nature portfolio

Peer Review File

Mixed effectiveness of global protected areas in resisting habitat loss

Reviewers' Comments:

Reviewer #1:
Remarks to the Author:
This is an impressive global study on the effectiveness of protected areas, which I very much enjoyed reading. The authors present a large dataset and draw some important conclusions from their analysis. Having said that, I have some concerns and questions about the study. These include questions about the novelty of the paper's findings, methodological questions as well as potential extensions.

First, it is unclear to me what the novelty is of the analysis. Please don't get me wrong this is an important paper but given where the authors are aiming to publish their study, the novelty angle is (in my opinion) key. There are other global studies on protected area effectiveness using counterfactual methods (see Geldmann et al. in PNAS) that come to very similar conclusions with a very similar, albeit smaller, dataset. The authors try to justify this by highlighting need for a comprehensive, multidimensional, and systematic assessment. What this means, isn't clear to me and their conclusions seem to incrementally improve our understanding rather than drastically change that.

Second, while I am familiar with counterfactual approaches I would have liked much more information about the authors approach. Specifically, I would like to see much more detail about how they conducted their study, diagnostics and robustness checks (how can we be sure that your study is not data-dependent). This all seem to be missing. Details that I would have like to see, for example, are: do they generate a synthetic control for each individual PA iteratively? What variables and information is used to build each individual synthetic control, are these data available globally, how do we know that it's a good counterfactual. What are the source of hidden bias? This might all require some additional work.

Third, not all PAs are created equal and there is substantial variation in how PAs are managed and by whom (e.g., Brazil's Strict and Sustainable Use PAs and Indigenous Territories). There's some differentiation (strict v non-strict), but at such a scale, the analysis fails to address a key component of heterogeneity. Here, I believe the authors could do much more to disentangle what works where and why. This understanding is really critical for the 30x30 agenda. Again, this might all require some additional thinking and work but might also help with the first point about novelty.

In summary, the paper is interesting and has lots of potential but falls a little short on novelty, contribution and ambition. I hope these comments are useful.

Reviewer #2:
Remarks to the Author:
Comments on "Mixed Effectiveness of Global Protected Areas in Resisting Habitat Loss"

The paper assesses the performance of over 160,000 protected areas in resisting habitat loss globally. While I found the findings compelling, it is important to understand the main drivers behind protected areas, which can help with policy decisions. For instance, the study finds that larger and stricter protected areas tend to have lower habitat loss. Several questions arise. Maybe it is because more important/diverse protected areas, which tend to have more resources or are established earlier, are more effective in reducing habitat loss? Or maybe it is the distance to cities/urban areas? (My prior is that older protected areas are located further away from cities, while newer protected areas are closer to cities, thus more likely to be deforested). Or maybe it is simply governance? Most of the positive results (for instance, in Table 4, are concentrated in Europe). If it is governance, what is happening with Africa? What are they doing correctly that can be extrapolated to other developing regions?

My second key comment is: How to converge individual/country level results of the impact of protected areas with the results found in Fig. 4. For instance, several studies in Latin America have found that protected areas are effective at reducing deforestation [e.g., Costa Rica (Andam et al. 2008; Pfaff et al. 2009), México (Blackman et al. 2011), Peru (Miranda et al. 2013), Bolivia (Hanauer y Canavire 2013), etc.] while the current results suggest the opposite (either significantly negative or non-significant). It is important to understand the approach and results and find channels for the differences.

I like the lag effects. Can we generalize more between stricter protected areas or continents? This can be an important factor to highlight strongly in the paper with important policy recommendations for policymakers.

In terms of methodological approach, as I understand, using the Hansen Global Forest Change dataset is not free of bias (e.g., Alix Garcia and Millimet, 2023). The algorithm tends to reduce deforestation in small areas. Does it have an impact in smaller protected areas? It would be great to see whether the effects are consistent when dividing the sample between small vs. large/medium PAs. Also, regarding the Propensity score matching approach, it is unclear which variables were used to select the correct matches, and more discussion is needed.

Reviewer #1 (Remarks to the Author):

This is an impressive global study on the effectiveness of protected areas, which I very much enjoyed reading. The authors present a large dataset and draw some important conclusions from their analysis. Having said that, I have some concerns and questions about the study. These include questions about the novelty of the paper's findings, methodological questions as well as potential extensions.

Response:

We thank the reviewer for their positive and constructive comments. The issues raised have strengthened the manuscript considerably and we believe we have addressed them all concerns and questions. Below are our responses to your concerns.

First, it is unclear to me what the novelty is of the analysis. Please don't get me wrong this is an important paper but given where the authors are aiming to publish their study, the novelty angle is (in my opinion) key. There are other global studies on protected area effectiveness using counterfactual methods (see Geldmann et al. in PNAS) that come to very similar conclusions with a very similar, albeit smaller, dataset. The authors try to justify this by highlighting need for a comprehensive, multidimensional, and systematic assessment. What this means, isn't clear to me and their conclusions seem to incrementally improve our understanding rather than drastically change that.

Response:

We understand the concern raised by the reviewer regarding the novelty aspects of our research and we do cite papers such as Geldmann et al. in PNAS and now explain the novelty of our research compared to these studies in the introduction section of our revised manuscript (lines 59-87 and lines 103-106).

There are four aspects to this study that we believe make the science novel while building on these previous efforts:

(1) Unlike previous studies that focuses on human pressures and human footprints (Geldmann *et al*. in PNAS and Jones *et al*. in Science), we pay greater attention to long time-series data and fine-grained habitat loss. In our view, the dynamic changes in habitats we show are more directly related to the conservation performance and should a core component of the assessment of protected area effectiveness. Compared to the 77 square kilometer and 1 square kilometer data mentioned above (Geldmann *et al*. in PNAS and Jones *et al*. in Science), thanks to much higher-resolution data (30m), we can assess the performance of protected areas smaller than 1 square kilometer. The fine-resolution performances have been significantly overlooked in previous studies.

(2) Previous research tended to focus on differences inside and outside protected areas, but we believe this is one-sided. We argue that, in addition to considering the spatial effectiveness of protected areas, it is also important to examine the dynamic changes (or time effectiveness) in protection performance before and after the establishment of protected areas. Utilizing causal inference methods enables a more thorough assessment of the actual impact resulting from the establishment of protected areas.

(3) We applied the integrated framework of metacoupling (human-nature interactions within a PA as well as between the PA and other places such as cities near and far) to understand factors influencing the PA effectiveness, such as governance and human footprints within the PA and distances from cities. Previous studies did not report this metacoupling framework for assessing the effectiveness of PAs.

(4) We found a time lag in the effectiveness of PAs, which was not reported in previous studies. The observed lag effect in global PAs' effectiveness in curbing habitat loss provides several policy implications. These findings also suggest that proactive and early establishment of PAs is essential to anticipate and mitigate habitat loss before it becomes critical.

When compared to existing single land cover and human pressure studies, we believe that our research represents the first comprehensive, fine scale evaluation of habitat changes in the largest number of protected areas over a long period of time. Furthermore, we applied the integrated framework of metacoupling (e.g., human-nature interactions within a PA as well as between the PA and other places such as cities near and far) to understand factors influencing the PA effectiveness, such as governance and human footprints within the PA and distances from cities. By doing the science in this way, our findings have uncovered new patterns that was unexpected and differ from previous published research. For instance, we have uncovered that protected areas exhibit varying effects on different types of habitat loss. They prove effective in restraining urban expansion but are not as effective in limiting cropland expansion. Additionally, we have identified distinct patterns of effectiveness both within and outside protected areas across different spatial extents. Notably, there is a rapid and significant increase in forest loss and pastureland expansion in the immediate vicinity of protected areas. These findings were not previously reported in published research, but they are of paramount importance for improving the effectiveness of protected areas, especially considering the bolder '30 by 30' agenda in the Global Biodiversity Framework.

Following your third comment, in order to highlight the novelty of this work, we have added a section in the main text. We have endeavored to explore further heterogeneity in the effectiveness of PAs based on the integrated framework of metacoupling, with the aim of informing the global 30×30 agenda. We have focused on five aspects of heterogeneity, including the establishment time of the PA, governance type (indigenous and community vs. government), the importance of biodiversity, the distance from cities, and human footprint. We also have revealed the reasons for the heterogeneity in different PAs based on their performance. We hope that these analyses will assist in enhancing the novelty of our work.

In the section "Differentiated effectiveness of protected areas and its implications for 30×30 agenda", we added the following additional thinking and work (in P11-13):

Differentiated effectiveness of protected areas and its implications for 30×30 agenda

Nations who have committed to the Kunming-Montreal Global Biodiversity Framework aim to designate 30% of Earth's land and water areas for protection by 2030. Given this ambition, an urgent need is to identify where setting up PAs will be most effective. The current effectiveness assessment of PAs may provide a reference for this purpose. We therefore examined the variation of the effectiveness of existing

PAs against five criterial: the establishment time of the PA, governance type (indigenous and community vs. government), the importance of biodiversity, the distance from cities, and human footprint (Fig. 5). We applied the integrated framework of metacoupling (human-nature interactions within a PA as well as between the PA and other places such as cities near and far) to understand factors influencing the PA effectiveness. We used proportion changes in combined total habitat loss in PAs as example of integrated habitat alteration.

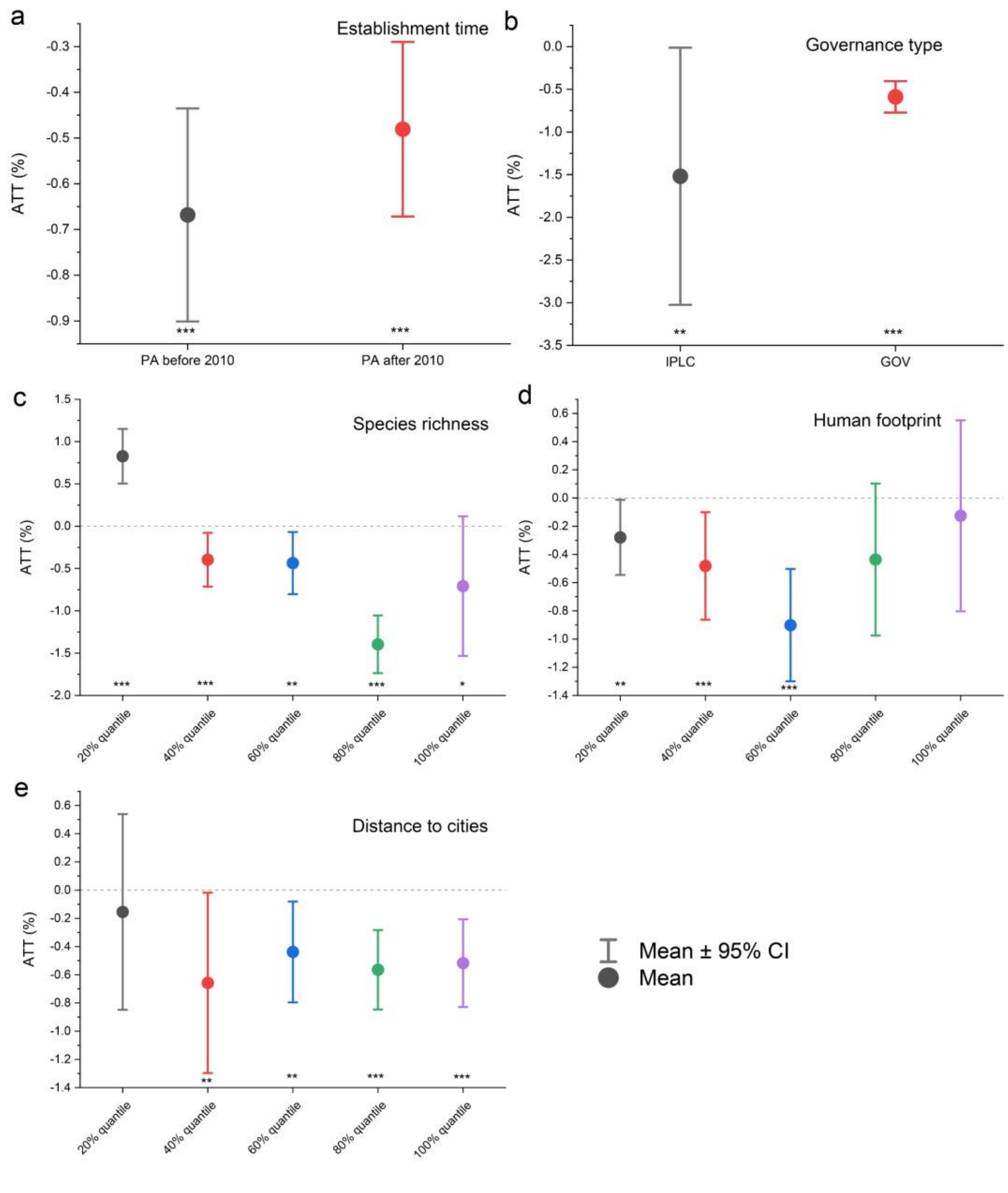

Fig. 5 | Differentiated effects of designated PAs on retarding habitat loss. ATT is the "average treatment effect on the treated" of PAs, namely the estimated annual average effect of established PAs on retarding the proportion of total habitat loss in PAs between 2003 and 2019. Error bars indicate 95% confidence intervals. Asterisks represent significance levels: * $P< 0.1$, ** $P< 0.05$, *** $P< 0.01$. The absence of an asterisk indicates statistical non-significance. If AAT is significantly negative, it indicates that PA establishment has a significant effect on reducing habitat loss. Here we consider the multiple heterogeneity of PAs, including establishment time of PAs **(a)**, governance type: indigenous peoples and local communities' lands (IPLC) and government (GOV) **(b)**, species richness **(c)**, human footprint **(d)**, distance to cities **(e)**. To further examine the heterogeneity of PAs, we divided the sample PAs into quintiles (e.g., 20% quantile in (**c**) represents PAs with the lowest 20% species richness).

We found that earlier established PAs are more effective in curbing the overall loss proportion of habitat over the time period assessed (Fig. 5a). The ATT values for PAs established before and after 2010 were -0.67% and -0.48%, respectively. This implies that to achieve the 30×30 agenda, it is crucial to act promptly, and the longer the establishment time, the more likely the PA will become effective. This is potentially because only after the establishment of PAs can there be corresponding resource support, and the effectiveness of PAs is not immediately apparent but requires a long-term process. Moreover, it is essential to recognize that newly established PAs tend to become "paper parks", as nations try and achieve numbers of PAs in terms of area and not outcomes. Moreover, the government type of PAs is an important factor that affects their effectiveness. Indigenous peoples and local communities' lands are more effective in controlling habitat loss in PAs compared to government and other entities (ATT were -1.52% vs. -0.59%, Fig. 5b). This is potentially because of their deep understanding of the local environment and traditional knowledge and practices for sustainable resource use and management[39]. Their participation in the management of PAs can help to improve the effectiveness of conservation efforts and promote the sustainable use of natural resources[39].

