasets may locally lack coherency (for example, gain may occur in pixels not mapped as forest cover in a later revision). However, this is the most comprehensive and reliable dataset currently available, especially at broader scales. More information on the differences between the datasets is available online (https://earthenginepartners.appspot.com/science-2013-global-forest/download_v1.6.html).

Protected areas. Natural regeneration potential has been shown to be higher when deforested areas neighbour or are newly incorporated into protected areas[70]. To represent protected areas, we used the World Database on Protected Areas[71], which we downloaded and rasterized according to the methods of a previous study[72]. We incorporated into the August 2020 version of WDPA 768 protected areas (1.43 million km^2) in China (sites that were available in the June 2017 version of WDPA, but not publicly available thereafter). Following the WDPA best practice guidelines (www.protectedplanet.net/c/calculating-protected-area-coverage) and similar global studies, we included in our analysis only protected areas from the WDPA database that have a status of 'designated', 'inscribed' or 'established', and removed all points and polygons with a status of 'proposed' or 'not reported'. We buffered the point feature class in accordance with the area of the points as stated in the 'REP AREA' field, and merged the buffered points with polygons to create one polygon layer. We then dissolved this layer by IUCN protection category.

Biomes. We include biome[63] as a categorical variable in our models, as biomes represent broad ecological systems within which the processes governing natural forest regeneration may vary. The analysis is restricted to the three tropical and subtropical forest biomes (tropical and subtropical dry broadleaf forests, tropical and subtropical moist broadleaf forests, and tropical and subtropical coniferous forests).

Global Soil Map (International Soil Reference and Information Centre). Soil properties can influence forest regeneration potential as they may affect seedlings' establishment, growth and survival[73]. Thus, 12 soil properties mapped at a 250-m resolution[74] were evaluated for use in the modelling, including: (1) total profile depth (cm); (2) plant exploitable (effective) soil depth (cm); (3) organic carbon (g per kg); (4) pH (×10); (5) sand (g per kg); (6) silt (g per kg); (7) clay (g per kg); (8) gravel (m^3); (9) ECEC ($cmol_c\ kg^{-1}$); (10) bulk density of the fine earth (< 2 mm) fraction (excludes gravel) ($t\ m^{-3}$); (11) bulk density of the whole soil in situ (includes gravel) ($t\ m^{-3}$); and (12) available water capacity (mm) were included. These are mapped at six depth intervals (0–5 cm; 5–15 cm; 15–30 cm; 30–60 cm; 60–100 cm; 100–200 cm). We calculated the mean of each of the 12 properties within the top 30 cm of soil (weighted by the depth of each layer).

Slope (derived from SRTM elevation). Slope can be positively associated with natural regeneration potential, especially for areas dominated by mechanized crop and pasture land uses[75], where slope areas are frequently abandoned to regenerate. The mean elevation above sea level was obtained from the Shuttle Radar Topography Mission (SRTM)[76]. The data are made available by NASA in raster format at a 1 arc-second (approximately 30 m) resolution, based on Digital Elevation Models and other open source data to fill the voids in that dataset using Google Earth Engine[77]. Slope was derived from this elevation dataset at 30-m resolution.

Net primary productivity (annual 2000–2015, global; MOD17A3). Natural regeneration potential is positively associated with net primary productivity. We used the mean net primary productivity estimate from the MODIS-based estimates of net primary productivity (gC m^{-2} yr^{-1}) at 1-km^2 resolution[78].

WorldClim v.2.1 climate data for 1970–2000. Climatic variables may influence natural regeneration potential, as some biophysical conditions, such as high annual precipitation, can promote tree regeneration and reduce fire frequency[79]. We used the WorldClim dataset[80] which includes 19 bioclimatic variables derived from the monthly temperature and rainfall values. They are 30-arc-second-resolution (approximately 1 km) rasters and include: (1) annual mean temperature; (2) mean diurnal range; (3) isothermality; (4) temperature seasonality; (5) max temperature of the warmest month; (6) min temperature of the coldest month; (7) temperature annual range; (8) mean temperature of the wettest quarter; (9) mean temperature of the driest quarter; (10) mean temperature of the warmest quarter; (11) mean temperature of the coldest quarter; (12) annual precipitation; (13) precipitation of the wettest month; (14) precipitation of the driest month; (15) precipitation seasonality (coefficient of variation); (16) precipitation of the wettest quarter; (17) precipitation of the driest quarter; (18) precipitation of the warmest quarter; and (19) precipitation of the coldest quarter.

