Peer Review File

Manuscript Title: The global potential for natural regeneration in deforested tropical regions

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referees' comments:

Referee #1 (Remarks to the Author):

This manuscript uses a random forests machine learning algorithm to predict the potential area available for natural regeneration of forests in the tropics at a 30m resolution. The authors estimate that 215 million ha of land could be available for natural regeneration, which has the potential to sequester 23.4 Gt CO2 over the next 30 years. Recent research has touch upon the carbon sequestration rates and distribution of naturally regenerating forests, however a pan-tropical analysis is lacking and therefore this paper presents important new findings that would be of interest to the forest restoration community both in academia and the wider NGO & private sector who are involved in development of forestry NBS projects. Natural regeneration of tropical forests is a hugely under-exploited nature based solution and therefore highlighting the importance of these forests as a solution to climate change mitigation is necessary and timely. However, there are a couple of key issues in this paper related to 1) the application of the methodology and 2) the framing of the findings into the wider context of forestry based NBS.

Overall the methodological approach is solid and I think the use of machine learning (ML) is appropriate in this context. However, I am concerned about the ability of the model to accurately predict the potential for natural regeneration (NR) across three highly variable biomes, which includes dry, moist and coniferous forest types (Tropical & Subtropical Dry Broadleaf Forests, Tropical & Subtropical Moist Broadleaf Forests, and Tropical & Subtropical Coniferous Forests). The training data for these different biomes will have highly variable values for the predictor variables used in analysis, in particular climate variables, for example the different in annual precipitation and precipitation seasonality between tropical dry forests and moist forests will be large and I wonder how well the ML model can predict NR with such a wide range of predictive values. I think there is potentially scope to run these models for the three different biomes separately and see if this improves the predictive ability of the model and the accuracy. Or alternatively, additional details about the validity of model over these varying biomes.

Leading on from this, the study presented here uses the map of naturally regenerating tropical forest produced by Fagan et al (2022) Nat Sust. to select training data for the ML model. However, the authors state that the Fagan study ‘regularly fails to detect regrowth patches’ (L97) and that ‘the final predicted map omitted the majority of natural regrowth patches in the humid biome’ quoting a producers accuracy of 18.7% (lines 348-350). The producers accuracy is a measure of how well the model predicts a given class outside of the training data range, therefore with such a low producers accuracy for one of the three biomes I am concerned that the training data selected in this study to predict NR in humid forests will be of poor quality and thus affect the accuracy of the map produced here. Some details of why this Fagan study is acceptable for use as training data with the model is needed.

The discussion and conclusions of the study are generally valid and compare the results against relevant recent research (although please see line by line comments for more specific feedback), however in places they are somewhat generic, and I think some work is needed to 1) reframe the study better within the current context of NBS and carbon offsetting, which is the likely route by which such NR projects would be funded, and 2) more accurately represent the carbon benefits of NR as a climate mitigation strategy.

Specifically, the discussion around how NR could be scaled up (L275-289) could be improved if it included details about some of the constraints of implementing NR that need to be overcome. To give one example, carbon offsetting projects via the voluntary carbon market would likely be a large source of funding for NR activities (as it is for REDD+ and reforestation), however there are currently not really suitable methodologies available to certify NR projects. Currently REDD+ methods require project areas to be forested for 10 years prior to a project starting, which is not valid in this case, ARR (afforestation, reforestation & revegetation) methods require project areas to be unforested for 10 years prior to a project starting, again not valid here if areas are already regenerating, and for IFM (improved forest management) methods project areas typically have to be within a logging concession also unlikely to be valid in most cases. Without the ability for a NR project to be certified and issue carbon credits it will be challenging for projects to receive income to continue activities or more importantly for land holders/ communities involved in NR activities to receive an income for their work thus incentivising them to participate. Whilst the authors state that this map is intended to help guide broader decision making processes, I think highlighting the challenges of implementing such NR activities would be helpful.

Regarding the potential carbon benefits, the numbers presented here reflect the max potential of NR if it takes place over the entire 215 Mha area for 30 years, such pantropical max potential estimates are useful but do not necessarily reflect what is likely to happen on the ground (e.g. Lewis et al. 2019 Nature) and therefore a more realistic assessment of the likelihood of NR happening is also required. This may be outside the scope of this analysis but is at least worthy of discussion in the paper.

Furthermore, the carbon numbers don't account for the leakage effects of potential NR activities. Whilst the training data is taken from areas that are already under NR, this predictive map is showing where NR could take place, but there is no discussion of what the currently land use of these areas is and what the transition to NR would mean in term of leakage impacts, which at a minimum will reduce some of the carbon benefits estimated and at worse has the potential to completely negate the benefits seen from NR (see: Filewood & McCarney 2023 One Earth https://doi.org/10.1016/j.oneear.2023.05.024, Streck 2022 Climate Policy https://doi.org/10.1080/14693062.2021.1920363, and Villoria et al 2022 Nat Comms https://www.nature.com/articles/s41467-022-33213-z). Again this may be outside the scope of the analysis but it warrants discussion as leakage is often overlooked but has potentially huge impacts of carbon sequestration benefits.

Line by line comments

L47/48: Need a reference for this statement

L58: ‘magnitude of benefits they provide’ seems vague, rephrase

L61: remove ‘of’

L93-96: not immediately clear your referring to the Fagan study here, avoid using ‘this’ as start of a sentence as its unclear, suggest rephrasing to improve clarity.

L97/98: referring to the Fagan study, here it states ‘…., regularly failing to detect regrowth patches and rarely predicting false positive regrowth.’ – should this say rarely failing rather than regularly failing? If the Fagan model regularly fails to detect regrowth patches then why is it suitable for use as training data?

L103: change ‘generation’ to ‘regeneration’

L119: Figure 1 is not reference in the main text

L181-183: this is the only reference in the main text to how the carbon numbers are calculated, it is not detailed in the methods or Appendix 1, I suggest that this is included in the methods section for clarity.

L227-229: This seems like an invalid point as the likelihood of intensive agriculture ceasing and natural regeneration taking over is very low, furthermore the leakage impacts of this sort of transition are not factored into the assumption, we still need to produce food so if this process were to take place then production would likely be displaced elsewhere leading to no net benefit of NR under these circumstances. Suggesting this is a means of underestimation seems fairly weak as it just won’t happen, and if it did the climate benefits would be limited. Suggest removing.

L251-253: This point that naturally regenerating forest have a high chance of re-clearance seems like a key limitation of this sort of forestry solution and needs to be discussed in more detail. What is driving the re-clearance of naturally regenerating forests? What policy interventions or PES/carbon financing schemes need to be put in place to prevent this? I suspect that counterproductive policies, which disincentivise the retention of regenerating forests, is leading to their re-clearance. This is the case in Brazil where re-clearance of regenerating forest shas been observed (Piffer et al 2022 Glob Change Bio) and is linked to policies related to the ability for landholders to clear forested land once it is >10 yrs old. Maintaining NR forest is essential if they are to contribute to climate mitigation so way to overcome the current issues need more consideration here in the discussion.

L279: mentions ‘enrichment planting’ this seems to contradict points made in introduction which state that NR (as opposed to enrichment planting) is a much cheaper solution. If this is being done to support livelihoods then the additional costs of planting need to be met.

L303/304: Are the estimates presented here over a comparative area to the Giscom & Cook-Patton studies? is it possible to scale the potential from the Cook-Patton & Giscom papers with the results presented here for a direct comparison?

L361: ‘areas that were already classed as tree cover’, when is this referring to? 2000? Please add.

L395: ‘the potential for natural regeneration in the future’, until when? The map presented here uses current day climate data so won’t be valid far in the future under different climate regimes, please add a timeframe over which these predictions are valid.

L463-467: Think there is some formatting problem here which is making these line hard to understand.

L479: The 1km forest density estimate is unclear. 1km around what? Is this a buffer around a predicted NR area or a moving window across all pixels? Is this done at the pixel level as with cropland area? Is it done for every pixel to create a 1km forest density layer or is it just done for NR areas/pixels?

L526: add resolution of NPP product.

Figure 2: Interesting way to display but the scale is hard to see particularly for neotropics where there is only values for Brazil with a broken bar. At minimum include a value on the Colombia bar (as with the 2 other regions), but maybe include values for top 5 countries for each country.
Further, why is the area available for 'all restoration' included here? ‘Restoration in general’ (L144) is not defined in text so its unclear what these other restoration options are. If what are only interested in natural regeneration, this scaling makes it really hard to see the actual contribution of NR for each country as the green bars are so small and dominated by the yellow which is somewhat redundant here. Maybe consider just showing the green area available for NR which is the focus of this study.

Figure 3: panel A/B, the scale for the potential for natural regeneration which runs from -2.5 to 0.2 is not that intuitive and somewhat confusing to understand as a reader. The figure needs to be understandable without refereeing to the text and currently this unit of measurement is unclear. Suggest either improving explanation of this is the caption, or using another unit of measurement to present the data

APPENDIX 1 comments below:

L13/L15: Use of ‘the former’ & ‘the latter' is confusing, state what you are referring to specifically for clarity

L18: refers to savannah biomes as a ‘target biome’ is this correct? The methods just refer to analysis including the dry, moist and coniferous forest biomes. Please clarify

Figure A1.3: The data for all these variables seems very clustered at the lower (<0.4) and higher

(>0.7) probability scores, so the use of a mean predicted probability (red lines) doesn't seem to reflect the actual data. Could this be explained a little more in the caption?

Referee #2 (Remarks to the Author):

A SUMMARY OF THE KEY RESULTS

This is a novel study of an important subject: how much of the tropics would regenerate naturally to forest if left (pretty much) alone.

The authors use a dataset that shows where regeneration occurred from 2000-2018, along with a set of environmental and anthropogenic correlate variables and machine learning, to predict the likelihood of natural regeneration for every pixel. They then multiply this likelihood map by area to obtain rough estimates of the area available for natural regeneration, and the amount of carbon that could be captured, coming up with a reasonable area number of 215 million hectares (approximately a quarter of the current tropical forest area), and a carbon sequestration potential number of around 23 Pg CO_2 over 30 years (again sounding reasonable).

B ORIGINALITY AND SIGNIFICANCE
I'm not aware of a similar paper in the literature. The paper is clearly novel, important, and timely.

It is an important result and map they produce. Natural regeneration indeed can be very/more effective than tree planting, and of course cheaper - but only at a local scale is it normally known where it is really possible. Global-scale analyses like these will not replace local maps and knowledge: despite the 30m resolution this map is not a tool for local managers, its accuracy will be too poor. But it is a very useful tool for enabling national and international decisions on the importance, potential, and funding of natural regeneration methods. It could for example give sensible numbers to input into climate change mitigation scenarios, or to target funding from international donors. It also answers important scientific questions on this subject: it was hard for a scientist like me to give a good estimate to a policy maker of how much area would just regenerate naturally if left alone.

C Data and methodology

I have various issues with the methodology. The overall maps look fairly reasonable, and the overall estimates seem plausible. But the methods, and in particularly the uncertainty analysis (or lack of it) need work to come up to a sufficiently robust scientific standard.

1. The output maps are a combination of 30m and 1km inputs. This leaves the maps a blocky look (see Figure 1), and means that their true resolution is not 1km. The authors are open about this, but it is a limitation that could have been improved on. 1km layers could have been co-krigged to 30m using associated 30m output layers - or even just upsampled using e.g. a bilinear kernal. And where they were not very significant in the model, 1km layers could have been removed from the analysis in preference to higher resolution input layers.

Doing the analysis at 30m not 1km resolution is clearly essential, given the importance of distance to forest patches in the hundred of meter ranges. But this justification for high resolution is diluted by having any 'blurring' 1km layers.

