

# Protected areas have a mixed impact on waterbirds, but management helps

<https://doi.org/10.1038/s41586-022-04617-0>

Received: 27 June 2021

Accepted: 4 March 2022

Published online: 20 April 2022

 Check for updates

Hannah S. Wauchope<sup>1,2✉</sup>, Julia P. G. Jones<sup>3</sup>, Jonas Geldmann<sup>1,4</sup>, Benno I. Simmons<sup>1,2</sup>, Tatsuya Amano<sup>5,6</sup>, Daniel E. Blanco<sup>7</sup>, Richard A. Fuller<sup>4</sup>, Alison Johnston<sup>8,9</sup>, Tom Langendoen<sup>10</sup>, Tiej Mundkur<sup>10✉</sup>, Szabolcs Nagy<sup>10</sup> & William J. Sutherland<sup>1</sup>

International policy is focused on increasing the proportion of the Earth's surface that is protected for nature<sup>1,2</sup>. Although studies show that protected areas prevent habitat loss<sup>3–6</sup>, there is a lack of evidence for their effect on species' populations: existing studies are at local scale or use simple designs that lack appropriate controls<sup>7–13</sup>. Here we explore how 1,506 protected areas have affected the trajectories of 27,055 waterbird populations across the globe using a robust before–after control–intervention study design, which compares protected and unprotected populations in the years before and after protection. We show that the simpler study designs typically used to assess protected area effectiveness (before–after or control–intervention) incorrectly estimate effects for 37–50% of populations—for instance misclassifying positively impacted populations as negatively impacted, and vice versa. Using our robust study design, we find that protected areas have a mixed impact on waterbirds, with a strong signal that areas managed for waterbirds or their habitat are more likely to benefit populations, and a weak signal that larger areas are more beneficial than smaller ones. Calls to conserve 30% of the Earth's surface by 2030 are gathering pace<sup>14</sup>, but we show that protection alone does not guarantee good biodiversity outcomes. As countries gather to agree the new Global Biodiversity Framework, targets must focus on creating and supporting well-managed protected and conserved areas that measurably benefit populations.

Protected areas have been the cornerstone of conservation practice for more than a century. Nearly 16% of land and 7% of the ocean are now designated as protected areas<sup>15</sup>, and there are prominent calls for the Convention on Biological Diversity to set an area-based target of 30% coverage for protected areas and other effective area-based conservation measures by 2030<sup>2</sup>. Given the importance to humanity of addressing biodiversity loss<sup>16</sup>, it is crucial that the next decade's biodiversity conservation targets are informed by evidence of the most effective conservation strategies and actions<sup>3,17</sup>.

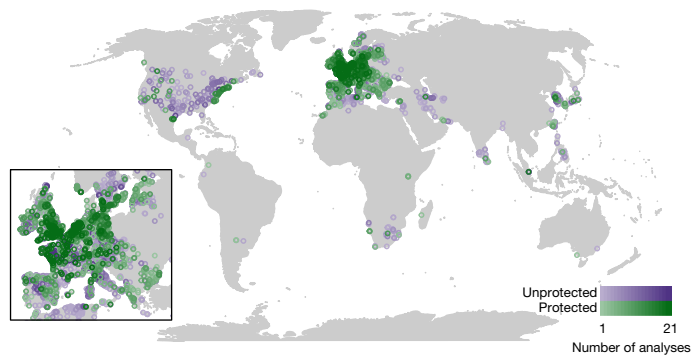
Optimizing where protected areas are placed to most efficiently conserve species and their habitats has been a major research theme in conservation science for decades<sup>18</sup>. However, until recently, robust attempts (those making an explicit effort to account for confounding factors) to evaluate the performance of protected areas have been lacking<sup>19,20</sup>. A number of studies have shown that protected areas slow habitat loss, particularly in forests<sup>3–6</sup>, however intact habitat does not guarantee the health of populations<sup>21</sup>. Studies attempting to address this problem by quantifying the impact of protected areas on population health and persistence have suffered from a lack of suitable controls<sup>19</sup>. To accurately estimate the impact of a protected area, it is

necessary to understand what would have happened in the absence of protection<sup>22</sup> and most studies do this by using either before–after or control–intervention study designs. Before–after studies compare populations pre- and post-protected area designation<sup>7,13</sup>, but cannot ascertain whether the observed difference was caused by the protected area or other factors that changed in the same time period. Control–intervention studies compare populations between protected and unprotected sites<sup>8–12</sup>, but cannot ascertain whether the observed difference was due to the effectiveness of the protected area, or because it was placed where populations were performing well to begin with.

Combining these designs into a before–after control–intervention (BACI) framework—where populations in protected and unprotected sites are compared before and after the date of protected area designation—can overcome these limitations<sup>23</sup>, and even approximate causality<sup>24</sup>. The recent emergence of large biodiversity databases in ecology provides an opportunity to test the effects of protected areas on populations under a BACI framework.

Here, using one of the largest global datasets of bird population counts, compiled from citizen science initiatives and non-governmental organization (NGO) and government-led monitoring programmes in 68

<sup>1</sup>Conservation Science Group, Department of Zoology, University of Cambridge, Cambridge, UK. <sup>2</sup>Centre for Ecology and Conservation, College of Life and Environmental Sciences, University of Exeter, Cornwall Campus, Penryn, UK. <sup>3</sup>School of Natural Sciences, College of Engineering and Environmental Science, Bangor University, Bangor, UK. <sup>4</sup>Center for Macroecology, Evolution and Climate, GLOBE Institute, University of Copenhagen, Copenhagen, Denmark. <sup>5</sup>School of Biological Sciences, University of Queensland, Brisbane, Queensland, Australia. <sup>6</sup>Centre for Biodiversity and Conservation Science, University of Queensland, Brisbane, Queensland, Australia. <sup>7</sup>Wetlands International LAC Argentina Office, Buenos Aires, Argentina. <sup>8</sup>Cornell Lab of Ornithology, Ithaca, NY, US. <sup>9</sup>Centre for Research into Ecological and Environmental Modelling, University of St Andrews, St Andrews, UK. <sup>10</sup>Present address: Wetlands International, Ede, The Netherlands. ✉e-mail: hannah.wauchope@gmail.com; tiej.mundkur@wetlands.org



**Fig. 1 | Map of study sites.** Locations of protected (green;  $n = 1,506$ ) and unprotected (purple;  $n = 3,343$ ) sites used across analyses. Darker colours mean a given site was used in a greater number of analyses, to a maximum of 21 (our focal analysis and 20 full parameter analyses; there are 864 protected sites in the focal analysis). See Supplementary Information for a map showing only the sites used in the focal analysis.

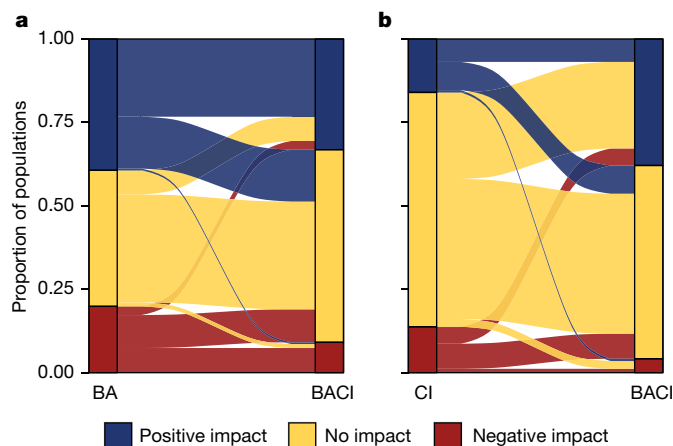
countries, we present the first robust, global-scale assessment of protected area impact on populations. We examined how 1,506 protected areas have impacted the population trajectories of 27,055 waterbird populations, where ‘population’ is defined as a particular species at a particular site (Fig. 1). Waterbirds are an appropriate taxonomic group with which to explore impact, given their broad distribution and ability to respond rapidly to changes in site quality<sup>25</sup>. We asked three questions: (1) how much do the study designs typically used to assess protected area effectiveness cause misleading conclusions, compared with a BACI study design?; (2) what is the impact of protected areas on waterbird populations?; and (3) what factors contribute to protected area impact?

We estimated impact using before–after, control–intervention and BACI study designs. For BACI and control–intervention analyses, we matched protected populations to similar unprotected populations using a combination of exact matching and Mahalanobis distance matching (see Methods). We considered the wide range of ways in which populations may respond to protection by counting cases where local immigrations or extinctions had occurred and using generalized linear models to assess both immediate changes in population numbers and longer-term changes in population trend (an extension of the traditional BACI study design that considers only average change in population size<sup>24</sup>). We used these measures to classify populations into three broad groups: positive, negative, or no impact from protection (the full range of population responses and how they were classified are described in Fig. 3, Extended Data Figs. 3, 4).

To explore the sensitivity of our results to different parameter decisions (such as years of sampling required, the maximum geographical distance between sites, or the strictness of Mahalanobis matching), we ran our entire analysis 21 times: one ‘focal analysis’ using our best guess parameter estimates, and 20 analyses using estimates sampled from a plausible range for each parameter (full parameter analyses) (Methods, Extended Data Table 1).

### Before–after and control–intervention estimates

Estimates of protected area effectiveness varied markedly on the basis of study design, and studies using before–after or control–intervention designs can lead to highly misleading conclusions. In our focal analysis, 37% of populations analysed using a before–after design, and 50% of populations analysed using a control–intervention design, had different outcomes to those in the BACI analysis (Fig. 2). These changes were not simply a result of BACI detecting positive or negative signals where other designs could not: 41% (before–after) and 57% (control–intervention) of populations that were apparently positively impacted were shown to be not impacted, or even negatively impacted



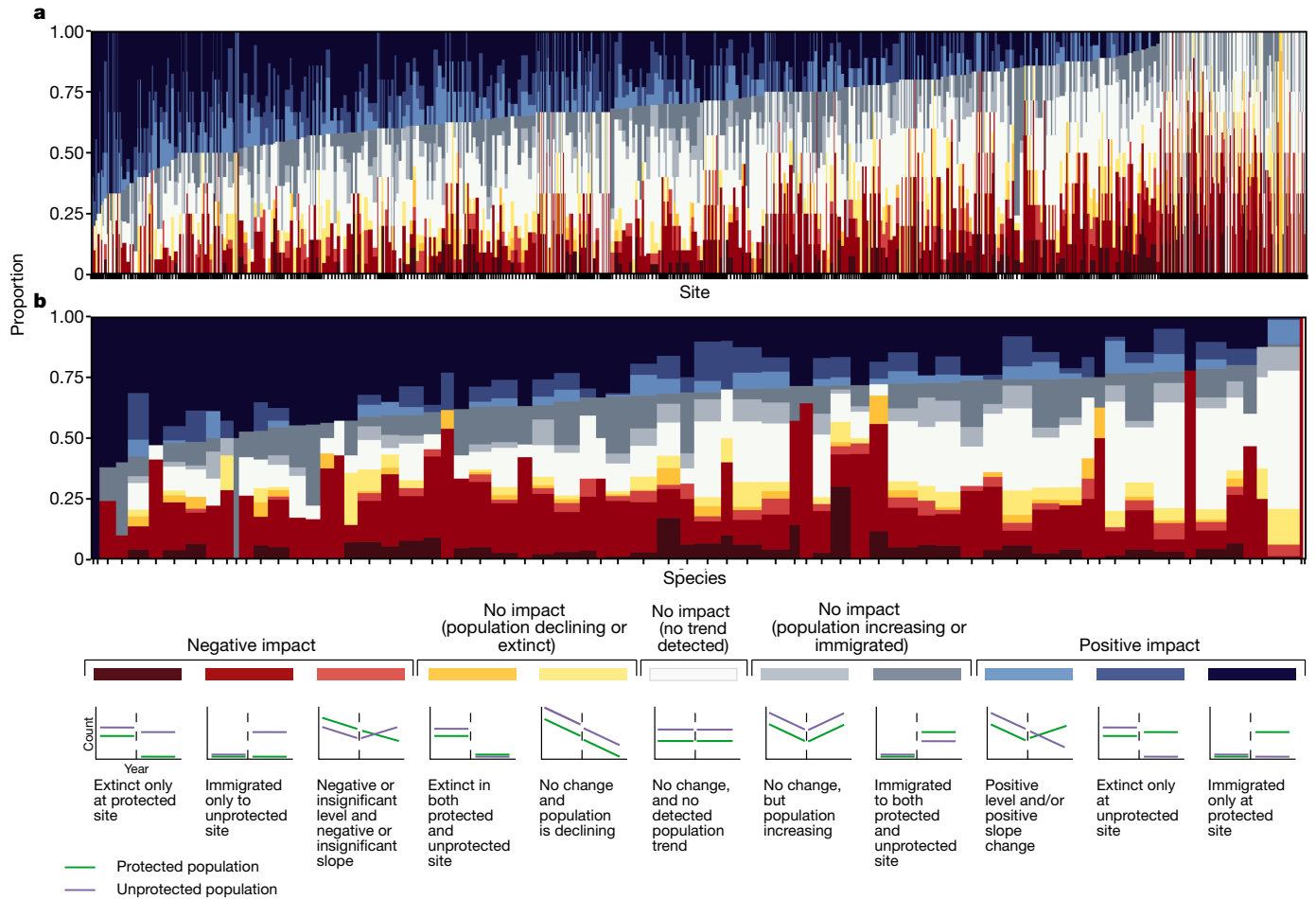
**Fig. 2 | Changes in estimates of protected area impact under different study designs.** **a, b.** The change in protected area impact when estimated under a before–after (BA) versus BACI framework (**a**) or a control–intervention (CI) versus BACI framework (**b**). y-axes show the proportion of populations in each category under before–after or control–intervention on the left, and BACI on the right. The shift of each colour shows how our estimate of the impact of protected areas on populations change between study designs. Note that these figures only contain populations where we could obtain both before–after and BACI ( $n = 6,006$ ) or control–intervention and BACI ( $n = 3,609$ ) estimates of protected area effectiveness. This figure is based on our focal analysis, Extended Data Fig. 1 shows changes in outcome across all full parameter analyses.

under a BACI analysis (Fig. 2). Changes to negative impacts were even more substantial, with 63% (before–after) and 92% (control–intervention) of apparently negatively impacted populations shown to be not impacted or positively impacted by protection under a BACI analysis (Fig. 2). The findings from our full parameter analyses were similar (Extended Data Fig. 1). Before–after models were also heavily impacted by regression to the mean (Supplementary Information 5), an additional reason to consider them unreliable. These results show that relying on before–after or control–intervention studies can distort the picture of a protected area’s impact.

### BACI estimates of protected area impact

We found a mixed impact of protected areas on populations when using a BACI approach. Within nearly all sites, populations showed a range of responses from positive to negative (in the focal analysis the proportion of positively impacted populations within a site was  $0.25 \pm 0.21$  (mean  $\pm$  s.d.) with a range of values from 0 to 1; Fig. 3a). Impacts on populations were similarly variable when grouped by species (in the focal analysis the proportion of positively impacted populations within a species was  $0.36 \pm 0.17$  with a range from 0 to 1; Fig. 3b). In our focal analysis, 27% of all populations were positively impacted by protected areas (blues), 21% were negatively impacted (reds) and for 48% we could detect no impact of protection (greys, white and yellows) (our full parameter analyses produced similar results; Extended Data Fig. 2). Four per cent of populations were excluded because of model failure. Of the 48% of cases where we could not detect any difference between protected and unprotected populations, 85% (41% of all populations; whites and greys, Fig. 3) were increasing, or had no trend. These cases are difficult to define as a success or failure as, while the protected area did not have a demonstrably positive impact compared to an unprotected area, the protected population appears to be healthy.

Regardless, over a quarter of populations showed a negative response (Fig. 3). These are formed from two groups: (1) negatively impacted populations, that is, those that perform worse in protected areas relative to matched controls (21%, reds) and (2) populations for which there was no



**Fig. 3 | Estimates of protected area impact under a BACI study design.** **a, b**, The proportion of populations ( $n = 7,313$ ) showing various responses to protection, according to site (**a**;  $n = 864$ ) and species (**b**;  $n = 67$ ), when calculated in a BACI framework. Each vertical bar comprises species or site, and the proportions of the populations in each category are shown on the y-axis. Bar width is scaled to the number of populations of that species or site in the

dataset (log scaled in the case of species) with a wider bar indicating that the species or site has more populations. Each colour represents a different way in which a population can respond to protection, and an example of each response is shown at the bottom. This figure is based on our focal analysis; Extended Data Fig. 2a shows the proportion of populations within each broad outcome category across all full parameter analyses.

positive or negative signal of protection and which were either declining in protected areas at a similar rate to unprotected populations, or where both protected and unprotected populations went locally extinct (7%, yellows). Of note, half of these negative responses (14% of populations overall), do not occur in sites designated for waterbirds or their habitat (that is, Ramsar Sites<sup>26</sup> or Special Protected Area–Birds Directive<sup>27</sup> sites) and so we might not necessarily expect a positive impact in these cases and thus should not consider these to be cases where protected areas have not worked.

We consider protected area impact exclusively in the context of how protected areas support the persistence of populations, which ignores the potential benefit of protection on the maintenance of the habitats in which these populations occur. Our dataset was restricted to sites where monitoring occurred: if habitat change meant that waterbirds were no longer found at a site, monitoring would likely cease<sup>28</sup>. Thus, we could not consider such sites as counterfactuals, and so could not account for protected areas having prevented complete habitat conversion. We also do not consider the potential for protected areas to defend against future threats, for instance, protecting a future climate refuge. In sum, it is important to remember that the results presented here about the impact of protected areas on populations are above and beyond these already-known benefits<sup>3–5,29,30</sup>.

Our results are also likely to underestimate the positive impact of protection as we were restricted to species for which we were able to

obtain adequate matches between protected and unprotected populations, resulting in a bias towards common species (Supplementary Information 10). Common species tend to have more generalist habitat requirements<sup>31</sup> and so may fare better in degraded sites than rarer species. They are also less likely to be the target of specific interventions, which in some cases could actively impede them; for instance, water could be kept at levels appropriate for rare waders, but not for common ducks. To explore whether this affected our results, we assessed whether outcomes varied between regionally threatened and non-threatened species in Europe (Supplementary Information 11; a global analysis was not possible due to data restrictions). We did not find any differences in the impact of protected areas between these groups, possibly because there was only a small set of threatened species in our data, though a recent study<sup>32</sup> similarly found no difference between rare and common species when studying population trends.

### Predictors of protected area impact

We show that the mere designation of a protected area does not necessarily bring benefits to populations. Given this, we used cumulative link mixed models, where the response variable was the impact (positive, no or negative), to investigate which species and protected area characteristics predict outcomes for populations, on the basis of our BACI framework (see Fig. 4). The models had random intercepts for

country, site, species and spatial grid cell. Our explanatory variables included a management variable, which broadly categorized sites as either ‘waterbird-managed’ (Ramsar or Special Protected Area–Birds Directive sites), or ‘mixed-management’ (sites either not designated for waterbirds or their habitat, or of unknown management status).

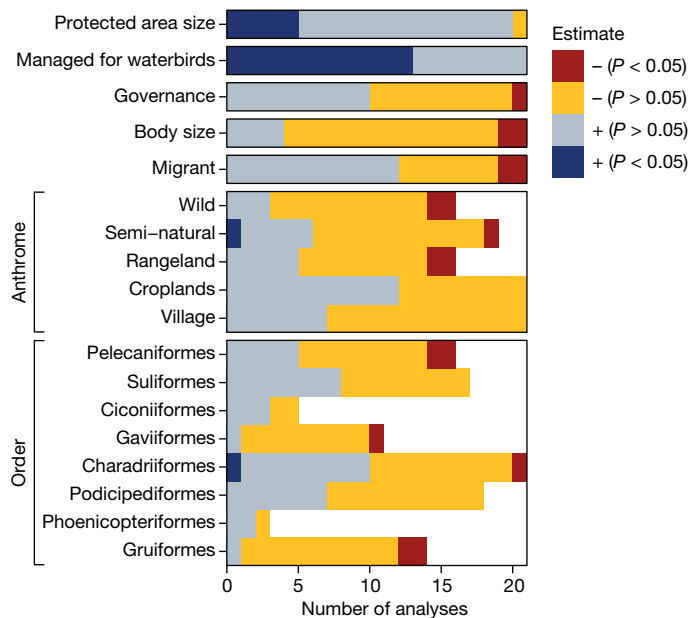
Management for waterbirds was consistently positively correlated with protected area success (Fig. 4). Larger protected areas were also almost always positively correlated with success, although significantly so in only a few analyses (Fig. 4). No other site or species-based characteristic was consistently positively or negatively associated with success (Fig. 4; Extended Data Fig. 5). Depending on the analysis, a large, waterbird-managed area could increase the likelihood of a positive impact on a population anywhere from 1 to 25 percentage points (mean weighted by model confidence = 9 percentage points; see Supplementary Information 13) compared to a small, mixed-management area.

These values are likely to underestimate the positive impact of management. Our classification of sites into waterbird-managed sites and mixed-management sites is a simple metric of diverse on-the-ground practices (a more nuanced classification was not possible at the global scale) and, inevitably, some mixed-management sites are likely to be managed for waterbirds, while management quality will vary within waterbird-managed sites<sup>33,34</sup>. Both these factors would reduce the observed difference between the two management classifications, meaning the true difference is likely higher. That waterbird-managed sites perform better emphasizes the need for effective management to avoid negative outcomes, and suggests that policy needs to focus on setting and adhering to ambitious management targets.

The weak positive association between protected area size and impact adds a new element to the ‘single large or several small’ protected area debate that considers which is better for conserving biodiversity. Studies have agreed that several smaller protected areas typically provide higher species richness than a few large areas<sup>35</sup>, but that larger areas are critical for persistence of larger species<sup>36</sup>. Our results demonstrate some importance of larger protected areas for supporting populations of waterbirds through time. This is concerning given many protected areas across the world are small and many are currently being downsized<sup>37</sup>.

While our analysis included data from 68 countries across 6 continents, the data are biased towards Europe, North America and East Asia; a common problem in large-scale ecological studies<sup>38</sup>. There are a number of initiatives in less-studied areas of the world to increase the supply and quality of ecological data<sup>39</sup> (for example, <https://african-conservation.org>, <https://www.avesargentinas.org.ar>, <https://www.birdscaribbean.org>, and <https://www.amazonteam.org/brazil/>); supporting and incorporating efforts such as these will be vital to informing truly global evaluations of conservation effectiveness.

Our results show a mixed impact of protected areas, supporting concerns raised over protected area efficacy in recent years<sup>40,41</sup>. We had expected that, given their ability to move between sites<sup>25</sup>, waterbirds would show a more immediate and positive signal of protection than other non-mobile taxa, such as reptiles, where positive signals might not be apparent until multiple generation cycles of improved breeding rate had occurred. The lack of signal could be due to poor or limited management of many protected areas, or it could be due to forces that cannot be controlled within the borders of a protected area. Waterbirds rely on water, and threats such as pollution, upstream dam installation and sea level rise cannot be managed by a protected area, and can have devastating consequences<sup>42–44</sup>. Terrestrial taxa will be less impacted by such threats and therefore may experience more positive responses to protection<sup>45</sup>, although beyond border threats are not limited to those affecting water: climate change, air pollution and disease have the potential to impact all species<sup>45</sup>. Finding solutions to conserving species in the face of these more ubiquitous threats is a key conservation challenge.



**Fig. 4 | Predictors of protected area impact.** Number of analyses (20 full parameter analyses plus one analysis with focal parameter estimates) that found significantly positive (+) or negative (-) (blue or red, respectively) or insignificantly positive or negative (grey or yellow, respectively) relationships between various predictors and protected area impact. Orders are measured relative to Anseriformes, and anthromes are measured relative to urban. For odds ratios of each estimate and confidence intervals, see Extended Data Fig. 5.

## Conclusions

The parties to the UN Convention on Biological Diversity will soon decide on the post-2020 Global Biodiversity Framework, which will set nature conservation policy for the decade ahead. It is likely to include a commitment to protect and conserve 30% of Earth protected by 2030 (and there are growing calls for this to reach 50% by 2050<sup>14</sup>). Researchers have warned that such calls must consider the social and political context in which conservation operates, or risk undermining conservation support<sup>46</sup>. Our results raise additional concerns about the ‘30 by 30’ approach by showing that protection alone does not guarantee optimal biodiversity outcomes. Halting biodiversity loss requires improvements to the performance of existing protected areas, and action to address ubiquitous threats beyond area borders. Ever-increasing area-based targets must be accompanied by equally ambitious targets that ensure protected area effectiveness.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-022-04617-0>.

1. High Ambition Coalition for Nature and People. *50 Countries Announce Bold Commitment to Protect at Least 30% of the World’s Land and Ocean by 2030* (Campaign for Nature, 2021).
2. Waldron A. et al. *Protecting 30% of the Planet for Nature: Costs, Benefits and Economic Implications* (Campaign for Nature, 2020).
3. Geldmann, J., Manica, A., Burgess, N. D., Coad, L. & Balmford, A. A global-level assessment of the effectiveness of protected areas at resisting anthropogenic pressures. *Proc. Natl Acad. Sci. USA* **116**, 23209–23223 (2019).
4. Nelson, A. & Chomitz, K. M. *Protected Area Effectiveness in Reducing Tropical Deforestation* (The World Bank, 2009).
5. Scharlemann, J. P. W. et al. Securing tropical forest carbon: the contribution of protected areas to REDD. *Oryx* **44**, 352–357 (2010).

6. Feng, Y. et al. Assessing the effectiveness of global protected areas based on the difference in differences model. *Ecol. Indic.* **130**, 108078 (2021).
7. Laurance, W. F. et al. The fate of Amazonian forest fragments: A 32-year investigation. *Biol. Conserv.* **144**, 56–67 (2011).
8. Laurance, W. F. et al. Averting biodiversity collapse in tropical forest protected areas. *Nature* **489**, 290–294 (2012).
9. Terraube, J., Van doninck, J., Helle, P. & Cabeza, M. Assessing the effectiveness of a national protected area network for carnivore conservation. *Nat. Commun.* **11**, 2957 (2020).
10. Barnes, M. D. et al. Wildlife population trends in protected areas predicted by national socio-economic metrics and body size. *Nat. Commun.* **7**, 12747 (2016).
11. Amano, T. et al. Successful conservation of global waterbird populations depends on effective governance. *Nature* **553**, 199–202 (2018).
12. Kleijn, D., Cherkaoui, I., Goedhart, P. W., van der Hout, J. & Lammertsma, D. Waterbirds increase more rapidly in Ramsar-designated wetlands than in unprotected wetlands. *J. Appl. Ecol.* **51**, 289–298 (2014).
13. Reyes-Arriagada, R. et al. Population trends of a mixed-species colony of Humboldt and Magellanic Penguins in Southern Chile after establishing a protected area. *Avian Conserv. Ecol.* **8**, 13 (2013).
14. Bukart, K. Motion 101 passes at IUCN, calls for protecting 50% of Earth's lands and seas. *One Earth* <https://www.oneearth.org/motion-101-passes-at-iucn-calls-for-protecting-50-of-earths-lands-and-seas/> (2021).
15. *Protected Planet Report 2020* (UNEP-WCMC and IUCN, 2021).
16. *Global Assessment Report on Biodiversity and Ecosystem Services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services* (IPBES Secretariat, 2019).
17. Barnes, M. D., Glew, L., Wyborn, C. & Craigie, I. D. Prevent perverse outcomes from global protected area policy. *Nat. Ecol. Evol.* **2**, 759–762 (2018).
18. Pressey, R. L., Cabeza, M., Watts, M. E., Cowling, R. M. & Wilson, K. A. Conservation planning in a changing world. *Trends Ecol. Evol.* **22**, 583–592 (2007).
19. Geldmann, J. et al. Effectiveness of terrestrial protected areas in reducing habitat loss and population declines. *Biol. Conserv.* **161**, 230–238 (2013).
20. Rodrigues, A. S. L. & Cazalis, V. The multifaceted challenge of evaluating protected area effectiveness. *Nat. Commun.* **11**, 5147 (2020).
21. Redford, K. H. The empty forest. *BioScience* **42**, 412–422 (1992).
22. Ferraro, P. J. Counterfactual thinking and impact evaluation in environmental policy. *N. Direct. Eval.* **2009**, 75–84 (2009).
23. Adams, V. M., Barnes, M. & Pressey, R. L. Shortfalls in conservation evidence: moving from ecological effects of interventions to policy evaluation. *One Earth* **1**, 62–75 (2019).
24. Wauchope, H. S. et al. Evaluating impact using time-series data. *Trends Ecol. Evol.* **36**, 196–205 (2021).
25. Kingsford, R. T., Roshier, D. A. & Porter, J. L. Australian waterbirds time and space travellers in dynamic desert landscapes. *Mar. Freshw. Res.* **61**, 875–884 (2010).
26. The Ramsar Convention Secretariat. Managing Ramsar Sites. [ramsar.org](https://www.ramsar.org/sites-countries/managing-ramsar-sites) <https://www.ramsar.org/sites-countries/managing-ramsar-sites> (2014).
27. European Commission. The Birds Directive. [https://ec.europa.eu/environment/nature/legislation/birdsdirective/index\\_en.htm](https://ec.europa.eu/environment/nature/legislation/birdsdirective/index_en.htm) (accessed 3 April 2022).
28. Zhang, W., Sheldon, B. C., Grenyer, R. & Gaston, K. J. Habitat change and biased sampling influence estimation of diversity trends. *Curr. Biol.* **31**, 3656–3662.e3 (2021).
29. Bruner, A. G., Gullison, R. E., Rice, R. E. & da Fonseca, G. A. B. Effectiveness of parks in protecting tropical biodiversity. *Science* **291**, 125–128 (2001).
30. Carranza, T., Balmford, A., Kapos, V. & Manica, A. Protected area effectiveness in reducing conversion in a rapidly vanishing ecosystem: the Brazilian Cerrado. *Conserv. Lett.* **7**, 216–223 (2014).
31. Rabinowitz, D. In *The Biological Aspects of Rare Plant Conservation* (ed. Synge, H.) 205–217 (John Wiley & Sons, 1981).
32. Daskalova, G. N., Myers-Smith, I. H. & Godlee, J. L. Rare and common vertebrates span a wide spectrum of population trends. *Nat. Commun.* **11**, 4394 (2020).
33. Hettiarachchi, M., Morrison, T. H. & McAlpine, C. Forty-three years of Ramsar and urban wetlands. *Glob. Environ. Change* **32**, 57–66 (2015).
34. Munishi, P., Chuwa, J., Kilungu, H., Moe, S. & Temu, R. Management effectiveness and conservation initiatives in the Kilombero Valley Flood Plains Ramsar Site, Tanzania. *Tanzania J. For. Nat. Conserv.* **81**, 1–10 (2012).
35. Fahrig, L. Why do several small patches hold more species than few large patches? *Glob. Ecol. Biogeogr.* **29**, 615–628 (2020).
36. Newmark, W. D. Extinction of mammal populations in western North American National Parks. *Conserv. Biol.* **9**, 512–526 (1995).
37. Mascia, M. B. & Pailler, S. Protected area downgrading, downsizing, and degazettement (PADDD) and its conservation implications. *Conserv. Lett.* **4**, 9–20 (2011).
38. Di Marco, M. et al. Changing trends and persisting biases in three decades of conservation science. *Glob. Ecol. Conserv.* **10**, 32–42 (2017).
39. Wetlands International. Asian Waterbird Census. <https://south-asia.wetlands.org/our-approach/healthy-wetland-nature/asian-waterbird-census/> (accessed 3 April 2022).
40. Gill, D. A. et al. Capacity shortfalls hinder the performance of marine protected areas globally. *Nature* **543**, 665–669 (2017).
41. Geldmann, J. et al. A global analysis of management capacity and ecological outcomes in terrestrial protected areas. *Conserv. Lett.* **11**, e12434 (2018).
42. Kingsford, R. T., Bino, G. & Porter, J. L. Continental impacts of water development on waterbirds, contrasting two Australian river basins: global implications for sustainable water use. *Glob. Change Biol.* **23**, 4958–4969 (2017).
43. Jia, Q., Wang, X., Zhang, Y., Cao, L. & Fox, A. D. Drivers of waterbird communities and their declines on Yangtze River floodplain lakes. *Biol. Conserv.* **218**, 240–246 (2018).
44. Lehtikoinen, A., Rintala, J., Lammi, E. & Pöysä, H. Habitat-specific population trajectories in boreal waterbirds: alarming trends and bioindicators for wetlands. *Animal Conserv.* **19**, 88–95 (2016).
45. Boyd, C. et al. Spatial scale and the conservation of threatened species. *Conserv. Lett.* **1**, 37–43 (2008).
46. Schleicher, J. et al. Protecting half of the planet could directly affect over one billion people. *Nat. Sustain.* **2**, 1094–1096 (2019).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2022

We published a pre-analysis plan for this paper laying out our planned analysis before we looked in detail at the data<sup>47</sup>. Pre-analysis plans are useful to reduce the risk of cherry picking or hypothesizing after results are known, which has led to a replication crisis in science<sup>48</sup>. As much as possible, we have followed the methods we set out, however we discovered a number of factors we had not considered (for instance, the potential for immigrations and extinctions and the fact that both trend and immediate change must be considered<sup>24</sup>). The conceptual basis of our revised methodology is described in detail elsewhere<sup>24</sup>, and Supplementary Information 7 describes the choices we have made that deviate from the pre-analysis plan and why.

### Overview

A brief summary of our workflow is as follows: we took yearly counts of 749 waterbird species at 45,745 sites across the world from the International Water Census and Christmas Bird Count. Of these, we wanted to find populations, here defined as a certain species at a certain site, that occurred in a protected area and where yearly counts had begun before the protected area was designated. For our before–after (BA) analysis, we then assessed how each population at each of those sites changed from before to after the protected area was designated. For control–intervention (hereafter CI) and BACI analysis, we matched each of these protected populations to unprotected populations surveyed over the same period, that were similar based on a number of site and species characteristics. For CI, we compared populations in the years after the protected area was designated between unprotected and protected population pairs. For BACI, we compared change in protected populations from before to after protected area designation, and then compared this to the BA change in matched unprotected populations over the same period.

Whether BA, CI or BACI, we then classified the impact to the population as positive, negative or no impact from protection. Next, we looked to see whether our conclusions about impact varied when we analysed a population in a BA, CI or BACI framework. We found BA and CI analyses to be unreliable, so discarded them at this point. Next, we looked to see whether there were correlates that predicted protected area impact, by running cumulative link models on BACI data. These correlated outcome (Positive, Negative or No impact) to a range of site and species level predictors such as protected area size, species body size, land use type and whether the site was managed for waterbirds. Finally, we ran sensitivity tests varying a range of parameters that were used to make analytical decisions to test the robustness of conclusions.

All analysis was completed using R v4.0.3<sup>49</sup> and QGIS v3.10<sup>50</sup>, data figures and base maps were produced using the R package ggplot2<sup>51</sup>, impact legends were produced using Inkscape.

### Time-series preparation

We took site-specific annual counts from two long term surveys: the International Waterbird Census (IWC), coordinated by Wetlands International, and the Christmas Bird Count (CBC), run by the National Audubon Society. We used Wetland International's definition of waterbird, and took any species from the corresponding families (list of families in Supplementary Information 2). Our initial dataset consisted of 749 species at 45,475 sites, spanning 1940 to 2018. We then restricted our data to only sites surveyed in December to February. We imputed zeroes, by taking any site where a species has been observed, and recording any years where the species was not mentioned as '0' years.

As CBC data is not standardized for effort, we required that these species showed a log-linear relationship with effort (that is, the rate of new individuals detected in a search slows with increased effort), in order to be able to include effort as a term in our models. For each species, we ran a simple negative binomial generalized linear model in R, using the glm.nb function from package MASS<sup>52</sup>, using all available CBC data for that species:

$$\log(E(\text{Count}_i)) = \beta \log(h_i) \quad (1)$$

Where Count is all counts of a species and  $h_i$  is the number of survey hours for each count. We retained CBC data for all species where there was a significant positive relationship between count and effort.

### Protected and unprotected area data

We first created a dataset of counts at protected sites. We took our protected area data from the World Database on Protected Areas ('Protected Planet')<sup>53</sup>, downloading the full dataset of all protected areas globally, and overlaying our sites to determine which fell in protected areas. Some coastal site coordinates fell just outside the land cover layer that protected areas are aligned to, so we snapped all sites to the base terrestrial layer used by Protected Planet<sup>54</sup>, but by no more than 10 km. We removed any sites where the designation status was proposed, and any United Nations Educational, Scientific and Cultural Organization (UNESCO) biosphere reserves as these are often not afforded formal protection<sup>55</sup>. We next removed any sites where there was no information about designation date. In some cases, there were multiple protected area data entries for a site, in these cases we took the earliest designation year given. Finally, we reduced the count dataset to only the 10 years before and after the designation date of whichever protected area the survey site fell within, requiring that at least 7 years before and after were surveyed (we tested the number of years restricted from 5–15 years, and number of years measured from 4–13; Extended Data Table 1a, b, respectively).

We next created a dataset of counts at unprotected sites for CI and BACI analysis. For Christmas Bird Count data, surveying is conducted in a circle with a radius of 12.07 km. If there is a protected site in this circle, we cannot be sure that the counts are not being biased by protection. Therefore, we only counted sites as unprotected if no protected area occurred in the entire circle. For IWC data, we included sites that were at least 1 km from a protected area, to avoid any confounding of results from spill-over effects<sup>56</sup> (we sensitivity tested this threshold from 500 m to 5 km; Extended Data Table 1c). We consider sites to be unprotected until the point in time when a protected area was designated at that site. For instance, a site, A, could be designated as a protected area in the year 2000, but this would mean that counts before this point, say, from the 1980s, would be of waterbirds at a site not experiencing any benefit of protection. We could therefore match a protected site from the 1980s to Site A's counts in the 1980s, and treat A's counts as unprotected at this time.

### BA, CI and BACI datasets

In all cases, we defined the 'after' period as being the years after, but not including, the designation date of the protected area. We also defined cases of 'all zeros' to account for local immigrations and extinctions. Waterbirds are highly mobile and can quickly immigrate to, or emigrate from a site. In these cases we cannot assess a change in trend between, for instance, a before period where there are individuals absent and an after period when they have immigrated to the site (for a detailed explanation of why immigrations and extinctions pose a problem for trend analysis, see ref. <sup>24</sup>). Theoretically, we should only consider those cases where there are zero counts in all before or after years as 'all zero' local immigrations or extinctions, but because waterbirds are able to appear as vagrants at a site, we chose to classify cases where at least 70% of years were zero counts as all zeroes. We tested this threshold from 60–80% (Extended Data Table 1d). Of note, any sites where the species had never occurred would not be included in the dataset, so even in cases of all zeroes the species is known to be able to occur at the site.

To create the BA dataset, we took all protected populations where there were cases of counts (as opposed to all zeroes) in either the before period, after period or both. We subset the BA dataset to only protected populations that also occur in the BACI dataset.

To create the CI dataset, we took all protected populations with counts (as opposed to all zeroes) in the after period, and matched these to unprotected populations also with counts over the same time period (see matching below). We subset the CI dataset to only protected populations that also occur in the BACI dataset.

To create the BACI dataset, we matched protected and unprotected populations, requiring that at least one period (either protected before, protected after, unprotected before, or unprotected after) had counts (as opposed to all zeroes).

## Matching

**Data preparation.** We developed a statistical matching method to achieve matching of BACI and CI analyses. The covariates we used for matching, how we prepared them and justification for their use are given in Extended Data Table 2, broadly they encompass variables related to climate, land use and human impact. We removed highly correlated variables by first calculating the variance inflation factor (using the VIF function from the `usdm` package in R<sup>57</sup>) of all covariates, and iteratively removing variables with a VIF greater than four until none were over four<sup>58</sup>. We next removed variables with a Pearson's correlation coefficient of over 0.7.

For BACI, we matched only on covariates in the years prior to designation (as protected and unprotected sites might be expected to differ in the years after protected area designation, especially on covariates related to human impact). For CI, we matched on covariates only in the years after designation, as we choose to be blind to the 'before' period in this analysis.

We then proceeded with matching, separately for each species. The following describes the procedure for one species.

**Mahalanobis distances.** We used Mahalanobis distance matching to evaluate how similar protected and unprotected sites were. Though Mahalanobis distance has been criticized in the past for performing poorly when matching on many covariates<sup>59,60</sup>, recent criticisms of the most commonly used matching method, Propensity Score Matching<sup>61</sup>, meant we were interested to test other options and found Mahalanobis distance matching to perform markedly better in comparisons (Supplementary Information 9).

Mahalanobis Distance ( $md$ ) computes the distance between points in multivariate space. The Mahalanobis distance between two sets of points is calculated as follows:

$$md_{(x,y)} = \sqrt{(\mathbf{x} - \mathbf{y})^T \mathbf{S}^{-1} (\mathbf{x} - \mathbf{y})} \quad (2)$$

Where  $\mathbf{x}$  and  $\mathbf{y}$  are vectors containing values for each covariate (in our case, therefore, the list of covariate values for sites  $x$  and  $y$ ) and  $\mathbf{S}$  is the covariance matrix of the covariates.

This formula requires each site to have one value for each covariate, so we took means of the values for the years pre- (BACI) or post- (CI) designation.

For each species, we created a large matrix with protected sites in columns and unprotected sites in rows, with Mahalanobis distance values populating the rows. Because we wanted to match exactly on the years only prior to protected area designation, we first created separate matrices (using function `mahal` from R package `DOS`<sup>62</sup>), each containing only protected areas designated in a certain year (see Extended Data Fig. 6a, b for an example). Mahalanobis distance requires at least two protected sites to work (to be able to calculate the covariance matrix), and so we could not build Mahalanobis distance matrices for years where only one protected area in our dataset was designated. This resulted in a minimal loss of sites.

These Mahalanobis distance matrices were then combined into the larger distance matrix containing all the sites across all designation years that the species occurred in (Extended Data Fig. 6c).

**Exact matching.** We required that sites were exactly matched on a number of criteria; where sites failed they were excluded from the

Mahalanobis distance matrix (Extended Data Fig. 6d). For each protected site, we removed unprotected sites not of the same anthrome category, continent, and migratory status. We also removed any sites greater than 500 km from the protected area (we tested this value from 100 km to 2,500 km; Extended Data Table 1e).

For BACI analysis, we needed to satisfy the parallel trends assumption<sup>24,63</sup>, which specifies that the trends of control and intervention populations in the 'before' period must be parallel. To test this, we modelled the difference in trends between each protected and potential unprotected matched population. We used a negative binomial glm (`glm.nb`, R package `MASS`<sup>52</sup>),:

$$\log(E(\text{Count}_{i,j})) = \alpha + \beta_1 Y_i + \beta_2 C_j + \beta_3 Y_i C_j + \text{offset}(\log(h_i)) + \epsilon \quad (3)$$

Where the count of the population in year  $i$  at site  $j$  is predicted by the Year ( $Y$ ), a binary term that is one for the protected site and zero for the unprotected site (CI) and the interaction between the two. Log of effort is included as an offset for CBC data (effort is held at one for IWC data). We also checked for temporal autocorrelation and adjusted the model if it was present (see 'Temporal autocorrelation' below). If the interaction coefficient ( $\beta_3$ ) was significant ( $P < 0.05$ ), then there was a significant difference between the trend of the two populations, and the unprotected population was discarded.

If no unprotected sites met the exact match criteria, the protected site did not have a match and was excluded (for example, Extended Data Fig. 6d, site E).

**Picking matches.** Next, we ran an optimized greedy nearest-neighbour algorithm to select, from the Mahalanobis distance matrix (with any sites not satisfying exact match criteria excluded), the unprotected site with the smallest Mahalanobis distance. We ran this without replacement, meaning each protected site could be matched to only one unprotected site, to ensure no pseudoreplication. A greedy algorithm works through the dataset, picking the best match for each successive protected site and removing the matched unprotected site from the potential matching pool as it goes. However, greedy algorithms have a tendency to get stuck in local optima<sup>64</sup>, so to account for this, we ran the greedy algorithm 1,000 times, each time randomizing the order of protected sites that the greedy algorithm would work through. We found the global distance for each iteration and used the set with the smallest global distance (Extended Data Fig. 6e, for example, with randomizations in the figure a smaller global distance would be detected).

**Evaluating match quality.** Once we had our matched sets for each species, we needed to ensure that the matches were of a high enough quality to be used. This was done by assessing the covariate balance between matched and unmatched sites for each species using the 'standardized difference in means' (SDiM), which is calculated using the following formula<sup>65</sup>:

$$d_{cov} = \frac{\bar{\mathbf{T}}_{cov} - \bar{\mathbf{C}}_{cov}}{\sqrt{\frac{\text{var}(\mathbf{T}_{cov}) - \text{var}(\mathbf{C}_{cov})}{2}}} \quad (4)$$

Where  $\mathbf{T}_{cov}$  is the values of covariate  $cov$  for protected sites (mean from the years before and equal to designation),  $\mathbf{C}_{cov}$  is the same for unprotected sites,  $\text{var}$  is the variance of each of these and  $d_{cov}$  is the standardized mean difference between protected and unprotected sites. We assessed the SDiMs to determine whether they were below 0.25 for all covariates<sup>60,66</sup> (we sensitivity tested this threshold from 0.1 to 0.25; Extended Data Table 1f). If they were not, the matched pair with the greatest distance was removed and the SDiM checked again. Once all covariates had a SDiM of  $< 0.25$  (or the relevant sensitivity value), the remaining matched pairs were considered the 'final' matched dataset for that species (Extended Data Fig. 6f). If fewer than 80% of the sites

# Article

that a species occurred in were remaining, we discarded the species, to ensure that the matched set was not biased to a certain subset of all sites for that species (we sensitivity tested this value from 50–90%; Extended Data Table 1g).

## Assessing protected area impact

Following the framework set out in ref.<sup>24</sup>, we defined a number of ways that a population could respond to protection. Broadly, populations can respond to a protected area by immigrating to the area, going locally extinct from the area, showing a change in trend over time, or by showing an immediate change, that is, an immediate increase or decrease in the number of individuals (see legends of Fig. 3, Extended Data Figs. 3, 4).

For comparing BA, a population could show an immediate change or change in trend, or the population could immigrate to the site or go locally extinct at the site (Extended Data Fig. 3). For comparing BACI, the BA changes were compared between protected and unprotected sites. For example, a population could be stable in the period before protection, and declining in the period after – this would be a negative BA trend change (Extended Data Fig. 3). But if a matched unprotected population was also stable in the before period, but declining at a faster rate in the after period, then the BACI trend change would be positive (Fig. 3), as the protected area had slowed the decline of the protected species, even if it hadn't halted it. If the unprotected population was declining at a similar rate to the protected population in the after period, this would be a case of no impact under a BACI framework (Fig. 3). For comparing CI, only the difference in trend between protected and unprotected populations was considered (Extended Data Fig. 4).

All BA, CI or BACI time periods with all zeroes were categorized as immigrations or extinctions, for instance, in BACI analysis, if a protected population had no counts in the before period, but did in the after period, while the matched unprotected population had no counts in the before and after period, this would be classified as a local immigration (and a positive impact of the protected area).

For time periods with all counts, we ran the following models. In all cases  $Y$  represents the year, centred around the year of protected area designation so that year of designation equals zero. BA is a binary term that is zero in the years before protected area designation, and one in the years after; note that this isn't included in the CI model as only 'after' years are used. CI is a binary term that is zero for the unprotected population and one for the protected population; note that the CI term is not included in the BA model as this model does not include unprotected populations. Finally, each model includes an offset term for effort ( $h$ ), to account for variable effort in CBC data. For IWC data, effort is always set to one and so does not contribute to the model. All models were negative binomial glms, run using R package MASS<sup>52</sup>; see ref.<sup>24</sup> for a more detailed explanation of these models.

For BA:

$$\log(E(\text{Count}_i)) - \beta_0 + \beta_1 \text{BA}_i + \beta_2 Y_i + \beta_3 \text{BA}_i Y_i + \text{offset}(\log(h_i)) + \epsilon \quad (5)$$

$\beta_1$  gives the immediate change and  $\beta_3$  gives the trend change<sup>24</sup>.

For CI:

$$\log(E(\text{Count}_i)) - \beta_0 + \beta_1 \text{CI}_i + \beta_2 Y_i + \beta_3 \text{CI}_i Y_i + \text{offset}(\log(h_i)) + \epsilon \quad (6)$$

$\beta_3$  gives the difference in trend between protected and unprotected sites.

For BACI:

$$\log(E(\text{Count}_{i,j})) \sim \beta_0 + \beta_1 \text{BA}_i + \beta_2 \text{CI}_j + \beta_3 Y_i + \beta_4 \text{BA}_i \text{CI}_j + \beta_5 \text{BA}_i Y_i + \beta_6 \text{CI}_j Y_i + \beta_7 \text{BA}_i \text{CI}_j Y_i + \text{offset}(\log(h_i)) + \epsilon \quad (7)$$

$B_4$  gives the immediate change and  $\beta_7$  gives the trend change<sup>24</sup>. We excluded any cases where  $\beta_6$  was significant as this indicates a

significant difference between protected and unprotected trends in the before period, meaning the parallel trends assumption is not satisfied. Though we checked for this in matching, running a full model containing 'after' data (compared with only 'before' data, as in matching) meant that very occasionally this term became significant, presumably because of an increase in power.

In a small proportion of populations, models failed to converge. In these cases, we removed the population from analysis.

**Temporal autocorrelation.** Time-series data are vulnerable to the effects of temporal autocorrelation, where counts in one year are impacted by counts in the years before, and as a result are not independent, as models assume. Being mobile, we expect less temporal autocorrelation in waterbird data than for sessile species (waterbird population numbers can change markedly at a site year to year), but nevertheless we checked for, and accounted for, temporal autocorrelation in our data. For each population model (whether BA, CI or BACI; and also for the models used to check for parallel trends in the matching stage), we checked for temporal autocorrelation using three implementations of the Durbin–Watson test in R: `durbinWatsonTest` from package `car`<sup>67</sup>, `testTemporalAutocorrelation` from package `DHARMA`<sup>68</sup>, and `dwtest` from package `lmtest`<sup>69</sup>. Though each of these implementations performs the same test, variations in methodology meant we found some population models had significant temporal autocorrelation under one, but not another. To be conservative, we decided that if a population had significant autocorrelation under any of the three tests, we considered there to be temporal autocorrelation. If this was the case, we re-ran the population model as a negative binomial generalized linear mixed model (using `glmer.nb` from package `lme4`<sup>70</sup>) including a random intercept for year for BA analyses, and `site:year` for CI and BACI analyses, to account for the autocorrelation.

**Classifying outcomes.** We then classified outcomes. We aimed to be generous for assigned positive outcomes, and so for BA and BACI, a significantly ( $P < 0.05$ ) positive immediate or trend change (even if the other was significantly negative) meant that the protected area was classed as having had a positive impact on the population. If both immediate and trend were insignificant, then the protected area was classed as having had no impact. And if either was negative and the other insignificant, or if both were significantly negative, the protected area was classed as having had a negative impact. We conducted a supplementary analysis to see whether relaxing this  $P$ -value would result in detecting more positive impacts (see Supplementary Information 12), results did not affect our conclusions.

For CI, a significantly positive difference between protected and unprotected trends was classed as a positive impact, significantly negative was a negative impact, and an insignificant difference no impact.

## Drivers of change

To explore the predictors of protected area effectiveness, we considered body mass, species migratory status, taxonomic order, the broad anthrome category (i.e. land use type) of the protected area, protected area size, and country governance<sup>11</sup>. See Extended Data Table 3 for details of how each covariate was obtained.

To test how these covariates might correlate to protected area effectiveness, we ran cumulative link mixed effects models that allow for ordinal predictors and random factors, with the response term being a three-level factor: negative impact, no impact, or positive impact. To account for spatial autocorrelation, we included a random intercept for 'grid cell', with sites each assigned to a gridcell of size  $2 \times$  the maximum distance between protected and unprotected sites, depending on the focal/full parameter analysis (Extended Data Table 1e). In this way errors are grouped by sites that are closer together. In some of the 21 analyses, typically those with smaller sample sizes, including both country and grid cell as random factors meant the model could



not converge; in these cases we retained only country as a random factor. We used the `clmm` function from R package `ordinal`<sup>71</sup>. The model specification was as follows:

$$\text{Impact}_{i,j,k} \sim \beta_1 \text{MigratoryStatus}_i + \beta_2 \log(\text{BodyMass})_i \\ + \beta_3 \text{Order}_i + \beta_4 \text{Anthrome}_j + \beta_5 \text{RamsarSPA}_j \\ + \beta_6 \log(\text{PA size})_j + \beta_7 \text{MeanGovernance}_k \\ + (1|i) + (1|k) + (1|k:j) + (1|m) + (1|m:j) \epsilon$$

Where  $i, j, k$  and  $m$  are species, site, country and gridcell, respectively. In some sensitivity tests some covariates did not have sufficient populations to be able to test them, in these cases certain levels of the covariate were removed (for example, if there were not enough populations of a particular taxonomic order) or in some cases the entire covariate was removed. Not all protected areas have area data reported, and so we had to run models only on the subset of data where area data was available. To ensure this reduced set was not misrepresenting the full dataset, we also ran models without the ‘protected area size’ covariate and on the full dataset; results were broadly similar (Supplementary Information 8), and in the case of BACI, waterbird-managed sites were more strongly positively associated with outcomes.

We estimated the effect size of management and protected area size using the function `ggpredict` from R package `ggeffects`<sup>72</sup>, which returns odds ratios from the cumulative link mixed models. We estimated effect size for water-bird managed vs mixed-managed sites, and for 5 quintiles of  $\log(\text{protected area size})$ : 0.05, 0.25, 0.5, 0.75 and 0.95. For the effect size reported in the manuscript, we compared the chance of a positive impact on a population in a mixed-management site in the 0.05th size quintile to the chance of a positive impact on a population in a waterbird-managed site in the 0.95th quintile.

Finally, some covariates violated the proportional odds ratio assumption upon which cumulative link models rest. To check for the impact of this we ran individual binomial generalized linear mixed-effects models (using function `glmer` from R package `lme4`<sup>70</sup>) to conduct pairwise comparisons of outcome levels. These models supported the general conclusions made in this paper (see Supplementary Information 13 for further details).

### Full parameter analyses

The focal analysis inevitably is based on somewhat arbitrary modeling choices. We therefore ran our models an additional 20 times with a range of parameter values for decisions such as: the number of years of counts required before and after protection, the threshold at which we classify ‘all zeroes’, the maximum distance between protected and unprotected sites for an acceptable match and how similar we required matched sites to be (Extended Data Table 1). Testing all parameter combinations was computationally impractical so we used a latin hypercube sampling method<sup>73</sup>. This is a way to adequately sample a high dimensional parameter space when random sampling is prohibitively inefficient; it creates multiple combinations of covariates that together evenly sample the entire  $n$  dimensional sample space. We randomly created 20 parameter combinations (using function `randomLHS` from the R package ‘`lhs`’<sup>74</sup>), which are displayed in Extended Data Table 1. We call these analyses our ‘full parameter’ analyses ([www.govindicators.org](http://www.govindicators.org)).

### Reporting summary

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

### Data availability

The waterbird count data used in this study are collated and managed by Wetlands International and the National Audubon Society, and are

available on request (<http://iwc.wetlands.org/index.php/> and <http://netapp.audubon.org/cbcobservation/>, respectively). We requested all data from both providers for the years 1900–2018, for all waterbird families (see Supplementary Information 2), and for sites in all available countries (though data from Russia was excluded as permissions were not given). All the data that pertain to explanatory variables are freely available, as specified in Extended Data Tables 2, 3.

### Code availability

The code used to produce all analysis and figures are archived on Zenodo at <https://doi.org/10.5281/zenodo.5794511>. Code are also available on GitHub at <https://github.com/hannahwauchope/PALImpact>; this is the recommended mode of access as it will contain any updates or clarifications.

47. Wauchope, H. et al. Quantifying the impact of protected areas on near-global waterbird population trends, a pre-analysis plan. Preprint at <https://doi.org/10.7287/peerj.preprints.27741v2> (2019).
48. Nosek, B. A., Ebersole, C. R., DeHaven, A. C. & Mellor, D. T. The preregistration revolution. *Proc. Natl Acad. Sci. USA* **115**, 2600–2606 (2018).
49. R Core Team. R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, 2020).
50. QGIS Geographic Information System (QGIS, 2021).
51. Hadley Wickham. *ggplot2: Elegant Graphics for Data Analysis* (Springer-Verlag, 2016).
52. Venables, W. N. & Ripley, B. D. *Modern Applied Statistics with S* (Springer, 2002).
53. The World Database on Protected Areas (WDPA)/The Global Database on Protected Areas Management Effectiveness (GD-PAME) [www.protectedplanet.net](http://www.protectedplanet.net) (UNEP-WCMC and IUCN, 2019).
54. Global Self-consistent, Hierarchical, High-resolution Geography Database (GSHHG) (NOAA, 2017).
55. Coetzer, K. L., Witkowski, E. T. F. & Erasmus, B. F. N. Reviewing Biosphere Reserves globally: effective conservation action or bureaucratic label? *Biol. Rev.* **89**, 82–104 (2014).
56. Ament, J. M. & Cumming, G. S. Scale dependency in effectiveness, isolation, and social-ecological spillover of protected areas. *Conserv. Biol.* **30**, 846–855 (2016).
57. Naimi, B., Hamm, N. A. S., Groen, T. A., Skidmore, A. K. & Toxopeus, A. G. Where is positional uncertainty a problem for species distribution modelling? *Ecography* **37**, 191–203 (2014).
58. Salmerón Gómez, R., García, Pérez, J., López Martín, M. D. M. & García, C. G. Collinearity diagnostic applied in ridge estimation through the variance inflation factor. *J. Appl. Stat.* **43**, 1831–1849 (2016).
59. Gu, X. S. & Rosenbaum, P. R. Comparison of multivariate matching methods: structures, distances, and algorithms. *J. Comput. Graph. Stat.* **2**, 405–420 (1993).
60. Stuart, E. A. Matching methods for causal inference: a review and a look forward. *Stat. Sci.* **25**, 1–21 (2010).
61. King, G. & Nielsen, R. Why propensity scores should not be used for matching. *Pol. Anal.* **27**, 435–454 (2019).
62. Rosenbaum, P. R. DOS: design of observational studies. <https://cran.r-project.org/web/packages/DOS/index.html> (2018).
63. Linden, A. A matching framework to improve causal inference in interrupted time-series analysis. *J. Eval. Clin. Pract.* **24**, 408–415 (2018).
64. Simmons, B. L., Hoeppeke, C. & Sutherland, W. J. Beware greedy algorithms. *J. Anim. Ecol.* **88**, 804–807 (2019).
65. Austin, P. C. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat. Med.* **28**, 3083–3107 (2009).
66. Rubin, D. B. Using propensity scores to help design observational studies: application to the tobacco litigation. *Health Serv. Outcomes Res. Methodol.* **2**, 169–188 (2001).
67. Fox, J. & Weisberg, S. *An R Companion to Applied Regression* (Sage, 2019).
68. Hartig, F. DHARMA: residual diagnostics for hierarchical (multi-level/mixed) regression models. <https://cran.r-project.org/web/packages/DHARMA/index.html> (2021).
69. Zeileis, A. & Hothorn, T. Diagnostic checking in regression relationships. *R News* **2**, 7–10 (2002).
70. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using `lme4`. *J. Stat. Softw.* **67**, 48 (2015).
71. Christensen, R. Ordinal-regression models for ordinal data. <https://cran.r-project.org/web/packages/ordinal/index.html> (2019).
72. Lüdtke, D. `ggeffects`: tidy data frames of marginal effects from regression models. *J. Op. Source Softw.* **3**, 772 (2018).
73. McKay, M. D., Beckman, R. J. & Conover, W. J. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* **21**, 239–245 (1979).
74. Carnell, R. `lhs`: latin hypercube samples. <https://cran.r-project.org/web/packages/lhs/index.html> (2020).
75. Harris, I., Jones, P. D., Osborn, T. J. & Lister, D. H. Updated high-resolution grids of monthly climatic observations—the CRU TS3.10 Dataset. *Int. J. Climatol.* **34**, 623–642 (2014).
76. Lu, C. & Tian, H. Global nitrogen and phosphorus fertilizer use for agriculture production in the past half century: Shifted hot spots and nutrient imbalance. *Earth Syst. Sci. Data* **9**, 181–192 (2017).
77. Joppa, L. N. & Pfaff, A. High and far: biases in the location of protected areas. *PLoS ONE* **4**, e8273 (2009).

# Article

78. Hurtt, G. C. et al. Harmonization of land-use scenarios for the period 1500-2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. *Clim. Change* **109**, 117–161 (2011).
79. Lloyd, C. T., Sorichetta, A. & Tatem, A. J. High resolution global gridded data for use in population studies. *Sci. Data* **4**, 17001 (2017).
80. Pekel, J.-F., Cottam, A., Gorelick, N. & Belward, A. S. High-resolution mapping of global surface water and its long-term changes. *Nature* **540**, 418–422 (2016).
81. Sandvik, B. World Borders Dataset. *Thematic Mapping* [http://thematicmapping.org/downloads/world\\_borders.php](http://thematicmapping.org/downloads/world_borders.php) (2009).
82. BirdLife International. *Species Distribution Data Download* <http://www.birdlife.org/datazone/info/spdownload> (accessed 25 February 2020).
83. Wilman, H. et al. EltonTraits 1.0: species-level foraging attributes of the world's birds and mammals. *Ecology* **95**, 2027 (2014).
84. WWF International. *Management Effectiveness Tracking Tool* [https://wwfeu.awsassets.panda.org/downloads/mett2\\_final\\_version\\_july\\_2007.pdf](https://wwfeu.awsassets.panda.org/downloads/mett2_final_version_july_2007.pdf) (2007).

**Acknowledgements** We thank the coordinators, thousands of volunteer counters, and funders of the International Waterbird Census. This data collection effort is funded by the Ministry of the Environment of Japan, Environment Canada, AEW Secretariat, EU LIFE+ NGO Operational Grant, MAVA Foundation, Swiss Federal Office for Environment and Nature, French Ministry of Environment and Sustainable Development, UK Department of Food and Rural Affairs, Norwegian Nature Directorate, Dutch Ministry of Economics, Agriculture and Innovation, DOB Ecology and Wetlands International members. CBC Data is provided by National Audubon Society and through the generous efforts of Bird Studies Canada and countless volunteers across the western hemisphere. H.S.W. was funded by a Cambridge–Australia Poynton Scholarship, Cambridge Department of Zoology J. S. Gardiner Studentship and Cambridge Philosophical Society Grant. H.S.W. and B.I.S. are funded by the Royal Commission for the

Exhibition of 1851. W.J.S. was funded by Arcadia, The David and Claudia Harding Foundation and MAVA. J.P.G.J. was supported by a visiting fellowship to Fitzwilliam College Cambridge. This work was performed using resources provided by the Cambridge Service for Data Driven Discovery (CSD3) operated by the University of Cambridge Research Computing Service ([www.csd3.cam.ac.uk](http://www.csd3.cam.ac.uk)), provided by Dell EMC and Intel using Tier-2 funding from the Engineering and Physical Sciences Research Council (capital grant EP/P020259/1), and DiRAC funding from the Science and Technology Facilities Council ([www.dirac.ac.uk](http://www.dirac.ac.uk)). Finally, the authors would like to acknowledge the use of the University of Exeter High-Performance Computing (HPC) facility in carrying out this work.

**Author contributions** H.S.W., J.P.G.J., J.G., B.I.S., T.A., R.A.F., A.J. and W.J.S. conceived the study. D.E.B., T.L., T.M. and S.N. provided waterbird count data, which H.S.W. and T.A. collated. H.S.W. performed analysis, produced figures and wrote the text with advice from all authors, especially J.P.G.J., J.G., B.I.S., T.A., A.J. and W.J.S. All authors contributed to the review of the manuscript before submission for publication and approved the final version.

**Competing interests** The authors declare no competing interests.

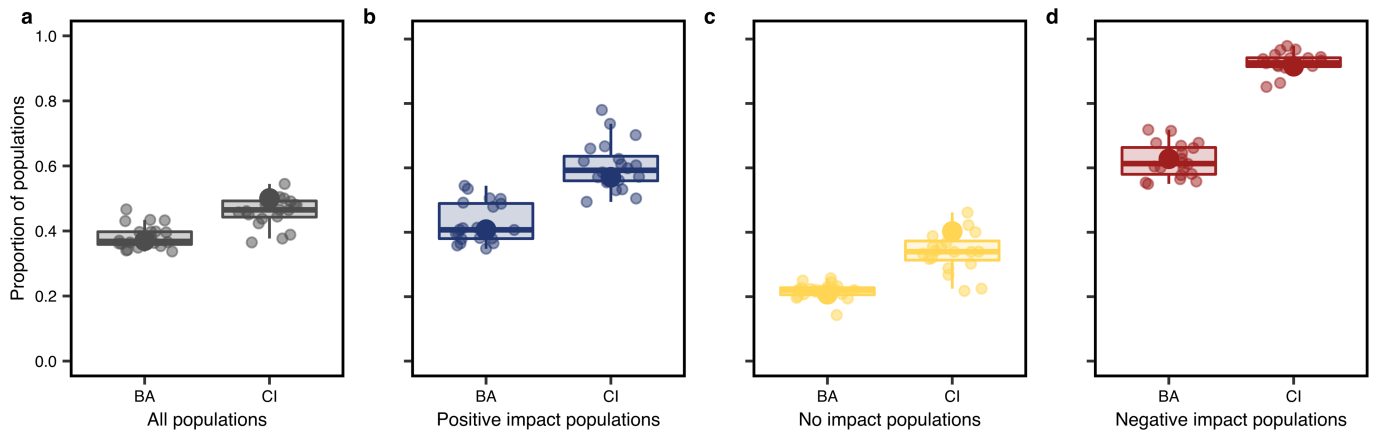
## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41586-022-04617-0>.

**Correspondence and requests for materials** should be addressed to Hannah S. Wauchope or Taej Mundkur.

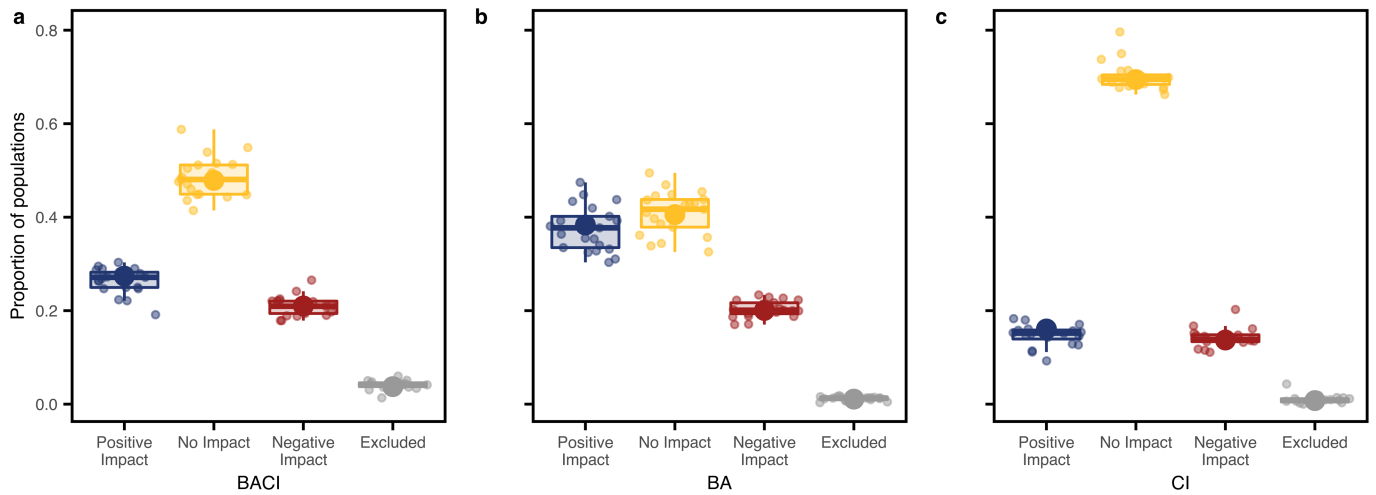
**Peer review information** *Nature* thanks Gergana Daskalova and the other, anonymous reviewers for their contribution to the peer review of this work. Peer review reports are available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.



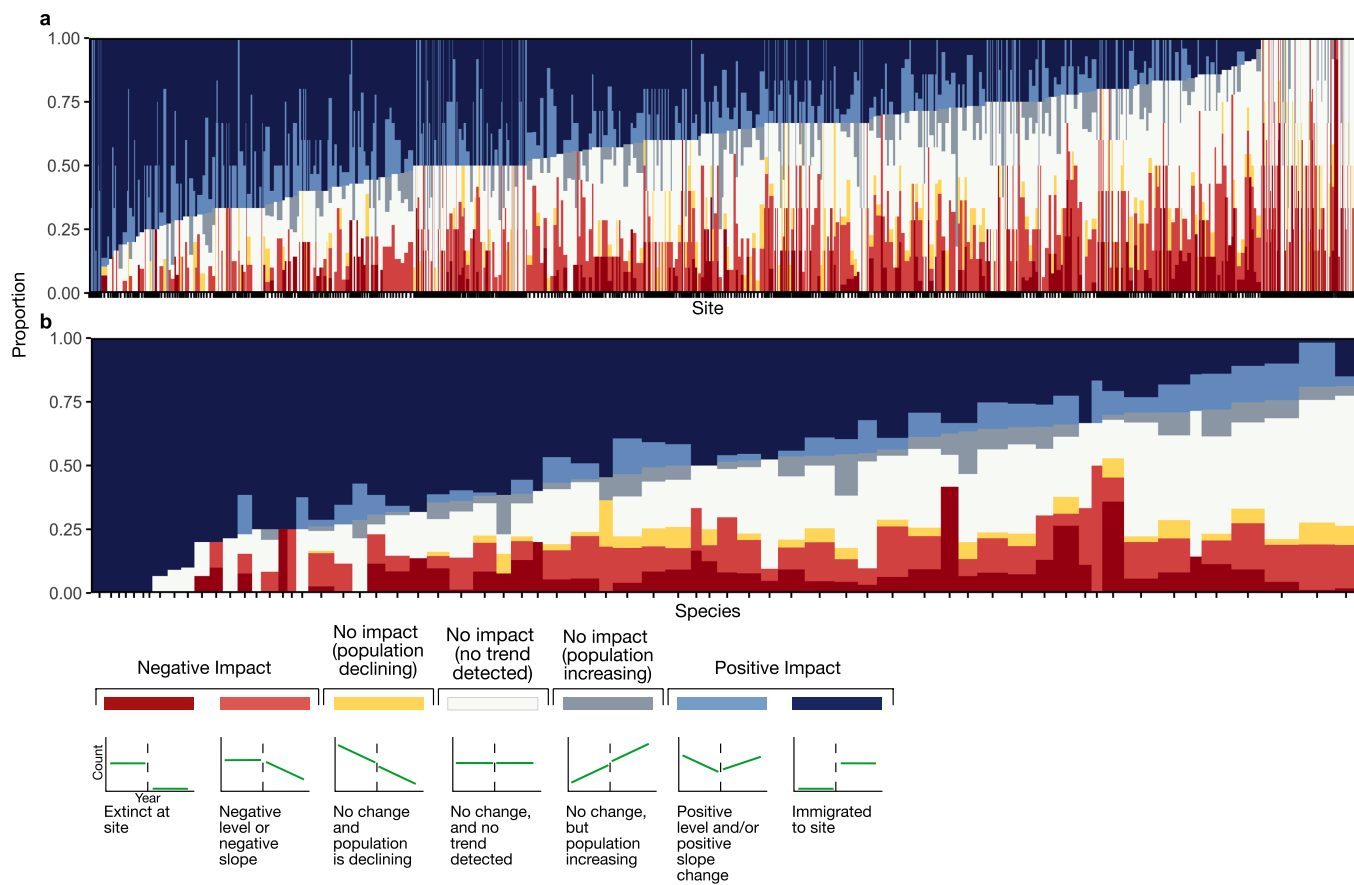
**Extended Data Fig. 1 | Changes in estimates of protected area impact under different study designs, for all analyses.** Proportion of Before-After (BA) or Control-Intervention (CI) populations that changed outcome when analysed under a BACI framework, by each analysis ( $n = 21$ ; 20 full parameter, plus one focal analysis). Shown for all populations (a), then the proportion of positive

(b), no (c) or negative impact populations (d) that changed in outcome. Each point is an analysis, with boxplots showing distribution (box bounded by 25<sup>th</sup> and 75<sup>th</sup> percentiles, centre shows 50<sup>th</sup> percentile, whiskers extend to 1.5\*IQR above 75<sup>th</sup> percentile, for maxima, or below 25<sup>th</sup> percentile, for minima). Large points show focal analysis estimates.



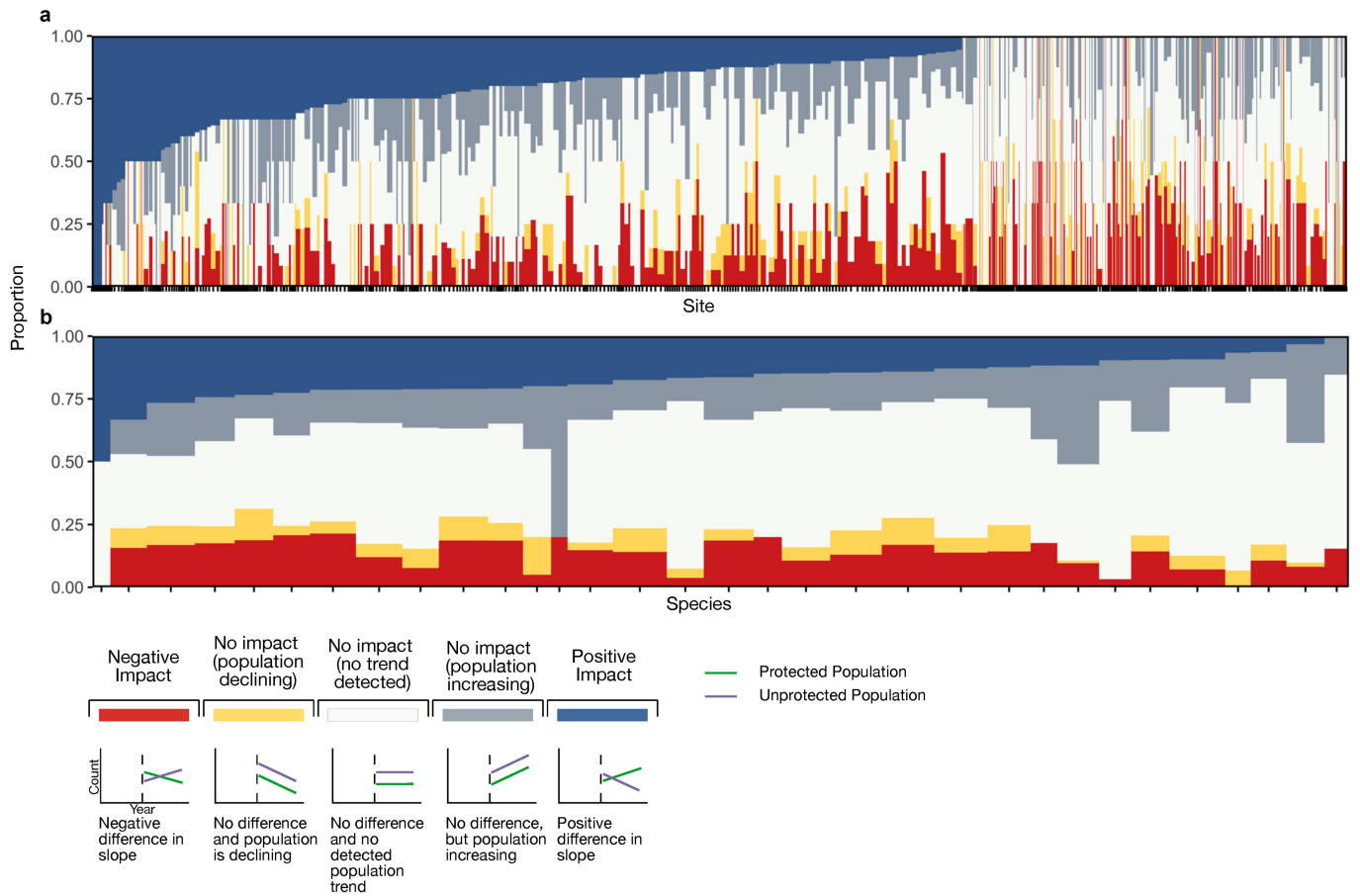
**Extended Data Fig. 2 | Estimates of protected area impact under a BACI study design, for all analyses.** Percentage of populations that have been positively, negatively or not impacted by protected areas, by each analysis (n = 21; 20 full parameter analyses, plus one focal analysis). Each point is an analysis, with boxplots showing distribution (box bounded by 25<sup>th</sup> and 75<sup>th</sup>

percentiles, centre shows 50<sup>th</sup> percentile, whiskers extend to 1.5\*IQR above 75<sup>th</sup> percentile, for maxima, or below 25<sup>th</sup> percentile, for minima). Large points show estimates from focal analysis. Panels show estimates under BACI (a), Before-After (b) or Control-Intervention (c) frameworks.



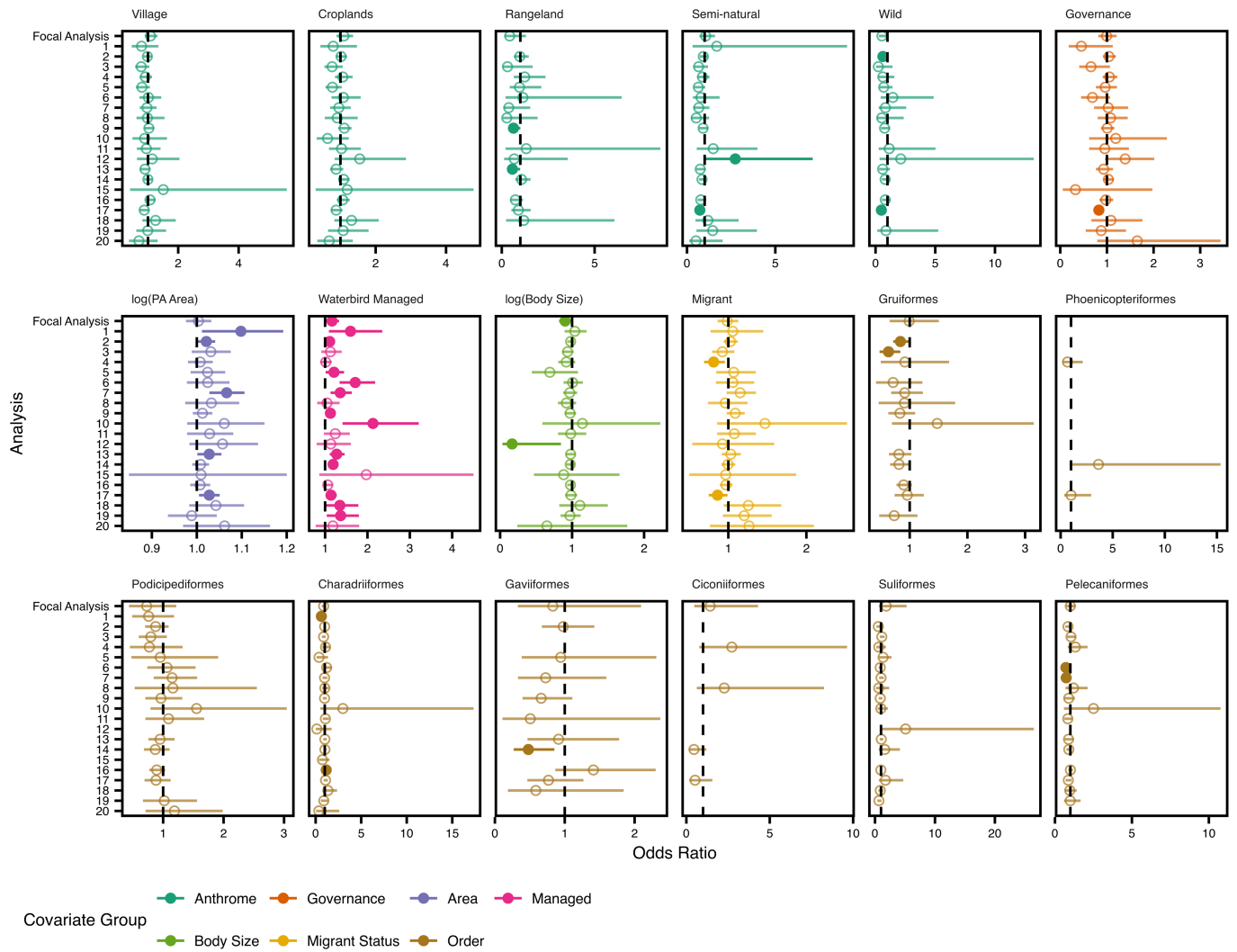
**Extended Data Fig. 3 | Estimates of protected area impact under a BA study design.** Proportion of populations ( $n = 6263$ ) showing various responses to protection, per site ( $n = 860$ ) and species ( $n = 66$ ), when response to protection is calculated in a BA framework. Each species/site is one bar, with the proportion of their populations in each category shown on the y axis. Bar width is scaled to the number of populations of that species/site in the dataset,

log scaled in the case of species, with a wider bar meaning the species/site has more populations. Each colour represents a different way a population can respond to protection, and an example of each is shown at the bottom. This figure is based on our focal analysis; Extended Data Fig. 2b shows the proportion of populations within each broad outcome category across all full parameter analyses.



**Extended Data Fig. 4 | Estimates of protected area impact under a CI study design.** Proportion of populations (n = 3783) showing various responses to protection, per site (a; n = 698) and per species (b; n = 32), when response to protection is calculated in a CI framework. Each species/site is one bar, with the proportion of their populations in each category shown on the y axis. Bar width is scaled to the number of populations of that species/site in the dataset, log

scaled in the case of species, with a wider bar meaning the species/site has more populations. Each colour represents a different way a population can respond to protection, and an example of each is shown at the bottom. This figure is based on our focal analysis; Extended Data Fig. 2c shows the proportion of populations within each broad outcome category across all full parameter analyses.

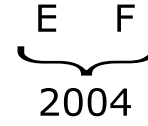
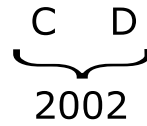
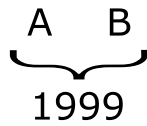


**Extended Data Fig. 5 | Predictors of protected area impact, with odds ratios and confidence intervals.** Odds ratios for covariates predicting protected area (PA) effectiveness under a BACI framework. Estimated using cumulative link mixed models, points show model estimates, tails show 95% confidence intervals, and significance is indicated by bold colours ( $P < 0.05$ ). Dashed line given at an odds ratio of one (ratios above one indicate a positive relationship,

and below one a negative relationship). Y axis shows all analyses (20 full parameter analyses, plus one focal analysis, with the focal analysis given in the first row). Colours show covariate grouping. Orders are measured relative to Anseriformes, and Anthromes relative to Urban. Note that we expect continuous variables (PA Area, Body Size, Governance) to have smaller coefficients as they express odds ratios per unit increment.

# Article

We have 6 protected sites, A to F, varying in their designation year from 1999 to 2004, and 3 unprotected sites, X, Y & Z.



a) Get covariate values. Each cell is the average value of each covariate for all years that are less than or equal to the designation year (for BACI matching), or all years greater than the designation year (for CI matching).

1999			
A			
B			
X			
Y			
Z			

2002			
C			
D			
X			
Y			
Z			

2004			
E			
F			
X			
Y			
Z			

b) Create a mahalanobis distance (M) matrix. Each cell is the distance in multivariate space between sites based on the covariate means, a greater value means greater distance.

M(1999)	A	B
X	2	9
Y	5	6
Z	4	1

M(2002)	C	D
X	8	3
Y	3	7
Z	2	6

M(2004)	E	F
X	4	2
Y	1	5
Z	9	2

c) Combine into a full matrix of distances, using the values from each designation year matrix.

M	A	B	C	D	E	F
X	2	9	8	3	4	2
Y	5	6	3	7	1	1
Z	4	1	2	6	9	2

d) Remove any sites that are not an exact match (e.g. in different anthrome, population showing a different trend direction).

M	A	B	C	D	E	F
X	2	9	8	3	<del>4</del>	<del>2</del>
Y	5	6	3	7	<del>1</del>	1
Z	<del>4</del>	1	<del>2</del>	6	<del>9</del>	2

E has no exact matches so is excluded.

e) Conduct matching from left to right. Repeat 1000 times, randomising column order. The order with the smallest sum of distances (in this case 1+2+3=6) is used for Step f.

For each protected site (columns) we pick the unprotected site (rows) with the smallest M.

M	A	B	C	D	F
X	②	9	8	3	2
Y	5	6	③	7	1
Z	<del>4</del>	①	2	6	2

A's closest match is X, removing it as an option for other sites. B takes Z, removing that as an option, etc.

f) Assess the distribution of covariates. If the standardised difference in means (SDiM) is >0.25 for any covariate, the worst match is removed and the SDiM calculated again.

	M		SDiM	
A	X	2	0.3	<del>X</del>
B	Z	1	0.24	✓
C	Y	3	0.13	✓

The SDiM of climate is >0.25, so the worst match (C-Y) is removed.

	M		SDiM	
A	X	2	0.19	✓
B	Z	1	0.23	✓
			0.14	✓

Now the SDiM is below 0.25 for all covariates, leaving A-X and B-Z as the final matched dataset.

**Extended Data Fig. 6 | Schematic demonstrating matching procedure.** Example of the matching procedure for one species, using a toy dataset of 6 protected sites (A to F) and 3 unprotected sites (X, Y and Z), with three dummy

example covariates, climate (cloud), land use (wheat) and human population (person). See methods, 'Matching' for a more detailed step by step walk through of this process.



## Extended Data Table 1 | Parameter estimates and sample sizes across analyses

Analysis	a) Total years to either side of PA designation	b) Min number of measured years to either side of PA designation	c) Min distance to PA for unprotected sites (km)	d) Proportion of counts that are zero for period to be classified as "All Zeroes"	e) Max distance between protected and unprotected sites (km, matching, BACI/CI)	f) Standardised difference in means threshold (matching, BACI/CI)	g) Proportion of species' populations that must be matched to retain species (matching, BACI/CI)	h) N Protected Sites	i) N Species	j) N Populations
Focal	10	7	1.00	0.70	500	0.25	0.70	864	67	7313
1	10	10	0.50	0.68	272	0.12	0.53	209	23	951
2	6	5	1.78	0.71	2091	0.19	0.72	933	77	12475
3	14	13	0.95	0.72	2500	0.25	0.50	282	63	4325
4	7	4	2.95	0.69	587	0.23	0.71	1328	68	6050
5	9	6	2.61	0.70	1986	0.15	0.87	953	17	2709
6	11	10	4.30	0.60	1542	0.21	0.68	395	34	1937
7	12	10	4.73	0.78	785	0.19	0.55	469	55	3784
8	5	4	1.36	0.76	100	0.24	0.63	492	51	1402
9	8	6	2.22	0.79	1390	0.20	0.84	952	66	11198
10	13	11	3.49	0.61	1121	0.14	0.53	309	11	677
11	15	10	3.37	0.62	454	0.17	0.90	493	32	1781
12	12	7	3.83	0.77	2199	0.11	0.81	543	4	686
13	9	8	1.95	0.64	1874	0.22	0.66	592	74	7465
14	5	4	1.56	0.74	1154	0.18	0.85	1115	74	13901
15	11	11	4.08	0.68	648	0.14	0.77	122	7	235
16	6	5	0.88	0.65	1676	0.16	0.64	930	74	11922
17	8	5	4.43	0.62	2402	0.23	0.60	1242	77	8738
18	10	9	3.13	0.80	1401	0.15	0.75	404	16	1184
19	15	10	5.00	0.66	192	0.10	0.57	392	33	1894
20	13	10	2.40	0.75	1010	0.13	0.80	334	5	519

Shows focal parameter estimates, plus 20 estimates from full parameter samples. Parameters are a) the maximum number of years of data the sample can have, to either side of protected area (PA) designation; b) the minimum number of years that must be sampled, to either side of protected area designation; c) the closest distance an unprotected site can be to a protected area before it is excluded from analysis; d) the proportion of counts that must be zeroes for the time period to be classified as "All Zeroes"; e) the maximum distance between paired protected and unprotected sites; f) the standardised difference in means threshold for BACI and CI matching; g) the proportion of populations that must be matched successfully to retain a species, for BACI and CI matching. h), i), j) show the number of protected sites/species/populations in that analysis run (note that BA and CI will generally be a subset of these). See Supplementary Information 4 for a further taxonomic break down of species in the focal analysis.

# Article

## Extended Data Table 2 | Covariates used to perform site matching

Category, reason for inclusion, and variables	Data source	Resolution	Data transformation
<p><b>Climate.</b> This is a key variable that can determine suitability of a site for a species as well as likelihood of being designated a PA.</p> <p><b>Variables:</b> Total annual precipitation (mm); total precipitation December – February (mm); mean, min, max annual temperature (°C); mean, min, max temperature December – February (°C).</p>	CRU TS4.01 <sup>75</sup>	0.5°, monthly (1961–2016)	For Dec to Feb values, takes value of Dec from previous year and Jan & Feb of current year.
<p><b>Fertiliser input.</b> Eutrophication can affect waterbird populations<sup>44</sup>. It also acts as a metric of distance to farming land and of potential land value for uses other than protection.</p> <p><b>Variables:</b> Nitrogen (g N/m<sup>2</sup> cropland/yr) &amp; Phosphorous (g P/m<sup>2</sup> cropland/yr)</p>	Lu & Tian <sup>76</sup>	0.5°, yearly (1961–2013)	NA
<p><b>Land use.</b> This is a direct measure of nearness to human impact, important for impacts to bird populations but also for likelihood of protected area designation – protected areas are less likely to be designated in areas suitable for agriculture and farming<sup>77</sup>.</p> <p><b>Variables:</b> Anthrome (categorical); grazing land (km<sup>2</sup>/gridcell); irrigated land (not rice, km<sup>2</sup>/gridcell); irrigated land (rice, km<sup>2</sup>/gridcell); pasture land (km<sup>2</sup>/gridcell); rangeland (km<sup>2</sup>/gridcell); rainfed crop land (no rice; km<sup>2</sup>/gridcell); rainfed crop land (rice; km<sup>2</sup>/gridcell)</p>	HYDE 3.2.001 <sup>78</sup>	5', centennial (10,000BC–1600AD) decadal (1700–2000), yearly (2001–2016)	For Anthromes, pre-2000 data taken from nearest decade. For all other variables, temporal linear interpolation to obtain yearly data between decades of 1960–2000
<p><b>Human presence.</b> Protected areas are more likely to be designated in areas far from humans<sup>77</sup>, and human presence can also affect waterbird numbers either directly through hunting or through habitat degradation.</p> <p><b>Variables:</b> Human population density (inhabitants/km<sup>2</sup>/gridcell); rural, urban human population count (inhabitants/gridcell); total built up area (km<sup>2</sup>/gridcell)</p> <p><b>Variable:</b> Travel time to nearest city</p>	WorldPop <sup>79</sup>	1km, yearly	Bilinear interpolation to 5' grid cells
<p><b>Governance.</b> Governance in a country is a significant predictor of protected area effectiveness<sup>11</sup>, meaning it is important we compare protected areas with similar governance.</p> <p><b>Variable:</b> Mean of the six World Governance Index metrics (Control of Corruption, Government Effectiveness, Political Stability and Absence of Violence/Terrorism, Rule of Law, Regulatory Quality, Voice and Accountability)</p>	World Bank (www.govindicat.ors.org)	By country, 1996, 1998, 2000, and yearly 2002–2016	Mean taken across all years because data is only available from 1996. Therefore, just one value per site for all years.
<p><b>Water.</b> Water presence is an important covariate for waterbirds, which rely on it for survival.</p> <p><b>Variable:</b> Surface water (presence/absence)</p>	Pekel et al <sup>80</sup>	30m, 1985–2005	Sum of 'presence' 30m <sup>2</sup> cells in each 5' grid cell
<p><b>Elevation.</b> Protected areas are biased towards where they can least prevent land conversion<sup>77</sup> which often results in them being in high elevation regions. Higher elevation sites are also likely to have less pressure and thus have lower biodiversity losses regardless of whether they are protected areas or not.</p>	WorldPop <sup>79</sup>	1km, NA	Bilinear interpolation to 5' grid cells
<p><b>Global Region.</b> We wanted populations to be in similar regions to reduce unknown variance in comparisons.</p>	TM World Borders <sup>81</sup>	NA	NA
<p><b>Migratory Status.</b> In some cases, species have some migratory and some resident populations. To ensure we were not comparing between populations of different migratory types we exact matched on migratory status.</p>	Birdlife.org <sup>82</sup>	Species range polygons by migratory types	Each population (site species combination) classified based on the polygon the site fell within

First, the three categorical variables (anthrome, region and migratory status) were used for exact matching. Next, all continuous variables were assessed for collinearity and highly collinear variables were removed. The remaining continuous variables were used to calculate mahalanobis distance. Data sources: refs. <sup>75–84</sup>.

### Extended Data Table 3 | Covariates used to assess what factors affect protected area impact

Category	Variable and reason for inclusion	Data source	Continuous/Categorical
Species	<b>Body Mass.</b> We expected larger species to respond better to protected areas <sup>10</sup> , due to the fact that larger bodied species are more vulnerable to hunting.	Wilman et al <sup>83</sup>	Continuous
	<b>Taxonomic group (Order).</b> Different taxonomic groups may respond differently to protection, so we looked for differences between orders.	Birdlife.org <sup>82</sup>	Categorical. Levels: Anseriformes, Pelecaniformes, Suliformes, Ciconiiformes, Gaviiformes, Charadriiformes, Podicipediformes, Phoenicopteriformes, Gruiformes.
Species (nested within Site)	<b>Migration Status.</b> Because migrants are affected by other stressors than just those in their wintering site, we expected migrants would show less responsiveness to protected areas (and it was beyond the scope of this study to consider migratory networks). Some species are migrants in parts of their range and non-migrant in others, so we categorised each population at each site separately. See Supplementary Information 1 for an extended discussion about migrants.	Birdlife.org <sup>82</sup>	Categorical. Levels: Non-migrant, Migrant
Site (nested in Country)	<b>Anthrome.</b> We expected that sites in more remote regions (i.e. semi-natural, wild) would show less responsiveness to protection, as these sites are less likely to have been being exploited in the absence of protection.	HYDE <sup>78</sup> (see Extended Data Table 2)	Categorical. Levels: Urban, Village, Croplands, Rangeland, Semi-natural, Wild
	<b>Protected area size.</b> We expected larger protected areas to perform better, because of reduced edge effects. In some cases, sites occurred in multiple protected areas that were of different sizes and had been designated at different times. In these cases, we used the size of the largest size protected area, if that area was designated earliest. If not, we took the mean of the size of all areas.	World Database on Protected Areas <sup>53</sup>	Continuous
	<b>Protected area Management.</b> The best way to assess management would be with the Management Effectiveness Tracking Tool (METT <sup>84</sup> ) but unfortunately this is biased away from Europe and the USA, unlike our dataset, and only a few of the protected areas in our dataset are included in the METT. Instead, we chose to compare sites we know to be managed for birds to other sites, acknowledging that some of the ‘other’ sites may also be managed for waterbirds, but not having the power to ascertain management status of all cases. We created a category comprised of Ramsar and Special Protected Area (Birds Directive) sites, which encompasses 55-57% of populations. A full list of waterbird-managed and other sites is given in Supplementary Information 6.	World Database on Protected Area <sup>53</sup>	Categorical. Levels: Waterbird-managed, mixed managed.
Country	<b>Governance.</b> We expected sites in better governed areas to respond better to protection <sup>11</sup> .	World Bank (www.govindicators.org; see Extended Data Table 2)	Continuous

## Reporting Summary

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

### Statistics

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

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

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

#### Data collection

Data was received from Wetlands Internation and the National Audobon Society as excel spreadsheets. These were then imported into R version 4.0.3 for cleaning and collation.

Predictor data were downloaded from various sources online (see Extended Data Table 2 and 3), and were then imported and cleaned in R version 4.0.3.

#### Data analysis

QGIS 3.10 with plugin NNJoin was used to snap points to data layers when assigning the various predictor variables to each site location, otherwise R was used for all analysis. Impact legends and other aesthetic figure edits were completed using Inkscape. All R code are available on GitHub, <https://github.com/hannahwauchope/PAImpact>, and final versions will be uploaded to Dryad upon acceptance of the paper. There are four R scripts used for analysis: 1) Cleaning and Collating Waterbird Data, 2) Extracting Covariate Data, 3) Conducting Matching and 4) Assessing protected area impact, producing figures. All analysis was conducted using R version 4.0.3. Full sessionInfo across the 4 code files (including all loaded package versions) is as follows:

R version 4.0.3 (2020-10-10)  
Platform: x86\_64-apple-darwin17.0 (64-bit)  
Running under: macOS Big Sur 10.16

Matrix products: default  
LAPACK: /Library/Frameworks/R.framework/Versions/4.0/Resources/lib/libRlapack.dylib

locale:  
en\_GB.UTF-8/en\_GB.UTF-8/en\_GB.UTF-8/C/en\_GB.UTF-8/en\_GB.UTF-8

attached base packages:

```
grid parallel stats graphics grDevices utils datasets methods base
```

other attached packages:

```
gridExtra_2.3 RColorBrewer_1.1-2 cowplot_1.1.0 ggeffects_1.0.1 ggalluvial_0.12.3 glmTMB_1.1.2.2 ggbeeswarm_0.6.0
ordinal_2019.12-10 lme4_1.1-27.1 nlme_3.1-149 lmtest_0.9-38 zoo_1.8-8 DHARMA_0.4.3 car_3.0-11 carData_3.0-4
fields_11.6 spam_2.6-0 dotCall64_1.0-0 DOS_1.0.0 maptools_1.0-2 scales_1.1.1 foreign_0.8-80 resample_0.4
StatMatch_1.4.0 lpSolve_5.6.15 survey_4.0 survival_3.2-7 Matrix_1.2-18 proxy_0.4-24 usdm_1.1-18 MASS_7.3-53
wdpar_1.0.5 sf_0.9-6 ClusterR_1.2.2 gtools_3.8.2 abind_1.4-5 forcats_0.5.0 purrr_0.3.4 readr_1.4.0 tidyr_1.1.2
tibble_3.0.4 tidyverse_1.3.0 rlist_0.4.6.1 chron_2.3-56 dplyr_1.0.2 rredlist_0.7.0 stringr_1.4.0 taxize_0.9.99 ggalt_0.6.2
maps_3.3.0 pbapply_1.4-3 ncd4_1.17 plyr_1.8.6 rgeos_0.5-5 reshape2_1.4.4 pbmcapply_1.5.0 data.table_1.13.4
raster_3.4-5 ggplot2_3.3.5 rgdal_1.5-18 sp_1.4-4
```

loaded via a namespace (and not attached):

```
readxl_1.3.1 uuid_0.1-4 backports_1.2.1 TMB_1.7.18 splines_4.0.3 gmp_0.6-1 foreach_1.5.1 fansi_0.4.1
magrittr_2.0.1 openxlsx_4.2.4 modelr_0.1.8 extrafont_0.17 extrafontdb_1.0 colorspace_2.0-0 rvest_0.3.6 mitools_2.4
haven_2.3.1 xfun_0.19 crayon_1.3.4 jsonlite_1.7.2 iterators_1.0.13 ape_5.4-1 glue_1.4.2 gtable_0.3.0
emmeans_1.5.3 proj4_1.0-10 Rttf2pt1_1.3.8 mvtnorm_1.1-1 DBI_1.1.0 Rcpp_1.0.5 xtable_1.8-4 units_0.6-7
bold_1.1.0 httr_1.4.2 ellipsis_0.3.1 pkgconfig_2.0.3 reshape_0.8.8 dbplyr_2.0.0 conditionz_0.1.0 crul_1.0.0
tidyselect_1.1.0 rlang_0.4.10 munsell_0.5.0 cellranger_1.1.0 tools_4.0.3 cli_2.2.0 generics_0.1.0 sjlabelled_1.1.7
broom_0.7.3 fs_1.5.0 zip_2.1.1 ash_1.0-15 xml2_1.3.2 compiler_4.0.3 rstudioapi_0.13 beeswarm_0.2.3 curl_4.3
e1071_1.7-4 reprex_0.3.0 stringi_1.5.3 lattice_0.20-41 classInt_0.4-3 nloptr_1.2.2.2 vctrs_0.3.5 pillar_1.4.7
lifecycle_0.2.0 ucminf_1.1-4 estimability_1.3 insight_0.11.1 R6_2.5.0 KernSmooth_2.23-17 rio_0.5.27 vipor_0.4.5
codetools_0.2-16 boot_1.3-25 assertthat_0.2.1 withr_2.3.0 httpcode_0.3.0 hms_0.5.3 coda_0.19-4 class_7.3-17
minqa_1.2.4 numDeriv_2016.8-1.1 lubridate_1.7.9.2 tinytex_0.28 e
```

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

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

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

The waterbird count data used in this study are collated and managed by Wetlands International and the National Audubon Society.

The following paragraph is taken from Wauchope et al., 2019, Methods in Ecology and Evolution 10(12):2067-2078, supplementary material. The Audubon Christmas Bird Count (CBC) has been conducted yearly since 1900, primarily in the US and Canada, and involves a count of all bird individuals within a 'site', defined as a circle 24.1km in diameter. Sites are accessed by a variety of means including foot, car, boat, and snow mobile. Effort varies between sites, but the number of survey hours spent per count is documented, allowing effort to be accounted for. All CBC data is available on request, from <http://netapp.audubon.org/cbcobservation/>.

The following paragraph is taken with thanks from Amano et al., 2018, Nature. 453:199-202, 10.1038/nature25139: Launched in 1967, the IWC is a scheme involving more than 15,000 observers that monitors waterbird numbers and covers more than 25,000 sites in over 100 countries. The IWC is divided into four regions, each of which corresponds to a major migratory flyway of the world: the African–Eurasian Waterbird Census (AEWC), Asian Waterbird Census (AWC), Caribbean Waterbird Census (CWC) and Neotropical Waterbird Census (NWC). The survey methodology is essentially the same across the four regional schemes. Population counts are typically carried out once every year in mid-January. Additional counts are conducted in other months, particularly in July in the Southern Hemisphere; for consistency, we used only counts from January and February. Our Northern Hemisphere data therefore relate to non-breeding populations, whereas those from the Southern Hemisphere also include some breeding populations. In each country that is covered by the survey, national coordinators manage an inventory of wetland sites (hereafter, survey sites) that include sites of international- or national-level recognition (for example, Ramsar sites, Important Bird Areas, national parks and so on). Each survey site is generally defined by boundaries so that observers know precisely which areas are to be covered in the surveys. The observers consist of a wide variety of volunteers, but national coordinators usually train them using materials produced by Wetlands International to ensure the quality of count data. Survey sites (normally up to a few km<sup>2</sup>) are typically surveyed by about two observers for up to four hours, but larger sites can require a group of observers to work over several days. The time of survey on any given day depends on the type of survey sites: inland sites are normally surveyed during the morning or late afternoon, whereas coastal sites are surveyed during high tide periods (mangrove areas and nearby mudflats are, however, surveyed during low tides). Surveys cover waterbirds, which are defined as bird species that are ecologically dependent on wetlands<sup>29</sup>. Counts are usually made by scanning flocks of waterbirds with a telescope or binoculars and counting each species. Zero counts are not always recorded and are thus inferred using a set of criteria (see below). Count records and associated information are submitted to the national coordinators, who compile the submitted records, check their validity and submit them to Wetlands International. Data is available upon request from <http://iwc.wetlands.org/index.php/requestingdata>

All the data that pertain to explanatory variables are freely available to download, as specified in Extended Data Tables 2 and 3.

## Field-specific reporting

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

- Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

# Ecological, evolutionary & environmental sciences study design

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

Study description	This study used yearly count data of waterbird species at sites throughout the world, taken from counts made by the International Waterbird Census and Christmas Bird Count. We analysed how the trend of populations (defined as a species at a site) changed before and after the designation of protected area, for populations occurring within protected areas, and compared these to how populations changed in unprotected areas. Protected area data was taken from the world database on protected areas. We used a number of covariates to identify similar protected and unprotected populations for comparison, these are listed in Extended Data Table 1. We assessed how various covariates might correlate with protected area impact on population trend, these are listed in Extended Data Table 2.
Research sample	We took all count data at all available sites and years for waterbird species (where a waterbird is defined as any species in the families Gaviidae, Podicipedidae, Pelecanidae, Phalacrocoracidae, Anhingidae, Ardeidae, Balaenicipitidae, Scopidae, Ciconiidae, Threskiornithidae, Phoenicopteridae, Anhimidae, Anatidae, Pedionomidae, Gruidae, Aramidae, Rallidae, Heliornithidae, Eurypygidae, Jacanidae, Rostratulidae, Dromadidae, Haematopodidae, Ibisornithidae, Recurvirostridae, Burhinidae, Glareolidae, Charadriidae, Scolopacidae, Thinocoridae, Laridae or Rynchopidae) from the International Waterbird Census, acquired by request from Wetlands International, and the Christmas Bird Count, acquired by request from Audubon
Sampling strategy	Our sample size was not predetermined and there was no explicit sampling procedure. We conducted site matching to pair protected and unprotected populations for BACI and Control-Intervention analyses, which could be considered a form of sampling. To do this, we matched sites using a combination of exact matching and mahalanobis distance matching, conducted by species. For each site by species combination occurring in a protected area, we identified unprotected sites with the same species that were of the same anthrome category, continent, and where the species was of the same migratory status (as some species are migratory in some regions, and resident in others). We then calculated the mahalanobis distance between the two sites, based on a range of climate, land use and human impact characteristics (see Extended Data Table 2). Within each species, we then removed matches with the largest mahalanobis distance until the standardised difference in means of matched covariates was below 0.25 (see methods). Any species for which we could not obtain matches for 80% of sites was discarded, so that we did not get a biased impression of the impact of protected areas on that species based on the sites we could match. This resulted in an overall bias towards more common species.
Data collection	Data was collected from the International Waterbird Census, acquired by request from Wetlands International, and the Christmas Bird Count, acquired by request from Audubon, for the years 1900 to 2018.
Timing and spatial scale	All available counts were taken from 1900 to 2018. Spatially, counts were taken from all available locations in the Christmas Bird Count and International Waterbird Census, with no filtering, except that counts from Russia (a small amount) were removed as permissions for use were not given. Further filtering as described in the methods of the paper (e.g. requiring a certain number of years of sampling before and after the designation of a protected area) meant that the spatial distribution of samples was biased towards Europe and North America
Data exclusions	<p>After acquiring our data we removed the following:</p> <ul style="list-style-type: none"> <li>- Samples without coordinate information.</li> <li>- Samples taken outside the wintering months of December to February (Northern Hemisphere) and June to August (Southern Hemisphere)</li> <li>- Christmas Bird Count samples where the effort was listed as zero (i.e. no recording of the amount of effort that went into collecting the sample)</li> <li>- Samples where multiple site names were attributed to the same coordinate (indicating confusion as to how multiple sites could occur in the same place)</li> <li>- International Waterbird Census samples where coordinates were clearly wrong (e.g. in the center of a city. As decided by Amano et al., 2018, Nature, 553.</li> <li>- Any counts of hybrid species</li> </ul> <p>This resulted in a dataset of 749 waterbird species at 45,745 sites</p>
Reproducibility	As our analysis was not experimental we could not conduct replicates of experiments. However, we did need to make a number of decisions regarding various parameter estimates, such as years of sampling required, the maximum geographical distance between sites, or the strictness of matching. To ensure that our results were not vulnerable to these various decisions, we ran our entire analysis 21 times: one 'focal analysis' using our best guess parameter estimates, plus 20 analyses using estimates sampled from a plausible range for each parameter using latin hypercube sampling. Because results from these 21 analyses agreed, we can be confident our results are robust to these parameter estimates.
Randomization	We did not perform randomisation or allocate samples into groups, due to the observational nature of our study. In order to approximate a random experimental design, we conducted a BACI (Before-After-Control-Intervention) analysis by matching protected and unprotected sites. This is described in detail in the methods.
Blinding	Before conducting analysis we wrote a pre-analysis plan in order to blind ourselves to our analysis, and prevent us from hypothesising after results were known or adjusting analysis to suit outcomes (Wauchope, H. et al. Quantifying the impact of protected areas on near-global waterbird population trends, a pre-analysis plan. PeerJ Prepr. 7, e27741v, (2019).) The pre-analysis plan details our intentions for analysis, and analytical decisions. Once we conducted analysis, we found that we had not anticipated a number of decisions. Any deviations that we made from the pre-analysis plan are specified in supplementary material section 7.

Did the study involve field work?  Yes  No

## Reporting for specific materials, systems and methods

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

### Materials & experimental systems

- | n/a                                 | Included in the study   |
|-------------------------------------|---|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies                             |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines                  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology          |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Human research participants            |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data                          |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern           |

### Methods

- | n/a                                 | Included in the study                           |
|-------------------------------------|---|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq               |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

## Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals

NA

Wild animals

We used observation data on 749 waterbird species, age and sex were not identified. This data was collected by the International Waterbird Census (managed by Wetlands International) for regions outside North America, and the Christmas Bird Count (managed by Audubon) for North America.

Field-collected samples

NA

Ethics oversight

No ethical approval was required as we were using already obtained data from established NGOs.

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