Because several of the bioclimatic variables are correlated, we used principal component analysis (PCA) to reduce the set of 19 bioclimatic variables to a set of 5 principal components that capture 99.4% (Supplementary Table 1) of the variation among the original 19 variables but that are orthogonal (uncorrelated) with each other. This was achieved by generating a random sample of 1 million points within land areas (the bioclimatic datasets have NoData values for marine areas, which were excluded) and extracting the cell values associated with those locations from each of the 19 bioclimatic variables into a matrix. Each raster contained over 309 million data values; it was therefore necessary to base the PCA on a sample of the values rather than all values. The 'prcomp' function[81] in R was used to calculate the principal components based on this matrix. This PCA model was then applied to all pixels in the rasters (using the 'predict.prcomp' function[81] in R) and the first five components were written as new raster datasets that were used in subsequent modelling.

Global Human Settlement dataset (GHS-POP). Human population density may negatively influence natural regeneration potential, as deforested areas may be subjected to further disturbances due to human populations[82]. The global human settlement dataset depicts the distribution of population, expressed as the number of people per cell[83]. Residential population estimates for target years 1975, 1990, 2000 and 2015 provided by CIESIN GPWv4.10 were disaggregated from census or administrative units to grid cells, informed by the distribution and density of built-up as mapped in the Global Human Settlement Layer (GHSL) global layer per corresponding epoch.

Gross domestic product and human development index. Natural regeneration potential is positively associated with gross domestic product[82] and human development index[84]. Kummu et al.[85] provide gridded global datasets for gross domestic product and human development index over 1990–2015.

Road density. Natural regeneration potential may be higher in areas of low road density, where agricultural land is more likely to be abandoned[86]. The Global Roads Inventory Project[87] gathered, harmonized and integrated nearly 60 geospatial datasets on road infrastructure into a global roads dataset. The resulting dataset covers 222 countries and includes over 21 million km of roads, which is two to three times the total length in the currently best available country-based global roads datasets. From this dataset, we obtained the value of road density in m per km^2.

Fire: burned area. While, under favourable ecological conditions, some recently burned areas may have a higher natural regeneration rate[88], humid tropical forests are often poorly adapted to fire, and therefore less likely to naturally regenerate after a fire event[89]. The monthly average of burned area due to wildfires over the period of 2001–2017 (January–December for all years) was obtained from ref. 90.

Distance to water. Natural regeneration potential can be influenced by distance to watercourses and waterbodies as riparian areas have higher water availability, higher fauna movement, and usually present remnant trees and forests supplying seeds for regeneration[75]. We used a layer which represents the Euclidean distance (metres) of each raster cell to the closest freshwater feature (lake or river). Detailed methodology on how the dataset was produced is given in ref. 91.

A summary of all data used in this analysis is available in Supplementary Table 2.

Areal calculations. Area values reported in this study were calculated as the area of the 30 m × 30 m grid cell multiplied by its relative potential for natural regeneration value. We provide a sensitivity analysis to this calculation where area is regarded as cells that have >50% potential for natural regeneration, which expands areal estimates slightly (globally from 215 million ha to 263 million ha; Supplementary Table 4). All calculations were carried out using the Mollweide projection.

