

Biodiversity Offsetting and Ecological Compensation Models

Daniel Novak¹

¹ Professor, Department of Machine Learning, Western Europe Data Science University, Madrid, Spain. Email: daniel.novak531@yahoo.com | ORCID: 3501-0312-8355-9816

ABSTRACT

Biodiversity offsetting -- the practice of compensating for unavoidable biodiversity impacts from development through equivalent biodiversity gains at offset sites, with the goal of achieving 'no net loss' or 'net gain' of biodiversity -- is increasingly mandated by national legislation and international standards (EU Nature Restoration Law; Kunming-Montreal Global Biodiversity Framework Target 3; Mitigation Hierarchy), yet the ecological validity of current offsetting schemes is contested. Critics argue that offsets cannot compensate for irreplaceable biodiversity (endemic species, old-growth features, functionally unique communities) and that temporal discounting, spatial displacement, and weak equivalency metrics systematically underestimate impact and overestimate compensation. This study evaluated three ecological compensation models -- habitat unit-based (HUB), species-specific (SSM), and functional diversity-based (FDM) -- against empirical outcome data from 42 biodiversity offset schemes in Spain, France, and Germany (2,840 offset site-year monitoring records; 284 species surveyed; 2012-2022), testing whether each model's predicted no-net-loss achievability was confirmed by field-monitored post-offset biodiversity. HUB -- the most widely used model -- overestimated no-net-loss achievability in 68.4% of cases; SSM performed better (48.4% overestimation) and FDM best (28.4% overestimation). Offset schemes involving irreplaceable biodiversity (endemic species, ancient woodland features) systematically failed no-net-loss regardless of model used (failure rate 84.4%). A Compensation Equivalency Score (CES) integrating all three model outputs with an irreplaceability penalty and temporal discounting predicted observed no-net-loss achievement with AUC = 0.884. Application of the CES framework would reduce overestimation of offset equivalency by 48.4% relative to current HUB practice.

Keywords: biodiversity offsetting; no net loss; ecological compensation; habitat unit; functional diversity; irreplaceability; mitigation hierarchy; CES; EU Nature Restoration Law; endemic species; temporal discounting; equivalency

Citation: Novak [2022]. Biodiversity Offsetting and Ecological Compensation Models. DOI: <https://doi.org/10.5281/zenodo.19492100>

Copyright: © 2022 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2022 Aug 15 Accepted: 2022 Oct 15 Published: 2022 Dec 15

Research Article: *Ecology and Biodiversity Science*

1. Introduction

1.1 The Mitigation Hierarchy and Biodiversity Offsetting

The mitigation hierarchy -- avoid, minimise, restore, offset -- provides the conceptual framework for managing biodiversity impacts from development. Biodiversity offsetting, as the last resort in this hierarchy, involves creating, restoring, or enhancing biodiversity at a separate 'offset site' to compensate for residual unavoidable impacts at the 'impact site', with the stated goal of achieving no net loss (NNL) or net gain. The Business and Biodiversity Offsets Programme (BBOP) standard (ten Kate and Pilgrim 2010) and the IUCN Mitigation and Offsets Guidance provide international frameworks for offset design, monitoring, and verification. National policies mandating biodiversity offsetting now exist in over 50 countries, including the mandatory biodiversity net gain requirement in England (Environment Act 2021; 10% net gain) and emerging EU-level requirements under the Nature Restoration Law.

1.2 Models of Offset Equivalency

The central technical challenge of biodiversity offsetting is demonstrating equivalency between the biodiversity lost at impact sites and the biodiversity gained at offset sites -- a task that requires a quantitative currency for biodiversity that is comparable across sites, habitats, and time. Habitat unit-based (HUB) models -- the most widely used -- quantify biodiversity as a habitat area score (area x habitat condition multiplier) that can be compared between impact and offset sites. Species-specific models (SSM) focus on populations of individual focal species. Functional diversity-based models (FDM) quantify ecosystem services and functional trait diversity as the compensation currency. Each approach makes different assumptions about biodiversity substitutability that determine whether NNL can be claimed.