2. I am unconvinced by the approach of multiplying relative potential by pixel area to produce actual area. Ignoring differences in pixel sizes (see below), does a 0.09 ha pixel that comes out of the model with a 90% potential really count 90x more to the area count than one with a 1% potential? Might it not be better to count the 90% potential pixl as a full 0.09 ha of restoration, and the 1% potential pixels as 0 ha?

I think counting every pixel with greater than say 50%, and everything below that as 0%, might be better - or at least I'd like to see those numbers.

This is especially true as Random Forest models tend to favour mid values, and do not push to extremes. There will normally be fewer 0s and 100s in your model output than the reality.

3. I'm concerned that the area calculations were done using pixels that were considered '30m x 30m', but are in fact variable across the grid as they are not in an equal area projection, but a simple lat/lon projection. This would give greater weight to pixels further from the equator in the analysis, as well as simply leading to incorrect area values. Can the authors confirm that an equal area projection was used to perform the calculations, or if not that area was corrected for in a different way?

4. When comparing the potential for carbon sequestration of these areas to the current biotropical sink, can you be certain that naturally regenerating forests are not already counted to some degree in global figures? I suspect they will be to some degree, making the percentage figures here an overestimate of the impact of leaving all these areas alone compared to Business As Usual.

5. I worry about the input Hansen/Fagan dataset on forest regrowth. This dataset has a very low accuracy, missing the vast majority of regrowth pixels: as the paper states, it detects less than 20% of natural regrowth pixels. While these were pretty accurate when detected (user's accuracy of 85%), the authors do not, in my opinion, sufficiently justify that the patches that were detected are an unbiased set. Further, by definition, the pixels that do not have patches detected in the Hansen/Fagan dataset are, according to the model, not regeneration. They will be treated as such and included as 'no regeneration' - even though in a large number of cases (cover 4x more pixels than the actual detected gains) they are in fact natural regeneration. So even if the pixels successfully picked up are a representative sample, they will be represented on both regrowth and non-regrowth sides of the model, causing confusion and likely underestimating regrowth potential globally.

In order for the publication to go ahead I think the authors need to show in far more detail, for example in supplementary figures in some key case studies areas, the nature of the undetected natural regrowth patches and explain through a comprehensive analysis how this does not add great uncertainty to their results. The current error analysis predicts very tight confidence intervals on the area of regrowth predicted (none) and narrow confidence intervals on the carbon stock change over 30 years - I do not think they can be justified given the low accuracy of the training data of forest gain they use.

6. Are the Hansen et al. 2000 tree cover and 2018 tree cover datasets comparable? The model is trained on the former and applied with the latter, but no analysis is presented to suggest any test has been performed to see if this is the case. As the 2018 dataset is referenced to the Hansen et al. (2013) paper, clearly the methods used have not been scrutinised or put through peer review. The satellites used are clearly different between the two (Landsat 7 vs. Landsats 8/9, and as the authors will know ETM+ and OLI are quite different sensors in their characteristics). Hansen et al. caution in the notes to their website giving out the data about a worry about continuity after 2012. Was this considered?

D. Statistics and uncertainties

As stated in 5. above, I do not think uncertainties in the area estimation are propagated through at all: no uncertainties are given for what are definitely very uncertain numbers. I would like to see these calculated prior to publication.

E. Conclusions

The map of the probability of natural regeneration is useful and probably fairly accurate in pattern (albeit a mix of 1km and 30m resolution making use difficult).

However I am unsure that the area and carbon numbers can be used to draw robust conclusions as currently calculated. At minimum, they have large uncertainty and this should be reported.

F. Improvements.

Beyond those above, I have some minor points:

1. Different units used - e.g. Gt C02 line 193, Pg C/yr line 300, Pg CO2e/yr line 304. Pg is identical to Gt, but should be used throughout (SI unit). C and CO2 should not be switched between.

2. Brazil/Indo/China/Mexico/Colombia between them account for 52% of potential regeneration - but what percentage of land between 25 degrees south and north do they cover? If around 52% then this is not surprising.

G/H - no comments.

Referee #3 (Remarks to the Author):

The manuscript titled "The global potential for natural regeneration in deforested tropical regions" aims to determine what is the surface extent around the tropics for deforested regions that has the potential for natural regeneration if allowed. They accomplish this by leveraging a recently released satellite-derived map of regenerated forest patches as was detected by a suite of satellites over the 2000-2012 time period. They use this dataset as the ground-truth of where forest regeneration occurred and then build a Random forest model using these data to map out where the regeneration occurred more generally over the tropical regions during this time period and more importantly to assess what is the potential for regeneration over the 2016-2030 time period. The relevance of this work is trying to determine how much extra carbon could go into natural regrowth of forests and thus being a fairly "easy" attainable goal. Overall, the paper is well written, however, I am not convinced that the results are sufficiently compelling and convincing for publication in Nature. The authors themselves list a large amount of assumptions and counter-arguments in the discussion section. I actually appreciate their honesty as many of the points that they bring up are ones that I thought of as I was reading through the article, but it does highlight that the results are prone to large uncertainties. For example, the assumption that forest regeneration will continue the same way from 2016-2030 as it did from 2000-2016 is an important assumption; other issues include the fact that the dataset that was used to train the RF was still a sub-sample and therefore we are left with questions about its performance over regions where these data don't exit. Other issues exist but overall my hesitation with the paper is not the methods or the writing, I think the authors do a great job. I just don't think it is at the level for Nature. Instead, I think it would be great for Nature Sustainability. And if that is the choice, I would support it being published there practically "as is".

Author Rebuttals to Initial Comments:

RESPONSE TO REVIEWERS COMMENTS

Overall response: We thank the Reviewers for finding novelty and merit in our work. We agree with their comments and have responded accordingly below. All adjustments have greatly improved the manuscript.

Referees' comments:

Referee #1 (Remarks to the Author):

This manuscript uses a random forests machine learning algorithm to predict the potential area available for natural regeneration of forests in the tropics at a 30m resolution. The authors estimate that 215 million ha of land could be available for natural regeneration, which has the potential to sequester 23.4 Gt CO2 over the next 30 years. Recent research has touch upon the carbon sequestration rates and distribution of naturally regenerating forests, however a pan-tropical analysis is lacking and therefore this paper presents important new findings that would be of interest to the forest restoration community both in academia and the wider NGO & private sector who are involved in development of forestry NBS projects. Natural regeneration of tropical forests is a hugely under-exploited nature based solution and therefore highlighting the importance of these forests as a solution to climate change mitigation is necessary and timely. However, there are a couple of key issues in this paper related to 1) the application of the methodology and 2) the framing of the findings into the wider context of forestry based NBS.

We thank the reviewers for their positive feedback and thoughtful comments. We have addressed all their concerns below accordingly.

Comment 1

Overall the methodological approach is solid and I think the use of machine learning (ML) is appropriate in this context. However, I am concerned about the ability of the model to accurately predict the potential for natural regeneration (NR) across three highly variable biomes, which includes dry, moist and coniferous forest types (Tropical & Subtropical Dry Broadleaf Forests, Tropical & Subtropical Moist Broadleaf Forests, and Tropical & Subtropical Coniferous Forests). The training data for these different biomes will have highly variable values for the predictor variables used in analysis, in particular climate variables, for example the different in annual precipitation and precipitation seasonality between tropical dry forests and moist forests will be large and I wonder how well the ML model can predict NR with such a wide range of predictive values. I think there is potentially scope to run these models for the three different biomes separately and see if this improves the predictive ability of the model and the accuracy. Or alternatively, additional details about the validity of model over these varying biomes.

Response 1

The reviewer raises an important point. For 68% of the tropical forested nations included in our analysis this isn't an issue as they only contain a single biome; however, for the other 32% (such as Brazil and most nations in South East Asia) accurately using and interpreting multiple models will complicate and hinder national-level restoration planning. For this reason we would much prefer to produce and release a single model. We do agree that our manuscript would benefit from details about the validity of model accuracy across biomes. We have therefore included at L143 (additions in bold):

"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 & Subtropical Moist Broadleaf Forests, Tropical & Subtropical Dry Broadleaf Forests, and Tropical & Subtropical Coniferous Forests, respectively.** The validation accuracy estimate of 87.9% was based on an independent set of 4.87 M random points (equally stratified with respect to the two levels of the dependent variable) (see Appendix 1 and Fig. A.1.4 for more details)."

Comment 2

Leading on from this, the study presented here uses the map of naturally regenerating tropical forest produced by Fagan et al (2022) Nat Sust. to select training data for the ML model. However, the authors state that the Fagan study 'regularly fails to detect regrowth patches' (L97) and that 'the final predicted map omitted the majority of natural regrowth patches in the humid biome' quoting a producers accuracy of 18.7% (lines 348-350). The producers accuracy is a measure of how well the model predicts a given class outside of the training data range, therefore with such a low producers accuracy for one of the three biomes I am concerned that the training data selected in this study to predict NR in humid forests will be of poor quality and thus affect the accuracy of the map produced here. Some details of why this Fagan study is acceptable for use as training data with the model is needed.

Response 2

The reported low producer's accuracy estimate in Fagan and colleagues 2022 arises from a known issue when reporting area-based producer's accuracy assessments. Producer's accuracy is defined as the inverse of omission error. Originally, Olofsson and colleagues 2013, 2014 advocated for the incorporation of map area into accuracy assessments as a 'best practice' in remote sensing. They highlighted the fact that attempts to estimate the area of a specific class by methods that sum values for map units assigned to that map class ("pixel-counting") produce erroneous area estimates because the effects of classification errors are ignored. They argued the area of land use or land cover change obtained directly from a map may differ greatly from the true area of change because of map classification error.

However, Olofsson and colleagues 2020 raised important shortcomings in their method for rare classes that make up a minority of a map, like deforestation (and in our study, regeneration). They detail how their area-correction practices lead to artificially low, inflated estimates of omission errors for rare classes, using case studies highlighting real-world examples of bias. For the specific rare class they focus on, deforestation, they propose a strata-based correction for error-inflation that takes advantage of the spatial concentration of deforestation on the perimeter of forests. In effect, if we buffer forests inwards by a set distance, we capture the majority of the area where deforestation can occur, constraining where omission error can occur and limiting error-inflation. While this solution works for deforestation, there is currently no applicable fix for less spatially-constrained rare classes (P. Olofsson, *pers. comm.*), like regeneration.

Because natural regrowth can theoretically occur anywhere that lacks forest within forested biomes, the relatively high producer's accuracy of the Fagan and colleagues 2022 model (78.9%) was artificially lowered by the vast area of non-forest in the humid biome, down to 18.7% (when based on estimated area, the error matrix and hence producer's accuracy are adjusted by area weights, calculated as the proportionate area occupied by the class, meaning that the complement of producer's accuracy measures potentially omitted area proportions, instead of model accuracy). In reality, natural regrowth is not possible across the majority of this omitted area due to human land-use, but pantropical maps of land-use are not currently accurate enough to remove this human-dominated non-forest area as a possibility. Thus, this "corrected" estimate omission error is low due to bias but was included in the original publication as a conservative estimate.

A recent publication in *Nature* report on the same issue. Wang and colleagues 2023, found a significant drop in their producer's accuracy, from 0.95 to 0.57, when it was based on estimated area (see https://doi.org/10.1038/s41586-023-06642-z the "Rubber map for Southeast Asia" section for more details).

Unfortunately, for both rubber (somewhat rare) and reforestation (highly rare, <2% of the tropical area), adjusting for area led to inflation of omission error estimates. Until best accuracy assessment practices are updated to address this issue, it is important to take into account a range of accuracy metrics. We emphasise that the producer's accuracy for the Humid tropical biome was high – 150/190 sampled points for the regrowth class were correct.