We also found PAs currently distributed in regions of high biodiversity richness and importance have stronger effectiveness in preventing habitat loss (Fig. 5c). One potential reason for this is that these areas with high species richness are often far from human activity areas, and the probability of human disturbance is therefore relatively low compared to other regions. This has been further supported by human footprint data, which shows that PAs distributed in the top 40% of human footprint areas do not significantly inhibit habitat loss (Fig. 5d). In addition, the distance of PAs from cities can provide a reference for the establishment of future PAs. PAs that are farther from cities were more effective in significantly curbing habitat loss, while the effectiveness of PAs near cities is not significant (Fig. 5e). This is potentially because areas near cities are more likely to be affected by intense human activities, such as infrastructure construction, urbanization, industrialization, and tourism, which can lead to habitat loss. Therefore, it is important to consider the distance from cities when establishing PAs to maximize their effectiveness in protecting biodiversity and habitats.

Second, while I am familiar with counterfactual approaches I would have liked much more information about the authors approach. Specifically, I would like to see much more detail about how they conducted their study, diagnostics and robustness checks (how can we be sure that your study is not data-dependent). This all seem to be missing. Details that I would have like to see, for example, are: do they generate a synthetic control for each individual PA iteratively? What variables and information is used to build each individual synthetic control, are these data available globally, how do we know that it's a good counterfactual. What are the source of hidden bias? This might all require some additional work.

Response:
We thank the reviewer for their insightful review of our manuscript. We appreciate the familiarity with counterfactual methods the reviewer shows. We have now added details on the methodological approach in the revision.

We note the method we employed in the study is not precisely equivalent to the synthetic control method the reviewer mentioned. In fact, we utilized a newly developed approach known as Synthetic Difference-in-Differences (SDID). SDID differs in some respects from the Synthetic Control (SC) and Difference-in-Differences (DID). SDID draws inspiration from both Difference-in-Differences and Synthetic Control, which brings advantages from both models. Like SC, SDID still works with multiple periods when pre-treatment trends are not parallel. However, unlike SC, SDID estimates unit weights to build a control unit which is only parallel to the treated group (it doesn't have to match its level). From DID, SDID leverages time and unit fixed effects, which helps to explain a lot of the variance in the outcome, which in turn reduces the variance of the SDID estimator. SDID also introduces some new ideas of its own. First, there is an additional penalty in the optimization of the unit weights which makes them more spread out across control units. Second, SDID allows for an intercept (and hence, extrapolation) when building such weights. Third, SDID introduces the use of time weights, which are not present in either DID nor SC. For this reason, we wouldn't say SDID is just merging SC and SDID. It is rather building something new, inspired by these two approaches. More information refers to Arkhangelsky et al. (2021) for theoretical derivations [Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W. & Wager, S. Synthetic Difference-in-Differences. Am. Econ. Rev. 111, 4088-4118 (2021)].

We recognize it is essential to present more details in the Methods section, particularly in terms of how we conducted the study. We have now expanded the methodology section to provide a detailed account of our research design and implementation process.

In line 598-603, we added a more detailed explanation:
"We implemented our analysis in Stata 17, employing the SDID package[61]. As SDID is based on a balanced panel of observations, we first make sure there are no missing values. If there is a missing value, we will delete this observation. In this analysis SDID handles a staggered adoption configuration, given that in the period under study, PA designation occurred in different yearly periods. We implement the SDID estimator using the bootstrap procedure to calculate standard errors (default 50 iterations)."

We also have introduced a model diagnostic for SDID and DID model. In fact, we have reported the details of diagnostics for SDID in Supplementary Information, including the ATT, Std. Err., t value, *P* value, and 95% Conf. Interval. Here, we emphasize why we chose the SDID model, based on the diagnostic information of the DID model. A substantial number of empirical studies more generally seek to estimate causal effects using DID style designs. Here impacts are inferred by comparing treated to control units, where time-invariant level differences between units are permitted as well as general common trends. However, the drawing of causal inferences requires a parallel trend assumption, which states that in the absence of treatment, treated units would have followed parallel paths to untreated units. In many cases, parallel trend test does not pass. To overcome this limitation, SDID seeks to optimally generate a matched control unit which considerably loosens the need for parallel trend assumptions.

In line 604-610, we added a more detailed explanation:

“To ensure the reliability of the model we chose, we followed the recommendation of Arkhangelsky et al.[34] and compared the performance of the SDID and DID models. For DID estimator we estimated the average effect of PA designation time on built-up land (we used the annual proportion and area of built-up land in PAs as example, see Table S8) based on a csdid model[62] (Difference-in-Differences with multiple time periods). The results show that the DID model does not pass the parallel trend test (using command of Pretrend Test in csdid). This means that if the DID model is used to analyze data here, it will be biased.”

To ensure the robustness of our findings and mitigate data-dependency concerns, we have introduced a placebo test to validate the robustness of our findings.

In line 611-615, we added a more detailed explanation:
“To validate the robustness of our findings, we conducted a placebo test. We randomly manipulated the establishment time of the PA to create another treatment group time. Through the placebo test, we found that our results are robust, as the ATT of both models (we also used the annual proportion and area of built-up land in PAs as example) are not statistically significant, which is significantly different from our analysis results, indicating that our results are not random (see Table S9)”.

We hope that these improvements have addressed your concerns, as they enhance the quality and reliability of our study, and provide readers with a more comprehensive understanding of our approach.

Below we focus on answering the reviewer’s specific questions:
(1) Details that I would have like to see, for example, are: do they generate a synthetic control for each individual PA iteratively?

In this study we used SDID model using a staggered adoption design to estimate the ATT rather than generating a synthetic control for each individual PA iteratively. Consider for example the treatment matrix below, consisting of 8 PA units, 2 PA of which (1 and 2) are untreated (protected areas that are not established in the time period we are concerned about (e.g. after 2019 or 2020)), while the other 6 PA are treated, however at varying points.

$$\mathbf{W} = \begin{pmatrix} & t_1 & t_2 & t_3 & t_4 & t_5 & t_6 & t_7 & t_8 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 4 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 5 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 6 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 7 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 8 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \qquad (1)$$

This single staggered treatment matrix $\mathbf{W}$ can be broken down into adoption date specific matrices, $\mathbf{W^1}$, $\mathbf{W^2}$ and $\mathbf{W^3}$, or generically, $\mathbf{W^1}, \ldots, \mathbf{W^A}$, where A indicates the number of distinct adoption dates. Additionally, a row vector **A** consisting of A elements contains these distinct adoption periods. In this specific setting where units first adopt treatment in period t_3 (units 7 and 8), t_4 (units 5 and 6), and t_7 (units 3 and 4), the adoption date vector consists sample of periods 3, 4 and 7.

$$\mathbf{A} = \begin{pmatrix} 3 & 4 & 7 \end{pmatrix} \tag{2}$$

Finally, adoption-specific matrices $\mathbf{W}^1$-$\mathbf{W}^3$ simply consist of pure treated units, and units which adopt in this specific period, as below:

$$\mathbf{W}^1 = \begin{pmatrix} & t_1 & t_2 & t_3 & t_4 & t_5 & t_6 & t_7 & t_8 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 4 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix},$$

$$\mathbf{W}^2 = \begin{pmatrix} & t_1 & t_2 & t_3 & t_4 & t_5 & t_6 & t_7 & t_8 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 5 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 6 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{pmatrix},$$

$$\mathbf{W}^3 = \begin{pmatrix} & t_1 & t_2 & t_3 & t_4 & t_5 & t_6 & t_7 & t_8 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 7 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 8 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

As laid out in Arkhangelsky et al. (2021, Appendix A), the average treatment effect on the treated can then be calculated by applying the SDID estimator to each of these 3 adoption-specific samples, and calculating a weighted average of the three estimators, where weights are assigned based on the relative number of treated units and time periods in each adoption group. Generically, this ATT is calculated based on adoption-specific SDID estimates as:

$$\widehat{ATT} = \sum_{for\ a \in A} \frac{T^a_{post}}{T_{post}} \times \hat{\tau}^{sdid}_a \tag{3}$$

where T_{post} refers to the total number of post-treatment periods observed in treated units. $\hat{\tau}^{sdid}_a$ denoted the average causal effect of exposure.

(2) What variables and information is used to build each individual synthetic control, are these data available globally, how do we know that it's a good counterfactual.

There is no need to use other variables to build each individual synthetic control in the SDID model. Arkhangelsky et al. (2021) introduced a new idea including two weights (unit weights and time weights) to estimate the ATT. Considering a balanced panel with N PA units and T time periods, where the outcome for unit i in period t is denoted by Y_{it}, and exposure to the binary treatment is denoted by $W_{it} \in \{0, 1\}$. Suppose moreover that the first N_{co} (control) PA units are never exposed to the treatment, while the last $N_{tr} = N - N_{co}$ (treated) PA units are exposed after time T_{pre}. Like with synthetic control methods, we start by finding unit weights $\hat{\omega}^{sdid}$ that align pre-exposure trends in the outcome of unexposed units with those for the exposed units, e.g., $\sum_{i=1}^{N_{co}} \hat{\omega}_i^{sdid} Y_{it} \approx N_{tr}^{-1} \sum_{i=N_{co}+1}^{N} Y_{it}$ for all $t = 1, \ldots, T_{pre}$. We also look for time weights $\hat{\lambda}_t^{sdid}$ that balance pre-exposure time periods with postexposure ones. Then we use these weights in a basic two-way fixed effects regression to estimate the average causal effect of exposure (denoted by τ):

$$\left(\hat{\tau}^{sdid}, \hat{\mu}, \hat{\alpha}, \hat{\beta}\right) = \underset{\tau,\mu,\alpha,\beta}{\operatorname{argmin}} \left\{ \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau)^2 \hat{\omega}_i^{sdid} \hat{\lambda}_t^{sdid} \right\} \tag{4}$$

The use of weights in the SDID estimator effectively makes the two-way fixed effect regression "local," in that it emphasizes (puts more weight on) units that on average are **similar** in terms of their past to the target (treated) units, and it emphasizes periods that are on average **similar** to the target (treated) periods. This localization can bring two benefits relative to the standard DID estimator. Intuitively, using only similar units and similar periods makes the estimator more robust. The use of the weights can also improve the estimator's precision by implicitly removing systematic (predictable) parts of the outcome.

The underlying idea for determining unit weights and time weights is that, on average, individuals more similar to the treatment group are assigned higher individual weights, and time periods more similar to the treatment period are assigned higher time weights. Individual weights are applied to control group individuals and are used to match the pretreatment trends of control group individuals with those of treatment group individuals. Individual weights can be obtained by solving the optimization problem in Equation (5). The basic idea behind this optimization problem is to make the weighted average outcomes of the control group and the arithmetic average outcomes of the treatment group as close as possible for each pretreatment period, differing by a constant term (intercept). The intercept is introduced to ensure parallel pretreatment trends and does not require a perfect match. The last term in the objective function of Equation (5) adds a regularization penalty term to constrain the values of the weights, ensuring that they are dispersed and unique.

$$\left(\hat{\omega}_0, \hat{\omega}^{sdid}\right) = \underset{\omega_0 \in \mathbb{R}, \omega \in \Omega}{\arg\min}\ \ell_{unit}(\omega_0, \omega),$$

where

$$\ell_{unit}(\omega_0, \omega) = \sum_{t=1}^{T_{pre}} \left(\omega_0 + \sum_{i=1}^{N_{co}} \omega_i Y_{it} - \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} Y_{it} \right)^2 + \zeta^2 T_{pre} \|\omega\|_2^2,$$

$$\Omega = \left\{ \omega \in \mathbb{R}_+^N : \sum_{i=1}^{N_{co}} \omega_i = 1,\ \omega_i = N_{tr}^{-1} \text{ for all } i = N_{co} + 1, \ldots, N \right\}, \tag{5}$$

Time weights are applied to the pretreatment periods to balance the periods before and after treatment.

Individual weights can be obtained by solving the optimization problem in Equation (6). The fundamental idea behind this optimization problem is to make the weighted average outcomes of the pretreatment periods for each individual in the control group as close as possible to the arithmetic average outcomes of the post-treatment periods, differing by a constant term. Since time weights may not be unique (the objective function may have multiple local minima), a small regularization penalty term is typically added in practice to constrain their values.

$$\left(\hat{\lambda}_0, \hat{\lambda}^{sdid}\right) = \underset{\lambda_0 \in \mathbb{R}, \lambda \in \Lambda}{\arg\min}\, \ell_{time}(\lambda_0, \lambda),$$

where

$$\ell_{time}(\lambda_0, \lambda) = \sum_{i=1}^{N_{co}} \left(\lambda_0 + \sum_{t=1}^{T_{pre}} \lambda_t Y_{it} - \frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} Y_{it} \right)^2, \tag{6}$$

$$\Lambda = \left\{ \lambda \in \mathbb{R}_+^T : \sum_{t=1}^{T_{pre}} \lambda_t = 1,\ \lambda_t = T_{post}^{-1} \text{ for all } t = T_{pre}+1, \ldots, T \right\}.$$

Optimizing the calculation to determine weights and obtain the best matching results can indeed enhance the accuracy of counterfactual matching.

(3) What are the source of hidden bias?

We believe that the hidden bias primarily stems from whether it is possible to find a suitable control group match for each individual in the treatment group. In this study, we used the combination of the three bias-reducing components in the SDID estimator, (i) double differencing, (ii) the unit weights, and (iii) the time weights, to reduce the bias to a sufficiently small level.

Third, not all PAs are created equal and there is substantial variation in how PAs are managed and by whom (e.g., Brazil's Strict and Sustainable Use PAs and Indigenous Territories). There's some differentiation (strict v non-strict), but at such a scale, the analysis fails to address a key component of heterogeneity. Here, I believe the authors could do much more to disentangle what works where and why. This understanding is really critical for the 30x30 agenda. Again, this might all require some additional thinking and work but might also help with the first point about novelty.

Response:

We thank the reviewer for this suggestion. We have considered the heterogeneity of IUCN categories of PAs in the previous version (see Fig.2 in P6). To avoid potential ambiguity, we have added a note in the caption to Fig. 2. "The total sample encompasses all the PAs examined in the study. PAs designated as IUCN I-II are those strictly protected according to the IUCN categories. Conversely, PAs classified as IUCN III-VI are not under strict protection. The term "Bottom 25% size" refers to PAs within the smallest 25% in terms of size, representing small-scale PAs. Meanwhile, "Top 25% size" pertains to the largest 25% of PAs, indicative of large-scale PAs."