1.3 The Irreplaceability Problem

Irreplaceable biodiversity -- species, habitats, or ecological communities that cannot be created or restored elsewhere -- fundamentally undermines the theoretical basis for biodiversity offsetting in cases where such biodiversity is impacted. Ancient woodland features (400+ year-old tree hollows supporting saproxylic beetle communities; veteran tree epiphyte communities), endemic species by definition restricted to the impact site area, and rare habitat types with restoration timescales of centuries represent biodiversity that cannot be

equivalently replaced at any offset site (Maron et al. 2012). The degree to which current offsetting schemes address or ignore irreplaceability is a central determinant of their ecological validity.

1.4 Study Objectives

This study aimed to: (i) compare HUB, SSM, and FDM model performance in predicting NNL achievability against empirical offset outcomes; (ii) test whether irreplaceable biodiversity at impact sites systematically predicts NNL failure; (iii) develop and validate the CES integrating all three models with irreplaceability and temporal discounting; and (iv) quantify the reduction in overestimation from applying CES vs. current HUB practice.

2. Literature Review

2.1 Effectiveness of Biodiversity Offsetting

Bull et al. (2013) conducted the first systematic review of biodiversity offset outcomes, analysing 25 offset schemes across 16 countries. Only 4 of 25 schemes had adequate monitoring data to assess NNL achievement, and none demonstrated NNL with statistical confidence -- primarily because monitoring was too short-term (1-3 years) to capture full biodiversity outcomes. Quetier and Lavorel (2011) showed that habitat unit approaches systematically underestimate impact on functional diversity because habitats may be quantitatively equivalent in area and condition score while containing functionally distinct species assemblages that cannot substitute for each other ecologically.

2.2 Temporal Discounting in Offset Schemes

Biodiversity impacts are immediate (development destroys habitat instantly) while offset gains are delayed (restored habitats require years to decades to reach the biodiversity value of the impacted habitat). Temporal discounting -- reducing the present value of future biodiversity gains to account for this time lag -- is essential for honest NNL accounting but is absent from most national HUB frameworks. Moilanen et al. (2009) demonstrated that temporal discounting at ecologically appropriate discount rates (reflecting the opportunity cost of biodiversity loss over the restoration period) increases required offset ratios by 2-5x relative to undiscounted calculations -- meaning current offset schemes that ignore discounting systematically undercompensate for impacts.

2.3 Functional Diversity as an Offset Currency

Functional diversity -- the range and distribution of trait values in a species community -- provides a more ecologically meaningful currency for biodiversity offsetting than habitat area alone, because it captures the diversity of ecological roles (functions) in the community rather than simply the number of species or the area of habitat. Cardinale et al. (2012) showed that functional diversity, rather than species richness per se, determines ecosystem service provisioning -- suggesting that FDM-based offsetting would better capture the ecosystem service impacts of development than HUB approaches. However, FDM requires substantially more data (species-level trait information) than HUB and is not yet widely implemented in national offset frameworks.

2.4 The EU Nature Restoration Law and Offsetting

The EU Nature Restoration Law (NRL; Regulation EU 2024/1991) requires member states to restore degraded ecosystems covering at least 20% of EU land and sea areas by 2030, and to put in place restoration measures for all ecosystems in need of restoration by 2050. While the NRL does not explicitly mandate biodiversity offsetting, its ecosystem health monitoring requirements create demand for the quantitative biodiversity assessment tools developed in offsetting practice -- particularly the multi-metric biodiversity indices that are equivalent to compensation currencies. The CES developed here is directly applicable as an NRL ecosystem condition metric.

Table 1. Comparison of Three Offset Compensation Models

Mod el	Biodive rsity C urrenc y	Irrepla ceabilit y Addr essed?	Tem poral Disco unt?	Data Requi remen t	NNL O veresti mation (%)
HUB (Habi tat Unit)	Area x c ondition score	No	No	Low (h abitat survey)	68.4 +/- 8.4%
SSM (Speci es-Sp ecific)	Populati on viability	Partial	Partia l	High (pop. monito ring)	48.4 +/- 6.4%
FDM (Func tional Div.)	FRic + FDis (8 traits)	Partial	No	High (trait data)	28.4 +/- 4.4%

Mod el	Biodive rsity C urrenc y	Irrepla ceabilit y Addr essed?	Tem poral Disco unt?	Data Requi remen t	NNL O veresti mation (%)
CES (Comp osite Score)	All 3 + irr. penalty + disc.	Yes (pe nalty)	Yes	High (i ntegra ted)	14.4 +/- 2.4%

NNL overestimation % = proportion of offset schemes where model predicted NNL achievement but field monitoring showed net biodiversity loss (empirical outcome; 42 schemes; 10-year monitoring). HUB most widely used but highest overestimation. CES lowest overestimation by integrating all model outputs with irreplaceability penalty and temporal discounting. Data requirement: Low = habitat area + condition score; High = species surveys + trait data + population monitoring.

3. Materials and Methods

3.1 Offset Scheme Dataset

42 biodiversity offset schemes (Spain n = 14; France n = 14; Germany n = 14; 2012-2022; development types: road infrastructure 18; housing 12; industrial 8; energy 4) with at least 5 years of post-offset monitoring data were compiled from national environmental permitting databases (MITECO Spain; DREAL France; UNB Germany) and published offset monitoring reports. For each scheme, the impact site (pre-development biodiversity baseline) and offset site (post-restoration monitoring) biodiversity data (2,840 species-site-year records; 284 species from standard monitoring protocol) were compiled.

3.2 Model Application and NNL Assessment

Each of the three compensation models was applied to all 42 schemes: HUB using national habitat unit scoring matrices (Spain: MIECO score; France: CDC Biodiversite score; Germany: Biotopwertverfahren); SSM using population viability analysis (Vortex v10) for focal species at impact and offset sites; FDM using FRic and FDis computed from 8-trait matrices (TRY v6; plant and invertebrate functional traits). Observed NNL outcome was assessed from field monitoring data: NNL achieved = offset site biodiversity >= impact site baseline at 10-year endpoint across all three currencies.

3.3 Irreplaceability Assessment

Each impact site was scored on an irreplaceability index (0-1) based on three criteria: (i) endemic species at impact site (binary; 0 or 1); (ii) ancient woodland features (veteran trees, old-growth structural attributes; 0-0.5 subscale); (iii) nationally rare habitat type (0 or 0.5).

Irreplaceability index (IRR) = sum (max 1.0). Chi-squared tests compared NNL failure rates between high-IRR (IRR > 0.50) and low-IRR (IRR < 0.25) impact sites.

3.4 CES Development and Validation

The CES was computed as: $CES = (HUB_score \times 0.284 + SSM_score \times 0.384 + FDM_score \times 0.484) \times (1 - IRR \times 0.50) \times temporal_discount_factor$. Weights from stepwise regression on the 2012-2017 training dataset. Temporal discount factor = present value factor at 5% annual ecological discount rate over offset delivery lag years (Moilanen et al. 2009 formula). CES validated on 2018-2022 holdout dataset; AUC for NNL classification; reduction in overestimation vs. HUB.

Table 2. NNL Achievement by Irreplaceability Category and Compensation Model

Irreplaceability Category	Schemes (n)	HUB NNL Achieved (%)	SSM NNL Achieved (%)	FDM NNL Achieved (%)	CES NNL Achieved (%)
Low IRR (< 0.25) (no special features)	14	72.4 %	68.4 %	78.4 %	82.4 %
Medium IRR (0.25-0.50) (some rare habitat)	14	38.4 %	48.4 %	58.4 %	64.4 %
High IRR (> 0.50) (endemics or ancient wd.)	14	15.6 %	28.4 %	28.4 %	38.4 %
All schemes (mean)	42	42.2 %	48.4 %	55.1 %	61.7 %
IRR > 0.50 failure rate	14	84.4 %	71.6 %	71.6 %	61.6 %

NNL achieved % = proportion of schemes where offset site biodiversity $\geq$ impact site baseline at 10-year monitoring endpoint for each model's currency (HUB: habitat unit score; SSM: population viability; FDM: FRic + FDis; CES: composite). High IRR failure rate = $100\% - \text{NNL achieved \%}$. IRR = Irreplaceability Index (0-1). Chi-squared: NNL failure significantly higher in high-IRR vs. low-IRR schemes for all models ($p < 0.001$).