Carbon accumulation potential calculations. We calculate carbon accumulation potential using the Cook-Patton et al.[10] dataset, which provides carbon accumulation in t of C yr^{-1} over 30 years of forest regrowth. We converted this layer to gross accumulation potential over the entire 30 year period by multiplying values by 30. Original values were in t per ha at a 1 km by 1 km resolution. To downscale to a 30 m by 30 m cell resolution, we first resampled the layer and multiplied it by 0.09, because there are 0.09 hectares in the 30 m by 30 m cell. This gave us a total value per cell, which we then multiplied by the potential for natural regeneration, and finally summarized (sum) at the global, and bioregional levels.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All input datasets are available from the references cited. The raster datasets arising from this work are available link below and include (1) the approximately 1-km-resolution overview raster representing the proportion of the area of each pixel that has the potential for natural regeneration ('prop pnv v1 1 km.tif'; 1.1 GB) for display purposes only (not used in this analysis); (2) the 30-m-resolution binary predictions (where a value of >0.5 is allocated a value of 1, acknowledging that users may prefer to use a different threshold for their context) of areas suitable for natural regeneration (files named with the prefix 'pnv bin 30 m'; 2.3 GB); and (3) the 30-m continuous probability predictions of the potential for natural regeneration, stored as integers representing percentages to minimize file sizes (files named with the prefix 'pnv pct 30 m'; 9.3 GB). The 30-m-resolution datasets have been tiled into 10-degree latitude/longitude tiles. The dataset is available at Zenodo[92] (https://zenodo.org/record/7428804).

Code availability

R and python codes developed for and used in this analysis are available on request from hawthornebeyer@gmail.com.

