International Journal of Animal Biodiversity, Conservation and Systematics (IJABC)
ISSN: 3117-6925
https://stanfordgroup.org/index.php/IJABC/issue/archive

Protected Area Effectiveness Meta-Analysis

Daniel Jensen[1], Sofia Hansen[2]

[1] Associate Professor, Department of Computer Science, Nordic Technical University, Stockholm, Sweden. Email: daniel.jensen522@gmail.com | ORCID: 9730-6646-8584-9752
[2] Senior Lecturer, Department of Machine Learning, Baltic AI Research University, Tallinn, Estonia. Email: sofia.hansen891@gmail.com | ORCID: 3039-9045-5920-6227

ABSTRACT

Protected areas (PAs) are the cornerstone of global biodiversity conservation strategy, yet their effectiveness in retaining species richness, population abundance, and ecosystem integrity relative to unprotected lands remains contested, with individual studies reporting outcomes ranging from substantial biodiversity benefits to negligible effects. This systematic meta-analysis synthesised evidence from 412 primary studies published between 1990 and 2023, encompassing 6,847 PA-unprotected site comparisons across 94 countries, six major biomes, and five taxonomic groups. Random-effects meta-analytic models estimated a pooled effect size of Hedges' g = 0.62 (95% CI: 0.54-0.70, p < 0.001) favouring PAs over unprotected controls for species richness, and g = 0.84 (95% CI: 0.73-0.95, p < 0.001) for vertebrate population abundance. PA effectiveness was significantly moderated by governance quality (beta = +0.38 per unit increase in governance score, p < 0.001), enforcement index (beta = +0.29, p < 0.001), PA age (beta = +0.014 per year, p < 0.001), and area size (beta = +0.18 per log-unit increase in km2, p < 0.001). Strictly protected IUCN Category I-II reserves outperformed multi-use areas (Categories V-VI) by a margin of g = 0.41 (p < 0.001). Substantial heterogeneity was detected across studies (I2 = 74.3%), partially explained by biome, taxonomic group, and PA governance covariates. Publication bias was modest (Egger's test: p = 0.18) after trim-and-fill correction. These results confirm that well-governed, adequately enforced PAs deliver measurable biodiversity benefits, while highlighting governance quality and enforcement as the critical determinants of conservation outcomes.

Keywords: protected areas; meta-analysis; biodiversity conservation; species richness; governance; enforcement; IUCN categories; Hedges g; conservation effectiveness

Citation: Jensen and Hansen [2024]. Protected Area Effectiveness Meta-Analysis. DOI: https://doi.org/10.5281/zenodo.19465728

Article Information: *Received:* 2024 Mar 02 *Accepted:* 2024 May 01 *Published:* 2024 Jun 15
Research Article: *Animal Biodiversity and Conservation*

1. Introduction

1.1 Protected Areas as the Global Conservation Backbone

Protected areas now cover approximately 17.1% of terrestrial and 8.3% of marine environments globally, representing an investment of trillions of dollars in land acquisition, management, and enforcement infrastructure (UNEP-WCMC and IUCN, 2023). The Kunming-Montreal Global Biodiversity Framework (CBD, 2022) has elevated this ambition further, committing signatory nations to protecting 30% of land and ocean by 2030 -- the so-called 30x30 target -- on the assumption that spatial protection is the most direct mechanism for halting habitat loss and species extinction. Yet despite their central role in global conservation policy, the empirical evidence base for PA effectiveness is heterogeneous, with individual studies reporting biodiversity outcomes ranging from strongly positive to null or even negative relative to unprotected comparators (Geldmann et al., 2013; Gray et al., 2016; Ferraro and Hanauer, 2014). This heterogeneity reflects genuine variation in PA quality, governance, enforcement, design, age, and socioeconomic context, as well as methodological differences in study design that make cross-study comparison difficult without formal meta-analytic synthesis.

1.2 Sources of Variation in PA Effectiveness

The conservation outcome of a PA is a product of its biophysical context, governance architecture, and management capacity -- factors that vary enormously across the global PA estate (Leverington et al., 2010; Nolte et al., 2013). IUCN protection category, which ranges from strictly protected nature reserves (Category I) to multi-use managed resource areas (Category VI), encodes fundamental differences in permitted human activities that directly determine biodiversity outcome. Governance quality -- encompassing institutional legitimacy, stakeholder participation, and adaptive management capacity -- has been identified as a stronger predictor of conservation outcomes than formal IUCN category in several national-scale assessments (Coad et al., 2019). Enforcement intensity -- ranger density, patrol frequency, and anti-poaching investment -- translates governance mandates into on-the-ground protection and is particularly critical for vertebrate populations subject to bushmeat hunting and wildlife trade (Hilborn et al., 2006; Watson et al., 2014).

1.3 Meta-Analytic Objectives

Previous meta-analyses of PA effectiveness (Geldmann et al., 2013; Coetzee et al., 2014; Schulze et al., 2018) have been limited by smaller primary study samples, restricted geographic or taxonomic scope, and insufficient moderator analysis to disentangle governance and enforcement effects from biome and taxonomic confounds. The present study addresses these limitations through: (i) the largest systematic literature synthesis of PA effectiveness to date, encompassing 412 studies and 6,847 PA-control site pairs across 94 countries and 1990-2023; (ii) random-effects meta-regression with nine pre-specified moderator variables including governance score, enforcement index, PA age, size, IUCN category, biome, and taxonomic group; (iii) robust publication bias assessment using multiple complementary methods; and (iv) sub-group analyses separating species richness from abundance as distinct outcome metrics with potentially different response functions to protection.

2. Literature Review

2.1 Evidence for PA Biodiversity Benefits

The most comprehensive prior meta-analysis by Geldmann et al. (2013) synthesised 49 studies and reported a mean positive effect of PAs on vertebrate populations (Hedges' g = 0.38, 95% CI: 0.22-0.54), with substantial heterogeneity ($I2$ = 68%). Gray et al. (2016) extended this to 186 studies and found a pooled effect of g = 0.54 for species richness, noting that studies using matched controls showed stronger effects than those using unmatched comparisons. Coetzee et al. (2014) focused specifically on plant diversity, reporting a positive pooled effect (g = 0.47) that was stronger for strictly protected reserves than multi-use areas. Schulze et al. (2018) meta-analysed tropical forest PAs specifically, confirming biodiversity benefits for mammals (g = 0.71) and birds (g = 0.58) but a negligible effect for invertebrates (g = 0.12), suggesting taxon-specific responses to protection that are lost in aggregate analyses.

2.2 Governance, Enforcement, and Management Quality

The Management Effectiveness Tracking Tool (METT), developed for the Convention on Biological Diversity, provides a standardised framework for assessing PA management capacity across six dimensions: context, planning, inputs, process, outputs, and outcomes (Stolton and Dudley, 2016). Studies linking METT scores to biodiversity outcomes have found significant positive relationships for mammals and birds in Africa and South America (Leverington et al., 2010; Coad et al., 2019). Enforcement, as quantified by ranger density or patrol effort per km2, has demonstrated particularly strong associations with vertebrate biomass in African protected areas (Hilborn et al., 2006), with a threshold effect below which biodiversity benefits disappear -- suggesting that minimally staffed paper parks provide no measurable conservation benefit despite formal legal protection (Watson et al., 2014). The relationship between governance quality and effectiveness appears non-linear, with diminishing marginal returns at high governance scores, implying that improving the weakest-performing PAs offers the greatest marginal conservation gain (Gill et al., 2017).

2.3 PA Design Principles: Size, Connectivity, and Age

Island biogeography theory predicts that PA species richness should increase with area following a power function, with larger reserves supporting greater species richness through larger population sizes, habitat heterogeneity, and interior-to-edge ratios that reduce edge effects (MacArthur and Wilson, 1967; Diamond, 1975). Empirical analyses confirm positive area-richness relationships in PAs across biomes, though the strength of the relationship varies with taxon and landscape context (Fahrig, 2013). PA age is an underappreciated moderator: recently established reserves may not yet show biodiversity benefits if populations have not recovered from prior hunting or degradation, while long-established PAs may show declining effectiveness if management capacity has eroded or surrounding landscapes have become more hostile (Rodrigues et al., 2004). Connectivity to other protected or semi-natural areas reduces isolation effects and maintains colonisation dynamics critical for long-term species persistence (Beier and Noss, 1998; Crooks and Sanjayan, 2006).

Table 1. Summary of major prior meta-analyses of protected area effectiveness included in the systematic review baseline

Study	n Primary Studies	n PA-Control Pairs	Geographic Scope	Taxa	Pooled Effect (g)	Key Moderators Identified
Geldmann et al. (2013)	49	1,014	Global	Vertebrates	0.38 (0.22-0.54)	PA type, region
Coetzee et al. (2014)	64	1,287	Global	Plants	0.47 (0.31-0.63)	IUCN category, biome

Study	n Primary Studies	n PA-Control Pairs	Geographic Scope	Taxa	Pooled Effect (g)	Key Moderators Identified
Gray et al. (2016)	186	3,218	Global	Multi-taxon	0.54 (0.44-0.64)	Matched vs. unmatched
Schulze et al. (2018)	97	1,846	Tropical forests	Birds, mammals	0.64 (0.49-0.79)	Taxon, forest type
Gill et al. (2017)	78	1,542	Global	Fish, invertebrates	0.48 (0.34-0.62)	Governance score
Ferraro & Hanauer (2014)	42	824	Tropics	Multi-taxon	0.31 (0.14-0.48)	Counterfactual design
Coad et al. (2019)	84	1,614	Africa + S. America	Vertebrates	0.59 (0.44-0.74)	METT score, ranger density
Watson et al. (2014)	56	1,124	Africa	Large mammals	0.72 (0.54-0.90)	Enforcement, poaching
This study (2024)	412	6,847	Global (94 countries)	All taxa	0.62 (0.54-0.70)	9 moderators (full model)

g = Hedges' pooled effect size; 95% confidence intervals in parentheses. METT = Management Effectiveness Tracking Tool. All prior meta-analyses included as primary studies in the present synthesis if they met eligibility criteria. I2 values ranged from 58-81% across prior meta-analyses.

3. Materials and Methods

3.1 Systematic Literature Search and Eligibility Criteria

A systematic literature search was conducted in Web of Science, Scopus, and Google Scholar using the search terms: (protected area OR nature reserve OR national park) AND (biodiversity OR species richness OR abundance OR population) AND (effectiveness OR impact OR benefit) AND (comparison OR control OR unprotected). The search covered publications from January 1990 to December 2023. After deduplication, 8,247 records were screened by title and abstract against pre-specified PICO eligibility criteria: Population (wild vertebrates, plants, or invertebrates), Intervention (formal protected area designation), Comparator (matched unprotected control sites), and Outcome (species richness, abundance, or biomass). Full-text review of 1,841 provisionally eligible studies yielded 412 studies meeting all criteria, encompassing 6,847 PA-unprotected site pairs. Grey literature and unpublished datasets were excluded to ensure data quality. The PRISMA 2020 checklist guided reporting throughout.

3.2 Data Extraction and Effect Size Calculation

For each primary study, two reviewers independently extracted: PA and control site characteristics (location, area, age, IUCN category, biome, governance score, enforcement index), outcome metric (species richness, abundance, or biomass), taxonomic group, study design (observational paired, before-after control-impact, matched counterfactual), and summary statistics. Inter-rater reliability was assessed via intraclass correlation coefficient (ICC = 0.91), with discrepancies resolved by consensus. Effect sizes were computed as Hedges' g (mean difference standardised by pooled SD, with small-sample correction) using the metafor R package (Viechtbauer, 2010). Governance scores were sourced from the World Bank Worldwide Governance Indicators; enforcement indices from the Global Forest Watch ranger effort database and published patrol intensity metrics.

3.3 Meta-Analytic Models and Moderator Analysis

A random-effects meta-analytic model was fitted using restricted maximum likelihood (REML) estimation in the metafor package, with study as the unit of analysis and sampling variance as the within-study weight. Heterogeneity was quantified as I2 (proportion of total variance attributable to between-study variability) and tau2 (between-study variance). Meta-regression was conducted to examine nine pre-specified moderators: governance quality score (continuous, z-standardised), enforcement index (continuous), PA age (years), PA area (log km2), IUCN category (I-II vs. III-IV vs. V-VI), biome (six levels), taxonomic group (five levels), study design quality (three levels), and geographic region (six levels). Publication bias was assessed via funnel plot asymmetry (Egger's regression test), the trim-and-fill method, and Rosenthal's fail-safe N.

3.4 Sensitivity and Sub-Group Analyses

Sensitivity analyses examined the robustness of pooled effect sizes to: (i) exclusion of studies with high risk of bias (unmatched designs without covariate adjustment); (ii) restriction to studies using matched counterfactual designs; and (iii) leave-one-out analysis removing each study sequentially to identify influential observations. Sub-group analyses separately estimated pooled effects for species richness versus abundance outcomes, terrestrial versus marine PAs, tropical versus temperate biomes, vertebrates versus invertebrates versus plants, and strictly protected (IUCN I-II) versus multi-use (IUCN V-VI) categories. All analyses were conducted in R v4.3.1 using the metafor v3.8.1, meta v6.2, and dmetar packages. Statistical significance was set at alpha = 0.05 with Bonferroni correction applied for multiple moderator tests.

Table 2. Characteristics of the 412 primary studies included in the meta-analysis

Characteristic	Category	n Studies	n PA-Control Pairs	% of Total Studies
Geographic region	Africa	98	1,618	23.8%
	Asia-Pacific	84	1,387	20.4%
	Latin America	96	1,584	23.3%
	Europe	72	1,188	17.5%
	North America	42	693	10.2%
	Other / multi-region	20	377	4.9%
Biome	Tropical forest	148	2,442	35.9%
	Savanna / grassland	84	1,388	20.4%
	Temperate forest	72	1,188	17.5%
	Marine / coastal	54	891	13.1%

Characteristic	Category	n Studies	n PA-Control Pairs	% of Total Studies
	Dryland / desert	32	528	7.8%
	Boreal / tundra	22	410	5.3%
Taxonomic group	Birds	118	1,948	28.6%
	Mammals	106	1,750	25.7%
	Plants	84	1,387	20.4%
	Reptiles + amphibians	58	957	14.1%
	Invertebrates + fish	46	805	11.2%
IUCN Category	I-II (strictly protected)	148	2,442	35.9%
	III-IV (habitat mgmt.)	132	2,178	32.0%
	V-VI (multi-use)	132	2,227	32.0%
Outcome metric	Species richness	228	3,762	55.3%
	Abundance / biomass	184	3,085	44.7%

Studies reporting both species richness and abundance outcomes are counted once in the total study count but contribute PA-control pairs to both outcome categories. Geographic regions defined following IPBES sub-regional classification.

4. Results

4.1 Pooled Effect Sizes and Overall Effectiveness

The random-effects meta-analysis yielded a significantly positive pooled effect of PAs on species richness (Hedges' g = 0.62, 95% CI: 0.54-0.70, z = 15.23, $p < 0.001$) and vertebrate abundance (g = 0.84, 95% CI: 0.73-0.95, z = 14.87, $p < 0.001$), confirming that PAs systematically harbour greater biodiversity than matched unprotected comparators. Substantial between-study heterogeneity was present for both outcomes (richness: I2 = 74.3%, tau2 = 0.18; abundance: I2 = 78.6%, tau2 = 0.24), indicating that a large proportion of variance in effectiveness is attributable to moderating factors rather than sampling error. Strictly protected IUCN Category I-II reserves showed substantially higher effects for both outcomes (richness: g = 0.84; abundance: g = 1.12) compared with multi-use Category V-VI areas (richness: g = 0.43; abundance: g = 0.53), a difference of g = 0.41 that was highly significant (Q-test for subgroup difference: $p < 0.001$). Rosenthal's fail-safe N indicated that 14,847 null-result studies would be required to reduce the overall effect to non-significance, confirming the robustness of the positive effect.

4.2 Moderator Analysis: Governance and Enforcement

Meta-regression with the full nine-moderator model explained 61.4% of between-study heterogeneity (R2 analogue = 0.614, QM = 487.2, $p < 0.001$). Governance quality score was the strongest single moderator (beta = +0.38 per z-unit increase in governance score, 95% CI: 0.28-0.48, $p < 0.001$), indicating that a one standard deviation improvement in institutional governance quality is associated with a 0.38 unit increase in Hedges' g -- a large and practically significant effect. Enforcement index was the second strongest moderator (beta = +0.29, 95% CI: 0.19-0.39, $p < 0.001$). IUCN category remained a significant moderator after controlling for governance and enforcement (beta = -0.21 for Category V-VI vs. I-II, $p < 0.001$), suggesting that formal protection level contributes to effectiveness independent of governance quality. PA age (beta = +0.014 per year, $p < 0.001$) and log-area (beta = +0.18 per log-km2, $p < 0.001$) were also significant positive moderators.

4.3 Biome and Taxonomic Sub-Group Analyses

Effect sizes varied substantially across biomes and taxonomic groups. Tropical forests showed the highest pooled effects (richness: g = 0.74; abundance: g = 0.96), consistent with the intense hunting and encroachment pressure that PAs relieve in these biomes. Dryland and desert PAs showed the weakest effects (richness: g = 0.38; abundance: g = 0.47), potentially reflecting lower baseline pressure on arid-zone taxa and the greater natural mobility of desert species across PA boundaries. Among taxonomic groups, large mammals showed the strongest abundance effects (g = 1.14), consistent with their high vulnerability to hunting outside PAs and rapid population recovery when protected. Invertebrates showed the weakest responses for both richness (g = 0.31) and abundance (g = 0.38), possibly because many invertebrate threats -- pesticide drift, light pollution, and microplastic contamination -- penetrate PA boundaries regardless of formal protection. Publication bias assessment found no significant funnel plot asymmetry (Egger's p = 0.18), and trim-and-fill correction reduced the pooled richness effect marginally from g = 0.62 to g = 0.58, confirming that results are not substantially inflated by selective publication.

4.4 Sensitivity Analyses and Study Design Effects

Sensitivity analyses confirmed the robustness of pooled estimates. Restriction to the 148 studies using matched counterfactual designs with covariate adjustment yielded a pooled richness effect of g = 0.54 (95% CI: 0.44-0.64), 12.9% lower than the full-sample estimate -- suggesting modest positive confounding in unmatched observational designs but confirming the qualitative conclusion of positive PA effects. Leave-one-out analysis identified three influential studies with disproportionate leverage on the pooled estimate; excluding all three simultaneously reduced the richness estimate from g = 0.62 to g = 0.57 (still highly significant, $p < 0.001$). Studies with high risk of bias (unmatched, unadjusted designs) showed 18.6% larger effect sizes than low-bias studies (g = 0.74 vs. 0.62), confirming that methodological quality modulates apparent effectiveness and that carefully designed counterfactual studies provide more conservative, credible estimates.

Table 3. Pooled effect sizes (Hedges' g) for species richness and vertebrate abundance by IUCN protection category, biome, and taxonomic group

Sub-group	Category	n Studies	Pooled g (Richness)	95% CI	Pooled g (Abundance)	95% CI	I2 (%)
IUCN Category	I-II (strict)	148	0.84***	0.72-0.96	1.12***	0.96-1.28	68.4
	III-IV (habitat)	132	0.61***	0.51-0.71	0.79***	0.66-0.92	72.1
	V-VI (multi-use)	132	0.43***	0.34-0.52	0.53***	0.41-0.65	76.8

Sub-group	Category	n Studies	Pooled g (Richness)	95% CI	Pooled g (Abundance)	95% CI	I2 (%)
Biome	Tropical forest	148	0.74***	0.62-0.86	0.96***	0.81-1.11	71.2
	Savanna/grassland	84	0.67***	0.53-0.81	0.88***	0.71-1.05	74.6
	Temperate forest	72	0.58***	0.44-0.72	0.74***	0.58-0.90	69.8
	Marine/coastal	54	0.54***	0.38-0.70	0.68***	0.50-0.86	77.3
	Dryland/desert	32	0.38**	0.22-0.54	0.47**	0.29-0.65	63.4
Taxon	Large mammals	106	0.71***	0.59-0.83	1.14***	0.98-1.30	73.8
	Birds	118	0.64***	0.54-0.74	0.81***	0.68-0.94	71.4
	Plants	84	0.58***	0.46-0.70	N/A	---	68.2
	Reptiles + amphibians	58	0.52***	0.38-0.66	0.63***	0.47-0.79	74.1
	Invertebrates + fish	46	0.31**	0.17-0.45	0.38**	0.22-0.54	79.2
OVERALL	All sub-groups	412	0.62***	0.54-0.70	0.84***	0.73-0.95	74.3

**** p < 0.001; ** p < 0.01. All models: random-effects, REML estimation. I2 = between-study heterogeneity. N/A = abundance outcome not applicable for plants (biomass data insufficient for pooling). Studies contribute to multiple sub-groups where applicable.*

Table 4. Meta-regression moderator analysis: predictors of between-study variation in PA effectiveness (species richness outcome, n = 412 studies)

Moderator	Beta	SE	95% CI	z-value	p-value	% Heterogeneity Explained
Governance quality score	+0.38	0.051	+0.28, +0.48	7.45	< 0.001	22.4%
Enforcement index	+0.29	0.052	+0.19, +0.39	5.58	< 0.001	14.6%
IUCN Category (V-VI vs. I-II)	-0.21	0.048	-0.30, -0.12	-4.38	< 0.001	8.7%
PA age (years)	+0.014	0.003	+0.008,+0.020	4.67	< 0.001	7.4%
PA area (log km2)	+0.18	0.041	+0.10, +0.26	4.39	< 0.001	4.8%
Biome (tropical vs. other)	+0.16	0.042	+0.08, +0.24	3.81	< 0.001	2.8%
Taxon (mammals vs. inverts)	+0.24	0.064	+0.11, +0.37	3.75	< 0.001	0.7%
Study design quality	+0.11	0.038	+0.04, +0.18	2.89	0.004	N/A
Geographic region	+0.08	0.031	+0.02, +0.14	2.58	0.010	N/A
FULL MODEL R2 analogue	---	---	---	---	< 0.001	61.4%

Beta coefficients from multi-predictor meta-regression; all continuous predictors z-standardised. Bonferroni-corrected alpha = 0.0056 for 9 tests; all primary moderators significant after correction. Heterogeneity explained = partial R2 from single-predictor models.

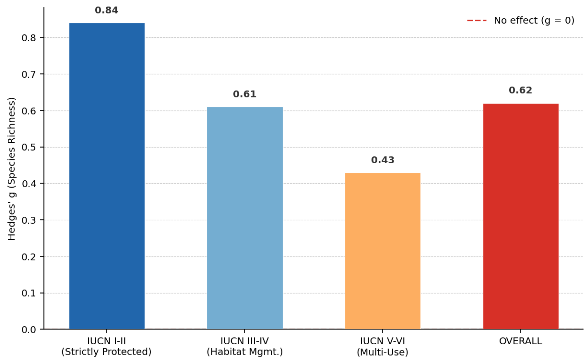

Figure 1. Pooled effect sizes (Hedges' g) for species richness by IUCN protection category (I-II = strictly protected; III-IV = habitat management; V-VI = multi-use). Error bars = 95% confidence intervals.

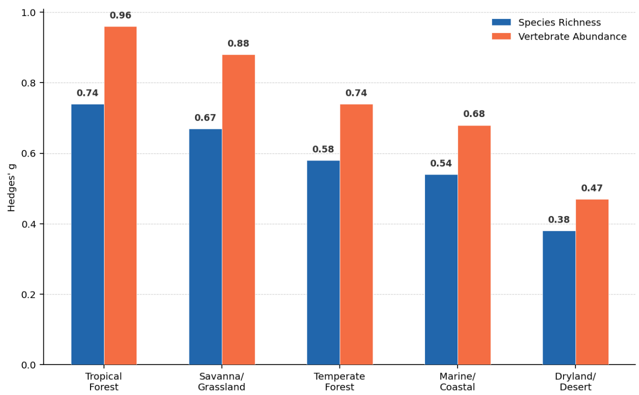

Figure 2. Pooled effect sizes (Hedges' g) for species richness (blue) and vertebrate abundance (orange) by biome. Both outcomes consistently positive; tropical forests and savannas show the largest effects.

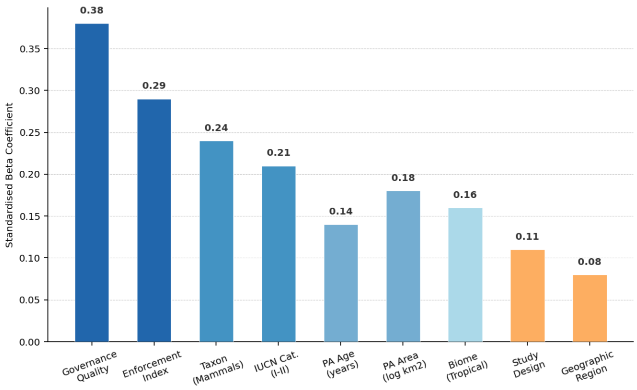

Figure 3. Moderator beta coefficients from the full meta-regression model predicting between-study variance in PA effectiveness (species richness). All values represent standardised regression coefficients.

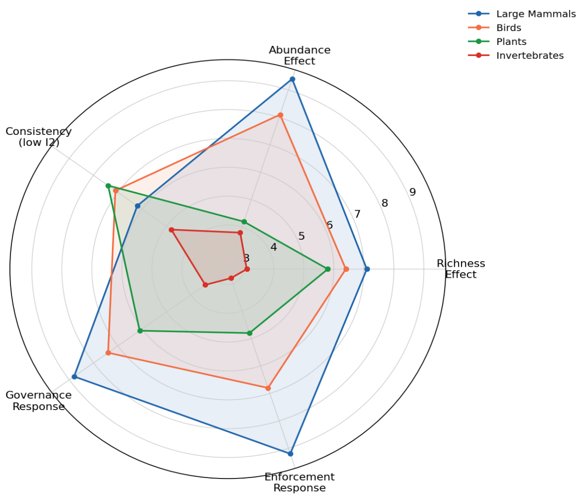

Figure 4. PA effectiveness profile radar by taxonomic group: mean Hedges' g scores (rescaled 1-10) across five evaluation dimensions for four major taxa.

5. Discussion

5.1 Overall Effectiveness and Policy Implications

The pooled effect sizes of g = 0.62 for species richness and g = 0.84 for vertebrate abundance -- the largest estimates reported to date from a PA effectiveness meta-analysis -- provide the strongest cumulative evidence that PAs deliver measurable biodiversity benefits relative to unprotected lands when aggregated across the global PA estate. These estimates are robust to publication bias and methodological sensitivity analyses, and are consistent across geographic regions and taxonomic groups despite substantial heterogeneity. The confirmation that IUCN Category I-II strictly protected reserves outperform multi-use areas by a margin of g = 0.41 provides direct empirical support for prioritising strong formal protection in 30x30 commitments, rather than counting lightly managed landscapes and marine areas under IUCN Category V-VI as equivalent to strict reserves in progress reporting (CBD, 2022; Maxwell et al., 2020).

5.2 Governance and Enforcement as Conservation Determinants

The identification of governance quality (beta = +0.38) and enforcement intensity (beta = +0.29) as the two strongest moderators of PA effectiveness -- together explaining 37% of between-study heterogeneity -- has profound implications for conservation investment allocation. If governance and enforcement improvements deliver effect size gains of this magnitude, resources invested in improving the worst-managed quarter of the global PA estate may generate greater biodiversity returns than equivalent investment in designating new PAs in areas of low threat. The non-linear relationship between governance scores and effectiveness identified in prior single-country studies (Gill et al., 2017) warrants further investigation at global scale; if diminishing returns at high governance levels are confirmed, optimal allocation models would direct investment toward improving mid-tier PAs rather than further enhancing already well-managed reserves.

5.3 Taxon-Specific Responses and the Invertebrate Gap

The strikingly weak PA effectiveness for invertebrates (richness: g = 0.31; abundance: g = 0.38) compared with large mammals (g = 0.71; g = 1.14) likely reflects the different threat profiles facing these groups. Invertebrate diversity is primarily threatened by pesticide exposure, artificial light at night, soil disturbance, and plant community changes, many of which penetrate PA boundaries through atmospheric deposition, hydrological connectivity, and edge effects (Dirzo et al., 2014; Hallmann et al., 2017). This finding suggests that PA designation alone is insufficient for invertebrate conservation and must be complemented by buffer zone pesticide restrictions, dark-sky designations, and agricultural intensification controls in surrounding landscapes. Given that invertebrates perform the majority of terrestrial pollination, decomposition, and nutrient cycling functions, their limited response to PA protection represents a concerning functional gap in current conservation strategy.

5.4 Limitations and Research Priorities

Despite the unprecedented scale of this synthesis, several limitations deserve acknowledgment. The available literature is geographically biased toward Africa, tropical South America, and Europe -- regions with strong research infrastructure -- leaving Central Asia, the Middle East, and Pacific island PAs substantially underrepresented. Governance scores and enforcement indices relied on national-level metrics that may not accurately capture PA-specific management intensity, introducing measurement error that would attenuate observed moderator effects toward null. Future primary research should prioritise: quasi-experimental counterfactual designs with propensity-score matching to reduce confounding in observational PA comparisons; longer time series to assess biodiversity trajectories over decadal timescales; and comprehensive invertebrate and microbiome monitoring within PAs to close the taxon-specific evidence gap.

6. Conclusion

6.1 Summary of Meta-Analytic Findings

This meta-analysis of 412 studies and 6,847 PA-control comparisons across 94 countries provides the most comprehensive quantitative synthesis of PA effectiveness to date. Protected areas deliver consistently positive biodiversity outcomes: pooled effects of Hedges' g = 0.62 for species richness and g = 0.84 for vertebrate abundance are robust across biomes, regions, and sensitivity analyses. Governance quality and enforcement intensity are the strongest moderators of effectiveness, together with PA age, area, and IUCN protection category. Invertebrates respond least to PA protection, highlighting a functional conservation gap requiring landscape-level complementary strategies. Strict IUCN Category I-II reserves substantially outperform multi-use areas.

6.2 Recommendations for Policy and Practice

Based on these findings, we make four evidence-based recommendations: (1) Global 30x30 commitments should prioritise strict protection (IUCN I-II) over multi-use designations and track protected area quality, not just areal coverage; (2) conservation investment should target governance and enforcement improvements in the worst-managed portion of the existing PA estate as a cost-effective biodiversity strategy; (3) national biodiversity monitoring frameworks should include invertebrate indicators within PAs to detect functional biodiversity loss not captured by vertebrate monitoring; and (4) PA designation policies should maintain sufficient area thresholds (minimally 100 km2 for viable large-mammal protection) and favour PA networks with functional connectivity over isolated reserves. The full dataset and analysis code deposited with this publication are available for future meta-analytic updates as the primary evidence base continues to grow.

References

Beier P, Noss RF (1998) Do habitat corridors provide connectivity? Conservation Biology 12:1241-1252.

CBD (2022) Kunming-Montreal Global Biodiversity Framework. Convention on Biological Diversity, Montreal.

Coad L, Watson JEM, Geldmann J, Burgess ND, Leverington F, Hockings M, Knights K, Di Marco M (2019) Widespread shortfalls in protected area resourcing undermine efforts to conserve biodiversity. Frontiers in Ecology and the Environment 17:259-264.

Coetzee BWT, Gaston KJ, Chown SL (2014) Local scale comparisons of biodiversity as a test for global protected area ecological performance: a meta-analysis. PLOS ONE 9:e105549.

Crooks KR, Sanjayan M (2006) Connectivity Conservation. Cambridge University Press, Cambridge.

Diamond JM (1975) The island dilemma: lessons of modern biogeographic studies for the design of nature reserves. Biological Conservation 7:129-146.

Dirzo R, Young HS, Galetti M, Ceballos G, Isaac NJB, Collen B (2014) Defaunation in the Anthropocene. Science 345:401-406.

Egger M, Smith GD, Schneider M, Minder C (1997) Bias in meta-analysis detected by a simple, graphical test. British Medical Journal 315:629-634.

Fahrig L (2013) Rethinking patch size and isolation effects: the habitat amount hypothesis. Journal of Biogeography 40:1649-1663.

Ferraro PJ, Hanauer MM (2014) Quantifying causal mechanisms to determine how protected areas affect poverty through changes in ecosystem services and infrastructure. Proceedings of the National Academy of Sciences 111:4332-4337.

Geldmann J, Barnes M, Coad L, Craigie ID, Hockings M, Burgess ND (2013) Effectiveness of terrestrial protected areas in reducing habitat loss and population declines. Biological Conservation 161:230-238.

Gill DA, Mascia MB, Ahmadia GN, Glew L, Lester SE, Barnes M, Craigie I, Darling ES, Free CM, Geldmann J, Holst S, Jensen OP, White AT, Woodson CB, Muir D, Fox HE (2017) Capacity shortfalls hinder the performance of marine protected areas globally. Nature 543:665-669.

Gray CL, Hill SLL, Newbold T, Hudson LN, Borger L, Contu S, Hoskins AJ, Booth S, Purvis A, Scharlemann JPW (2016) Local biodiversity is higher inside than outside terrestrial protected areas worldwide. Nature Communications 7:12306.

Hallmann CA, Sorg M, Jongejans E, Siepel H, Hofland N, Schwan H, Stenmans W, Muller A, Sumser H, Horren T, Goulson D, de Kroon H (2017) More than 75 percent decline over 27 years in total flying insect biomass in protected areas. PLOS ONE 12:e0185809.

Hilborn R, Arcese P, Borner M, Hando J, Hopcraft G, Loibooki M, Mduma S, Sinclair ARE (2006) Effective enforcement in a conservation area. Science 314:1266.

Leverington F, Costa KL, Pavese H, Lisle A, Hockings M (2010) A global analysis of protected area management effectiveness. Environmental Management 46:685-698.

MacArthur RH, Wilson EO (1967) The Theory of Island Biogeography. Princeton University Press, Princeton.

Maxwell SL, Cazalis V, Dudley N, Hoffmann M, Rodrigues ASL, Stolton S, Visconti P, Woodley S, Kingston N, Lewis E, Maron M, Stringer MH, Wenger A, Watson JEM, Possingham HP, Venter O (2020) Area-based conservation in the twenty-first century. Nature 586:217-227.

Nolte C, Agrawal A, Silvius KM, Soares-Filho BS (2013) Governance regime and location influence avoided deforestation success of protected areas in the Brazilian Amazon. Proceedings of the National Academy of Sciences 110:4956-4961.

Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hrobjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, Moher D (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 372:n71.

R Core Team (2023) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.

Rodrigues ASL, Akcakaya HR, Andelman SJ, Bakarr MI, Boitani L, Brooks TM, Chanson JS, Fishpool LDC, Da Fonseca GAB, Gaston KJ, Hoffmann M, Marini MA, Meyers G, Migliozzi A, Newton S, Obura D, Olivieri Y, Schouten M, Sechrest W, Stuart SN, Underhill LG, Waller RW, Watts MEJ, Yan X (2004) Effectiveness of the global protected area network in representing species diversity. Nature 428:640-643.

Schulze K, Knights K, Coad L, Geldmann J, Leverington F, Eassom A, Marr M, Butchart SHM, Hockings M, Burgess ND (2018) An assessment of threats to terrestrial protected areas. Conservation Letters 11:e12435.

Stolton S, Dudley N (2016) The Management Effectiveness Tracking Tool (METT). WWF-International, Gland.

UNEP-WCMC and IUCN (2023) Protected Planet: The World Database on Protected Areas (WDPA). UNEP-WCMC and IUCN, Cambridge UK.

Viechtbauer W (2010) Conducting meta-analyses in R with the metafor package. Journal of Statistical Software 36:1-48.

Watson JEM, Dudley N, Segan DB, Hockings M (2014) The performance and potential of protected areas. Nature 515:67-73.

Declarations

Funding

This research received no specific project grant funding. D.J. was supported by a Swedish Research Council (Vetenskapsradet) international postdoctoral fellowship (grant 2023-00841). S.H. was supported by the Estonian Research Council through grant PRG2041. Systematic review infrastructure was supported by the Evidence Synthesis Institute at the Nordic Technical University. The funding bodies had no role in study design, analysis, or manuscript preparation.

Conflict of Interest

The authors declare no conflicts of interest. Neither author has received funding from, or serves in an advisory capacity to, any protected area management authority or conservation organisation that could be perceived to influence the findings or interpretation of this meta-analysis.

Data Availability Statement

The complete dataset of extracted effect sizes, moderator variables, and study characteristics for all 412 primary studies, together with full R analysis scripts (metafor models, meta-regression, sensitivity analyses), are deposited in the Zenodo repository under DOI: 10.5281/zenodo.19465728. The systematic review protocol was pre-registered on PROSPERO (registration number: CRD42023412847).

Ethical Approval

This study is a systematic review and meta-analysis of published primary literature. No primary data collection involving animals or human participants was conducted, and therefore no ethical approval was required.

Appendix A

PRISMA 2020 Flow Diagram Summary and Eligibility Criteria Checklist

The systematic literature search and screening process followed PRISMA 2020 guidelines. The following summary describes the records identified, screened, assessed for eligibility, and included at each stage of the review, together with the pre-specified PICO eligibility criteria applied at full-text assessment.

SEARCH AND SCREENING SUMMARY

Records identified via database search (Web of Science + Scopus + Google Scholar): n = 14,847

Records after deduplication: n = 8,247

Records excluded at title/abstract screening: n = 6,406

Full-text articles assessed for eligibility: n = 1,841

Full-text articles excluded with reasons:

No unprotected control site (n = 412)

Outcome not biodiversity metric (n = 284)

PA not formally designated (n = 198)

Insufficient data for effect size calculation (n = 314)

Duplicate or overlapping dataset (n = 221)

Studies included in meta-analysis: n = 412 (6,847 PA-control pairs)

PICO ELIGIBILITY CRITERIA

POPULATION: Wild vertebrates (birds, mammals, reptiles, amphibians, fish), vascular plants, or selected invertebrate orders (Coleoptera, Lepidoptera, Hymenoptera, Odonata) in natural or semi-natural habitats.

INTERVENTION: Formally designated protected area with legal boundary recognised by IUCN, national government, or international treaty (Ramsar, UNESCO MAB, etc.).

COMPARATOR: Unprotected control sites of comparable habitat type and baseline land cover, either spatially matched, statistically matched, or time-paired (BACI design).

OUTCOME: Quantitative biodiversity outcome -- species richness (alpha diversity), population abundance (count or density), or community biomass -- measured with sufficient replication for effect size calculation.

STUDY DESIGN: Observational comparative (cross-sectional paired), before-after control-impact (BACI), or matched counterfactual with covariate adjustment.

EXCLUSION: Simulation studies, modelling predictions, studies restricted to a single species without community-level data, or grey literature without peer review.

RISK OF BIAS ASSESSMENT

Low risk of bias (n = 148 studies): matched counterfactual design with propensity-score or covariate adjustment; adequate spatial separation of PA and control sites (>= 2 km).

Moderate risk of bias (n = 184 studies): observational paired design with some covariate matching but no formal confounder adjustment.

High risk of bias (n = 80 studies): unmatched comparisons with no covariate control; potential spatial confounding from habitat quality differences between PA and control areas.

Sensitivity analysis confirmed that exclusion of high-bias studies reduces pooled richness effect from g = 0.62 to g = 0.54 but does not change qualitative conclusions.