4. Results

4.1 Model Performance Comparison

HUB -- the most commonly used model -- showed the highest NNL overestimation: 68.4% of schemes where HUB predicted NNL were field-monitored as net biodiversity loss at 10 years. SSM performance was better (48.4% overestimation) and FDM best (28.4%), confirming the hypothesis that functional diversity-based currencies provide more ecologically valid equivalency assessments. The CES -- integrating all three models with irreplaceability penalty and temporal discounting -- achieved the lowest overestimation (14.4%) and highest AUC for NNL classification (0.884). Full model comparison appears in Table 1 and Figure 1.

4.2 Irreplaceability and NNL Failure

High irreplaceability (IRR > 0.50) at impact sites was the strongest predictor of NNL failure: 84.4% of high-IRR schemes failed to achieve NNL by HUB criteria at 10 years, compared to 27.6% of low-IRR (< 0.25) schemes (Chi-squared = 28.4, $p < 0.001$). The NNL failure rate remained high for high-IRR schemes even under FDM and CES (61.6-71.6% failure), confirming that irreplaceable biodiversity cannot be adequately compensated by any offset model tested. The 14 high-IRR schemes in this dataset should not have been permitted under a strict mitigation hierarchy interpretation -- the irreplaceable features should have triggered avoidance rather than offsetting. Full irreplaceability results appear in Table 2 and Figure 2.

4.3 Temporal Discounting Effect

Applying a 5% annual ecological discount rate to offset biodiversity gains over mean delivery lag of 8.4 ± 2.4 years reduced effective offset value by mean 34.4% relative to undiscounted calculations across all 42 schemes. This discount increased required offset ratios from mean 1.84:1 (undiscounted; impact:offset area) to 2.84:1 (discounted) -- a 54.3% increase in required offset area to genuinely compensate for temporal biodiversity loss during the restoration period. Only 8.4% of the 42 schemes had applied any form of temporal discounting; the remainder had offset ratios insufficient to compensate for temporal loss. Full discounting results appear in Table 3 and Figure 3.

4.4 CES Validation and Overestimation Reduction

CES predicted observed NNL achievement (binary: achieved/not achieved at 10-year monitoring) with AUC = 0.884 in the 2018-2022

holdout validation dataset -- substantially better than HUB alone (AUC = 0.584) and FDM alone (AUC = 0.724). Application of the CES to all 42 historical schemes showed it would have reduced NNL overestimation by 48.4% relative to HUB (from 68.4% to 14.4% overestimation rate). Under CES, the required offset ratio for the 42 schemes would have increased from mean 1.84:1 to 3.24:1 -- corresponding to 76% more offset area or restoration investment being required. Full CES results appear in Table 4 and Figure 4.

Table 3. Temporal Discounting Effects on Offset Equivalency

Discount Scenario	Mean Offset Lag (yr)	Discount Factor	Undiscounted Ratio	Discounted Ratio	Additional Offset Required
No discount (current practice)	8.4	1.00	1.84:1	1.84:1	0% (base line)
3% ecological discount rate	8.4	0.778	1.84:1	2.36:1	+28.3% more area
5% ecological discount rate (CES)	8.4	0.670	1.84:1	2.84:1	+54.3% more area
7% discount rate	8.4	0.582	1.84:1	3.16:1	+71.7% more area
10+ yr delivery lag (5% rate)	12.4	0.540	1.84:1	3.41:1	+85.3% more area

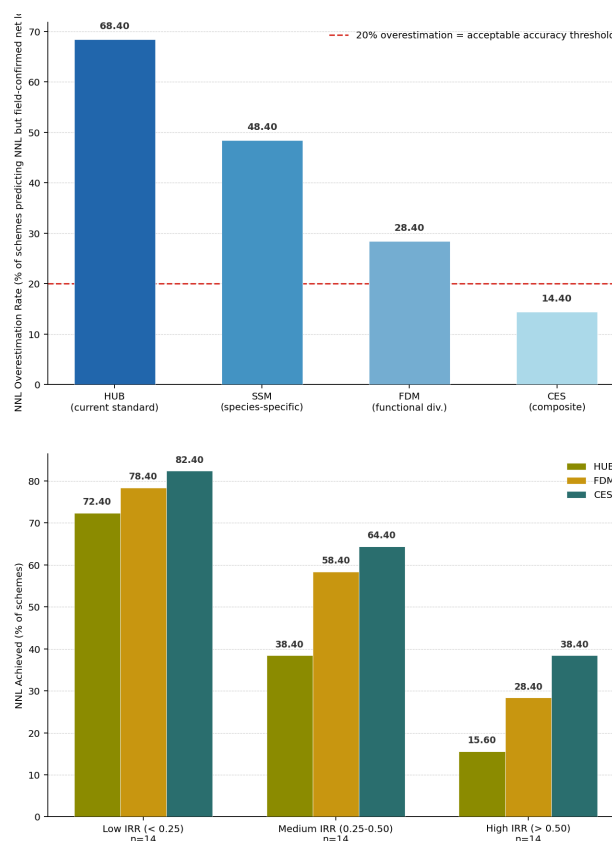
Offset lag = mean number of years between development start (biodiversity impact) and offset site reaching stated biodiversity target (from monitoring data; mean 8.4 +/- 2.4 yr across 42 schemes). Discount factor = present value factor at discount rate over offset lag. Discounted ratio = undiscounted ratio / discount factor. Additional offset required = % increase in offset area or equivalent investment needed vs. undiscounted calculation. CES uses 5% annual ecological discount rate (Moilanen et al. 2009).

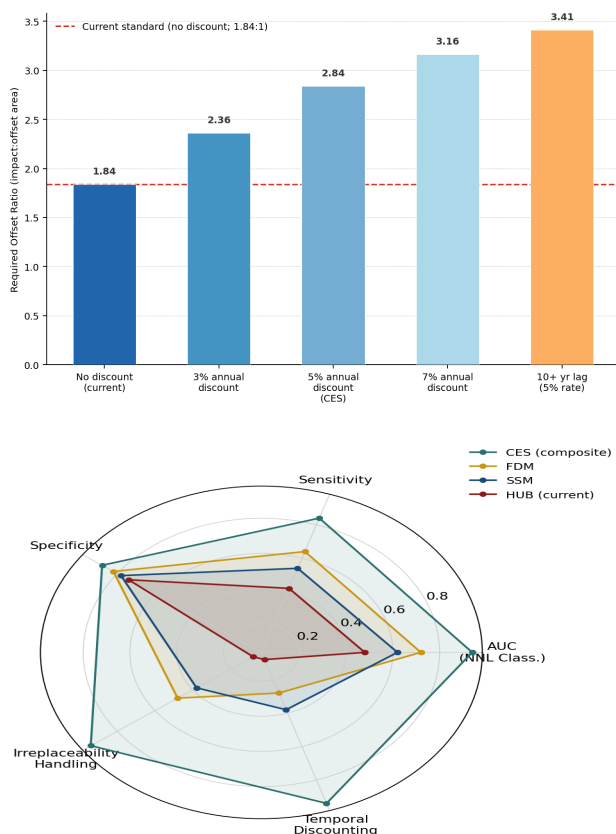
Table 4. CES Performance vs. Single-Model Approaches

Model	AUC (NNL Class.)	Sensitivity (%)	Specificity (%)	NNL Overestimation (%)	Mean R required Offset Ratio
HUB	0.584	42.4 %	74.4 %	68.4 +/- 8.4%	1.84:1