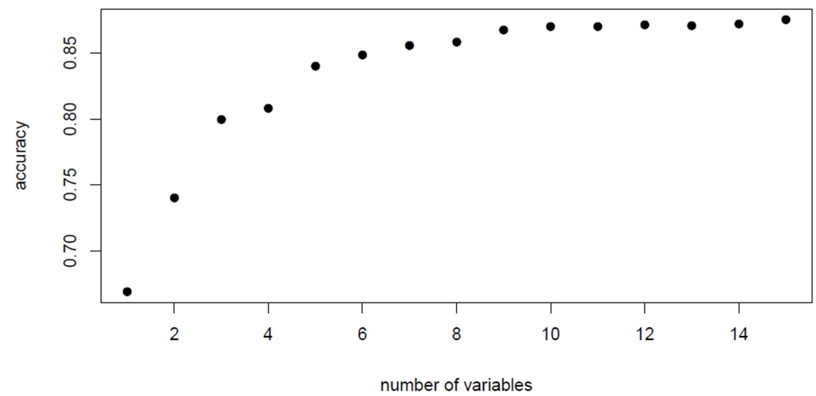
61. Wang, Y. et al. High-resolution maps show that rubber causes substantial deforestation. *Nature* **623**, 340–346 (2023).
62. Olofsson, P. et al. Mitigating the effects of omission errors on area and area change estimates. *Remote Sens. Environ.* **236**, 111492 (2020).
63. Dinerstein, E. et al. An ecoregion-based approach to protecting half the terrestrial realm. *Bioscience* **67**, 534–545 (2017).
64. *Land Cover CCI Product User Guide Version 2* Technical Report (ESA, 2017).
65. Liaw, A. & Wiener, M. Classification and regression by randomForest. *R News* **2**, 18–22 (2002).
66. Ploton, P. et al. Spatial validation reveals poor predictive performance of large-scale ecological mapping models. *Nat. Commun.* **11**, 4540 (2020).
67. *Database of Global Administrative Areas* (GADM, 2022).
68. Cubina, A. & Aide, T. M. The effect of distance from forest edge on seed rain and soil seed bank in a tropical Pasture1. *Biotropica* **33**, 260–267 (2001).
69. ArcGIS (GIS software) v.10.8. (ESRI, 2022).
70. Algeet-Abarquero, N., Sánchez-Azofeifa, A., Bonatti, J. & Marchamalo, M. Land cover dynamics in Osa Region, Costa Rica: secondary forest is here to stay. *Reg. Environ. Change* **15**, 1461–1472 (2015).
71. *Protected Planet: The World Database on Protected Areas (WDPA) v.1.6* www.protectedplanet.net/en (UNEP-WCMC, IUCN, 2020).
72. Maxwell, S. L. et al. Area-based conservation in the twenty-first century. *Nature* **586**, 217–227 (2020).
73. Heneghan, L. et al. Integrating soil ecological knowledge into restoration management. *Restor. Ecol.* **16**, 608–617 (2008).
74. Hengl, T. et al. SoilGrids250m: global gridded soil information based on machine learning. *PLoS ONE* **12**, e0169748 (2017).
75. Molin, P. G., Chazdon, R., Frosini de Barros Ferraz, S. & Brancalion, P. H. S. A landscape approach for cost-effective large-scale forest restoration. *J. Appl. Ecol.* **55**, 2767–2778 (2018).
76. Farr, T. G. et al. The shuttle radar topography mission. *Rev. Geophys.* **45**, RG2004 (2007).
77. Gorelick, N. et al. Google Earth Engine: planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **202**, 18–27 (2017).
78. Zhao, M., Heinsch, F. A., Nemani, R. R. & Running, S. W. Improvements of the MODIS terrestrial gross and net primary production global data set. *Remote Sens. Environ.* **95**, 164–176 (2005).
79. Yackulic, C. B. et al. Biophysical and socioeconomic factors associated with forest transitions at multiple spatial and temporal scales. *Ecol. Soc.* **16**, 15 (2011).
80. Fick, S. E. & Hijmans, R. J. WorldClim 2: new 1 km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).
81. R Core Team. *R: a Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2021).
82. Piffer, P. R., Rosa, M. R., Tambosi, L. R., Metzger, J. P. & Uriarte, M. Turnover rates of regenerated forests challenge restoration efforts in the Brazilian Atlantic forest. *Environ. Res. Lett.* **17**, 045009 (2022).
83. Schiavina, M., Freire, S. & MacManus, K. *GHS Population Grid Multitemporal (1975, 1990, 2000, 2015)* R2019A (European Commission JRC, 2019).
84. Redo, D. J., Grau, H. R., Aide, T. M. & Clark, M. L. Asymmetric forest transition driven by the interaction of socioeconomic development and environmental heterogeneity in Central America. *Proc. Natl Acad. Sci. USA* **109**, 8839–8844 (2012).
85. Kummu, M., Taka, M. & Guillaume, J. H. A. Data from: Gridded global datasets for gross domestic product and human development index over 1990–2015, v2. *Dryad* https://doi.org/10.5061/dryad.dk1j0 (2020).
86. Thomlinson, J. R. et al. Land-use dynamics in a post-agricultural Puerto rican landscape (1936-1988). *Biotropica* **28**, 525 (1996).
87. Meijer, J. R., Huijbregts, M. A. J., Schotten, K. C. G. J. & Schipper, A. M. Global patterns of current and future road infrastructure. *Environ. Res. Lett.* **13**, 064006 (2018).
88. Maillard, O. Post-fire natural regeneration trends in Bolivia: 2001–2021. *Fire* **6**, 18 (2023).
89. Scheper, A. C., Verweij, P. A. & van Kuijk, M. Post-fire forest restoration in the humid tropics: a synthesis of available strategies and knowledge gaps for effective restoration. *Sci. Total Environ.* **771**, 144647 (2021).
90. Artés, T. et al. A global wildfire dataset for the analysis of fire regimes and fire behaviour. *Sci. Data* **6**, 296 (2019).
91. Kummu, M., de Moel, H., Ward, P. J. & Varis, O. How close do we live to water? A global analysis of population distance to freshwater bodies. *PLoS ONE* **6**, e20578 (2011).
92. Williams, B. A. et al. Data for 'Global potential for natural regeneration in deforested tropical regions'. *Zenodo* https://doi.org/10.5281/zenodo.7428803 (2024).

Acknowledgements We acknowledge financial support from Conservation International, and the Walton Family Foundation (B.W.G.). The mapping work from Fagan and colleagues 2022, the basis of this analysis, was supported by NASA grant 80NSSC21K0297. J.R.R. was supported by an ARC Future Fellowship (FT200100096).