To ensure this area-based issue with producer's accuracy is made explicit to non-remote sensing readers, we have added the confusion matrix with both types of producer's accuracy calculations to Appendix 1 (Table A1.1):

Table A1.1. Confusion matrix depicting the independent model accuracy assessment for the humid tropics, and for each subregion (TP = tree plantations, REG = natural regrowth, NG = no-gain, NGP = no gain possible), using only reference polygons predicted with medium to high confidence. Percent accuracy estimates are presented for overall accuracy (mean +/- confidence interval), user's accuracy (100-commission error), and producer's error (100-omission error). Confusion matrix counts represent both the number of polygons, and the area-weighted estimates.

Map Accuracy Testing Dataset							
Humid tropical biome	**Reference data points (n=2981)**						
Predicted	*TP*	*REG*	*NG*	Total	Map Area (Mha)	*% User Acc.*	*Corrected Area (Mha)*
TP	160	12	1	173	9.9	*90.7 ±6.0*	*25.0 ±6.0*
REG	18	150	6	174	4.7	*85.1 ±5.6*	*21.6 ±6.1*
NG	25	28	1432	1485	875.4	*98.2 ±0.5*	*1711.0 ±8.5*
NGP	0	0	1149	1149	867.5		
Total	203	190	2588			***OA %***	
% Prod. Acc.	*78.8 ±5.6*	*78.9 ±5.8*	*99.7 ±0.20*			***97.0** ±0.5*	
% Prod. Acc. (area-based)	*35.8 ±8.7*	*18.7 ±5.4*	*99.99 ±0.01*			***98.1** ±0.5*	

To explain this issue we have also added the following text to L394:

"The class producers 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 (Table A1.1). The low producer's accuracy when based on estimated area is in part due to Fagan and colleagues 2022[33] 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 emphasise the higher (mapped) estimate given this is a common and known issue with reporting on producers accuracy (see[92,93]). Further,…"

Additionally, our follow-up nearest-neighbour analyses have indicated that the mean of nearest-neighbour distances for the observed omission error (6.38) is significantly more distant than nearest-neighbour distances for the full regrowth validation dataset (4.26, CI 5/95% of means 3.09/5.75 (10,000 bootstrapped samples using the same number of omission error points)). In other words, our omission errors are more spread out, on average, than by random chance. This indicates that the model was on average more spatially accurate across different regions than would be expected by random chance, ie. omission error was less clustered than by random chance. Further, we found that the mean nearest-neighbour distance for commission error (5.11) is similar to the means of the full regrowth dataset (5.98, CI 5/95% of mean 0.51/9.30 (10,000 random samples using the same number of commission error points)) indicating that there is no clustering present (ie. patterns are random). Finally, regrowth displays very little apparent correlation with a number of spatial environmental variables. Using binomial logistic regressions (with Bon Ferroni error correction for repeated tests) we found that the occurrence of omission errors in the validation data displays very little apparent correlation with a number of spatial environmental variables as most were non-significant (14/15) apart from slope ($p = 0.001$**). This is to be expected because, as Fagan and colleagues 2022 point out, their product was derived from both optical and microwave data and as such local variation in topography and forest structure affected classification accuracy - their product tended to have lower accuracy in mountainous areas (where microwave returns vary). Similarly, we found that the occurrence of commission errors in the validation data displays very little apparent correlation with the same spatial environmental variables as most were non-significant (14/15) apart distance to urban areas ($p = 0.001$**). Synthetic aperture radar (a key input to the dataset of Fagan and colleagues 2022) has elevated and variable values in urban areas (Ferro and colleagues 2011; Henderson and Xia 1997; Lee 2001), decreasing the likelihood of predicting forest in these regions. We also note that urban areas are a very small percentage of the humid tropics, and therefore not an issue for the present study. We can therefore conclude that there is minimal detectable spatial bias in the Fagan and colleagues 2022 model's producer's accuracy. In this study, our spatial predictive model includes topography and distance to urban areas as predictors, so it is able to correct for this minimal bias in omission error we observed in the input data.

We have added these details to Appendix 1:

"**Clustering analysis**

To check for potential spatial bias in the Fagan and colleagues 2022[2] regrowth patches, we conducted a clustering analysis. We first generated 10,000 bootstrapped samples from the full regrowth validation dataset from Fagan and colleagues 2022[2], using the same observed number of omission and commission error points in each bootstrap sample. Then, separately for the omission and commission bootstrap samples, we assessed potential spatial clustering of error by measuring nearest neighbour distances using the nndist tool from the spatstat package[7] in R[8]. These nearest-neighbour analyses indicated that the mean of nearest-neighbour distances for the original, observed omission error (6.38) is significantly more distant than nearest-neighbour distances for the full, bootstrapped regrowth validation dataset (4.26, CI 5/95% of means 3.09/5.75 (10,000 bootstrapped samples using the same number of omission error points)). In other words, the omission errors are more spread out, on average, than by random chance. This indicates that the Fagan and colleagues 2022 model was on average more spatially accurate across different regions than would be expected by random chance, i.e. omission error was less clustered than by random chance. Further, we found that the mean nearest-neighbour distance for the original commission error (5.11) is similar to the means of the full, bootstrapped regrowth dataset (5.98, CI 5/95% of mean 0.51/9.30

(10,000 random samples using the same number of commission error points)) indicating that there is no clustering present (i.e. patterns are random).

To test for potential correlations between the spatial environmental variables used in our potential for natural regeneration analysis and the validation dataset from the Fagan and colleagues 2022[2] analysis we used glm from the R base library[8] to conduct binomial logistic regressions (with Bon Ferroni error correction for repeated tests). We found that the occurrence of omission errors in the validation data displays very little apparent correlation with a number of spatial environmental variables as most variables were non-significant (14/15) apart from slope ($p = 0.001^{**}$). This is to be expected because, as Fagan and colleagues 2022[2] highlight, their product was derived from both optical and microwave data and as such local variation in topography and forest structure affected classification accuracy - their product tended to have lower accuracy in mountainous areas (where microwave returns vary). Similarly, we found that the occurrence of commission errors in the validation data displays very little apparent correlation with the same spatial environmental variables as most were non-significant (14/15) apart distance to urban areas ($p = 0.001^{**}$). Synthetic aperture radar (a key input to the dataset of Fagan and colleagues 2022[2]) has elevated and variable values in urban areas[9–11], decreasing the likelihood of predicting forest in these regions. We note that urban areas are a very small percentage of the humid tropics, and therefore not an issue for the present study. We can therefore conclude that there is minimal detectable spatial bias in the Fagan and colleagues 2022[2] model's producer's accuracy. In this study, our spatial predictive model includes topography and distance to urban areas as predictors, so it is able to correct for this minimal bias in omission and commission error observed in the input data."

References

Ferro, A., Brunner, D., Bruzzone, L. & Lemoine, G. (2011) On the relationship between double bounce and the orientation of buildings in VHR SAR images. *IEEE Geoscience and Remote Sensing Letters,* 8, 612–616.

Henderson, F. M. & Xia, Z.-G. SAR applications in human settlement detection, population estimation and urban land use pattern analysis: a status report. IEEE Trans. *Geosci. Remote Sens*. 35, 79–85 (1997).

Lee, W. K. (2001) Analytical investigation of urban SAR features having a group of corner reflectors. *IGARSS Scanning the Present and Resolving the Future. Proceedings. IEEE 2001 International Geoscience and Remote Sensing Symposium (Cat. No. 01CH37217),* 3, 1282 - 1284.

Olofsson, P., Foody, G. M., Stehman, S, V., Woodcock, C. W. (2013) Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sensing of Environment,* 129, 122-131. https://doi.org/10.1016/j.rse.2012.10.031

Olofsson, P., Foody, G. M., Herold, M., Stehman, S, V., Woodcock, C. W., Wulder, M. A. (2014) Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment,* 148, 42-57. https://doi.org/10.1016/j.rse.2014.02.015

Olofsson, P., Arévalo, P., Espejo, A. B., Green, C., Lindquist, E., McRoberts, R. E., Sanz, M. J. (2020) Mitigating the effects of omission errors on area and area change estimates, Remote Sensing of Environment, 236, 111492, https://doi.org/10.1016/j.rse.2019.111492

Olofsson, P. (2020) Personal Communication.

Wang, Y., Hollingsworth, P.M., Zhai, D., West, C. D., Green, J. M. H., Chen, H., Hurni, K., Su, Y., Warren-Thomas, E., Xu, J., Ahrends, A. (2023) High-resolution maps show that rubber causes substantial deforestation. *Nature*, 623, 340–346. https://doi.org/10.1038/s41586-023-06642-z

Comment 3

The discussion and conclusions of the study are generally valid and compare the results against relevant recent research (although please see line by line comments for more specific feedback), however in places they are somewhat generic, and I think some work is needed to 1) reframe the study better within the current context of NBS and carbon offsetting, which is the likely route by which such NR projects would be funded, and 2) more accurately represent the carbon benefits of NR as a climate mitigation strategy.

Response 3

We agree that framing the discussion and conclusions around NBS and the carbon offset market would improve the manuscript and lead to more concrete recommendations and insights. We have therefore included the following text at the end of the first paragraph of the Discussion L217:

"It 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, thus improving effectiveness of climate mitigation strategies[54]."

We have also expanded on the point that the voluntary carbon market could be a major source of funds for enabling large scale forest restoration through assisted natural regeneration of forests in a new fourth paragraph of the Discussion – also relevant to Comment 4 (see L266):

"A number of pathways exist for achieving large-scale forest regeneration, but each come with specific challenges that have thus far limited implementation. 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 via the voluntary carbon market could be a large source of funding for natural regeneration activities (as it is for REDD+ 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[66]. For example, currently REDD+ methods require project areas to be forested for 10 years prior to a project starting, while afforestation, reforestation & revegetation methods require project areas to be *unforested* for 10 years prior to a project starting, which would invalidate recently cleared areas or areas already under regeneration[66,67]. 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 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[66,68]."

Finally, we believe that the manuscript already emphasises the potential for natural regeneration to mitigate climate change based on the estimates of potential carbon sequestration (see L180 of results section).

Comment 4

Specifically, the discussion around how NR could be scaled up (L275-289) could be improved if it included details about some of the constraints of implementing NR that need to be overcome. To give one example, carbon offsetting projects via the voluntary carbon market would likely be a large source of funding for NR or ANR activities (as it is for REDD+ and reforestation), however there are currently not really suitable methodologies available to certify NR projects. Currently REDD+ methods require project areas to be forested for 10 years prior to a project starting, which is not valid in this case, ARR (afforestation, reforestation & revegetation) methods require project areas to be unforested for 10 years prior to a project starting, again not valid here if areas are already regenerating, and for IFM (improved forest management) methods project areas typically have to be within a logging concession also unlikely to be valid in most cases. Without the ability for a NR project to be certified and issue carbon credits it will be challenging for projects to receive income to continue activities or more importantly for land holders/ communities involved in NR activities to receive an income for their work thus incentivising them to participate. Whilst the authors state that this map is intended to help guide broader decision making processes, I think highlighting the challenges of implementing such NR activities would be helpful.

Response 4

We agree with the reviewer that it is important to highlight the challenges for putting our results into practice. We have therefore added to the discussion at L266:

"A number of pathways exist for achieving large-scale forest regeneration, but each come with specific challenges that have thus far limited implementation. 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 via the voluntary carbon market could be a large source of funding for natural regeneration activities (as it is for REDD+ 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[66]. For example, currently REDD+ methods require project areas to be forested for 10 years prior to a project starting, while afforestation, reforestation & revegetation methods require project areas to be *unforested* for 10 years prior to a project starting, which would invalidate recently cleared areas or areas already under regeneration[66,67]. 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 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[66,68].

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[26]. However, this would likely require additional interventions (and costs) to ensure that the appropriate species return (for example, through enrichment planting[69]), and ongoing support to access and build market connections."