We have added a section (P11-13) to respond to your concerns and have endeavored to explore further heterogeneity in the effectiveness of PAs, with the aim of informing the global 30×30 agenda. We now focus on five aspects of heterogeneity, including the establishment time of the PA, governance type (indigenous and community vs. government), the importance of biodiversity, the distance from cities, and human footprint. We also have revealed the reasons for the heterogeneity in different PAs based on their performance.

Differentiated effectiveness of protected areas and its implications for 30×30 agenda

Nations who have committed to the Kunming-Montreal Global Biodiversity Framework aim to designate 30% of Earth's land and water areas for protection by 2030. Given this ambition, an urgent need is to identify where setting up PAs will be most effective. The current effectiveness assessment of PAs may provide a reference for this purpose. We therefore examined the variation of the effectiveness of existing PAs against five criterial: the establishment time of the PA, governance type (indigenous and community vs. government), the importance of biodiversity, the distance from cities, and human footprint (Fig. 5). We applied the integrated framework of metacoupling (human-nature interactions within a PA as well as between the PA and other places such as cities near and far) to understand factors influencing the PA effectiveness. We used proportion changes in combined total habitat loss in PAs as example of integrated habitat alteration.

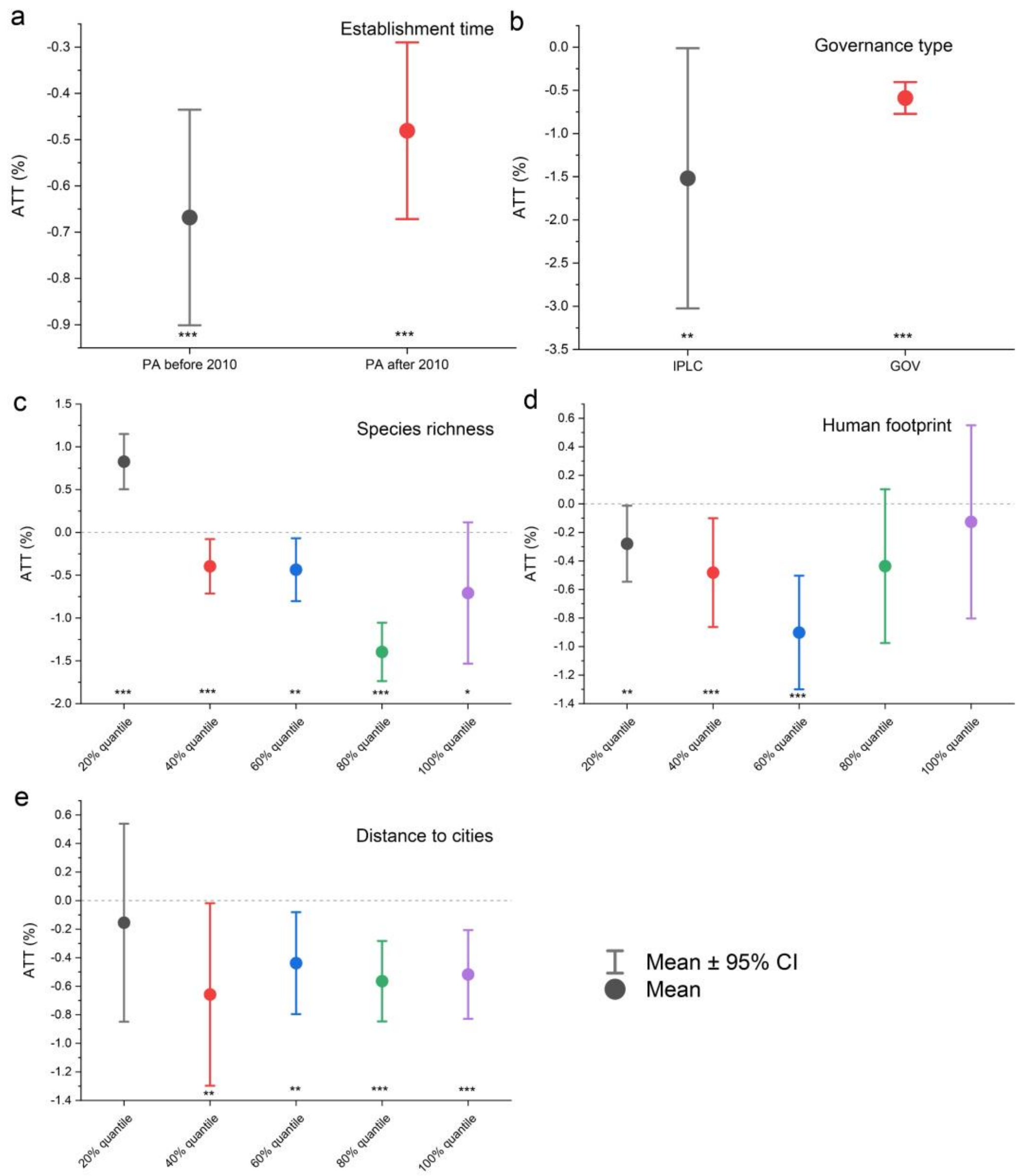

Fig. 5 | Differentiated effects of designated PAs on retarding habitat loss. ATT is the "average treatment effect on the treated" of PAs, namely the estimated annual average effect of established PAs on retarding the proportion of total habitat loss in PAs between 2003 and 2019. Error bars indicate 95% confidence intervals. Asterisks represent significance levels: * P< 0.1, ** P< 0.05, *** P< 0.01. The absence of an asterisk indicates statistical non-significance. If AAT is significantly negative, it indicates that PA establishment has a significant effect on reducing habitat loss. Here we consider the multiple heterogeneity of PAs, including establishment time of PAs **(a)**, governance type: indigenous peoples and local communities' lands (IPLC) and government (GOV) **(b)**, species richness **(c)**, human footprint **(d)**, distance to cities **(e)**. To further examine the heterogeneity of PAs, we divided the sample PAs into quintiles (e.g., 20% quantile in (c) represents PAs with the lowest 20% species richness).

We found that earlier established PAs are more effective in curbing the overall loss proportion of habitat over the time period assessed (Fig. 5a). The ATT values for PAs established before and after 2010 were -0.67% and -0.48%, respectively. This implies that to achieve the 30×30 agenda, it is crucial to act promptly, and the longer the establishment time, the more likely the PA will become effective. This is potentially because only after the establishment of PAs can there be corresponding resource support, and the effectiveness of PAs is not immediately apparent but requires a long-term process. Moreover, it is essential to recognize that newly established PAs tend to become "paper parks", as nations try and achieve numbers of PAs in terms of area and not outcomes. Moreover, the government type of PAs is an important factor that affects their effectiveness. Indigenous peoples and local communities' lands are more effective in controlling habitat loss in PAs compared to government and other entities (ATT were -1.52% vs. -0.59%, Fig. 5b). This is potentially because of their deep understanding of the local environment and traditional knowledge and practices for sustainable resource use and management[39]. Their participation in the management of PAs can help to improve the effectiveness of conservation efforts and promote the sustainable use of natural resources[39].

We also found PAs currently distributed in regions of high biodiversity richness and importance have stronger effectiveness in preventing habitat loss (Fig. 5c). One potential reason for this is that these areas with high species richness are often far from human activity areas, and the probability of human disturbance is therefore relatively low compared to other regions. This has been further supported by human footprint data, which shows that PAs distributed in the top 40% of human footprint areas do not significantly inhibit habitat loss (Fig. 5d). In addition, the distance of PAs from cities can provide a reference for the establishment of future PAs. PAs that are farther from cities were more effective in significantly curbing habitat loss, while the effectiveness of PAs near cities is not significant (Fig. 5e). This is potentially because areas near cities are more likely to be affected by intense human activities, such as infrastructure construction, urbanization, industrialization, and tourism, which can lead to habitat loss. Therefore, it is important to consider the distance from cities when establishing PAs to maximize their effectiveness in protecting biodiversity and habitats.

In summary, the paper is interesting and has lots of potential but falls a little short on novelty, contribution and ambition. I hope these comments are useful.

Response:

Thanks for your comments. Following your suggestions, we have expanded our research in a more targeted manner, increasing the innovation, contribution and ambition. We try to make our study more practical, especially for 30×30 agenda. We hope these revisions and extensions are meaningful.

Reviewer #2 (Remarks to the Author):

Comments on "Mixed Effectiveness of Global Protected Areas in Resisting Habitat Loss"

The paper assesses the performance of over 160,000 protected areas in resisting habitat loss globally. While I found the findings compelling, it is important to understand the main drivers behind protected areas, which can help with policy decisions. For instance, the study finds that larger and stricter protected areas tend to have lower habitat loss.

Response:

We thank the reviewer for this assessment.

Several questions arise. Maybe it is because more important/diverse protected areas, which tend to have more resources or are established earlier, are more effective in reducing habitat loss? Or maybe it is the distance to cities/urban areas? (My prior is that older protected areas are located further away from cities, while newer protected areas are closer to cities, thus more likely to be deforested). Or maybe it is simply governance? Most of the positive results (for instance, in Table 4, are concentrated in Europe). If it is governance, what is happening with Africa? What are they doing correctly that can be extrapolated to other developing regions?

Response:

These are excellent and constructive comments. We believe that the factors the reviewer mentioned could all potentially affect the effectiveness of the protected area. We have now endeavored to explore further heterogeneity in the effectiveness of PAs, with the aim of informing the global 30×30 agenda. We have focused on five aspects of heterogeneity, including the establishment time of the PA, governance type (indigenous and community vs. government), the importance of biodiversity, the distance from cities, and human footprint. We also have revealed the reasons for the heterogeneity in different PAs based on their performance.

In the section "Differentiated effectiveness of protected areas and its implications for 30×30 agenda", we added the following additional thinking and work:

Differentiated effectiveness of protected areas and its implications for 30×30 agenda

Nations who have committed to the Kunming-Montreal Global Biodiversity Framework aim to designate 30% of Earth's land and water areas for protection by 2030. Given this ambition, an urgent need is to identify where setting up PAs will be most effective. The current effectiveness assessment of PAs may provide a reference for this purpose. We therefore examined the variation of the effectiveness of existing PAs against five criterial: the establishment time of the PA, governance type (indigenous and community vs. government), the importance of biodiversity, the distance from cities, and human footprint (Fig. 5). We applied the integrated framework of metacoupling (human-nature interactions within a PA as well as between the PA and other places such as cities near and far) to understand factors influencing the PA effectiveness. We used proportion changes in combined total habitat loss in PAs as example of integrated habitat alteration.

We found that earlier established PAs are more effective in curbing the overall loss proportion of habitat over the time period assessed (Fig. 5a). The ATT values for PAs established before and after 2010 were -0.67% and -0.48%, respectively. This implies that to achieve the 30×30 agenda, it is crucial to act promptly, and the longer the establishment time, the more likely the PA will become effective. This is potentially because only after the establishment of PAs can there be corresponding resource support, and the effectiveness of PAs is not immediately apparent but requires a long-term process. Moreover, it is essential to recognize that newly established PAs tend to become "paper parks", as nations try and achieve numbers of PAs in terms of area and not outcomes. Moreover, the government type of PAs is an important factor that affects their effectiveness. Indigenous peoples and local communities' lands are more effective in controlling habitat loss in PAs compared to government and other entities (ATT were -1.52% vs. -0.59%, Fig. 5b). This is potentially because of their deep understanding of the local environment and traditional knowledge and practices for sustainable resource use and management[39]. Their participation in the management of PAs can help to improve the effectiveness of conservation efforts and promote the sustainable use of natural resources[39].

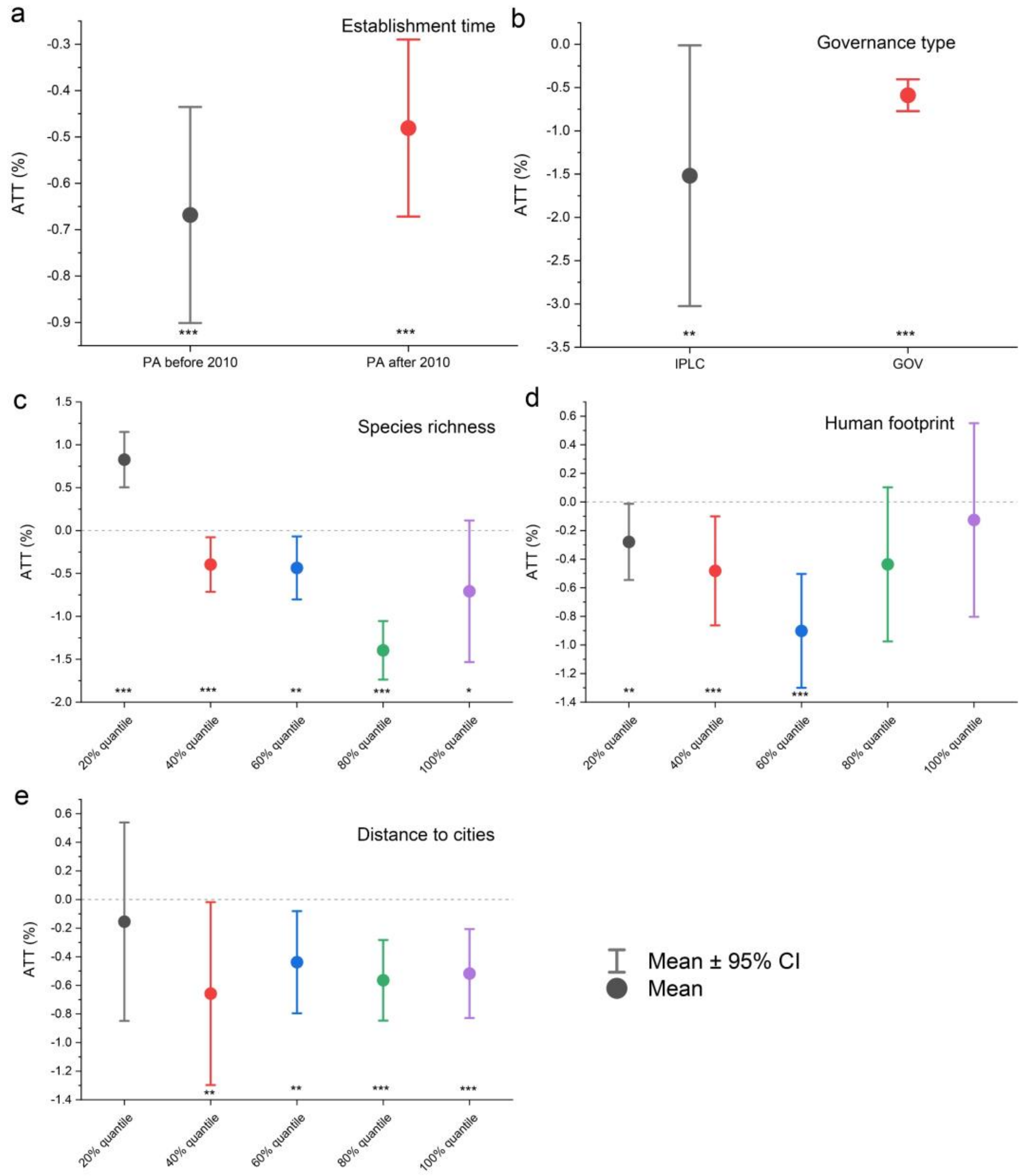

Fig. 5 | Differentiated effects of designated PAs on retarding habitat loss. ATT is the "average treatment effect on the treated" of PAs, namely the estimated annual average effect of established PAs on retarding the proportion of total habitat loss in PAs between 2003 and 2019. Error bars indicate 95% confidence intervals. Asterisks represent significance levels: * $P< 0.1$, ** $P< 0.05$, *** $P< 0.01$. The absence of an asterisk indicates statistical non-significance. If AAT is significantly negative, it indicates that PA establishment has a significant effect on reducing habitat loss. Here we consider the multiple heterogeneity of PAs, including establishment time of PAs **(a)**, governance type: indigenous peoples and local communities' lands (IPLC) and government (GOV) **(b)**, species richness **(c)**, human footprint **(d)**, distance to cities **(e)**. To further examine the heterogeneity of PAs, we divided the sample PAs into quintiles (e.g., 20% quantile in (c) represents PAs with the lowest 20% species richness).