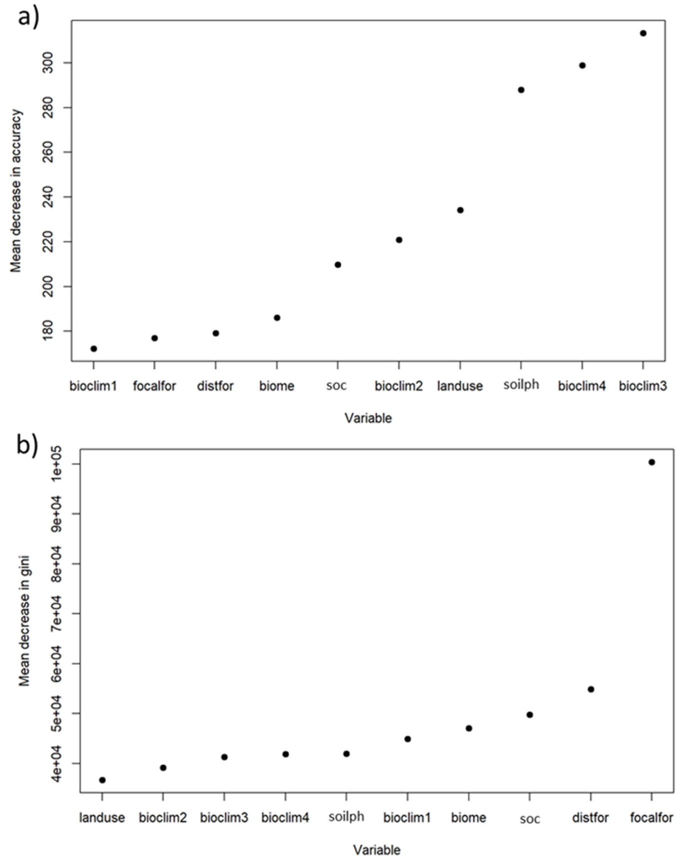
Author contributions B.A.W. and H.L.B. contributed equally. H.L.B. conceived the study with the support of R. Crouzeilles and N.S.A. H.L.B. coordinated the initial development of the modelling approach with the support of B.A.W., R. Crouzeilles, M.E.F., R. Chazdon, M.S., S.S.-H., B.W.G. and M.G.-R. and led the analysis with the support of B.A.W. H.L.B. led the first draft of the paper, and B.A.W. led the writing, synthesis, figure development and coordination of all subsequent versions. B.A.W. led the additional analyses and synthesis conducted during the review process with the support of M.E.F. All of the authors provided input into subsequent versions of the manuscript.

Competing interests The authors declare no competing interests.