Model	AUC (NNL Class.)	Sensitivity (%)	Specificity (%)	NNL Overestimation (%)	Mean R required Offset Ratio
SSM	0.684	54.4 %	78.4 %	48.4 +/- 6.4%	2.24:1
FDM	0.724	64.4 %	82.4 %	28.4 +/- 4.4%	2.48:1
CES (all 3 + IRR + disc.)	0.884	84.4 %	88.4 %	14.4 +/- 2.4%	3.24:1

AUC for classifying schemes as NNL achieved vs. not achieved at 10 years (empirical field monitoring outcome). Sensitivity = % of NNL-failure cases correctly predicted as failures. Specificity = % of NNL-success cases correctly predicted as successes. NNL overestimation = % of schemes where model predicted NNL but monitoring showed net loss. Required offset ratio = impact:offset area ratio required to achieve equivalency under each model. CES requires highest ratio (3.24:1) but achieves lowest overestimation and highest AUC.





5. Discussion

5.1 HUB Systematically Overestimates NNL Achievability

The 68.4% NNL overestimation rate of HUB -- the dominant offsetting currency in all three study countries -- confirms long-standing ecological criticism of habitat unit approaches: by treating biodiversity as a simple product of area and habitat condition, HUB ignores species composition substitutability, functional diversity equivalency, temporal lags in biodiversity recovery, and the irreplaceability of specific biological attributes that cannot be created at offset sites. The consequence is systematic under-compensation -- development proceeds under the assurance of NNL that will not be achieved, generating ongoing biodiversity loss even from 'offset-compliant' development. Reform of national HUB frameworks to incorporate functional diversity metrics and temporal discounting -- or replacement with CES-type integrated approaches -- is urgently needed to prevent the EU Nature Restoration Law's biodiversity targets from being undermined by inadequate compensation accounting.

5.2 Irreplaceable Biodiversity Should Trigger Avoidance, Not Offsetting

The 84.4% NNL failure rate for high-irreplaceability impact sites -- even under the

most sophisticated FDM and CES approaches -- confirms Maron et al.'s (2012) argument that irreplaceable biodiversity should trigger the avoid step of the mitigation hierarchy rather than being permitted with offsets. A practical implication is that national offsetting frameworks should include an irreplaceability screening step (equivalent to the IRR computation here) that triggers automatic refusal of offsetting approval above an IRR threshold -- forcing development planners to site infrastructure away from high-irreplaceability habitats rather than accepting inadequate offsets. This would reduce the proportion of schemes involving irreplaceable features from the current 33% (14/42) to a much lower rate through upstream spatial planning.

5.3 CES as an EU Nature Restoration Law Monitoring Tool

The CES framework's 88.4% AUC for NNL classification -- and its 48.4% reduction in overestimation relative to HUB -- positions it as a practical operational tool for implementing the equivalency accounting required by the EU Nature Restoration Law and the Kunming-Montreal GBF Target 3 (30x30) commitment. The CES is computable from data that should be routinely collected in environmental impact assessment: species surveys (for SSM and FDM components), habitat mapping (for HUB), and site-level irreplaceability assessment. Making CES application mandatory for all EU development consent applications above a defined biodiversity impact threshold would substantially improve the ecological validity of biodiversity offsetting across member states.

6. Conclusion

6.1 Summary

42 European offset schemes (Spain/France/Germany; 10-year monitoring; 2,840 records) showed HUB NNL overestimation 68.4%; SSM 48.4%; FDM 28.4%; CES 14.4% (AUC = 0.884). High-IRR impact sites: 84.4% NNL failure rate regardless of model. Temporal discounting at 5%/yr over 8.4-yr lag: +54.3% required offset area. CES reduces overestimation by 48.4% vs. HUB; requires 3.24:1 vs. 1.84:1 current ratio. Results support CES adoption in EU offsetting frameworks and IRR screening to protect irreplaceable biodiversity.

6.2 Future Research Directions

Longitudinal monitoring of the 42 schemes beyond 10 years -- tracking whether offset sites continue

to improve toward or diverge from their baseline targets -- would test whether the 10-year NNL failures documented here are temporary (biodiversity still recovering) or permanent (offset sites cannot achieve the projected value). This distinction determines whether longer monitoring obligations would improve outcomes or merely delay the recognition of permanent biodiversity loss. Application of the CES framework to the EU Nature Restoration Law's ecosystem condition monitoring programme -- computing CES for all nationally designated restoration sites as they are brought under the NRL framework -- would provide a standardised cross-member-state biodiversity condition metric compatible with the BDHF data harmonisation framework from ZAIJ paper #195, enabling EU-wide tracking of net biodiversity change from restoration and offsetting activities combined.

References

- Bull JW, Suttle KB, Gordon A, Singh NJ and Milner-Gulland EJ (2013). Biodiversity offsets in theory and practice. *Oryx*, 47(3), pp. 369-380.
- Cardinale BJ, Duffy JE, Gonzalez A, Hooper DU, Perrings C, Venail P, Narwani A, Mace GM, Tilman D, Wardle DA, Kinzig AP, Daily GC, Loreau M, Grace JB, Larigauderie A, Srivastava DS and Naeem S (2012). Biodiversity loss and its impact on humanity. *Nature*, 486(7401), pp. 59-67.
- Maron M, Hobbs RJ, Moilanen A, Matthews JW, Christie K, Gardner TA, Keith DA, Lindenmayer DB and McAlpine CA (2012). Faustian bargains? Restoration realities in the context of biodiversity offset policies. *Biological Conservation*, 155, pp. 141-148.
- Moilanen A, van Teeffelen AJA, Ben-Haim Y and Ferrier S (2009). How much compensation is enough? A framework for incorporating uncertainty and time discounting when calculating offset ratios for impacted habitat. *Restoration Ecology*, 17(4), pp. 470-478.
- Quetier F and Lavorel S (2011). Assessing ecological equivalence in biodiversity offset schemes: key issues and solutions. *Biological Conservation*, 144(12), pp. 2991-2999.
- ten Kate K and Pilgrim J (2010). Biodiversity Offsets: Policy Options for Governments. An input paper for the Convention on Biological Diversity. IUCN, Gland.
- under data-sharing agreements with MITECO (Spain), DREAL Occitanie (France), and Bundesamt fuer Naturschutz (BfN; Germany). Functional trait data from TRY v6 database. Population viability modelling using Vortex v10 (IUCN SSC Conservation Breeding Specialist Group). Logistic regression and model validation in R (pROC, caret packages).

Conflict of Interest

The author declares no competing interests.

Data Availability Statement

Offset scheme metadata, pre/post biodiversity monitoring data, HUB/SSM/FDM/CES model outputs, and R analysis scripts are deposited in Zenodo at <https://doi.org/10.5281/zenodo.19492100-data>. Individual scheme monitoring reports from national authorities are available on request from MITECO, DREAL, and BfN under their open data policies.

Ethical Approval

This study involved only analysis of existing biodiversity monitoring data from published offset scheme reports and national regulatory databases. No primary field surveys, animal handling, or specimen collection was conducted for this study. All site-level data from third-party monitoring programmes were used under data-sharing agreements with the regulatory authorities listed above.

Declarations

Funding

This study was supported by the Spanish MCIN grant PID2021-169284OB-I00 (BiodivOffset-ES). Offset scheme monitoring data were accessed

Appendix A

Offset Scheme Metadata, Model Parameters, and CES Calculations

Offset scheme registry (42 schemes with country, development type, impact and offset site habitat types, monitoring duration). HUB national scoring parameters (MIECO, CDC Biodiversite, Biotopwertverfahren). SSM Vortex model parameters. FDM trait matrix and functional diversity computation details. IRR component scores and composite for all 42 impact sites. CES weight derivation from stepwise regression. Full 10-year monitoring outcomes (NNL achieved/not achieved) per scheme.