We also found PAs currently distributed in regions of high biodiversity richness and importance have stronger effectiveness in preventing habitat loss (Fig. 5c). One potential reason for this is that these areas with high species richness are often far from human activity areas, and the probability of human disturbance is therefore relatively low compared to other regions. This has been further supported by human footprint data, which shows that PAs distributed in the top 40% of human footprint areas do not significantly inhibit habitat loss (Fig. 5d). In addition, the distance of PAs from cities can provide a reference for the establishment of future PAs. PAs that are farther from cities were more effective in significantly curbing habitat loss, while the effectiveness of PAs near cities is not significant (Fig. 5e). This is potentially because areas near cities are more likely to be affected by intense human activities, such as infrastructure construction, urbanization, industrialization, and tourism, which can lead to habitat loss. Therefore, it is important to consider the distance from cities when establishing PAs to maximize their effectiveness in protecting biodiversity and habitats.

Here we will focus on answering your following questions:

Maybe it is because more important/diverse protected areas, which tend to have more resources or are established earlier, are more effective in reducing habitat loss?

Response:

Our analysis shows that high biodiversity (species richness) and earlier establishment of protected areas can indeed be more effective in reducing habitat loss. These protected areas tend to have more investment and resources to control habitat loss. Our regression analysis (OLS regression) further indicates that there was a significant negative correlation between the proportion of habitat loss within PAs and species richness (Table S7). And PAs with high and important biodiversity are also often remote from human activity.

Source	SS	df	MS			
Model	982288.559	1	982288.559	Number of obs	=	294,995
Residual	161103015	294,993	546.124875	F(1, 294993)	=	1798.65
				Prob > F	=	0.0000
Total	162085304	294,994	549.452884	R-squared	=	0.0061
				Adj R-squared	=	0.0061
				Root MSE	=	23.369

Proportion	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
Species_richness	-.0233816	.0005513	-42.41	0.000	-.0244621	-.022301
_cons	14.98308	.0918461	163.13	0.000	14.80307	15.1631

Or maybe it is the distance to cities/urban areas? (My prior is that older protected areas are located further away from cities, while newer protected areas are closer to cities, thus more likely to be deforested).

Response:

As the reviewer mentioned, our findings confirm that the distance to cities is indeed one of the key factors affecting the effectiveness of protected areas. The proximity to cities may lead to more intense human activities, which in turn may cause greater pressure on the protected area to curb habitat loss. This pressure can come from factors such as urbanization, industrialization, pollution, and habitat fragmentation. Our regression analysis (OLS regression) further indicates that the proximity to urban centers (ln_dis_urban) significantly influences the effectiveness of PAs (proportion of combined total habitat loss) (Table S7).

Source	SS	df	MS			
Model	7423792.16	1	7423792.16	Number of obs	=	298,125
Residual	156039115	298,123	523.405154	F(1, 298123)	=	14183.64
				Prob > F	=	0.0000
Total	163462907	298,124	548.305091	R-squared	=	0.0454
				Adj R-squared	=	0.0454
				Root MSE	=	22.878

Proportion	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
ln_dis_urban	-3.112115	.0261313	-119.10	0.000	-3.163331	-3.060898
_cons	39.29144	.2369519	165.82	0.000	38.82703	39.75586

Our regression analysis (OLS regression) results showed a significant increase in habitat loss within PAs (proportion of combined total habitat loss) as human footprints increased (Human_footprint) (Table S7).

Source	SS	df	MS			
Model	8134112.89	1	8134112.89	Number of obs	=	294,305
Residual	153609126	294,303	521.942099	F(1, 294303)	=	15584.32
				Prob > F	=	0.0000
				R-squared	=	0.0503
				Adj R-squared	=	0.0503
Total	161743238	294,304	549.578798	Root MSE	=	22.846

Proportion	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
Human_footprint	.4729956	.0037889	124.84	0.000	.4655695	.4804218
_cons	5.359021	.0650879	82.34	0.000	5.23145	5.486591

Or maybe it is simply governance? Most of the positive results (for instance, in Table 4, are concentrated in Europe). If it is governance, what is happening with Africa? What are they doing correctly that can be extrapolated to other developing regions?

Response:

We have considered the heterogeneity of IUCN categories of PAs in the previous version (see Fig.2). To avoid potential ambiguity, we have added a note in the caption to Fig. 2. "The total sample encompasses all the PAs examined in the study. PAs designated as IUCN I-II are those strictly protected according to the IUCN categories. Conversely, PAs classified as IUCN III-VI are not under strict protection. The term "Bottom 25% size" refers to PAs within the smallest 25% in terms of size, representing small-scale PAs. Meanwhile, "Top 25% size" pertains to the largest 25% of PAs, indicative of large-scale PAs." We have now examined the impact of governance types on the effectiveness of protected areas, focusing on the differences between indigenous and local community governance vs. government and other entity governance. We found that indigenous and local community governance was more effective in curbing habitat loss. This is because indigenous and local communities have a deep understanding of the local ecosystem and have developed traditional knowledge and practices for conservation. They also have a strong sense of ownership and responsibility for the land, which motivates them to protect it.

In traditional understanding, Europe is considered to have high human activity levels. However, we found that some European countries' protected areas have performed well in curbing habitat loss. This suggests that human governance can play an important role in the effectiveness of protected areas. This type of governance can be a useful reference and promotion value for developing countries. Many African countries' protected areas may not have effective governance, possibly due to a lack of sufficient investment. This also suggests that increasing investment in protected areas and enhancing governance to prevent habitat loss is crucial. In line 180-183 we added these statements.

My second key comment is: How to converge individual/country level results of the impact of protected areas with the results found in Fig. 4. For instance, several studies in Latin America have found that protected areas are effective at reducing deforestation [e.g., Costa Rica (Andam et al. 2008; Pfaff et al. 2009), México (Blackman et al. 2011), Peru (Miranda et al. 2013), Bolivia (Hanauer y Canavire 2013), etc.] while the current results suggest the opposite (either significantly negative or non-significant). It is important to understand the approach and results and find channels *for the differences.*

Response:

We have now rechecked the results in Fig. 4 and found that we may not have adequately explained the meaning of the symbols. In fact, our analysis results show a significantly negative coefficient for South America (significant at the 95% level, see Fig.4 k and i), indicating that the proportion of forest loss in protected areas in South America is lower compared to unprotected regions, making protected areas more effective at reducing deforestation. Our results are consistent with other individual/country-level results.

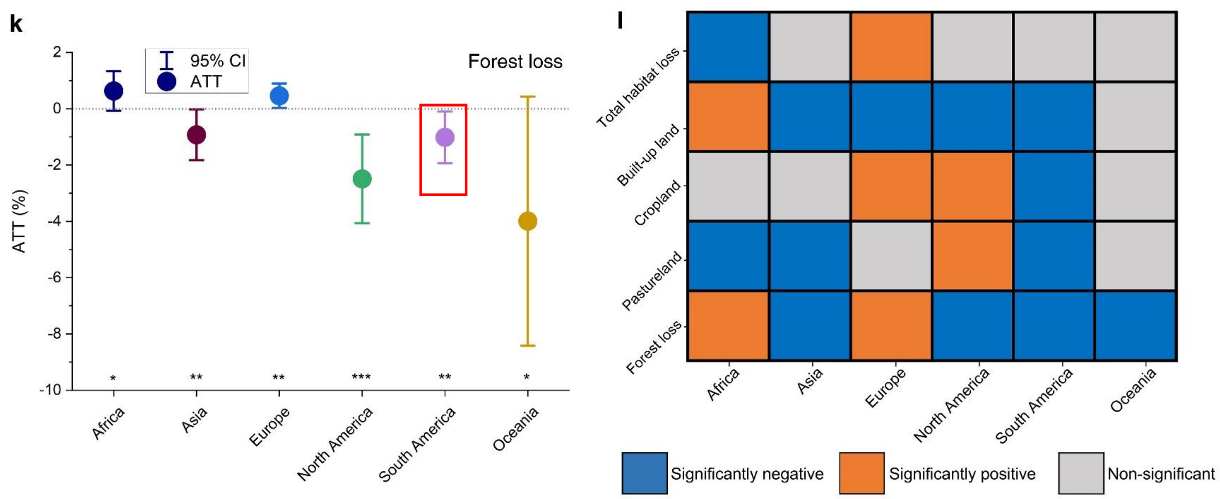

Fig. 4 | Estimates of PA impact on habitat loss proportion using PSM-DDD model. ATT is the "average treatment effect on the treated" of PAs, namely the estimated average effect of established PAs on retarding the proportion of habitat loss in PAs. The estimated annual average effect of established PAs on retarding the proportion of habitat loss in PAs, including the proportion of combined forest loss (**k**). A summary of influence direction and significance level were shown in **l**. ATTs were estimated by DDD model using matched panel data derived from the PSM. Asterisks represent significance levels: * $P< 0.1$, ** $P< 0.05$, *** $P< 0.01$. The absence of an asterisk indicates statistical non-significance. Significant negative values indicate that the establishment of the PA can significantly reduce habitat loss, while positive values indicate the opposite.

In fact, we did find that protected areas in South America are not effective in curbing the total habitat loss within these protected areas (the coefficient is non-significant). However, deforestation is only a part of the total habitat loss (42%), so our results are not directly comparable to existing studies on deforestation.

Perhaps another possible reason is that the protected area examples you mentioned in previous studies are just some of the PAs. This paper includes many more PAs, and some of the PAs were not included in the previous studies cited in the previous examples, which might have opposite results and offset the results in the examples from previous studies. We considered 5,723 PA in Latin America (1,544 PA in Central America, 565 PA in the Caribbean, and 3,614 PA in South America) in this paper, which is higher than the sample PAs in the previous studies.

Number of protected areas were considered in previous studies:
Costa Rica (Andam et al. 2008): 73 PA;
Costa Rica (Pfaff et al. 2009): 132 PA;
México (Blackman et al. 2015): 56 PA;
Peru (Miranda et al. 2016): 29 PA;
Bolivia (Canavire-Bacarreza & Hanauer 2013): 23 PA.

References:
Andam, K. S., Ferraro, P. J., Pfaff, A., Sanchez-Azofeifa, G. A. & Robalino, J. A. Measuring the effectiveness of protected area networks in reducing deforestation. Proceedings of the National Academy of Sciences of the United States of America 105, 16089-16094 (2008).
Pfaff, A., Robalino, J., Sanchez-Azofeifa, G. A., Andam, K. S. & Ferraro, P. J. Park Location Affects Forest Protection: Land Characteristics Cause Differences in Park Impacts across Costa Rica. The B.E. Journal of Economic Analysis & Policy 9 (2009).
Blackman, A., Pfaff, A. & Robalino, J. Paper park performance: Mexico's natural protected areas in the 1990s. Global Environmental Change 31, 50-61 (2015).
Miranda, J. J., Corral, L., Blackman, A., Asner, G. & Lima, E. Effects of Protected Areas on Forest Cover Change and Local Communities: Evidence from the Peruvian Amazon. World Development 78, 288-307 (2016).
Canavire-Bacarreza, G. & Hanauer, M. M. Estimating the Impacts of Bolivia's Protected Areas on Poverty. World Development 41, 265-285 (2013).

I like the lag effects. Can we generalize more between stricter protected areas or continents? This can be an important factor to highlight strongly in the paper with important policy recommendations for policymakers.
Response:

We are pleased to hear the reviewer appreciates the discussion of lag effects in our paper. The suggestion to further generalize our findings across different strictly protected areas or continents, and to emphasize its significance for policy recommendations, is indeed an insightful and important point.

Regarding the possibility of extending our findings to stricter protected areas or different continents, we agree that this is a direction worth exploring in more depth. We have revisited our data and analysis to provide a more comprehensive examination in this regard, enrich the perspective of our paper, and potentially have profound implications for policy. In the Supplementary Information Table 6 we have showed the differentiated results of lag effects for non-strict PAs, stricter PAs, and different continents. We also have added additional results in main text in lines 247-264.

"We delved further into the investigation of the heterogeneity of this lag effect across different continents and between strict and non-strict PAs (Table S6). Our findings reveal that PAs in North America, Europe, and under non-strict protection exhibit a more pronounced time-lagged effect in relation to the total habitat loss. Surprisingly, PAs in North America demonstrate significant efficacy in curbing the growth of built-up areas and cropland within their bounds, even with a lag of six years.

In addition, PAs in Asia and those under non-strict protection also manifest this effect in curbing the increase in cropland within their boundaries. However, this lag effect is less observed for pastureland and forest loss. Notably, PAs under strict protection do not exhibit this effect. The observed lag effect in global PAs' effectiveness in curbing habitat loss provides several policy implications. Establishing PAs proactively and promptly is vital to effectively mitigate habitat loss before it becomes severe. Tailoring strategies to regional specifics, such as climate, biodiversity, and socio-economic factors, enhances the effectiveness of these areas. The balance between strict and non-strict conservation measures should be carefully considered, recognizing that flexible approaches can sometimes yield significant benefits, especially in regions where strict conservation might be impractical. Long-term monitoring and adaptive management are crucial, as the impacts of conservation efforts on habitat loss may not be immediately apparent. This involves regular assessment and the willingness to adjust strategies based on ecological changes and evolving scientific understanding."