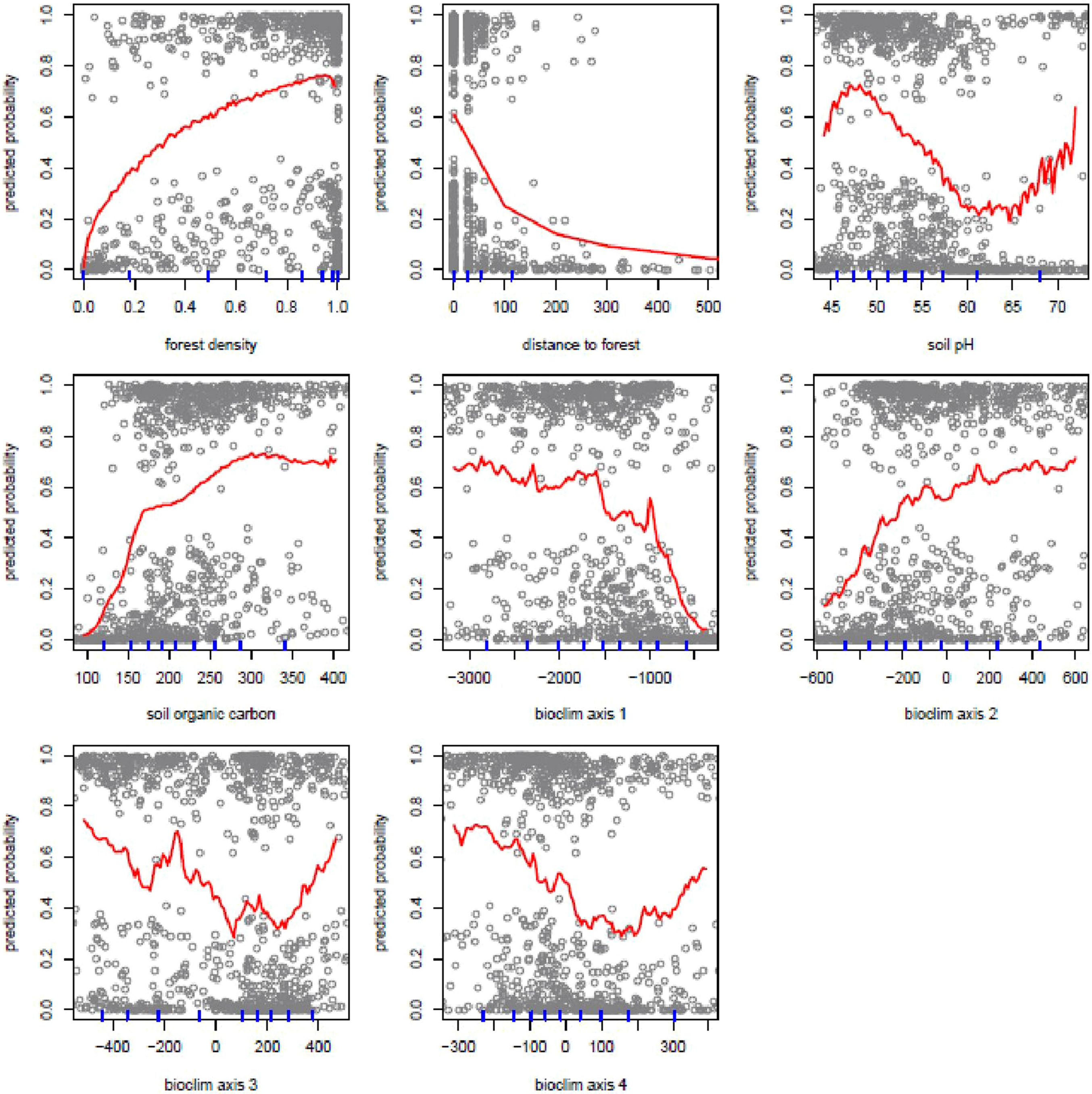
Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-024-08106-4.
Correspondence and requests for materials should be addressed to Brooke A. Williams.
Peer review information *Nature* thanks Edward Mitchard, Charlotte Wheeler and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.
Reprints and permissions information is available at http://www.nature.com/reprints.