Comment 5

Regarding the potential carbon benefits, the numbers presented here reflect the max potential of NR if it takes place over the entire 215 Mha area for 30 years, such pantropical max potential estimates are useful but do not necessarily reflect what is likely to happen on the ground (e.g. Lewis et al. 2019 Nature) and therefore a more realistic assessment of the likelihood of NR happening is also required. This may be outside the scope of this analysis but is at least worthy of discussion in the paper.

Response 5

We agree with the Reviewer that this is an important point and warrants discussion. While we agree that a scenario-based analysis is outside of the scope of our contribution, we have added to L341 of the discussion:

"The total pantropical potential carbon sequestration values we present here reflect the maximum potential of natural regeneration if it takes place over the entire 215 M ha area for 30 years."

And following on from this addition have added to L362 in response to Comment 6.

"The realised potential is likely to be much lower, as reforestation may include commercial or exotic species[80], not persist[81] or cause leakage[82]. Leakage is sometimes considered at the project level (although not necessarily mitigated against), but rarely at broad scales, likely leading to a reduced overall additionality of restoration projects[82]. In the context of restoration, leakage occurs when a restoration intervention has an effect outside the accounting boundary used to track mitigation effects (e.g., an action causing emissions reductions in one place may also cause increases elsewhere), and leakages frequently result from restoration activities[82–84]. That said, regional forest expansions are commonly observed across the tropics when deforestation pressures decline[85–87], indicating that societal and policy changes are achievable to increase reforestation."

Comment 6

Furthermore, the carbon numbers don't account for the leakage effects of potential NR activities. Whilst the training data is taken from areas that are already under NR, this predictive map is showing where NR could take place, but there is no discussion of what the currently land use of these areas is and what the transition to NR would mean in term of leakage impacts, which at a minimum will reduce some of the carbon benefits estimated and at worse has the potential to completely negate the benefits seen from NR (see: Filewood & McCarney 2023 One Earth https://doi.org/10.1016/j.oneear.2023.05.024, Streck 2022 Climate Policy https://doi.org/10.1080/14693062.2021.1920363, and Villoria et al 2022 Nat Comms https://www.nature.com/articles/s41467-022-33213-z). Again this may be outside the scope of the analysis but it warrants discussion as leakage is often overlooked but has potentially huge impacts of carbon sequestration benefits.

Response 6

This is an excellent point, and while we agree it is outside of the scope of our analysis, we have added to the discussion at L343:

"The realised potential is likely to be much lower, as reforestation may include commercial or exotic species[80], not persist[81] or cause leakage[82]. Leakage is sometimes considered at the project level (although not necessarily mitigated against), but rarely at

broad scales, likely leading to a reduced overall additionality of restoration projects[82]. In the context of restoration, leakage occurs when a restoration intervention has an effect outside the accounting boundary used to track mitigation effects (e.g., an action causing emissions reductions in one place may also cause increases elsewhere) and frequently results from restoration activities[82–84]. That said, regional forest expansions are commonly observed across the tropics when deforestation pressures decline[85–87], indicating that societal and policy changes are possible to increase reforestation."

Line by line comments

Comment 6

L47/48: Need a reference for this statement

Response 6

Thank you, we have referenced IPCC 2022 for this statement (now L47).

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 University Press, 2014).

Comment 7

L58: 'magnitude of benefits they provide' seems vague, rephrase

Response 7

We have changed 'magnitude of benefits they provide' to "the magnitude of economic, cultural, and recreational services they provide to people" (see L55).

Comment 8

L61: remove 'of'

Response 8

Done

Comment 9

L93-96: not immediately clear your referring to the Fagan study here, avoid using 'this' as start of a sentence as its unclear, suggest rephrasing to improve clarity.

Response 9

We have rephrased from "This analysis" to "Analysis by Fagan and colleagues 2022 distinguished".

Comment 10

L97/98: referring to the Fagan study, here it states '...., regularly failing to detect regrowth patches and rarely predicting false positive regrowth.' – should this say rarely failing rather than regularly failing? If the Fagan model regularly fails to detect regrowth patches then why is it suitable for use as training data?

Response 10

We have now deleted this sentence as have dealt with the issue of accuracy in Response 2 (see above), and added supporting text accordingly.

Comment 11

L103: change 'generation' to 'regeneration'

Response 11

Done

Comment 12

L119: Figure 1 is not reference in the main text

Response 12

We now refer to Figure 1 at L118.

Comment 13

L181-183: this is the only reference in the main text to how the carbon numbers are calculated, it is not detailed in the methods or Appendix 1, I suggest that this is included in the methods section for clarity.

Response 13

We have added the methodology used for calculating carbon values to the manuscript at L655 under the heading of "*Carbon accumulation potential calculations*":

"*Carbon accumulation potential calculations*

We calculate carbon accumulation potential using the Cook-Patton and colleagues 2020[11] dataset which provides carbon accumulation in Mg 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 Mg per ha at a 1km by 1km resolution. To downscale to a 30m by 30m cell resolution we first resampled the layer and then multiplied it by 0.09, because there are 0.09 hectares in a 30m by 30m cell. This gave us a total value per cell, which we then multiplied by the potential for natural regeneration, and finally summarised (sum) at the global, and bioregional levels."

Comment 14

L227-229: This seems like an invalid point as the likelihood of intensive agriculture ceasing and natural regeneration taking over is very low, furthermore the leakage impacts of this sort of transition are not factored into the assumption, we still need to produce food so if this process were to take place then production would likely be displaced elsewhere leading to no net benefit of NR under these circumstances. Suggesting this is a means of underestimation seems fairly weak as it just won't happen, and if it did the climate benefits would be limited. Suggest removing.

Response 14

We agree and have deleted this point.

Comment 15

L251-253: This point that naturally regenerating forest have a high chance of re-clearance seems like a key limitation of this sort of forestry solution and needs to be discussed in more detail. What is driving the re-clearance of naturally regenerating forests? What policy interventions or PES/carbon financing schemes need to be put in place to prevent this? I suspect that counterproductive policies, which disincentivise the retention of regenerating forests, is leading to their re-clearance. This is the case in Brazil where re-clearance of regenerating forest shas been observed (Piffer et al 2022 Glob Change Bio) and is linked to policies related to the ability for landholders to clear forested land once it is >10 yrs old. Maintaining NR forest is essential if they are to contribute to climate mitigation so way to overcome the current issues need more consideration here in the discussion.

Response 15

We agree that this point warrants more explanation, and we have provided additional discussion on what prevents longevity and how these challenges could potentially be overcome at L248:

"Forests undergoing post-agricultural natural regeneration are highly vulnerable to reclearance[60–62]. 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[62,63]. Where forests are recovering on unused agricultural land, global forces, like market shifts in forest-risk commodity prices, such as cattle, coffee, cacao, or palm oil, can push landholders to re-clear 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 want to retain control of their land[62,63]. At local levels, young regenerating forests are often perceived as unused or wasteland, and may be re-cleared to claim or retain control of land[64]. In all cases, taking appropriate measures to protect young regenerating forest is important for them to persist (e.g., firebreaks, fencing and/or patrolling at the local level, along with adequate incentives or compensation, appropriate legislation, and effective resources for enforcement). Changing perceptions at local and national levels around the value of these forests through education and/or market opportunities is also important in many places[65]. 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."

Comment 16

L279: mentions 'enrichment planting' this seems to contradict points made in introduction which state that NR (as opposed to enrichment planting) is a much cheaper solution. If this is being done to support livelihoods then the additional costs of planting need to be met.

Response 16

The Reviewer raises and excellent point. We have added to L285 of the discussion:

"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[26]. However, this would likely require additional interventions (and costs) to ensure that the appropriate species return (for example, through enrichment planting[69]), and ongoing support to access and build market connections."

Comment 17

L303/304: Are the estimates presented here over a comparative area to the Giscom & Cook-Patton studies? is it possible to scale the potential from the Cook-Patton & Giscom papers with the results presented here for a direct comparison?

Response 17

It is difficult to scale the potential C sequestration from the Cook-Patton & Griscom studies to compare with the results presented here. We use the Cook-Patton and colleagues 2020 dataset to derive our C estimates, which already makes comparing our estimates directly to the values presented in the Cook-Patton and colleagues 2020 paper a circular exercise. Furthermore, our predictions represent the carbon sequestration values over a different study region and are scaled by modelled potential. We have therefore deleted the comparison with Cook-Patton and colleagues 2020 from the discussion.

The main difference between the study regions of the Griscom and colleagues 2017 and 2020, and the one used in our analysis is that ours is restricted to +/-25 degrees latitude while the other studies included all tropical countries, defined as those having greater than 50% of their land area between 23.5°N and 23.5°S. Further, there are key differences in our analysis that make direct comparison to the Griscom and colleagues 2017 and 2020 paper convoluted.

Griscom and colleagues 2017 quantified the benefits of a variety of "natural climate solutions" including restoration, and used a standard natural forest growth rate for their calculations, failing to account for the biophysical heterogeneity of this potential. Cost-effective reforestation from the Griscom and colleagues 2020 paper summarised country-level mitigation potential values from Busch and colleagues 2019 https://doi.org/10.1038/s41558-019-0485-x 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 to ours, that provide comparable estimates of natural regeneration and its C sequestration potential.

Regardless, we have now scaled the values reported in this section to match our area of potential for natural regeneration (by deriving a per ha per year value from Griscom and colleagues 2017, and 2020 and multiplying it by our 215 M ha), and added discussion on why values are difficult to compare across methodologies L301 (edits and additions in **bold**):

"Previous studies[11,73,74] used general potential areas for reforestation (i.e., Forest Landscape Restoration Opportunities[75]) 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 by[73,74]). Griscom and colleagues 2017[73] estimated the maximum additional global reforestation mitigation potential to be **1.03 Gt C yr^{-1} over 215 Mha (scaled to our area with potential for natural regeneration based on Table S9[73]**). To determine this area, Griscom and colleagues 2017[73] 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 and colleagues 2020[74] update provided an estimated cost-effective potential of reforestation for tropical regions of **0.28 Gt C yr^{-1} (based on values from Busch and colleagues 2019[76])**. Similar to Griscom and colleagues 2017 and 2020, our spatially explicit estimates equate to 0.8 Gt C yr^{-1}. However, we note that **each of these analyses are quite distinct, and key differences in our analyses make direct comparison to previously published estimates convoluted. For example, cost-effective reforestation from Griscom and colleagues 2020[74] summarised country-level**

mitigation potential values from Busch and colleagues 2019[76], 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 (i.e., informed by data on actual forest regeneration) to provide directly comparable estimates of natural regeneration or C sequestration potential. Given that our model (and thus estimates of C sequestration potential) accounts for the heterogeneity of potential for natural regeneration across nations, they are an advance compared to estimates derived from flat sequestration rates."

Comment 18

L361: 'areas that were already classed as tree cover', when is this referring to? 2000? Please add.

Response 18

We have added "for the year 2000" (L414).

Comment 19

L395: 'the potential for natural regeneration in the future', until when? The map presented here uses current day climate data so won't be valid far in the future under different climate regimes, please add a timeframe over which these predictions are valid.

Response 19

It is correct that our predictions are based on present day data, we have therefore changed, "the potential for natural regeneration in the future" to "in the near future (2030), assuming that overall conditions from 2000-2016 apply to future scenarios." (L446). Our choice to develop predictions under present day conditions was to help support action today and in the near future.

Comment 20

L463-467: Think there is some formatting problem here which is making these line hard to understand.

Response 20

Hopefully this issue has not been carried over to the revised document. It seems to be an artifact of the journals submission system.

Comment 21

L479: The 1km forest density estimate is unclear. 1km around what? Is this a buffer around a predicted NR area or a moving window across all pixels? Is this done at the pixel level as with cropland area? Is it done for every pixel to create a 1km forest density layer or is it just done for NR areas/pixels?

Response 21

The reviewer is correct, this is unclear, but indeed followed the same methodology used for cropland area. We have therefore changed the original phrasing (L526):

"forest density in a 1-km area and the distance of each cell to the nearest forest patch"

to

"mean forest area within a 1 km circular buffer for each pixel within the study region using the focal statistics tool in ArcMap 10.8[100]."

Comment 22

L526: add resolution of NPP product.

Response 22

Done.

Comment 23

Figure 2: Interesting way to display but the scale is hard to see particularly for neotropics where there is only values for Brazil with a broken bar. At minimum include a value on the Colombia bar (as with the 2 other regions), but maybe include values for top 5 countries for each country.

Response 23

We have now added values for the top five countries for each region.

Comment 24

Further, why is the area available for 'all restoration' included here? 'Restoration in general' (L144) is not defined in text so its unclear what these other restoration options are. If what are only interested in natural regeneration, this scaling makes it really hard to see the actual contribution of NR for each country as the green bars are so small and dominated by the yellow which is somewhat redundant here. Maybe consider just showing the green area available for NR which is the focus of this study.

Response 24

'Restoration in general' refers to area available for restoration, or our study region. We agree this is not clear, and have therefore changed "restoration in general" to "area available for restoration" in the legend of Figure 2.

Comment 25

Figure 3: panel A/B, the scale for the potential for natural regeneration which runs from -2.5 to 0.2 is not that intuitive and somewhat confusing to understand as a reader. The figure needs to be understandable without refereeing to the text and currently this unit of measurement is unclear. Suggest either improving explanation of this is the caption, or using another unit of measurement to present the data

Response 25

We agree this is not clear and have therefore added an explanation to the legend of Figure 3, "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-axis of partial dependence plots do not typically represent the true value of the prediction because they visualise 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."

Comment 26

APPENDIX 1 comments below:

L13/L15: Use of 'the former' & 'the latter' is confusing, state what you are referring to specifically for clarity

Response 26

Former has now been replaced with "regeneration points were randomly" and latter has been replaced with "non-regeneration points".

Comment 27

L18: refers to savannah biomes as a 'target biome' is this correct? The methods just refer to analysis including the dry, moist and coniferous forest biomes. Please clarify

Response 27

The Reviewer is correct, we have deleted "and savannah" accordingly (this was accidently included as an artefact of a previous version of the manuscript).

Comment 28

Figure A1.3: The data for all these variables seems very clustered at the lower (<0.4) and higher (>0.7) probability scores, so the use of a mean predicted probability (red lines) doesn't seem to reflect the actual data. Could this be explained a little more in the caption?

Response 28

We have now added to the figure description, "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."

Referee #2 (Remarks to the Author):

A SUMMARY OF THE KEY RESULTS

This is a novel study of an important subject: how much of the tropics would regenerate naturally to forest if left (pretty much) alone.

The authors use a dataset that shows where regeneration occurred from 2000-2018, along with a set of environmental and anthropogenic correlate variables and machine learning, to predict the likelihood of natural regeneration for every pixel. They then multiply this likelihood map by area to obtain rough estimates of the area available for natural regeneration, and the amount of carbon that could be captured, coming up with a reasonable area number of 215 million hectares (approximately a quarter of the current tropical forest area), and a carbon sequestration potential number of around 23 Pg CO_2 over 30 years (again sounding reasonable).

B ORIGINALITY AND SIGNIFICANCE
I'm not aware of a similar paper in the literature. The paper is clearly novel, important, and timely.

It is an important result and map they produce. Natural regeneration indeed can be very/more effective than tree planting, and of course cheaper - but only at a local scale is it normally known where it is really possible. Global-scale analyses like these will not replace local maps and knowledge: despite the 30m resolution this map is not a tool for local managers, its accuracy will be too poor. But it is a very useful tool for enabling national and international decisions on the importance, potential, and funding of natural regeneration methods. It could for example give sensible numbers to input into climate change mitigation scenarios, or to target funding from international donors. It also answers important scientific questions on this subject: it was hard for a scientist like me to give a good estimate to a policy maker of how much area would just regenerate naturally if left alone.

C Data and methodology

I have various issues with the methodology. The overall maps look fairly reasonable, and the overall estimates seem plausible. But the methods, and in particularly the uncertainty analysis (or lack of it) need work to come up to a sufficiently robust scientific standard.

We thank the reviewer for their positive feedback and thoughtful comments. We have addressed all feedback below accordingly.

Comment 29

1. The output maps are a combination of 30m and 1km inputs. This leaves the maps a blocky look (see Figure 1), and means that their true resolution is not 1km. The authors are open about this, but it is a limitation that could have been improved on. 1km layers could have been co-krigged to 30m using associated 30m output layers - or even just upsampled using e.g. a bilinear kernal. And where they were not very significant in the model, 1km layers could have been removed from the analysis in preference to higher resolution input layers.

Doing the analysis at 30m not 1km resolution is clearly essential, given the importance of distance to forest patches in the hundred of meter ranges. But this justification for high resolution is diluted by having any 'blurring' 1km layers.

Response 29

The analysis was done at a 30m resolution which we agree is essential. We only created and made the 1km layer available to make global analyses achievable (as loading in all of the tiles at the 30 by 30m resolution is a significant computational challenge). We have now removed mention of the 1km raster from the manuscript, but have left it in the repository https://doi.org/10.5281/zenodo.7428804 now highlighted that it is for "for display purposes only (not used in the Williams et al. analysis)".

Comment 30

2. I am unconvinced by the approach of multiplying relative potential by pixel area to produce actual area. Ignoring differences in pixel sizes (see below), does a 0.09 ha pixel that comes out of the model with a 90% potential really count 90x more to the area count than one with a 1% potential? Might it not be better to count the 90% potential pixl as a full 0.09 ha of restoration, and the 1% potential pixels as 0 ha?

I think counting every pixel with greater than say 50%, and everything below that as 0%, might be better - or at least I'd like to see those numbers.

This is especially true as Random Forest models tend to favour mid values, and do not push to extremes. There will normally be fewer 0s and 100s in your model output than the reality.

Response 30

We believe multiplying the potential for natural regeneration by area of each pixel and adding up is an appropriate way of estimating expected area of natural regeneration (it is the standard method in the field, for example see Mo and colleagues 2023 and Bastin and colleagues 2019). But we also agree with the reviewer that it is important to provide to the reader an alternative method of estimating area. We have now added a sensitivity analysis to Appendix 3 where we have changed all areal-based calculations to reflect the areas of cells that have >50% PNR potential. This expanded areal estimates slightly. We have changed the methods at L650 to read:

"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 215Mha to 263Mha; Appendix 3 - Table A3.2)."

References

Mo, L., Zohner, C.M., Reich, P.B. et al. Integrated global assessment of the natural forest carbon potential. *Nature* **624**, 92–101 (2023). https://doi.org/10.1038/s41586-023-06723-z

Bastin et al. ,The global tree restoration potential. *Science* **365**,76-79 (2019). https://www.science.org/doi/epdf/10.1126/science.aax0848

Comment 31

3. I'm concerned that the area calculations were done using pixels that were considered '30m x 30m', but are in fact variable across the grid as they are not in an equal area projection, but a simple lat/lon projection. This would give greater weight to pixels further from the equator in the analysis, as well as simply leading to incorrect area values. Can the authors confirm that an equal area projection was used to perform the calculations, or if not that area was corrected for in a different way?

Response 31

The values in the manuscript were calculated using an equal-area projection (Mollweide). We have added this to L653: "All calculations were carried out using the Mollweide projection."

Comment 32

4. When comparing the potential for carbon sequestration of these areas to the current biotropical sink, can you be certain that naturally regenerating forests are not already counted to some degree in global figures? I suspect they will be to some degree, making the percentage figures here an overestimate of the impact of leaving all these areas alone compared to Business As Usual.

Response 32

The Reviewer is correct and "natural regrowth removal factors" were accounted for in the global Harris and colleagues 2021 estimates that we refer to in the discussion. However, our potential for natural regeneration model is run only on currently deforested lands whereas the Harris and colleagues 2021 estimates are based on currently forested lands. Therefore, our calculations of C are entirely additional. We have changed L185 of the discussion to read (additions in bold and underlined):

"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 yr^{-1}), **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**[46]."

Comment 33

5. I worry about the input Hansen/Fagan dataset on forest regrowth. This dataset has a very low accuracy, missing the vast majority of regrowth pixels: as the paper states, it detects less than 20% of natural regrowth pixels. While these were pretty accurate when detected (user's accuracy of 85%), the authors do not, in my opinion, sufficiently justify that the patches that were detected are an unbiased set. Further, by definition, the pixels that do not have patches detected in the Hansen/Fagan dataset are, according to the model, not regeneration. They will be treated as such and included as 'no regeneration' - even though in a large number of cases (cover 4x more pixels than the actual detected gains) they are in fact natural regeneration. So even if the pixels successfully picked up are a representative sample, they will be represented on both regrowth and non-regrowth sides of the model, causing confusion and likely underestimating regrowth potential globally. While these were pretty accurate when detected (user's accuracy of 85%), the authors do not, in my opinion, sufficiently justify that the patches that were detected are an unbiased set.

Response 33

It is possible that we are underestimating regrowth potential globally, if the undetected regeneration points are a biased subset of all natural regeneration. But before addressing that issue, we first wish to clarify that our predictions of regeneration potential come from a stratified random sample of the regrowth dataset and the global non-regrowth tropics, respectively. Because natural forest gain is rare globally (making up 1.96% of the Fagan and colleagues 2022 validation data), the number of undetected regeneration points that would be included on the non-regrowth side of the model is a small proportion of the overall non-regrowth sample. They would not be expected to dramatically alter model performance unless they contained significant bias in terms of spatial location or along environmental gradients.

With regards to potential bias, we have now conducted a clustering analysis to show that the patches that were detected are an unbiased set – see Response 2 above. We repeat it for convenience here:

*Additionally, our follow-up nearest-neighbour analyses have indicated that the mean of nearest-neighbour distances for the observed omission error (6.38) is significantly more distant than nearest-neighbour distances for the full regrowth validation dataset (4.26, CI 5/95% of means 3.09/5.75 (10,000 bootstrapped samples using the same number of omission error points)). In other words, our omission errors are more spread out, on average, than by random chance. This indicates that the model was on average more spatially accurate across different regions than would be expected by random chance, ie. omission error was less clustered than by random chance. Further, we found that the mean nearest-neighbour distance for commission error (5.11) is similar to the means of the full regrowth dataset (5.98, CI 5/95% of mean 0.51/9.30 (10,000 random samples using the same number of commission error points)) indicating that there is no clustering present (ie. patterns are random). Finally, regrowth displays very little apparent correlation with a number of spatial environmental variables. Using binomial logistic regressions (with Bon Ferroni error correction for repeated tests) we found that the occurrence of omission errors in the validation data displays very little apparent correlation with a number of spatial environmental variables as most were non-significant (14/15) apart from slope (p = 0.001**). This is to be expected because, as Fagan and colleagues 2022 point out, their product was derived from both optical and microwave data and as such local variation in topography and forest structure affected classification accuracy - their product tended to have lower accuracy in mountainous areas (where microwave returns vary). Similarly, we found that the occurrence of commission errors in the validation data displays very little apparent correlation with the same spatial environmental variables as most were non-significant (14/15) apart distance to urban areas (p = 0.001**). Synthetic aperture radar (a key input to the dataset of Fagan and colleagues 2022) has elevated and variable values in urban areas (Ferro and colleagues 2011; Henderson and Xia 1997; Lee 2001), decreasing the likelihood of predicting forest in these regions. We also note that urban areas are a very small percentage of the humid tropics, and therefore not an issue for the present study. We can therefore conclude that there is minimal detectable spatial bias in the Fagan and colleagues 2022 model's producer's accuracy. In this study, our spatial predictive model includes topography and distance to urban areas as predictors, so it is able to correct for this minimal bias in omission error we observed in the input data.*

Further, the reported low producer's accuracy estimate in Fagan and colleagues 2022 arises from a known issue when reporting area-based producer's accuracy assessments (in fact, the forest gain estimates are highly accurate), which we have elaborated on extensively and addressed in Response 2 (please see above).

Comment 34

In order for the publication to go ahead I think the authors need to show in far more detail, for example in supplementary figures in some key case studies areas, the nature of the undetected natural regrowth patches and explain through a comprehensive analysis how this does not add great uncertainty to their results. The current error analysis predicts very tight confidence intervals on the area of regrowth predicted (none) and narrow confidence intervals on the carbon stock change over 30 years - I do not think they can be justified given the low accuracy of the training data of forest gain they use.

Response 34

We have addressed the issue in Response 2 (please see above) and conducted formal tests of spatial error in the training data, which are now described in Appendix 1 (please see Response 2). The reported low producer's accuracy estimate in Fagan and colleagues 2022 arises from a known issue when reporting area-based producer's accuracy assessments (in fact, the forest gain estimates are highly accurate), which we have elaborated on extensively and addressed in Response 2 (please see above).

Comment 35

6. Are the Hansen et al. 2000 tree cover and 2018 tree cover datasets comparable? The model is trained on the former and applied with the latter, but no analysis is presented to suggest any test has been performed to see if this is the case. As the 2018 dataset is referenced to the Hansen et al. (2013) paper, clearly the methods used have not been scrutinised or put through peer review. The satellites used are clearly different between the two (Landsat 7 vs. Landsats 8/9, and as the authors will know ETM+ and OLI are quite different sensors in their characteristics). Hansen et al. caution in the notes to their website giving out the data about a worry about continuity after 2012. Was this considered?

Response 35

The Reviewer is correct that the sample points used in our analysis were derived from natural regeneration patches identified by Fagan and colleagues 2022 (based on the Hansen and colleagues 2013 reporting period of 2000 - 2012). We note that some of the predictor variables used for future predictions (forest density and distance to forest patch) were made using the 2018 layer available from https://earthenginepartners.appspot.com/science-2013-global-forest/download_v1.6.html.

Although we understand the reviewer's concerns, we used the best data available for this analysis. The 2000-2012 forest gain data have not been updated to 2018, and the forest gain and loss products produced by Hansen and colleagues (2013) are distinct entities, created by separate models. If we were comparing forest loss prior to 2012 with that after 2012, the Reviewer's concerns would be most cogent. However, we feel our reforestation-focused analysis is minimally affected by the changes in the 2011–2018 update (including the use of Landsat 8 OLI data for 2013 onward, the reprocessing of data from 2011 onward in measuring loss, improved training data for calibrating the loss model, improved per sensor quality assessment models to filter input data, and improved input spectral features for building and applying the loss model). While these changes altered the forest cover and loss predictions in 2018, they did not affect the forest gain dataset.

In addition, even if a subtle bias vis-à-vis tree cover and gain were introduced by the Hansen update, we do not necessarily see this as a limitation because the Hansen tree cover data is the best temporally explicit data of tree cover currently available, the most widely used and accepted by the scientific community, and the use of data derived from similar satellite detection techniques is common place (for example see Busch and Amarjargal 2022 and Ward and colleagues 2020).

While we believe comprehensive validation of the Hansen datasets is beyond the scope of our manuscript, we have now made this caution transparent to users by addition to L534 of the methods:

"We caution users that forest-based predictor (2000–2012 dataset) and explanatory variables (2011–2018 dataset) from the Global Forest Watch[12] used to generate the layers used in our analysis may differ in accuracy given that for the 2011–2018 dataset Landsat 8 OLI data was used, loss data was reprocessed, improved training data for calibrating the loss model was used, improved per sensor quality assessment models to filter input data was used, and improved input spectral features for building and applying the loss model was applied. The variety of mapping methods and attribution approaches means that the input datasets may locally lack coherency (e.g., 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. For more information on the differences between datasets please see https://earthenginepartners.appspot.com/science-2013-global-forest/download_v1.6.html"

References

Busch, J., Amarjargal, O. (2022) 100 global bright spots of green growth: Co-occurrence of nighttime light gain and forest gain, 1990–2015. *Global Environmental Change* **75**, 102556. https://doi.org/10.1016/j.gloenvcha.2022.102556

Ward, M., Tulloch, AIT., Radford, JQ., Williams, BA., Reside, AE., Macdonald, SL., Mayfield, HJ., Maron, M., Possingham, HP., Vine, SJ., O'Connor, JS., Massingham, EJ., Greenville, AC., Woinarski, JCZ., Garnett, ST., Lintermans, M., Scheele, BC., Carwardine, J., Nimmo, DG., Lindenmayer, DG., Kooyman, RM., Simmonds, JS., Sonter, LJ., Watson, JEM. (2020) Impact of 2019–2020 mega-fires on Australian fauna habitat. *Nat Ecol Evol* **4**, 1321–1326. https://doi.org/10.1038/s41559-020-1251-1

D. Statistics and uncertainties

Comment 36

As stated in 5. above, I do not think uncertainties in the area estimation are propagated through at all: no uncertainties are given for what are definitely very uncertain numbers. I would like to see these calculated prior to publication.

Response 36

We have now justified that the patches that were detected are an unbiased set through a clustering analysis – see Response 2 and Response 33. Further, the reported low producer's accuracy estimate in Fagan and colleagues 2022 arises from a known issue when reporting area-based producer's accuracy assessments (in fact, the forest gain estimates are highly accurate), which we have elaborated on extensively and addressed in Response 2 (please see above).

Given we are confident in the Fagan and colleagues 2022 estimates, we propagated the most important uncertainty in our analysis (error from the potential for natural regeneration model) through to our area-based estimates. This propagation was done using an independent accuracy assessment of our predicted potential for natural regeneration (see Response 1, above). See L117 – additions in bold:

"We estimate that biophysical conditions can support natural regeneration in tropical forests over 215 **(CI 214.78 – 215.22)** M ha globally (Fig. 21; Neotropics: 98 **(CI 97.80 – 98.20)** M ha; Indomalayan tropics: 90 **(CI 89.82 – 90.18)** M ha; and Afrotropics: 25.5 **(CI 25.47 – 25.53)** M ha) until 2030."

We have added the confusion matrix for the accuracy assessment of the potential for natural regeneration model to Appendix 1 (see Table A1.2 – A1.5), and note that it is subject to the same issue as the Fagan and colleagues 2022 when reporting area-based producer's accuracy assessments and have therefore provided both point-based and area-based assessments (Olofsson and colleagues 2022 – see Response 2). In addition, we have added a sensitivity analysis to Appendix 3 where we have changed all areal-based calculations to reflect the areas of cells that have >50% PNR potential (Response 30).

E. Conclusions

Comment 37

The map of the probability of natural regeneration is useful and probably fairly accurate in pattern (albeit a mix of 1km and 30m resolution making use difficult).

However I am unsure that the area and carbon numbers can be used to draw robust conclusions as currently calculated. At minimum, they have large uncertainty and this should be reported.

Response 37

We have now removed the 1km resolution product from our paper (which was not used for any of the analysis – just available for users who may wish to undertake a global-scale analysis where loading all of the 30m by 30m resolution tiles is infeasible) (Response 29).

All area values now have uncertainty estimates associated with them, propagated through from the model accuracy estimates (Response 36) (see L117 of manuscript), "We estimate that biophysical conditions can support natural regeneration in tropical forests over 215 (CI 214.78 – 215.22) M ha globally (Fig. 21; Neotropics: 98 (CI 97.80 – 98.20) M ha; Indomalayan tropics: 90 (CI 89.82 – 90.18) M ha; and Afrotropics: 25.5 (CI 25.47 – 25.53) M ha) until 2030." Further, we have added a sensitivity analysis to Appendix 3 where we have changed all areal-based calculations to reflect the areas of cells that have >50% PNR potential (Response 30).

All carbon values have uncertainty estimates reported with them (see L181 of manuscript), "could sequester 23.4 Gt of C (CI 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); Indomalayan tropics: 5.42 Gt (CI 4.87-5.96); Afrotropics: 3.1 Gt (CI 2.83-3.37))".

F. Improvements.

Beyond those above, I have some minor points:

Comment 38

1. Different units used - e.g. Gt C02 line 193, Pg C/yr line 300, Pg CO2e/yr line 304. Pg is identical to Gt, but should be used throughout (SI unit). C and CO2 should not be switched between.

Response 38

Agreed. We now use consistent units throughout.

Comment 39

2. Brazil/Indo/China/Mexico/Colombia between them account for 52% of potential regeneration - but what percentage of land between 25 degrees south and north do they cover? If around 52% then this is not surprising.

Response 39

The Reviewer has raised an important point. This finding was not intended to be an interesting empirical finding, but rather to highlight opportunities to take policy action. To ensure that this is clear we have now added to the end of this sentence (L123): "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 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 M ha each (see Appendix 3)."

G/H - no comments.

Referee #3 (Remarks to the Author):

Comment 40

The manuscript titled "The global potential for natural regeneration in deforested tropical regions" aims to determine what is the surface extent around the tropics for deforested regions that has the potential for natural regeneration if allowed. They accomplish this by leveraging a recently released satellite-derived map of regenerated forest patches as was detected by a suite of satellites over the 2000-2012 time period. They use this dataset as the ground-truth of where forest regeneration occurred and then build a Random forest model using these data to map out where the regeneration occurred more generally over the tropical regions during this time period and more importantly to assess what is the potential for regeneration over the 2016-2030 time period. The relevance of this work is trying to determine how much extra carbon could go into natural regrowth of forests and thus being a fairly "easy" attainable goal. Overall, the paper is well written, however, I am not convinced that the results are sufficiently compelling and convincing for publication in Nature. The authors themselves list a large amount of assumptions and counter-arguments in the discussion section. I actually appreciate their honesty as many of the points that they bring up are ones that I thought of as I was reading through the article, but it does highlight that the results are prone to large uncertainties. For example, the assumption that forest regeneration will continue the same way from 2016-2030 as it did from 2000-2016 is an important assumption; other issues include the fact that the dataset that was used to train the RF was still a sub-sample and therefore we are left with questions about its performance over regions where these data don't exit. Other issues exist but overall my hesitation with the paper is not the methods or the writing, I think the authors do a great job. I just don't think it is at the level for Nature. Instead, I think it would be great for Nature Sustainability. And if that is the choice, I would support it being published there practically "as is".

Response 40

We thank the Reviewer for seeing the value in our work. We address one of their concerns about subsampling across regions in a previous comment (page 2 Response 1, this review).

Our work is highly relevant and timely because it provides evidence for planning and decision-making to address Target 2 of the Kunming-Montreal Global Biodiversity Framework, which calls for the restoration of 30% of degraded areas by 2030. It also provides the evidence base needed for components of Target 1 of the Kunming-Montreal Global Biodiversity Framework, in particular around spatial planning and Target 11, which aims to "restore…. nature's contributions to people, including ecosystem functions and services, such as the regulation of air, water and climate, soil health, pollination and reduction of disease risk…for the benefit of all people and nature".

Tropical forest reforestation has recently been identified as one of the pathways with the highest current carbon market activity and mitigation potential, and rated as a high-confidence pathway in an expert elicitation survey (Buma and colleagues 2024). Our analysis directly feeds into resolving the vast uncertainties around Nature Based Climate Solutions that are highlighted by Buma and colleagues 2024.

Further, although there are other models that predict carbon and restoration potential at global to regional scales, all of these models make their estimates using unrealistic assumptions about "potential tree cover" (e.g., Bastin and colleagues 2019, Mo and colleagues 2023, Veldman and colleagues 2019) and do not predict the potential for pantropical natural forest regeneration (it is instead inferred via subtraction from an unrealistic potential tree cover). Our estimate is the first to use actual data on natural regeneration to make a large-scale estimate of natural regeneration potential, a much needed advance. The resulting dataset, methodology, and conclusions are of global relevance and can be used for both direct input into planning processes and for further research into the potential benefits of promoting natural regeneration. Therefore, we believe our article is suitable, and will be of interest to the wide readers of Nature.

References

Bastin, J. F., Finegold, Y., Garcia, C., Mollicone, D., Rezende, M., Routh, D., Zohner, C. M & Crowther, T. W. The global tree restoration potential. *Science*, 365(6448), 76-79 (2019). https://www.science.org/doi/full/10.1126/science.aax0848

Buma, B., Gordon, D.R., Kleisner, K.M., Bartuska, A., Bidlack, A., DeFries, R., Ellis, P., Friedlingstein, P., Metzger, S., Morgan, G., Novick, K., Sanchirico, J.N., Collins, J.R., Eagle, A.J., Fujita, R., Holst, E., Lavallee, J.M., Lubowski, R.N., Melikov, C., Moore, L.A., Oldfield, E.E., Paltseva, J., Raffeld, A.M., Randazzo, N.A., Schneider, C., Uludere Aragon, N and Hamburg, S.P. (2024) Expert review of the science underlying nature-based climate solutions. *Nature Climate Change*, **14**, 402–406. https://doi.org/10.1038/s41558-024-01960-0

Mo, L., Zohner, C.M., Reich, P.B. et al. Integrated global assessment of the natural forest carbon potential. *Nature* **624**, 92–101 (2023). https://doi.org/10.1038/s41586-023-06723-z

Veldman, J. W. et al. The global tree restoration potential. *Science* **366**, 6463 (2019). https://www.science.org/doi/full/10.1126/science.aay7976

Reviewer Reports on the First Revision:

Referees' comments:

Referee #1 (Remarks to the Author):

The authors have addressed all the concerns raised in the review. It is good to see the additional details regarding clustering analysis, the accuracy of the Fagan dataset and the carbon estimation methods. This has addressed all concerns raised and improves clarity of the methods used.
The changes to the discussion in line with comments are also welcome and now make a much stronger argument for the use of natural regeneration as a climate mitigation strategy.
My only outstanding concern is related to figure 2. the additional values are welcome but I note 2 things. 1) the units of the values are not in the caption, 2) on looking at Table A3.1 I can see the values correspond with total area available for restoration. as this study focuses on natural regeneration would it not make sense to report the values for potential area available for natural regeneration?
Other than that I am happy with the changes made to the manuscript and consider it worthy of publication in Nature

Referee #2 (Remarks to the Author):

The authors have done a good job with the responses, and solved many of my concerns. I retain one primary worry though, regarding the resolution of the input layers not matching the resolution of the analysis. Details below:

Comment 29. I do not think the authors understood my point here. I was not worried about the 1km summaries product - I agree that could be useful for global analysis.

However, layers coarser than 30m resolution are included in the layers used for prediction, and as a result large square blocks are clearly visible in the output maps. E.g. See clear square artefacts in A, B in Figure 1. I think these come from the soil map (1km resolution) and ESA CCI Landcover (300m resolution). These are not realistic.

Ideally only 30m inputs would be used. e.g. by using a 30m (likely coarsened 10m, e.g. ESA's new 10m landcover map?) in order to remove mixing resolutions in the input layers.

Otherwise the outputs can be provided at a 30m pixel spacing, but are clearly a hybrid and really of a much coarser resolution. This will seriously undermine the accuracy of the outputs.

Comment 30. Thanks for doing the sensitivity analysis - I'm happy now with the analysis in the paper (though I probably trust the slightly higher numbers in Appendix 3 more).

Comment 31. Excellent - that's fine.

Comment 32. Fine

Comment 33. This is well argued: I'm convinced.

Comment 34. Okay - yes I agree this is unlikely to be a major issue.

Fine with the rest of the responses, except as regards the resolution above.

Author Rebuttals to First Revision:

RESPONSE TO REVIEWERS COMMENTS

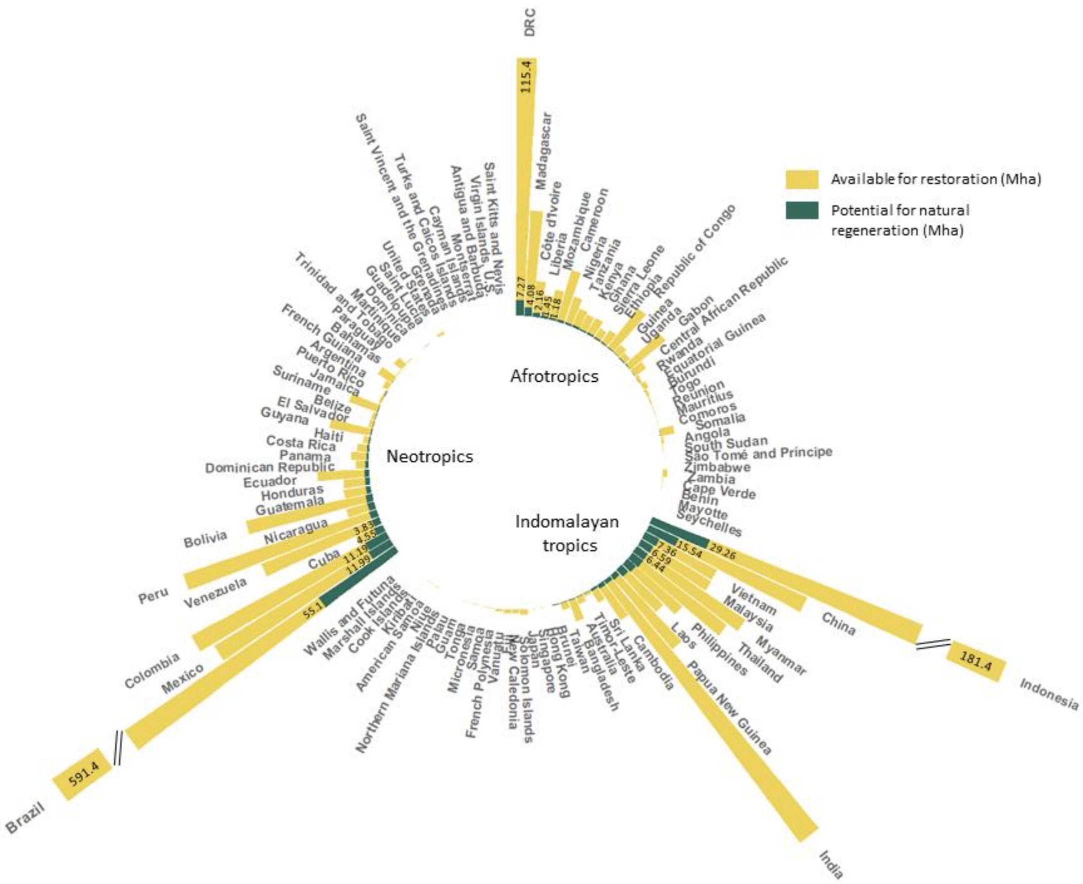
Overall response: We thank the Reviewers for their additional comments. We agree with their comments and have responded accordingly below – and have now justified why undertaking reanalysis at a finer resolution would reduce the accuracy of our product. All adjustments have greatly improved the manuscript.

Referees' comments:

Referee #1 (Remarks to the Author):

My only outstanding concern is related to figure 2. the additional values are welcome but I note 2 things. 1) the units of the values are not in the caption, 2) on looking at Table A3.1 I can see the values correspond with total area available for restoration. as this study focuses on natural regeneration would it not make sense to report the values for potential area available for natural regeneration?

We thank the reviewer for the suggestion. We have changed it so that the values in the plot represent the potential for natural regeneration (the green bars), and the area available for forest restoration (in yellow) is reported only for the countries with largest areas available for restoration in each biogeographic realm. The units are reported in the figure caption "(in millions of ha)" and we have also added these to the figure legend "Mha".



We have also added further detail to the figure caption (additions in bold):

"Figure 2. Summary of the area of land **that can support** natural forest regeneration (PNR) **shown in green**, with total land area available for **forest** restoration **shown in yellow** (in millions of ha) **for countries located at least in part within humid tropical & subtropical forest biomes (tropical and subtropical dry broadleaf forests, tropical and subtropical moist broadleaf forests, and tropical and subtropical coniferous forests)** (see Table A3.1 for full table)."

Referee #2 (Remarks to the Author):

The authors have done a good job with the responses, and solved many of my concerns. I retain one primary worry though, regarding the resolution of the input layers not matching the resolution of the analysis. Details below:

Comment 29. I do not think the authors understood my point here. I was not worried about the 1km summaries product - I agree that could be useful for global analysis.

However, layers coarser than 30m resolution are included in the layers used for prediction, and as a result large square blocks are clearly visible in the output maps. E.g. See clear square artefacts in A, B in Figure 1. I think these come from the soil map (1km resolution) and ESA CCI Landcover (300m resolution). These are not realistic.

Ideally only 30m inputs would be used. e.g. by using a 30m (likely coarsened 10m, e.g. ESA's new 10m landcover map?) in order to remove mixing resolutions in the input layers.

Otherwise the outputs can be provided at a 30m pixel spacing, but are clearly a hybrid and really of a much coarser resolution. This will seriously undermine the accuracy of the outputs.

While we agree that higher resolution data would be preferable, it simply does not exist at the global scale for many of our predictors - bioclimatic variables and soil, for instance. Because bioclimatic variables, soil organic carbon density, and soil pH were found to be significant variables in our models for predicting the potential of forest regeneration, we believe that excluding this data would decrease the accuracy of our model. Similarly, the ESA land cover data is only available at a 300 m resolution in 2015, during the time period of the natural regeneration patches (2000-2016; Fagan et al. 2022). We expand on these points in more detail below, and raise additional ones.

Land cover

Our patches of natural regeneration (delineated by Fagan and colleagues 2022) covers the time period 2000-2016. The land cover map we use and its derivatives are from the

2015 ESA CCI dataset. The reviewer is correct that there are ESA maps available at a 10m resolution for the years 2020 and 2021; however, despite their higher resolution, the 2015 dataset is more appropriate as a predictor variable for the natural regeneration patches as it better aligns with the time period of the natural regeneration data (2000-2016). Given the rapid rate of agricultural expansion across the tropics (Pendrill et al. 2022), temporal alignment is more important than assumed spatial improvement in accuracy.

Additionally, we disagree that mixing resolutions has degraded our spatial product accuracy, as our current accuracy assessment show the product is valid and robust. We note our validation points are, on average, 1 km apart, well beyond the scale of the ESA CCI land cover data in question (300 m).

Soil grids and bioclimatic variables

While we agree that using soil and bioclimatic data at a 30 m resolution would be preferable, unfortunately such data do not exist at the global scale. Since soil organic carbon density, soil pH, and bioclimatic variables were found to be significant (see Figure A1.2) we believe that including these data increases the accuracy of our model rather than decreases it, even when considering the slightly courser resolution. To acknowledge the potential importance of local finer scale data, we have added the following guidance to the Discussion at L225:

"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[55]."

Additionally, we discovered a small error in the text – the SoilGrids dataset is actually at a 250 m spatial resolution rather than a 1 km resolution. We have corrected this in the text.

Mixing resolutions

Our methodological process was to identify the predictor variables most likely to explain the potential for natural regeneration (the rationale for each included under its heading in the "Predictor variables" section), then to find and use the best and most reputable dataset available in the tropics to represent this predictor. While these datasets may not all be at a 30 m resolution, they are the best data available. This methodological process represents best practice, and has been used in countless studies prior to ours.

The reality of global analyses, and science in general, is to use the best data available - which is rarely perfect – but as increasingly high resolution and improved data becomes available, these exercises can be updated over time. For example, the first land cover map was at 1° latitude by 1° longitude (approx. 111 kilometres) resolution (Matthews 1983). In our manuscript we present a first product predicting potential for natural regeneration for the tropics, which we (or others) will no doubt update and improve upon in the future.

We note that mixing resolutions of input data to produce a global product for analysis is commonplace (Mo et al. 2023; Johnson et al. 2023; Bastin et al. 2019; Neugarten et al. 2024). In these analyses, the final output variable is almost always at a finer resolution than the coarsest input variables. In these mixed resolution models, the output resolution models most closely matches the most important variable(s) in the predictor model, which in our case were tree cover density and distance to forest, both at 30 m (Appendix 1). Three common examples of this practice are ecosystem service models (such as InVest), climate models, and species distribution models.

Ecosystem service models (for example, InVEST (Integrated Valuation of Ecosystem Services and Tradeoffs)) regularly use spatial data inputs at different resolutions, including high-resolution land use/land cover maps, medium-resolution digital elevation models, and coarse-resolution soil maps. These diverse data sources are used to generate spatial maps of ecosystem service provision, showing where services are being supplied and in what quantities (Chaplin-Kramer et al. 2023).

Similarly, climate models represent the components of Earth's climate system using the best available data, which are at different input resolutions. This includes the IPCC's Coupled Model Intercomparison Project (CIMP), inclusive of the most recent phase 6, known as CMIP6 (Eyring et al. 2016). For technical details on input dataset resolutions or "components" of the CMIP6 project see https://confluence.ecmwf.int/display/CKB/CMIP6%3A+Global+climate+projections. The data from the CMIP6 project is being used globally to inform on climate change policy, and answer important scientific questions, for example, informing on the IPCC sixth assessment report (AR6) (IPCC 2021).

In species distribution modelling, input datasets take the form of best available data at multiple resolutions. Records of where a species has been observed are taken, and then information about the environmental conditions – usually at different resolutions - at each occurrence location is collated. This can include climate data (temperature, precipitation), topography (elevation, slope), soil types, and vegetation, for example. The occurrence data and environmental variables are then used to fit the chosen model, the model learns the relationship between species presence and environmental conditions, and the model is applied to the entire study area to predict the potential distribution of the species (Reside et al. 2011; VanDerWal et al. 2012; Xu et al. 2023; Graham et al. 2019). This modelling process is well-established, and a functionally similar approach was used to develop our potential for natural regeneration predictions.

Mixing resolutions in input data: a standard practice

Here we point to a further 20 examples of reputable global analyses from some of the world's top scientists where mixing of resolutions has been used to create a higher-resolution product that was subsequently published in high-impact journals (Nature, Science, PNAS, Nature Communications, Nature Ecology and Evolution, Nature Climate Change, Science Advances):

Example using 49 covariates at resolutions spanning 30 m to 10km to produce a product at 30-arcsec resolution (1 km2 at the equator) resolution:

- Mo, L. et al. Integrated global assessment of the natural forest carbon potential. *Nature*, **624**, 92–101 (2023). https://doi.org/10.1038/s41586-023-06723-z

Example using multiple datasets for developing 14 models of nature's contributions to people (otherwise known as ecosystem services) spanning 30 m to 10km to produce products at a 300m resolution:

- Chaplin-Kramer, R. et al. (2023) Mapping the planet's critical natural assets. *Nature Ecology and Evolution*, **7**, 51–61. https://doi.org/10.1038/s41559-022-01934-5

Other examples, ordered by publication date:

- Huang, Y. et al. (2024) Size, distribution, and vulnerability of the global soil inorganic carbon. *Science*, **384**, 233-239. https://www.science.org/doi/10.1126/science.adi7918
- Neugarten, R A. et al. (2024) Mapping the planet's critical areas for biodiversity and nature's contributions to people. *Nature Communications*, **15**, 261. https://doi.org/10.1038/s41467-023-43832-9
- Qin, J. et al. (2024) Global energy use and carbon emissions from irrigated agriculture. *Nature Communications,* **15**, 3084. https://doi.org/10.1038/s41467-024-47383-5
- Johnson, J A. et al. (2023) Investing in nature can improve equity and economic returns. *Proceedings of the National Academy of Sciences,* **120**, 27, e2220401120. https://www.pnas.org/doi/abs/10.1073/pnas.2220401120
- Roebroek, C T J. et al. (2023) Releasing global forests from human management: How much more carbon could be stored? *Science*, **380**, 749-753. https://www.science.org/doi/10.1126/science.add5878
- Schug, F. et al. (2023) The global wildland–urban interface. *Nature,* **621**, 94–99. https://doi.org/10.1038/s41586-023-06320-0
- Harris, N L et al. (2021) Global maps of twenty-first century forest carbon fluxes. *Nature Climate Change*, **11**, 234–240. https://doi.org/10.1038/s41558-020-00976-6
- Cook-Patton, S C. et al. (2020) Mapping carbon accumulation potential from global natural forest regrowth. *Nature*, **585**, 545–550. https://doi.org/10.1038/s41586-020-2686-x
- Grantham, H S. et al. (2020) Anthropogenic modification of forests means only 40% of remaining forests have high ecosystem integrity. *Nature Communications*, **11**, 5978. https://doi.org/10.1038/s41467-020-19493-3

- Mokany, K. et al. (2020) Reconciling global priorities for conserving biodiversity habitat. *PNAS*, **117**, 9906-9911. https://www.pnas.org/doi/full/10.1073/pnas.1918373117
- Strassburg, B B N. et al. (2020) Global priority areas for ecosystem restoration. *Nature*, **586**, 724–729. https://doi.org/10.1038/s41586-020-2784-9
- Bastin, J F. et al. (2019) The global tree restoration potential. *Science*, **365**, 76-79. https://www.science.org/doi/10.1126/science.aax0848
- Brancalion, P H S. et al. (2019) Global restoration opportunities in tropical rainforest landscapes. *Science Advances*, **5**, eaav3223. https://www.science.org/doi/10.1126/sciadv.aav3223
- Manoli, G., Fatichi, S., Schläpfer, M. *et al.* Magnitude of urban heat islands largely explained by climate and population. *Nature*, **573**, 55–60 (2019). https://doi.org/10.1038/s41586-019-1512-9
- Venter, O. et al. (2016) Sixteen years of change in the global terrestrial human footprint and implications for biodiversity conservation. *Nature Communications*, **7**, 12558. https://doi.org/10.1038/ncomms12558
- Crowther, T. et al. (2015) Mapping tree density at a global scale. *Nature*, **525**, 201–205. https://doi.org/10.1038/nature14967
- Halpern, B. et al. (2015) Spatial and temporal changes in cumulative human impacts on the world's ocean. *Nature Communications*, **6**, 7615. https://doi.org/10.1038/ncomms8615
- Halpern, B S. et al. (2008) A Global Map of Human Impact on Marine Ecosystems. *Science*, **319**, 948-952. https://www.science.org/doi/10.1126/science.1149345

Visual artefacts in Figure 1

To conclude, although we believe we have made the best choices to improve the accuracy of our model, have been explicit about our input data choices, and that there is no further analysis we could do to improve model accuracy, we certainly agree with the reviewer that these details need to be made explicit to the reader – and highlight the fact that these artefacts are visually present in our product as seen in Figure 1 (although not an issue given how we assessed accuracy (given our 1 km assessment spacing on average)). We have therefore added to L351 of the manuscript:

"We 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 (see Appendix 2 for list of data and sources used in the analysis). 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 Figure 1). This emphasises the need for our product to be improved over time, and for implementation activities that take into account the local context."

References

Bastin, Jean-Francois, Yelena Finegold, Claude Garcia, Danilo Mollicone, Marcelo Rezende, Devin Routh, Constantin M. Zohner, and Thomas W. Crowther. 2019. "The Global Tree Restoration Potential." *Science (New York, N.Y.)* 365 (6448): 76–79.

Chaplin-Kramer, Rebecca, Rachel A. Neugarten, Richard P. Sharp, Pamela M. Collins, Stephen Polasky, David Hole, Richard Schuster, et al. 2023. "Mapping the Planet's Critical Natural Assets." *Nature Ecology & Evolution* 7 (1): 51–61.

Eyring, Veronika, Sandrine Bony, Gerald A. Meehl, Catherine A. Senior, Bjorn Stevens, Ronald J. Stouffer, and Karl E. Taylor. 2016. "Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) Experimental Design and Organization." *Geoscientific Model Development*. https://doi.org/10.5194/gmd-9-1937-2016.

Graham, Erin M., April E. Reside, Ian Atkinson, Daniel Baird, Lauren Hodgson, Cassandra S. James, and Jeremy J. VanDerWal. 2019. "Climate Change and Biodiversity in Australia: A Systematic Modelling Approach to Nationwide Species Distributions." *Australasian Journal of Environmental Management* 26 (2): 112–23.

IPCC. 2021. "Sixth Assessment Report." https://www.ipcc.ch/assessment-report/ar6/.

Johnson, Justin Andrew, Uris Lantz Baldos, Erwin Corong, Thomas Hertel, Stephen Polasky, Raffaello Cervigni, Toby Roxburgh, Giovanni Ruta, Colette Salemi, and Sumil Thakrar. 2023. "Investing in Nature Can Improve Equity and Economic Returns." *Proceedings of the National Academy of Sciences of the United States of America* 120 (27): e2220401120.

Matthews, Elaine. 1983. "Global Vegetation and Land Use: New High-Resolution Data Bases for Climate Studies." *Journal of Climate and Applied Meteorology* 22 (3): 474–87.

Mo, Lidong, Constantin M. Zohner, Peter B. Reich, Jingjing Liang, Sergio de Miguel, Gert-Jan Nabuurs, Susanne S. Renner, et al. 2023. "Integrated Global Assessment of the Natural Forest Carbon Potential." *Nature*, November. https://doi.org/10.1038/s41586-023-06723-z.

Neugarten, Rachel A., Rebecca Chaplin-Kramer, Richard P. Sharp, Richard Schuster, Matthew Strimas-Mackey, Patrick R. Roehrdanz, Mark Mulligan, et al. 2024. "Mapping the Planet's Critical Areas for Biodiversity and Nature's Contributions to People." *Nature Communications* 15 (1): 261.

Pendrill, Florence, Toby A. Gardner, Patrick Meyfroidt, U. Martin Persson, Justin Adams, Tasso Azevedo, Mairon G. Bastos Lima, et al. 2022. "Disentangling the Numbers behind Agriculture-Driven Tropical Deforestation." *Science (New York, N.Y.)* 377 (6611): eabm9267.

Reside, April E., Ian Watson, Jeremy VanDerWal, and Alex S. Kutt. 2011. "Incorporating Low-Resolution Historic Species Location Data Decreases Performance of Distribution Models." *Ecological Modelling* 222 (18): 3444–48.

VanDerWal, Jeremy, Helen T. Murphy, Alex S. Kutt, Genevieve C. Perkins, Brooke L. Bateman, Justin J. Perry, and April E. Reside. 2012. "Focus on Poleward Shifts in Species' Distribution Underestimates the Fingerprint of Climate Change." *Nature Climate Change* 3 (3): 239–43.

Xu, Wu-Bing, Shane A. Blowes, Viviana Brambilla, Cher F. Y. Chow, Ada Fontrodona-Eslava, Inês S. Martins, Daniel McGlinn, et al. 2023. "Regional Occupancy Increases for Widespread Species but Decreases for Narrowly Distributed Species in Metacommunity Time Series." *Nature Communications* 14 (1): 1463.

Reviewer Reports on the Second Revision:

Referees' comments:

Referee #2 (Remarks to the Author):

I'm happy the authors have addressed my concerns, and for the manuscript to be published in this form.