In terms of methodological approach, as I understand, using the Hansen Global Forest Change dataset is not free of bias (e.g., Alix Garcia and Millimet, 2023). The algorithm tends to reduce deforestation in small areas. Does it have an impact in smaller protected areas? It would be great to see whether the effects are consistent when dividing the sample between small vs. large/medium PAs. Also, regarding the Propensity score matching approach, it is unclear which variables were used to select the correct matches, and more discussion is needed.
Response:

Use of the Hansen Global Forest Change Dataset:

We appreciate the reviewer's concerns regarding potential biases in the Hansen Global Forest Change dataset, as discussed in Alix-García and Millimet (2023). Their research highlights the nonclassical measurement errors in satellite data, particularly in the binary classification of deforestation. To address the specific concern about the impact of these biases on smaller protected areas, we have conducted an additional analysis.

We divided the sample PAs into small, medium, and large sizes based on a principle of tripartition according to the area of the PA. This subdivision allowed us to assess whether the potential underestimation of deforestation in smaller areas significantly alters our findings. We then calculated the cumulative proportion of forest loss within each category of PAs from 2000 to 2020. This approach, focusing on the proportion rather than the absolute area of forest loss, offers better comparability. Our findings reveal that the percentage of forest loss in small PAs has not been underestimated. The respective percentages of forest loss in small, medium, and large PAs were 9.38%, 3.79%, and 4.21%, respectively. In our view, even though there may be underestimation issues with the Hansen Global Forest Change dataset, the quantification of forest loss within PAs should remain consistent. The size of a PA is unlikely to significantly skew the underestimation of forest loss in any particular category of PAs. The additional analysis indicates that the results support our original findings, thus reinforcing the robustness of our conclusions even when considering potential biases in the Hansen dataset.

Clarification on Propensity Score Matching Approach:

Regarding the Propensity Score Matching (PSM) approach, we apologize for any lack of clarity in our main text. We have added this information in the Methods section of our revised manuscript. The matching was based on five covariates potentially related to habitat loss within the PAs: elevation, slope, initial human footprint (1993), country, and ecoregion.

In our PSM analysis, we ensured that the matched groups (protected vs. unprotected areas) were balanced on these covariates, thereby reducing selection bias and allowing for a more accurate estimation of the causal effect of PA on curbing habitat loss. In Supplementary Information Fig. 4 we added the diagnostics graphs to confirm the balance of covariates in the matched sample. These graphs all show that the standardized % bias across covariates are less than 10% between unmatched and matched. We have added a more detailed discussion of these aspects in the revised manuscript to enhance the understanding of our methodological approach.

Reviewers' Comments:

Reviewer #1:
Remarks to the Author:
I would like to thank the author's for carefully considering my comments. However, my main concern with this paper continues to be novelty. The authors refer to Geldmann et al. and Jones et al. (published in PNAS and Science, respectively). Yet, their assessment comes to very similar conclusions; essentially that protected areas are not as effective as they need to be. The paper's empirical contribution, the use of better data to answer a question that we have already partially answered, is an incremental contribution to science but doesn't really stretch our understanding in a significant enough way.

The authors mention heterogeneity. But here, I feel the authors could have studied heterogeneity in much greater detail. The focus continues to be on a global assessment. Yet, we know that protected areas perform very differently in different parts of the world. In the Amazon, for example, Indigenous territories perform better than strict protected areas (see Nolte et al. in PNAS for example). Such nuance gets lost in the current analysis. Could we identify regions where different types of PAs perform better/worse? Can we gain any insights into why this might be? Are some biomes better protected than others? If so, why? Do we get different types of threats in different places?

Similarly, the theoretical contribution is not clear? The authors talk about using an integrated metacoupling framework. However, what this framework is, how it is used and applied, and how it helps us to better understand the effectiveness of protected areas isn't clear.

I feel this paper has potential given the amount of data that the authors have collated. However, I believe that it requires substantial reframing for it to represent a significantly novel contribution to our understanding.

Reviewer #2:
Remarks to the Author:
Dear Authors,
I appreciate the detailed response to the comments. This revised version is clearer, and no additional comments from me.

Reviewer #1 (Remarks to the Author):

I would like to thank the author's for carefully considering my comments. However, my main concern with this paper continues to be novelty. The authors refer to Geldmann et al. and Jones et al. (published in PNAS and Science, respectively). Yet, their assessment comes to very similar conclusions; essentially that protected areas are not as effective as they need to be. The paper's empirical contribution, the use of better data to answer a question that we have already partially answered, is an incremental contribution to science but doesn't really stretch our understanding in a significant enough way.

Response:

We thank the reviewer for their constructive comments. Based on their previous comments and the second point raised, we have reorganized our manuscript to enhance the innovative nature of our study, in particular undertaking detailed analyses to assess issues around heterogeneity, as described below.

We believe our research is distinct and novel because we have comprehensively evaluated the effectiveness of each protected area, including both the overall impact of protected areas on habitat loss and specific types of habitat loss (built-up land, cropland, pastureland, and forest loss). This exploration was not done in either of the papers raised by the reviewer. We also applied the PSM method to sample over 3.15 million 1km^2 grids globally, inside and outside of protected areas, to compare the differences in habitat loss rate. This also provided new insights into the heterogeneous performance of protected areas across different scales (site, ecoregion and biome) which has not been done before. As a consequence, our results provide far more context and specific information that are found in other studies and should, in our opinion, be considered an advance.

For example, by aggregating protected areas to the ecoregion and biome scales, and focusing on the performance of different types of protected areas in different regions (ranging from strict protection areas, sustainable use areas, indigenous peoples' land, local community conservation areas, and detailed IUCN protected area categories (Ia, Ib, II, III, IV, V, and VI)), we provide a series of new results (see lines 453-485). We thank the reviewer for raising the heterogeneity issue as it has strengthened the manuscript and has helped demonstrate the innovativeness of our research approach. Please refer to the second point below for detailed revisions.

The authors mention heterogeneity. But here, I feel the authors could have studied heterogeneity in much greater detail. The focus continues to be on a global assessment. Yet, we know that protected areas perform very differently in different parts of the world. In the Amazon, for example, Indigenous territories perform better than strict protected areas (see Nolte et al. in PNAS for example). Such nuance gets lost in the current analysis. Could we identify regions where different types of PAs perform better/worse? Can we gain any insights into why this might be? Are some biomes better protected than others? If so, why? Do we get different types of threats in different places?

Response:

We thank the reviewer for their constructive comments. Regarding the heterogeneity issue, we believe this is a great suggestion and we have now conducted meticulous revisions and re-analysis to present a more detailed heterogeneity analysis of the effectiveness of protected areas. In reference to the example from the Brazilian Amazon mentioned in Nolte et al., PNAS, we conducted a global investigation. Our goal was to identify the differences in the spatial effectiveness of different types of protected areas in different places. The detailed method is as follows:

(1) To maintain the consistency of spatial sampling, we first subdivided 161,248 protected areas into 1km^2 grids. Subsequently, we randomly sampled 1.05 million grids within the protected areas. To avoid potential spatial autocorrelation, we ensured that the distance between each sampled grid was greater than 3 km.

(2) Subsequently, we also randomly sampled a double-sized 1km^2 grid outside the protected areas, totaling 2.10 million grids. The purpose was to ensure that the grids inside the protected area could be matched to similar grids. At the same time, considering that the areas near the protected area might be directly influenced by the protected area, we restricted these sampling grids outside the 3 km buffer zones of the protected areas. Finally, we randomly sampled 3.15 million grids.

(3) We used the propensity score matching (PSM) in the Stata 'psmatch2' command to match each sample grid within protected areas. The matching was based on eight covariates potentially related to habitat loss within the PAs: elevation, slope, initial human footprint (1993), travel time to cities, tree cover in 2000, country, biome, and ecoregion. We used one-to-four nearest neighbor matching with calipers of width equal to 0.05 of the standard deviation of the logit of the propensity score to implement PSM, as this approach produces good matching results, reduces confounding between treatment and control grids, and allows for a relatively large sample size. Additionally, it can help to balance covariates between treatment and control groups.

(4) Using the matched data from 1.05 million grid cells, we further summarized the spatial effectiveness of three scales of protected areas, namely protected area scale, ecoregion scale, and biome scale. At the same time, we also paid special attention to different types of protected areas, including strict protection areas, sustainable use areas, indigenous peoples' land, local community conservation areas, and detailed IUCN protected area categories (Ia, Ib, II, III, IV, V, and VI).

Based on the PSM analysis mentioned above, we have made in-depth revisions to the paper from four aspects. To our knowledge, we have achieved the most detailed global-scale assessment of the spatial effectiveness of protected areas to date.

(1) We first estimated the spatial effectiveness of all sample points, followed by analyzing the differences between different types of protected areas. At the same time, we identified which protected areas performed well and which performed poorly (see Fig. 5). In the supporting documents, we demonstrated the differences in different types of protected areas globally (focusing on indigenous territories PA, strict protected PA, sustainable use PA, and local communities PA, see Fig.S6-S11). We found that the establishment of protected areas is more effective in restraining cropland expansion and deforestation. Further evidence confirmed that PAs in indigenous territories are the most effective, followed by strict PAs, but local community PAs are not effective. Specifically, wilderness protection areas (IUCN Ib) are not effective, and Category IV-VI protection areas are even less effective. Protected areas in southern Saharan Africa, the Amazon Basin, Europe, Central America, the central and southeastern United States, central Asia, and Southeast Asia have lower effectiveness. Protected areas in the Amazon Basin, North Africa, West Asia, central and northwestern Australia, China, and the Far East have higher effectiveness.

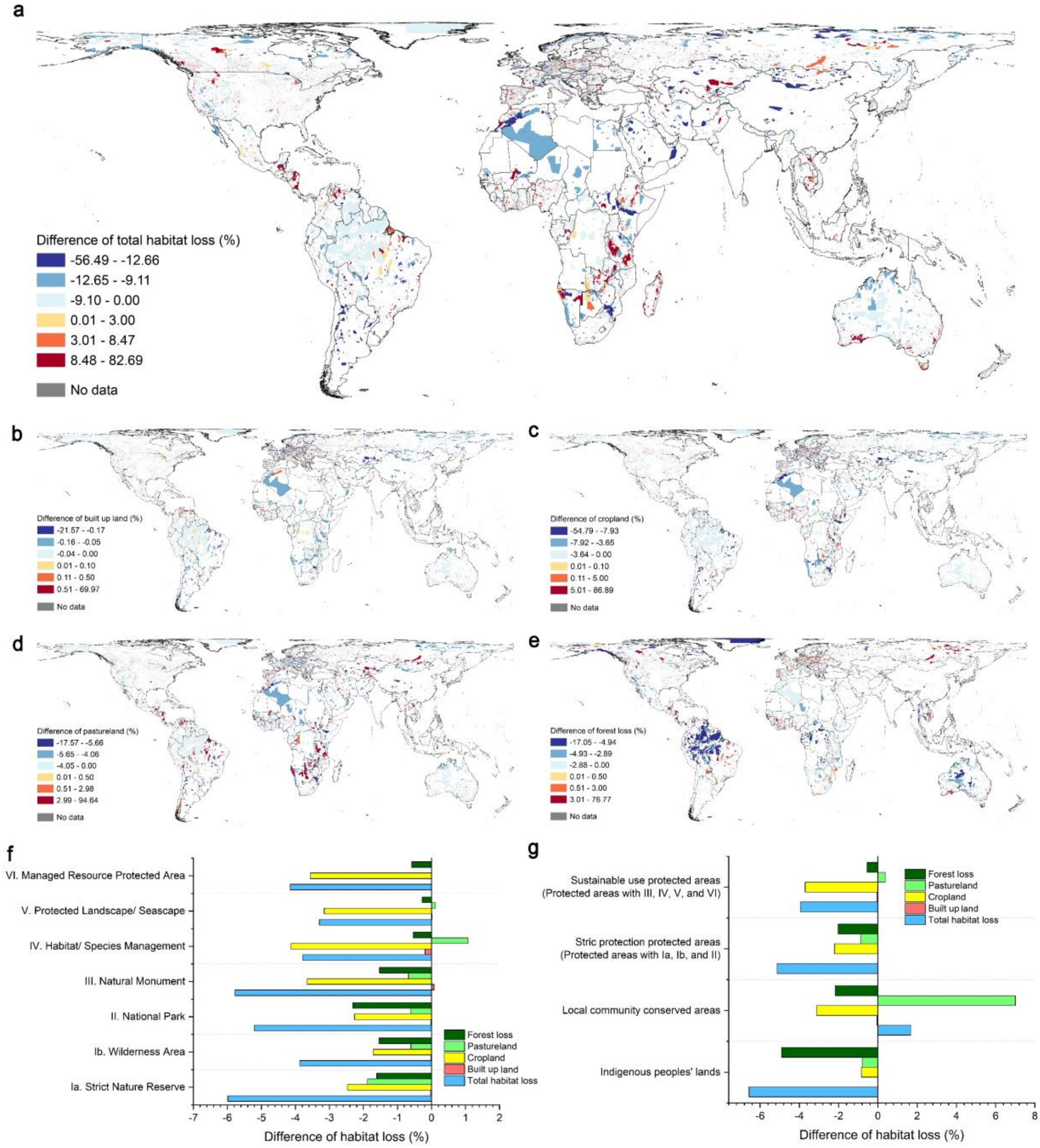

Fig. 5 | Habitat loss rate comparison with PAs and matched unprotected areas at the PA scale. a, Rate difference in combined total habitat loss with PAs and matched unprotected areas. **b**, Rate difference in built-up land. **c**, Rate difference in cropland. **d**, Rate difference in pastureland. **e**, Rate difference in forest loss. **f**, Rate differences in IUCN's protected areas management categories. **g**, Rate differences in aggregated four types of protected areas.

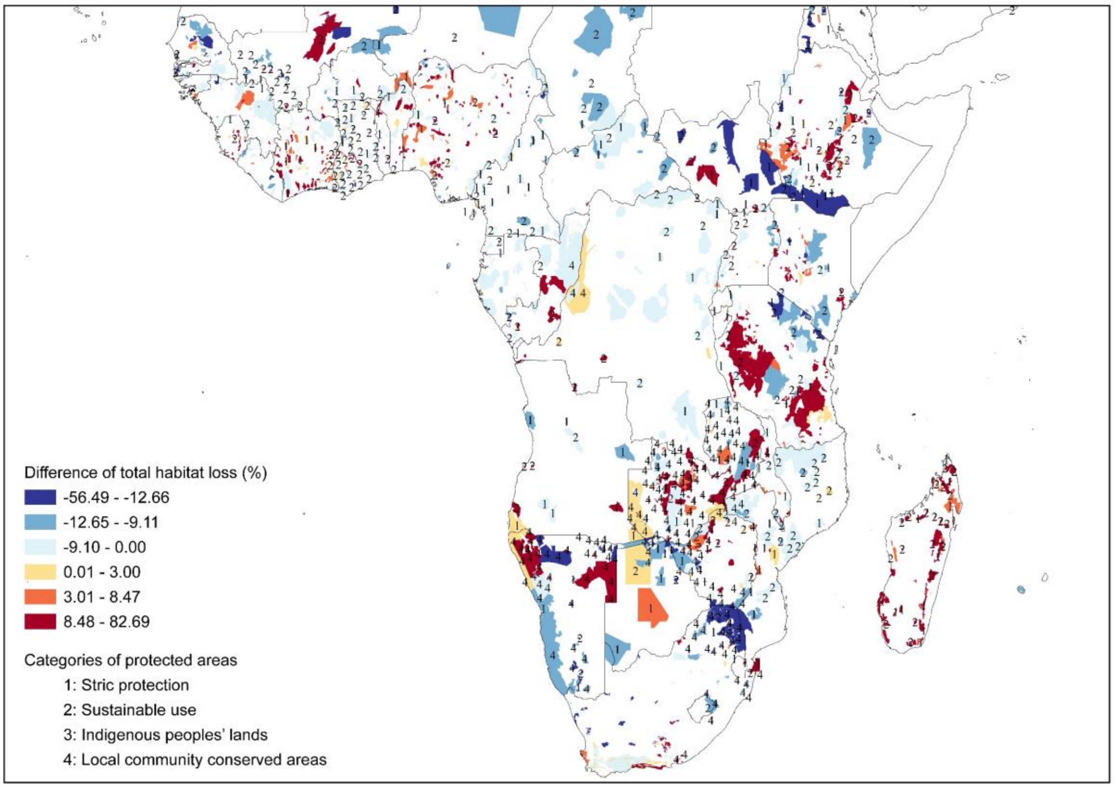

Fig. S6. Spatial effectiveness of different type PAs in Africa.

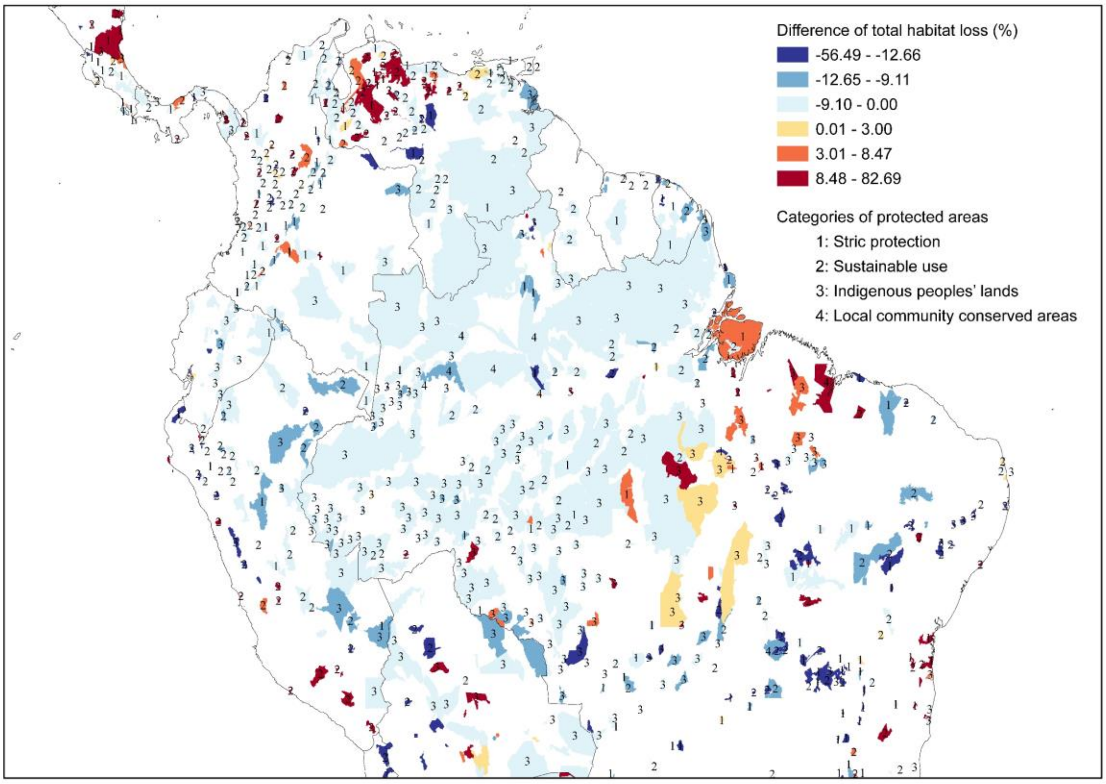

Fig. S7. Spatial effectiveness of different type PAs in South America.

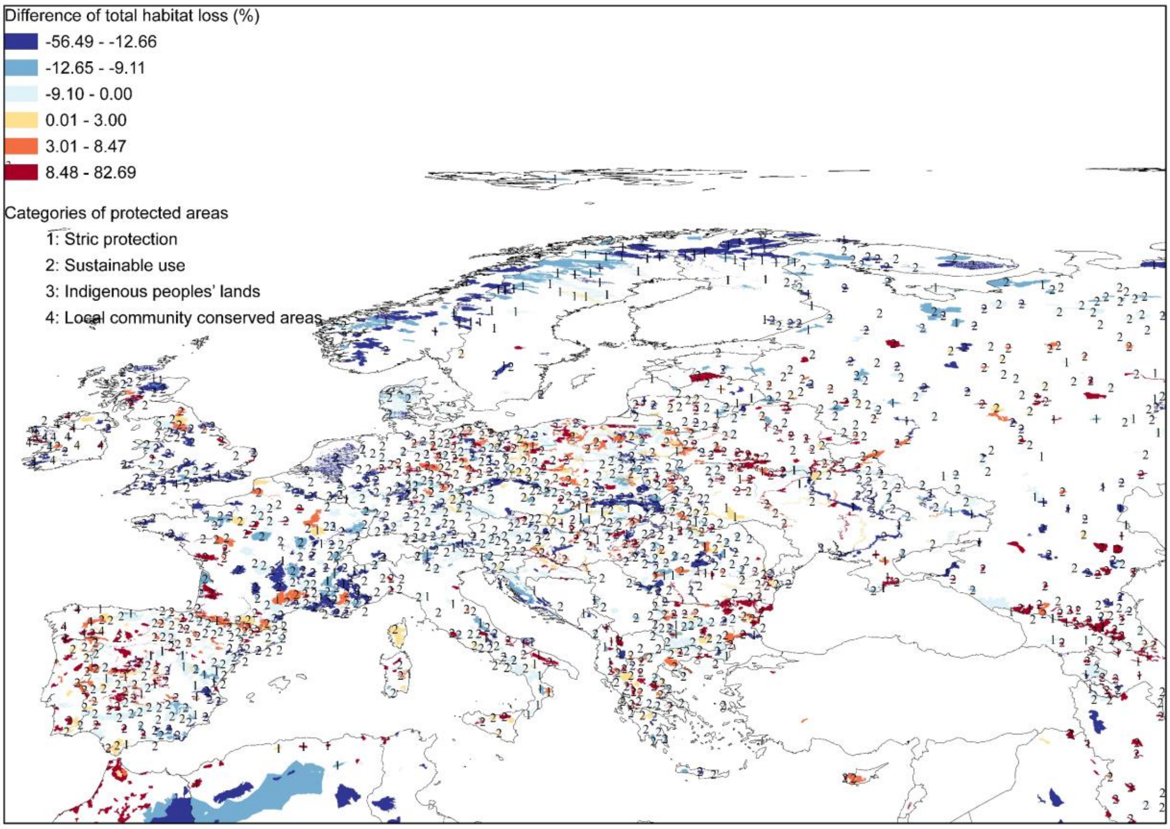

Fig. S8. Spatial effectiveness of different type PAs in Europe.

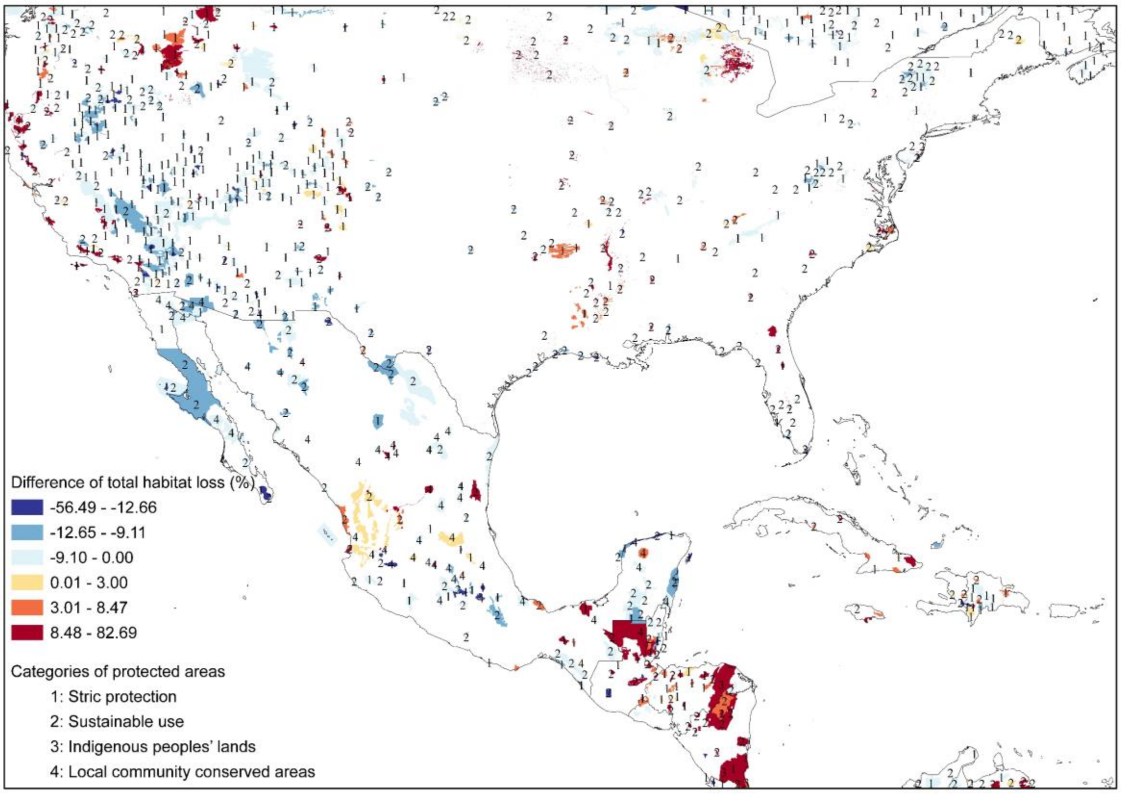

Fig. S9. Spatial effectiveness of different type PAs in North America.

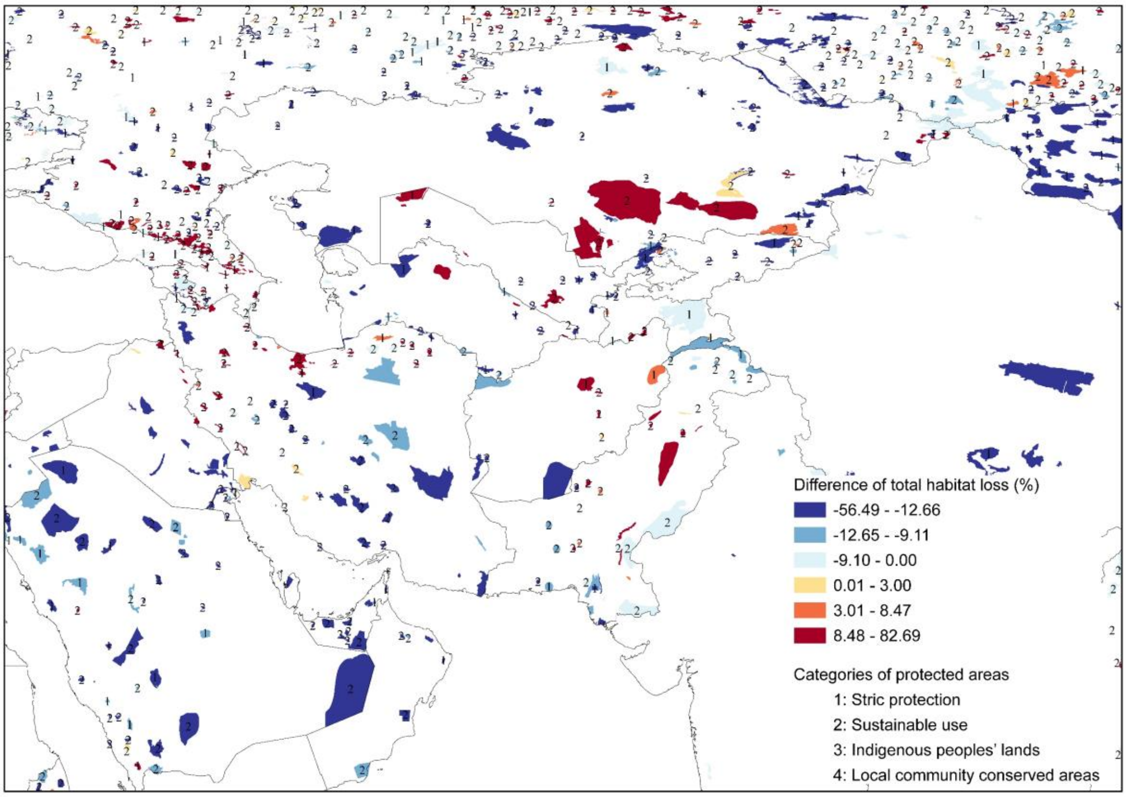

Fig. S10. Spatial effectiveness of different type PAs in Central and Western Asia.

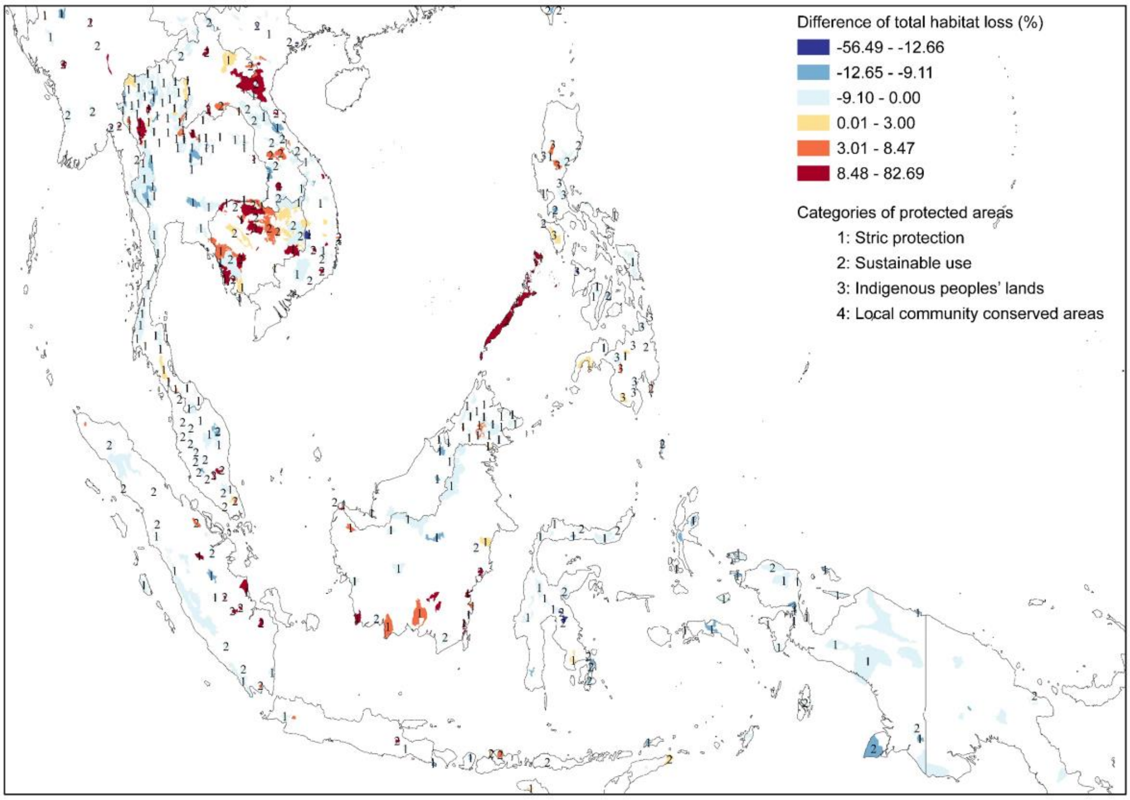

Fig. S11. Spatial effectiveness of different type PAs in Southeast Asia.

(2) We also summarized the effectiveness of protected areas at higher scales such as ecoregions and biomes (Fig.6). At the biome scale, Temperate grasslands, savannas, and shrublands, Tundra, Montane grasslands and shrublands, and Deserts and xeric shrublands have shown excellent performance. However, tropical biomes have exhibited poor performance.

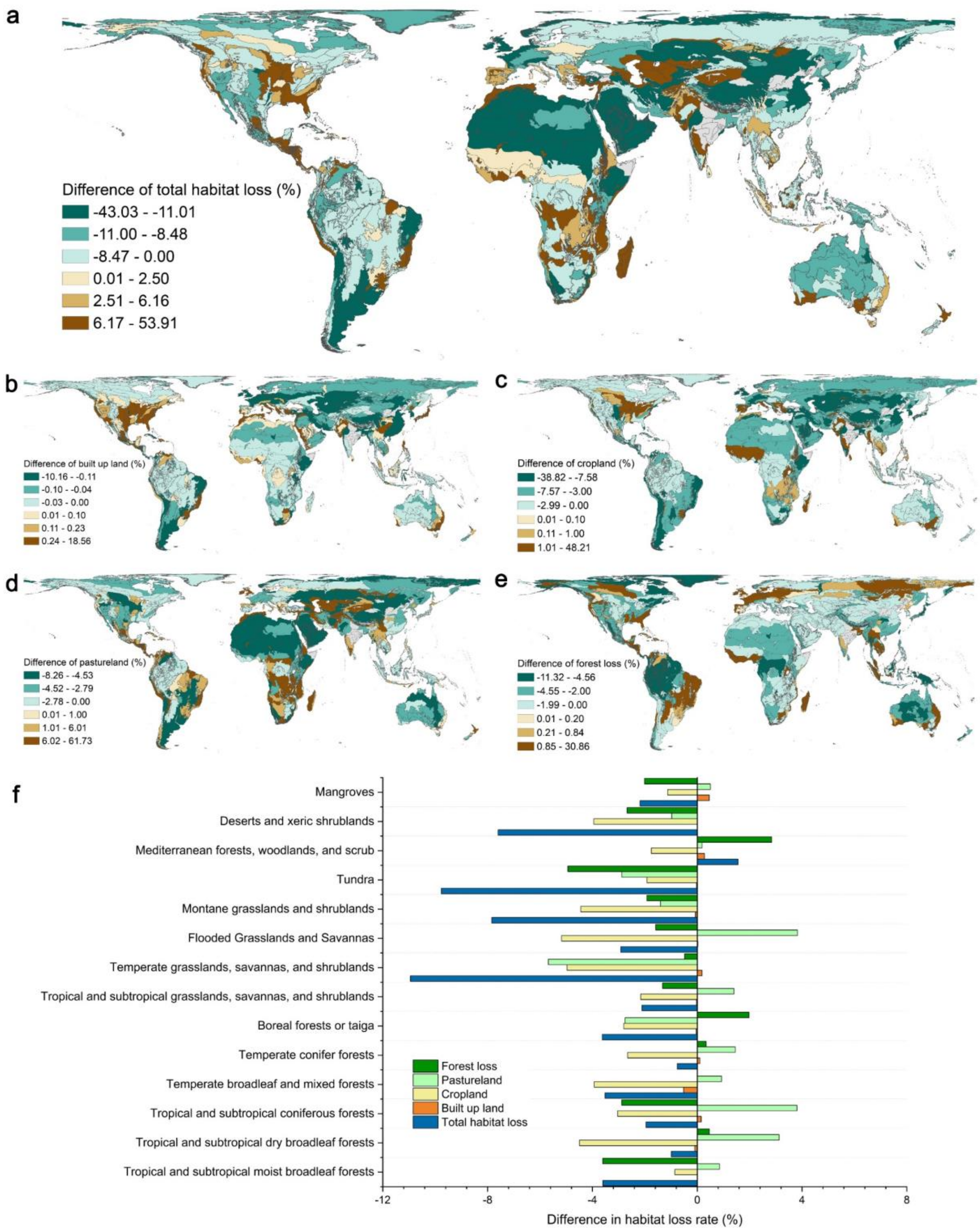

Fig. 6 | Habitat loss rate comparison with PAs and matched unprotected areas at the ecoregion scale and biome scale. a, Rate difference in combined total habitat loss with PAs and matched unprotected areas. **b**, Rate difference in built-up land. **c**, Rate difference in cropland. **d**, Rate difference in pastureland. **e**, Rate difference in forest loss. **f**, Rate differences in biomes.

(3) In addition, we identified the primary (Fig.7) and secondary (Fig.S14) threats facing each protected area. The primary threat to protected areas in Africa is cropland expansion, while Europe, South America, and Southeast Asia are facing forest loss. North America is primarily threatened by cropland expansion and forest loss, and Asia, Oceania, and South America are also significantly affected by pastureland expansion.

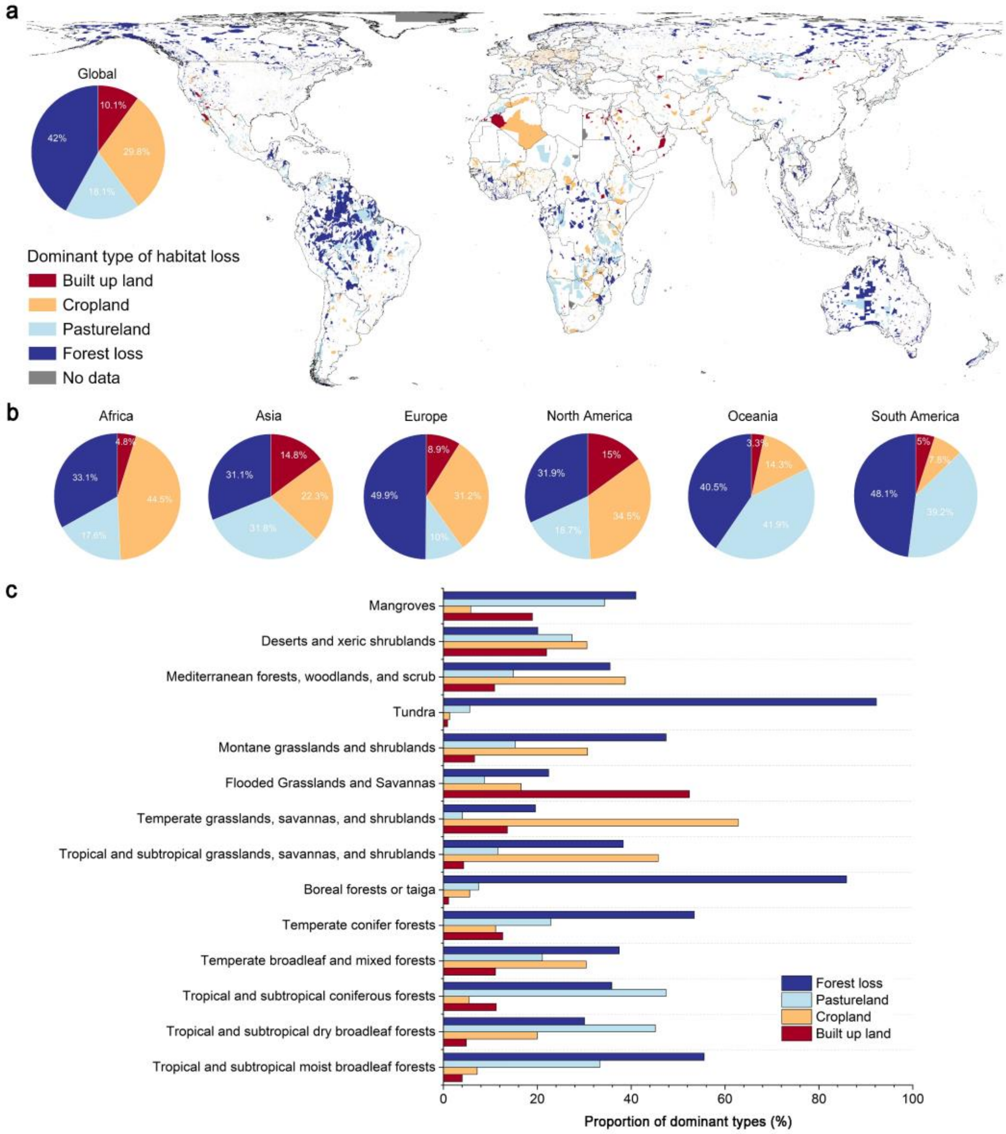

Fig. 7 | Primary types of threats to habitat loss for PAs. a, Four dominant types of habitat loss or threats in PAs characterized by the largest proportion of habitat loss type within a PA. Pie chart shows the proportion of PAs under different primary threats. **b**, The proportion of PAs under different primary threats for six continents. **c**, The proportion difference between the biomes.

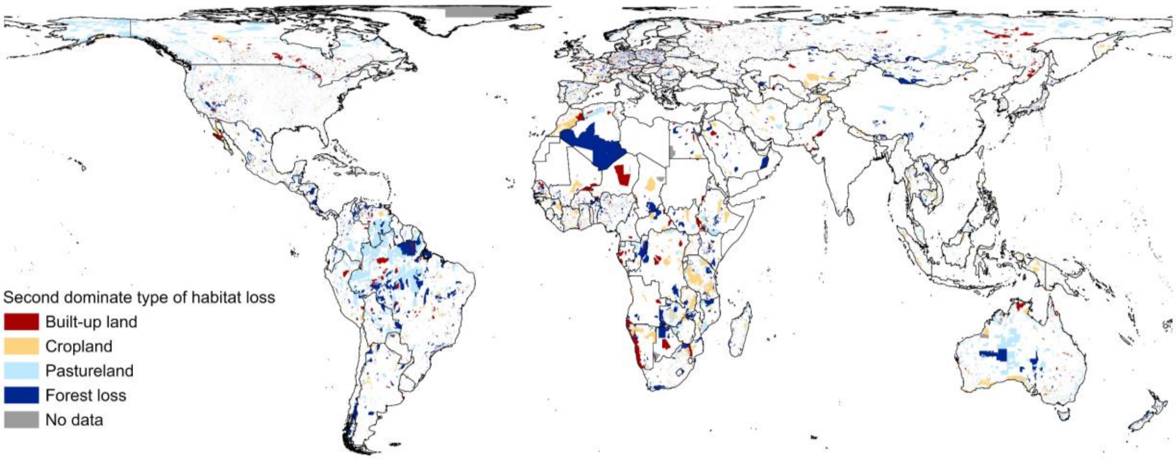

Fig. S14. Second dominate type of habitat loss in each PA.

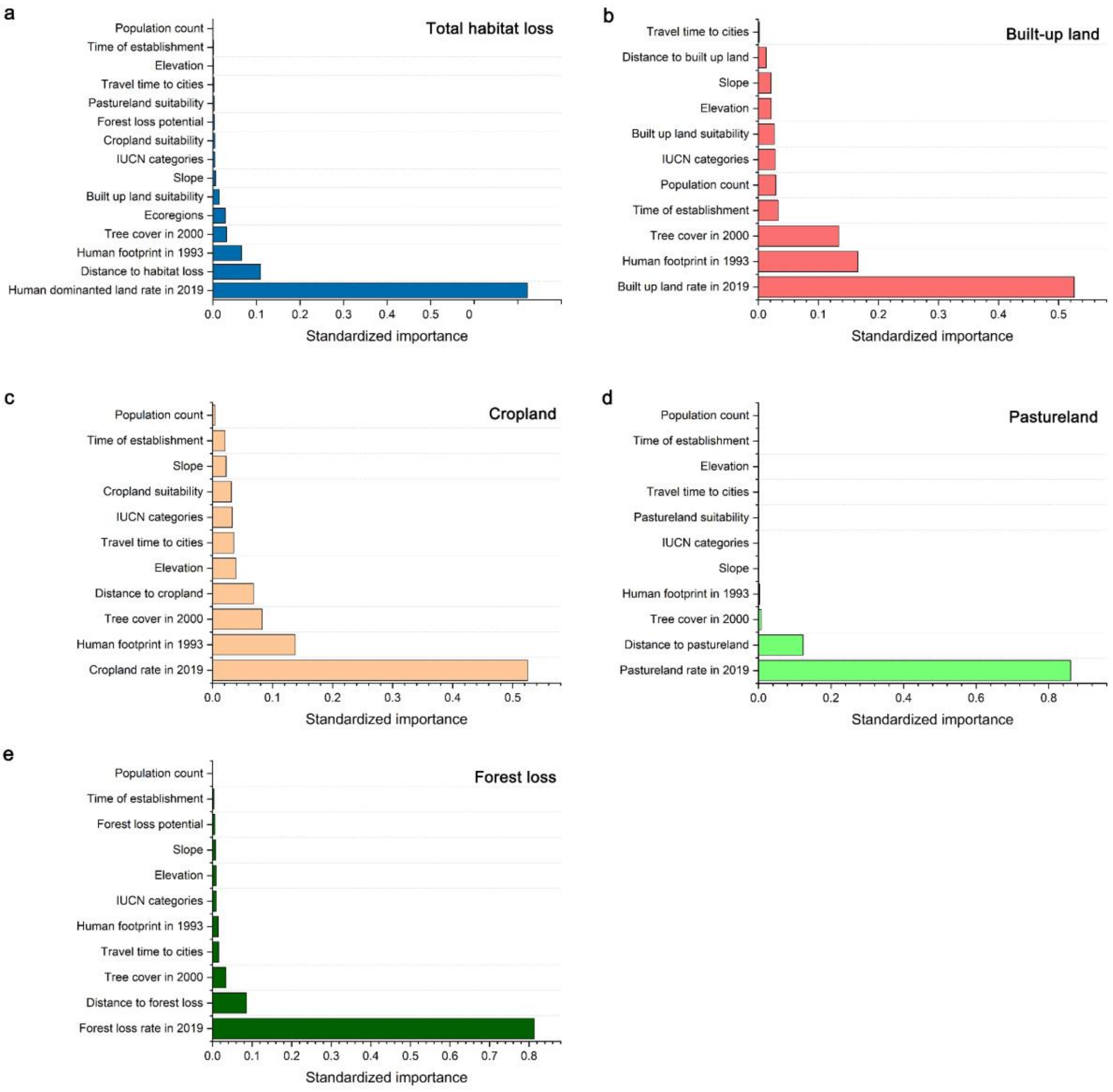

Fig. S12. Importance of variables in explaining different type of habitat loss.

(4) Finally, we revealed the influencing factors of protected area spatial effectiveness, including current habitat loss proportion (in 2019), distance to habitat loss events, human footprint in 1993, tree cover in 2000, ecoregion, elevation, slope, PA IUCN categories, travel time to cities, PA establishment time, population count, and land use suitability (Fig. S12). We found that path dependence and spatial spillover effects have a significant impact on the effectiveness of protected areas, and human footprint and tree cover are also significant variables. However, the management type of protected areas did not play a statistically significant role in explaining change within the protected area.

Similarly, the theoretical contribution is not clear? The authors talk about using an integrated metacoupling framework. However, what this framework is, how it is used and applied, and how it helps us to better understand the effectiveness of protected areas isn't clear.

Response:

We thank the reviewer for their constructive comments. We have clarified these issues in the Introduction (lines 103-114). The integrated metacoupling framework provides a theoretical conceptual foundation for our quantitative analysis. It treats a PA as an 'open system'. Besides the impact of human activities within a PA, the framework explicitly considers the impacts of a PA's immediate surroundings and distant regions. In other words, the framework enables us to have a more comprehensive understanding of the effectiveness of PAs than being limited to internal factors. Thus, our quantitative analysis integrates both internal factors (e.g. establishment time of the PA, governance type) and external factors near and far such as the PA's distances to cities. The results can help develop more effective strategies for PA management than considering only internal factors.

I feel this paper has potential given the amount of data that the authors have collated. However, I believe that it requires substantial reframing for it to represent a significantly novel contribution to our understanding.

Response:

We thank the reviewer for their positive comments. We have reframed the manuscript to emphasize the heterogeneity of protection at different scales in PAs, identify differences between different types of PAs, and distinguish the primary threats faced by each PA. In addition, we have also utilized machine learning models to reveal the importance of different factors in influencing the effectiveness of PAs. These analyses were not conducted before and have provided unprecedented insights. For example, we estimated the spatial effectiveness of each protected area, which can significantly enhance the targeting of relevant policies, making policies for PAs more targeted and goal-oriented.

Reviewer #2 (Remarks to the Author):

Dear Authors,

I appreciate the detailed response to the comments. This revised version is clearer, and no additional comments from me.

Response:

We thank the reviewer for their positive comments.

Reviewers' Comments:

Reviewer #1:
Remarks to the Author:
I’d like to thank the authors again for taking my comments so seriously and trying to address them. However, my concerns still stand and I am sorry for being that stubborn reviewer.

The authors’ main claim to novelty is bigger and better data – this is emphasised throughout the paper repeatedly. However, little is done to highlight what this bigger and better data allows them to do. The heterogeneity effects really add depth to our understanding but little is done to frame the introduction in terms of what bigger and better data allow us to do. In other words, what novel question can we answer with these data? Similarly, the authors highlight that their meta-coupling framework links PA analyses to more distant threats, yet the covariates used are similar to the covariates used in all other PA evaluations and so it’s still unclear to me what this new theoretical concept brings to the empirical analysis. Are there novel covariates that they are able to select, measure and link to treatment selection?

One final concern that I have, and I am sorry I did not picked it up earlier – but their distinction of drivers of PA habitat loss are to a certain degree overlapping – doesn’t urbanisation, cropland and pastureland expansion lead to deforestation (in forested areas)?

I truly appreciate the efforts that the authors have made but I still think that the novelty of the paper is not sufficiently highlighted and I cannot recommend it for publication.

Reviewer #1 (Remarks to the Author):

I'd like to thank the authors again for taking my comments so seriously and trying to address them. However, my concerns still stand and I am sorry for being that stubborn reviewer.

Response:

We thank the reviewer for their continued engagement with our manuscript and for their thoughtful feedback. We appreciate the time and effort the reviewer has dedicated to reviewing our work. We understand that some of their concerns remain, and we have carefully re-examine the areas they have highlighted. The reviewer's feedback is crucial to us, and we are committed to making the necessary improvements to address their concerns fully.

The authors' main claim to novelty is bigger and better data – this is emphasised throughout the paper repeatedly. However, little is done to highlight what this bigger and better data allows them to do. The heterogeneity effects really add depth to our understanding but little is done to frame the introduction in terms of what bigger and better data allow us to do. In other words, what novel question can we answer with these data? Similarly, the authors highlight that their meta-coupling framework links PA analyses to more distant threats, yet the covariates used are similar to the covariates used in all other PA evaluations and so it's still unclear to me what this new theoretical concept brings to the empirical analysis. Are there novel covariates that they are able to select, measure and link to treatment selection?

Response:

We thank the reviewer for their constructive comments. We have added relevant discussions in the introduction to address this concern (see lines 72-74). Compared to previous studies, our use of bigger and better data contributes to three key areas:

1. Higher Resolution Data: We employed high-resolution global data at 30-meter resolution, which significantly reduces the data uncertainty caused by insufficient resolution and enhances the accuracy of our analysis. This represents a 33 times improvement over the 1km^2 resolution datasets and more than a million-fold improvement over the 77 km^2 resolution used in the PNAS paper by Geldmann et al. (2019). By doing so, we greatly mitigate the issue of mixed pixels in land cover data and human pressures, allowing us to obtain more precise information on land cover and human pressures within and outside protected areas, thereby reducing uncertainty.

2. Comprehensive Protected Area Data: To the best of our knowledge, we used the largest dataset of protected areas available, which offers a significant advantage in comparing the effectiveness of protected areas across different sizes. While previous studies primarily focused on protected areas larger than 5 km^2 or 10 km^2, our research fills the gap in global studies by evaluating the effectiveness of smaller protected areas. Specifically, we are able to analyze protected areas ranging from 900 m^2 to 5 km^2, encompassing approximately 128,000 protected areas. Excluding these protected areas would omit a substantial portion of the data, likely leading to biased results.

3. Broader and Deeper Insights: The use of larger and better data enables us to obtain insights that were previously unattainable, providing international organizations and governments with more comprehensive and holistic information on the effectiveness of protected areas. Unlike earlier studies that focused on the loss of a single type of habitat, such as cropland or forest, our study recognizes that multiple types of habitat loss often occur simultaneously and exhibit different characteristics. Focusing solely on one type of habitat loss may lead to misleading conclusions, whereas providing decision-makers with a more integrated and comprehensive dataset offers better guidance for policy-making.

Regarding the second question about the innovation of the meta-coupling framework, we would like to clarify our position:

1. We believe that our framework indeed offers a novel framework to analyzing the effectiveness of PAs. We would like to emphasize that, to the best of our knowledge, no one has previously applied the meta-coupling theoretical framework, as developed by one of our co-corresponding authors, Jianguo Liu, to the evaluation of PA effectiveness. We applied the meta-coupling theoretical framework specifically because it is crucial to highlight that the effectiveness of PAs is influenced not only by factors within the PAs themselves but also by pressures from surrounding areas and distant regions. The framework allows us to emphasize that PA effectiveness is interconnected with broader, more complex influences beyond the local environment.

2. The meta-coupling theoretical framework is applied through the following dimensions in the paper specifically reflecting telecoupling, pericoupling, and intracoupling:

(1) Telecoupling (Global and Distant Interactions):
We used this framework to analyze how economic centers and human pressures from distant regions influence the effectiveness of PAs. For instance, the distance to cities and the travel time to cities are considered as critical factors influencing the extent of habitat loss within PAs. This reflects the global-to-local interactions where distant markets and economic centers exert pressure on local ecosystems, impacting the conservation outcomes within PAs.

(2) Pericoupling (Regional and Nearby Interactions):
The framework is employed to study the impact of nearby habitat loss, such as urban expansion and cropland encroachment, on the effectiveness of PAs. The research includes analyses of habitat loss within buffers around PAs, highlighting how regional land use changes near the boundaries of PAs can lead to habitat compression and loss within these areas. This emphasizes the role of nearby human settlements and land use in influencing the effectiveness of PAs.

(3) Intracoupling (Internal Dynamics within PAs):
The framework is applied to assess the internal management and governance factors that affect the effectiveness of PAs. The study looks at how internal characteristics of PAs, such as governance type, size, and establishment time, contribute to their ability to curb habitat loss. It reflects how the internal dynamics of PAs, including their management practices and internal regulations, interact with external pressures to determine their overall effectiveness.

These applications of the meta-coupling framework demonstrate its utility in providing a comprehensive understanding of the multiple scales of interactions—both within and beyond the boundaries of PAs—that influence their conservation effectiveness. This approach allows the research to address the complexity of human-nature interactions and the varying impacts these have on biodiversity conservation efforts.

3. We appreciate the reviewer's suggestion to consider adding new covariates to enhance the innovation of our study. However, we believe that introducing additional covariates may not necessarily increase the novelty of our work. Through our analysis, we have already identified the most critical factors within the existing set of covariates. In particular, we found that path dependence is a crucial factor influencing the spatial effectiveness of PAs. Regardless of the overall habitat loss rate or the specific types of habitat loss, the rate of current habitat loss within PAs consistently emerges as the most significant explanatory variable. This finding suggests that the most effective strategy for enhancing PA effectiveness is to prioritize reducing habitat loss within the PAs themselves. Furthermore, the distance to habitat loss events within PAs also plays a key role, as indicated by potential spatial spillover effects where habitat loss in one area can trigger chain reactions in surrounding areas. Baseline human footprint and forest cover are additional important factors, reflecting how the intensity of early human activities and natural endowments impact the current effectiveness of PAs. Interestingly, the management categories of PAs did not demonstrate the expected level of importance, ranking only mid-tier in terms of influence. Population count within PAs significantly affects built-up land expansion, but has limited impact on other types of land use changes. For forest loss, travel time to cities is a critical factor due to its correlation with market accessibility, which is crucial for logging activities. Through these analyses, we demonstrated that remote factors currently have a significant impact only on forest loss, indicating that remote factors may not be the most critical determinants of protected area effectiveness.

Given these findings, we believe our study's innovation lies in the application and interpretation of existing covariates within our meta-coupling framework, rather than in the addition of new ones. We have also clarified this point in the discussion section.

One final concern that I have, and I am sorry I did not picked it up earlier – but their distinction of drivers of PA habitat loss are to a certain degree overlapping – doesn't urbanisation, cropland and pastureland expansion lead to deforestation (in forested areas)?

Response:

We thank the reviewer for raising this important concern. We would like to clarify that in the Methods section of our manuscript, we have already addressed this issue. Specifically, we have defined forest loss as a distinct type of habitat loss, which excludes cases where forests were converted to built-up lands, croplands, or pasturelands. The remaining forest loss includes transitions to natural grasslands, shrublands, bare land, and other land uses.

P19 lines 625-637: "We used four datasets to characterize habitat loss (including four types: conversion to built-up land, cropland and pastureland, and forest loss) in PAs. These datasets have high spatial resolution and spatial consistency, which facilitate long-term and fine-scale evaluation. **However, overlaying these datasets can lead to double counting of habitat loss in PAs, particularly since forest loss may lead to conversion into any of the other three types of habitat**

loss. To address this issue, we combined the four types of habitat loss to produce a global 30-m resolution habitat loss dataset for 2003-2019. We used the 2003-2019 time period mainly because the cropland dataset covered only that period, which was further divided into five shorter time periods that were not interannual (2000-2003, 2004-2007, 2008-2011, 2012-2015, and 2016-2019). **We used the ArcGIS spatial overlay method to address potential double-counting among these four datasets. We first decomposed forest loss into different types for each pixel that overlapped with the other three types of habitat loss. Then, if there was an overlapping pixel, we determined the final type of this pixel in order of priority, i.e., built-up land, cropland, pastureland, and forest loss. This approach effectively prevents the double counting of forest loss and land cover conversions.**"

By clearly distinguishing these categories, we have ensured that there is no overlap in the analysis of the drivers of habitat loss within PAs. Therefore, the concern about overlapping drivers, such as urbanization, cropland, and pastureland expansion leading to deforestation, has been carefully considered and accounted for in our methodology.

I truly appreciate the efforts that the authors have made but I still think that the novelty of the paper is not sufficiently highlighted and I cannot recommend it for publication.

Response:

We thank the reviewer for recognizing the efforts we have made. In our study, we have utilized larger and more comprehensive datasets, which have allowed us to uncover findings that were previously unobserved, particularly in assessing the effectiveness of small-scale protected areas on a global scale. To our knowledge, we are the first to apply causal inference methods to examine the impact of protected area establishment on habitat changes both before and after their designation. This approach goes beyond the traditional focus on spatial effectiveness comparisons between protected and non-protected areas, although we have also conducted spatial effectiveness analyses in our study.

Moreover, we have performed extensive analyses on the heterogeneity of protected area effectiveness, offering valuable insights that can guide decision-makers in different regions around the world. These contributions represent significant advancements in the field of conservation. As outlined above, we have added new text to make clear the novelty of this research.