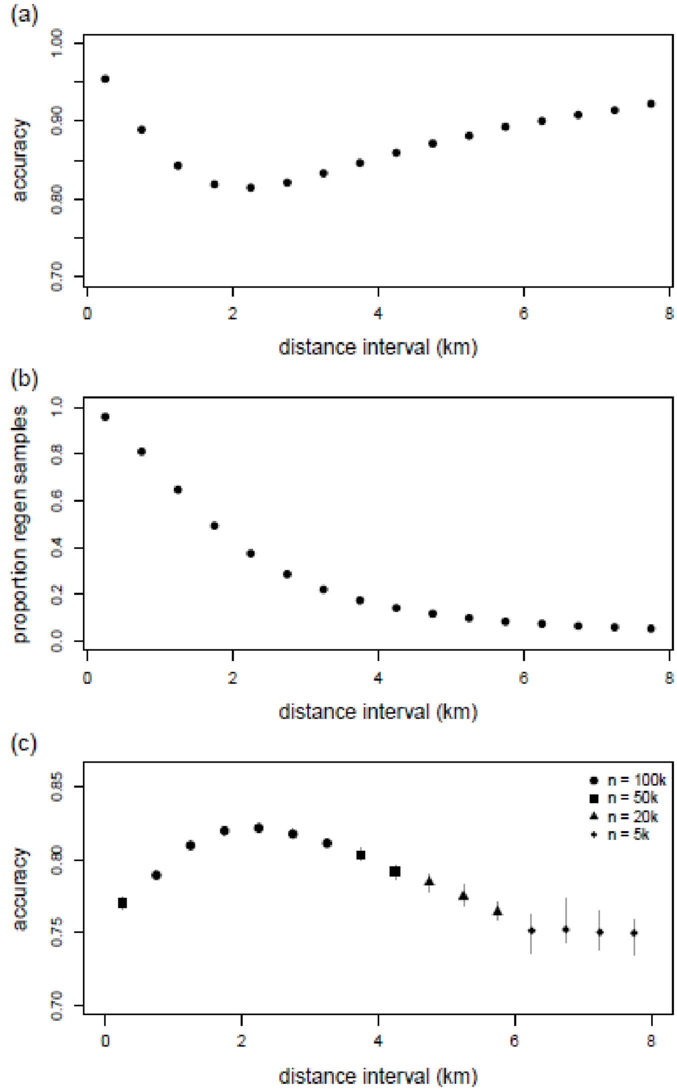
Extended Data Fig. 1 | The change in model accuracy (y axis) among a set of increasingly complex models (x axis). Covariates were added sequentially in order of importance (mean decrease in accuracy metric) to identify the point at which the addition of any additional variables no longer improved model accuracy (here, ten variables).



Extended Data Fig. 2 | Variable importance plots for the model with ten biophysical variables. Mean decrease in accuracy (a) reflects the loss of accuracy that arises from the omission of a variable, whereas the Gini coefficient (b) reflects the contribution of a variable to the purity of the classification (how cleanly the levels of the dependent variable can be separated). Variables are: bioclim1, 2, 3, and 4 = bioclim principle component axes 1, 2, 3, and 4, biome = biome, distfor = distance to forest, focalfor = forest density, landuse = land use, soc = soil organic carbon content, soilph = soil pH.



Extended Data Fig. 3 | Plots indicating how each covariate (x axis) is related to the predicted probability of natural regeneration (y axis). Red lines are the mean predicted probability value calculated over a range of 100 intervals of the covariate. Given that the predicted probabilities of the random sample are skewed towards lower and higher probability scores (a bimodal distribution) and the mean considers all values equally, the mean results in a value that doesn't correspond to any of the probability scores. Open circles are a random sample of 1000 records from the training dataset to provide an indication of the distribution and variation in the data. The blue rug marks along the bottom of each plot represent the deciles of the distribution of values of that covariate. The covariate include: forest density in a 1 km^2 area, distance to nearest forest, soil pH (PHIHOX), soil organic carbon content (OCDENS), and four bioclimatic variables representing the first four axes of a principal components analysis based on 19 original bioclimatic variables (see Data Sources for details). The distance to nearest forest graph is plotted over a limited x axis range to improve legibility.



Extended Data Fig. 4 | Model accuracy based on a random sample of 4.87 M points with equal number of regeneration and non-regeneration observations was 87.9%. To evaluate evidence of autocorrelation effects in this estimate accuracy was estimated in 16 0.5-km intervals between the training and validation points. (a) Accuracy was highest in the shortest distance interval, declined to 81.4% in the 2.0-2.5 km interval, and then increased. This pattern was driven by the combined effect of different accuracies for each outcomes (regeneration vs. non-regeneration) and the change in the relative frequency of each outcome as the distance interval increases (b). Accuracies were lower when forcing balanced samples in each interval and using a bootstrapping approach with 50 replicates of sampling with replacement. Sample sizes varied among intervals (c - see legend) because of the relative rarity of some outcomes at some distance-intervals (b). Vertical lines represent the range of accuracies among all 50 replicates.

nature portfolio

Corresponding author(s): brooke.a.williams@qut.edu.au

Last updated by author(s): 28/07/2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	All code was developed by the authors. R and python code developed for and used in this analysis are available upon request from hawthornebeyer@gmail.com.
Data analysis	The majority of geoprocessing in this analysis used open source libraries in R and Python (the 'osgeo', 'numpy' and 'multiprocessing' Python packages, and the 'sp', 'raster' and 'sf' libraries in R), with some additional processing and map-making in ESRI ArcGIS. References ESRI. ArcGIS [GIS software]. Version 10.8. (2022). GDAL/OGR contributors. GDAL/OGR Geospatial Data Abstraction software Library. Open Source Geospatial Foundation (2023) doi:10.5281/zenodo.5884351. R Core Team. A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria (2013) doi:ISBN3-900051-07-0.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All input datasets are available from the references cited. The raster datasets arising from this work are available link below and include: (i) the approximately 1km resolution overview raster representing the proportion of the area of each pixel that has the potential for natural regeneration ('prop pnv v1 1km.tif'; 1.1 GB); (ii) the 30m resolution binary predictions (where a value of >0.5 is allocated a value of 1, acknowledging users may prefer to use a different threshold for their context) of areas suitable for natural regeneration (files named with the prefix 'pnv bin 30m'; 2.3 GB); and (iii) the 30m continuous probability predictions of the potential for natural regeneration, stored as integers representing percentages to minimise file sizes (files named with the prefix 'pnv pct 30m'; 9.3 GB). The 30m resolution datasets have been tiled into 10 degree latitude/longitude tiles. https://zenodo.org/record/7428804

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	NA
Reporting on race, ethnicity, or other socially relevant groupings	NA
Population characteristics	NA
Recruitment	NA
Ethics oversight	NA

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Based on a sample natural forest regrowth, we use machine learning methods to distinguish areas where natural regeneration did or did not occur across the global tropics as a function of a suite of geospatial biophysical and socio-economic variables. To predict potential tropical forest generation across Earth, we select predictor variables spanning local (site) and landscape scales (from 30 m to country level) that are known to influence the potential for tropical forest regrowth. Spatially explicit values of predictor variables were then used to model the relative potential for natural regeneration in the present (2015), and in the near future (2030), assuming that overall conditions from 2000-2016 apply to future scenarios.
Research sample	The analysis was based on a sample of 5.4 M ha of natural regrowth detected by Fagan and colleagues 2022 (https://doi.org/10.1038/s41893-022-00904-w). Based on this sample use machine learning methods to distinguish areas where natural regeneration did or did not occur across the global tropics as a function of a suite of geospatial biophysical and socio-economic variables.
Sampling strategy	The statistical modelling and validation is based on a set of 6 million random points generated within the regeneration and non-regeneration spatial domains defined above, stratified so that each level of the dependent variable was equally represented but with no other form of spatial stratification. These random points were then intersected with the geospatial covariate datasets (see Predictor Variables section in Methodology in manuscript). Points falling outside of the target biomes (tropical and subtropical forest and savannah biomes), falling within excluded land cover categories (water, bare areas, urban areas, sparse vegetation; ESA CCI Landcover for year 2000), or falling within the regeneration or forestry polygons delineated in the Fagan and colleagues 20222 dataset were excluded.

Data collection	NA - predictor variables where selected for their known influence on the potential for tropical forest regrowth. See Predictor variables section of Methodology in main manuscript.
Timing and spatial scale	Data collection was not relevant to our study. The datasets used in this analysis span from 30 m to the country level spatial resolution.
Data exclusions	Any random points having NoData values for any of the covariates were removed, resulting in a final set of 5.8 million records. NoData values can arise due to mismatches in the coverage of datasets, such as along coastlines.
Reproducibility	NA
Randomization	The points within regeneration spatial domains were sampled as spatially random points generated within the forest regeneration polygons using the 'sf' library in R (function 'st sample'). The points in non-regeneration spatial domains were sampled as spatially random points within latitude limits of +/- 25 degrees using the 'random points GCS' function in R, which accounts for the elliptical shape of Earth.
Blinding	Blinding refers to the concealment of group allocation from one or more individuals involved in a study, and is therefore, not relevant to our study which was desktop-based.

Did the study involve field work? ☐ Yes ☒ No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging