UNIVERSITY OF HELSINKI

https://helda.helsinki.fi

Integrating species metrics into biodiversity offsetting calculations to improve long-term persistence

Marshall, Erica; Visintin, Casey; Valavi, Roozbeh; Wilkinson, David P.; Southwell, Darren ...

2022-04
Wiley Blackwell
http://hdl.handle.net/10138/357616
Marshall, E, Visintin, C, Valavi, R, Wilkinson, D P, Southwell, D, Wintle, B A & Kujala, H 2022, 'Integrating species metrics into biodiversity offsetting calculations to improve long-term persistence', Journal of Applied Ecology, vol. 59, no. 4, pp. 1060-1071. https://doi.org/10.1111/1365-2664.14117

Downloaded from Helda, University of Helsinki institutional repository. https://helda.helsinki.fi
This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.
Please cite the original version.

Marshall Erica (Orcid ID: 0000-0002-7297-5777)
Visintin Casey (Orcid ID: 0000-0003-2245-8998)
Valavi Roozbeh (Orcid ID: 0000-0003-2495-5277)
KUJALA HEINI (Orcid ID: 0000-0001-9772-3202)

Journal of Applied Ecology

Check for updates

Handling Editor: Don Driscoll

Integrating species metrics into biodiversity offsetting calculations to improve long-term persistence

Erica Marshall[1,2,3]*, Casey Visintin[1], Roozbeh Valavi[1], David P. Wilkinson[1], Darren Southwell[1,2], Brendan Wintle[1,2], Heini Kujala[1,2,4]

[1]University of Melbourne, School of Biosciences, VIC 3010, Australia

[2]National Environmental Science Program, Threatened Species Recover Hub

[3]University of Melbourne, School of Ecosystem and Forest Sciences, VIC 3010, Australia

[4]Finnish Museum of Natural History, University of Helsinki, FI-00140 Helsinki, Finland

*Corresponding author

Email: erica.marshall@unimelb.edu.au (EM)

Abstract

1. Several methods of measuring biodiversity in development-offset trades exist. However, there is little consensus on which biodiversity metrics should be used for quantifying development impacts and assigning offsets.
2. We simulated development impacts in a virtual landscape and offset these impacts using six biodiversity metrics: vegetation area, vegetation condition, habitat suitability, species abundance, metapopulation connectivity and rarity weighted richness. We tested long-term impacts of metric choice during offsetting by combining simulated landscapes with population viability analyses.
3. No net loss or net gains in habitat were achieved using all metrics except vegetation area and condition. Limited habitat and like-for-like requirements resulted in offsets exhausting available habitat in each vegetation class before offset requirements were met when using vegetation-based metrics. We also found that impact avoidance was an important driver in how much compensation offsets could deliver.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/1365-2664.14117

This article is protected by copyright. All rights reserved.

Accepted Article

When impacts avoided high-suitability habitats, all six metrics achieved no net loss or net gains for most species. However, when core habitats were developed, none of the metrics were able to consistently prevent population declines.

4. *Synthesis and application:* When impacts on high quality habitat were avoided, and assuming the protection and restoration benefits can occur in practice, vegetation-based metrics may produce offsets which deliver gains in species abundance equivalent to species-specific metrics. However, species-specific metrics outperformed vegetation-based metrics when core habitats were lost. Applying avoidance measures as a first step to minimise biodiversity impacts during development will significantly improve offset outcomes for species and result in greater long-term population benefits delivered through offsetting.

Key Words: biodiversity offsets, biodiversity metrics, population viability analysis, spatial simulations, species persistence, no net loss

Accepted Article

Accepted Article

Introduction

Biodiversity offsets aim to compensate the loss of biodiversity caused by a development impact through commensurate conservation gains (Maron et al., 2016). Accurate metrics for measuring biodiversity at development and offset sites are required to deliver the requisite conservation benefits (Maron et al., 2012; Marshall, Wintle, Southwell, & Kujala, 2020). Directly measuring and quantifying species outcomes is challenging, so habitat-based surrogates are commonly used in lieu of measuring outcomes for all species (Martin, Evans, Rice, Lodhia, & Gibbons, 2016). These surrogates generally consist of simple, easily assessed habitat features which contribute towards an overall site condition score (Harwood et al., 2016). This approach may simplify the task of assessing biodiversity outcomes, but such metrics often fail to accurately account for all levels of biodiversity in offset calculations and could lead to inadequate trades (Cristescu, Rhodes, Frére, & Banks, 2013; Hanford, Crowther, & Hochuli, 2016).

Failure to accurately account for biodiversity when assessing development impacts and designing offsets means that immediate losses caused by development may go unaccounted for and the potential long-term gains from offsetting are unclear (Maron et al., 2012, 2016; Weissgerber, Roturier, Julliard, & Guillet, 2019). Over time this could exacerbate already high biodiversity declines and perpetuate ineffective offset practices (McKenney & Kiesecker, 2010). Yet offset policies still rely heavily on habitat and area-based surrogates for biodiversity and have largely failed to incorporate even simple alternative biodiversity metrics such as habitat suitability or abundance, which are common in other fields of environmental research (Marshall, Wintle, et al., 2020). This is due to the ease with which habitat or area-based offset metrics can be measured (Gibbons et al., 2016) and the difficulties associated with data collection for alternative species driven approaches (Boykin et al., 2012).

Recent research has demonstrated that habitat and area-based offset metrics do little to mitigate development impacts on populations or species and do not ensure population persistence of target species

Accepted Article

(Marshall, Valavi, et al., 2020). Ideally we should also utilise metrics that account for the landscape and population dynamic processes driving species occurrence and abundance (Moilanen & Kotiaho, 2018), including connectivity and measures of habitat suitability to improve the quality of offset actions delivered (Buschke & Sinclair, 2019; Marshall, Valavi, et al., 2020). Metrics directly affected by ecological processes, such as abundance or metapopulation connectivity, could improve our understanding of how offsets function long-term.

Here, we test six metrics to understand how their application withing offsetting influences biodiversity persistence and to determine if offset outcomes could be improved by incorporating species-specific information.

Methodological framework

We use a methodological framework which combines a simulation tool we developed in R v3.6 (R Development Core Team, 2017) and population viability analyses. The framework involves four steps: 1) simulating habitat clearing in a virtual landscape; 2) calculating offset requirements and identifying offset locations; 3) simulating restoration to match offset requirements and deliver benefits, and; 4) predicting population trajectories of species under each scenario (Fig.1).

Data Requirements

Our tool simulates habitat clearance in a virtual gridded landscape. Simulations require user-provided raster layers representing development scenarios, species habitats and vegetation condition. Species' habitat was weighted using vegetation condition to account for past degradation or land clearing. This gives the simulation tool a current habitat condition value for each pixel and a target habitat value for restoration.

Step one: simulating development impacts

Development sites can be selected randomly or weighted towards areas of interest such as high development likelihood regions or high-quality sites. The number and size of development impacts can be specified, and in all development pixels condition is reduced to zero. Developments are created iteratively and at each

step a raster map is produced (Fig.S.10). Maps are incorporated sequentially into population models so trajectories can be tracked across development and offset events through time.

Step two: calculate offset requirements

How offset requirements are calculated during simulations is specified by the user. This can be either number of pixels (area), or summed pixel values lost at each development site. Offsets which match area will restore habitat until an equivalent area is achieved. Offsets restoring pixel values will match the summed value regardless of area. Offset metrics can be re-defined by replacing input data. For example, instead of habitat maps, abundance data can be used to calculate development impacts and simulate increased abundance.

Step three: restore habitat

After offset requirements are calculated, restoration is simulated (Fig.1) and is assumed to be immediately effective. This is a simplification of restoration processes in which time-lags in attainment of benefit and uncertainty about attainment would usually exist. However, as our aim is to compare relative performance of offset metrics, we purposefully assume all restoration efforts were successful so as to give each metric an equal opportunity to produce benefits. This ensures the impacts on species persistence are due to the metric itself rather the capacity of restoration to produce benefits within the simulation period. Pixels for restoration are selected from a buffer zone around development impacts and added incrementally until offset requirements are met or there is no more available habitat (Table S.1).

Step four: population modelling (*steps*)

We use the R package *steps* v1.0 (Visintin et al. 2020) to run temporally- and spatially-explicit population viability models (Fig.1). Population viability analyses use habitat requirements, dispersal capability and demographic variables for species to predict persistence probabilities (Visintin et al., 2020). Our approach in *steps* uses raster data on habitat suitability to develop a grid-based patch structure and estimates trajectories by simulating population dynamics across the landscape.

Accepted Article

Accepted Article

Case study: Hunter Region, New South Wales, Australia

We applied our simulation framework to predict impacts of metric choice on four species in the Hunter Region, New South Wales (NSW), Australia (see Fig.1 in Marshall, Valavi, et al. [2020]). Our focal species were: squirrel gliders (*Petaurus norfolcensis*), northern brown bandicoots (*Isoodon macrourus*), tiger quolls (*Dasyurus maculatus*) and yellow-bellied gliders (*Petaurus australis*). These four species were selected because they are sufficiently well-sampled to build spatially-explicit population models and they would be accounted for in offsets required by state legislation (Office of Environment and Heritage, 2017).

Data Requirements

We specified 100m resolution species distribution models (SDM), using *MaxEnt* (Phillips, Dudik, & Dudík, 2008), to include environmental variables, alongside a vegetation condition layer detailing current vegetation extent and land clearing (Fig.S.1, S.3). We multiplied SDM predictions and vegetation condition (ranging between 1=intact and 0=cleared) to derive current habitat suitability (Fig.S.4). Currently cleared and degraded areas that were predicted as environmentally suitable by the SDM acted as targets for restoration efforts.

Step one: simulating development impacts

We simulated two development scenarios in each species' current habitat suitability maps. Scenario one focused development impacts within the least suitable habitat (strict avoidance), which is pursuant to offsetting best-practices (Phalan et al., 2017). We used a species-specific Maximum Training Sensitivity plus Specificity (MTSS) threshold from our *MaxEnt* outputs (Liu, White, & Newell, 2013), to ensure developments only occurred in regions deemed unsuitable for species (Fig.S.18). These thresholds were also used to direct development impacts towards areas of high suitability for species in our worst-case scenario (S2; targeted development). Development impacts in both scenarios were 5000ha and each scenario included 20 developments, therefore the overall development footprint was 100,000ha (Fig.S.10). We repeated each development scenario 50 times to account for spatial stochasticity.

Accepted Article

Step two: calculate offset requirements

We calculated offset requirements using six offset metrics; 1) *vegetation area*; 2) *vegetation condition*; 3) *habitat suitability*; 4) *species abundance*; 5) *metapopulation connectivity* (Hanski, 1994); and 6) *rarity-weighted species richness* (richness; Table S.1; Albuquerque & Beier 2015). These metrics represent simplified versions of those commonly used in policies but do not necessarily encapsulate the breadth of data generally required for their calculation. For example, most offset policies using vegetation area metrics would require the use of multiple types of habitat data to calculate an overall score. We knowingly excluded such composite metrics from our analysis for two reasons: first, we did not have access to the spatial data commonly used to build composite metrics for offset calculations in NSW (Office of Environment and Heritage, 2017). Second, building our own composite metric would have required subjective decisions on what components to include and how to weight these against each other, making our results susceptible to those decisions. Our vegetation condition and habitat suitability metrics, although single metrics, are holistic measures of vegetation and habitat quality, and themselves based on multiple inputs (Figure S.2 to S.4). Therefore, while they do not include all the data potentially used in a composite metric, they mimic the broad categories of measures often included in composite calculations which range from entirely habitat focused to more species-specific approaches (Appendix S7). By testing this subset of potential measures included in composite calculations we aim to determine how their application within a broader offsetting framework may impact species persistence.

Vegetation area, vegetation condition and habitat suitability

Vegetation area and condition were calculated using maps of current habitat suitability and vegetation type (Fig.S.2, S.4) and offsets matched impacts on specific vegetation classes. Vegetation area restored species' habitat suitability until an equivalent area of each vegetation type lost to development was recovered. Vegetation condition offsets matched development impacts on the summed habitat suitability values lost in the species' current habitat suitability map for each vegetation class. Habitat suitability matched losses in

species' current habitat suitability values through equivalent gains elsewhere, restoring pixels to their predicted SDM values based on environmental variables (Fig.S.3; Table S.1).

Species Abundance

Species abundance maps (Fig.S.5, S.6) were generated using our habitat suitability maps and species-specific carrying capacity values (Thuiller et al., 2014) to get estimated abundance per pixel (Eq 1):

$$N_{ij} = H_{ij} * k_j \quad \text{(Eq 1)}$$

Where the estimated abundance *N* of species *j* for each pixel *i* is equal to species' habitat suitability in that pixel (H_{ij}) multiplied by the species maximum carrying capacity (k_j) to get the estimated number of individuals per cell (per ha). We assumed that local carrying capacity is linearly correlated with predicted *MaxEnt* values, giving the fraction of maximum carrying capacity attainable for the species in each pixel (Merow, Smith, & Silander, 2013).

Metapopulation Connectivity

Metapopulation connectivity maps for each species were produced using a measure derived from metapopulation models (Eq 2; Fig S.7, S.8; Hanski 1994). The connectivity of patch *i* for species *j* is calculated as $I_{ij} = \sum_k exp(-\propto_j d_{ik})A_k$, where d_{ik} is the distance between patch *i* and patch *k*, and A_k is the size of the source population in patch *k*. Parameter $\propto_j$ gives the dispersal capacity of species *j*. We utilised two-dimensional kernel smoothing to produce a map of connectivity values for each species, where the width of the smoothing kernel is defined by species' dispersal capability ($\propto_j$). The value for species *j* in grid cell *i* is calculated as:

$$H'_{ij} = \sum_x \sum_y exp\left(-\propto_j d(x-u, y-r)\right) H_{ij} \quad \text{(Eq 2)}$$

where H_{ij} is the original habitat suitability value of species *j* in grid cell *i*. We note that the summation is across all grid cells, including the focal cell. Values *u* and *r* denote the location of the focal cell *i* and $d(x-u, y-r)$ is the distance between cell *i* and all grid cells in the landscape, including itself. The final

Accepted Article

value H'_{ij}therefore reflects not only on the habitat suitability of pixels but also the connectivity between them (Schooley and Branch, 2011).

Rarity weighted species richness

A single rarity-weighted species richness map (hereafter rarity-weighted richness; Fig.S.9; Albuquerque & Beier 2015) was calculated across the four species habitat suitability maps as follows:

$$R_i = \sum_j \left(\frac{H_{ij}}{\sum_i H_{ij}} \right) \quad \text{(Eq 3)}$$

Where $\frac{H_{ij}}{\sum_i H_{ij}}$gives for each species *j* the fraction of the total habitat suitability (sum of cells) found in pixel *i*. These were then summed to get a rarity-weighted richness for each pixel $i(R_i)$. In the resulting map (Fig.S.9) pixel values represent the sum rarity scores across species (Albuquerque & Beier, 2015; Williams et al., 1996).

Step three: restore habitat

For each metric, we produced two raster layers in R using the *raster* package (v3.0-12): one representing current landscapes with varying vegetation condition and land-use, and another assuming no land-clearing and fully intact vegetation (Appendix S.1). Offset requirements were calculated based on current values but restoration efforts returned values to those in the predicted layers (e.g. predicted SDM values). If a cell had a current condition of zero but a high restoration potential, habitat could be restored (if it was not a developed site) representing a simplified approach to habitat creation. We used an offset multiplier of one across all metrics to calculate offset requirements.

Step four: population modelling (*steps*)

We tracked changes in species abundance through a 100-year period in *steps* (Appendix S.3). For each species we built a baseline model as our counterfactual where landscape structure remained static (species' current habitat suitability). We subsequently built models with habitat suitability maps altered by simulated developments and offsets. All models used species' baseline current habitat suitability maps for the first ten

Accepted Article

Accepted Article

years. The first development impact was introduced in year 11 and the corresponding offset in year 12. Subsequent development impacts were introduced every two years, followed by their corresponding offset gains. This continued iteratively until all 20 developments and offsets had been introduced at year 50 (Fig.S.11). Population modelling continued for another 50 years, completing at the 100-year time period. Population simulations for each development, offset and baseline scenarios were replicated 500 times. Demographic and dispersal parameters for each species were collected from the literature and parameters were tested through sensitivity analyses (Appendix S.4; Table S.4).

Scenario analysis and evaluation

We evaluated three measures to determine if no net loss (NNL) or net gains (NG) were achieved through offsetting; 1) percentage change in values measured by the metric; 2) percentage change in total habitat suitability from the species' baseline suitability maps; and 3) percentage change in final population size compared to the species' baseline population model. We defined NNL as percentage change not statistically different from zero (two-sided t-test). Statistically significant changes greater or less than zero were defined as NG and net loss (NL), respectively. Final population size was taken as the average abundance in the final time-step across all replicates. We also calculated the average expected minimum abundance (EMA) and final population size for three time-periods; 15, 30 and 50 years after the last offset (i.e. time-steps 65, 80 and 100; Appendix S.5) to assess potential time-lags.

Results

Percentage change in metric and habitat suitability

Most of the metrics we tested achieved a small, but statistically significant NG in the metric values (upward pointing triangles, Fig. 2; Table S.4), particularly when developments avoided core habitats (S1).When impacts were targeted (S2), offsets for the yellow-bellied glider failed to achieve NNL in metric values when using habitat suitability, abundance and metapopulation connectivity to quantify offset trades. Developments removed large amounts of high-quality habitat for this species, making it difficult to match

losses in these metrics in the remaining low-quality habitat. Simulations therefore restored primarily low-quality habitat (Fig. S.19, S.20), resulting in statistically significant declines of the metric itself and declines in total habitat suitability (average of 6.6% across all metrics; Fig.2).

Development impacts caused on average a 1.9% and 11.8% decline in habitat suitability across species when impacts were avoided (S1) or targeted (S2), respectively. In scenario one, identifying offsets using habitat suitability, abundance, metapopulation connectivity and richness resulted in NNL or small NG in habitat suitability for all species except yellow-bellied gliders, which had a small but statistically significant NL under metapopulation connectivity (downward pointing triangle, Fig 2; Table S.4). In scenario two, habitat suitability, abundance and richness still delivered NG across most species, whereas metapopulation connectivity resulted in NL for all species despite often achieving statistically significant NG in metric values (Fig.2; Table S.4). These differences between metric and habitat suitability outcomes are likely explained by the fact that restoring population connectivity emphasises locations between patches, which by nature are of lower habitat suitability (Hodgson et al. 2009; Fig.S.19, S.20).

Vegetation area and condition failed to match development impacts or meet offset requirements of the metric through gains in habitat suitability for all species. The vegetation area metric resulted in an average NL of 1.7% in scenario one and 11.4% in scenario two. Vegetation condition resulted in a NL of 1.4% and 10.9% on average in S1 and S2 respectively. These losses were predominantly caused by strict like-for-like requirements for specific vegetation classes and exhaustion of all the habitat available for restoration, resulting in net declines.

Accepted Article

Accepted Article

Accepted Article

Percentage change in population

S1: Strict avoidance

When impacts on high-suitability areas were avoided, all six metrics were able to achieve statistically significant NNL or NG in final population size for all species except tiger quolls (Fig.3; Table S.4). For squirrel gliders and yellow-bellied gliders, using metrics that resulted in habitat suitability losses (e.g., metapopulation connectivity; Fig.2) still resulted in NNL or NG in population size. Small population gains for these species were also observed in development only models. These small gains are probably due to strategic placement of development impacts within low suitability habitat, which may have resulted in removal of sink patches (Devictor et al., 2010). The largest NG in final population size were observed when using rarity weighted richness, followed by habitat suitability and metapopulation connectivity, which all produced NG in scenario one for every species except the tiger quoll. Despite observed declines in habitat suitability (Fig.2), vegetation area and vegetation condition generally produced offsets which achieved NNL or NG in final population size (Fig.3).

S2: Targeted developments

Scenario two developments resulted in significant declines in final population size for all four species (Fig.3, S.21). Benefits delivered by offset metrics varied between species, but none of the metrics we tested consistently achieved positive outcomes when impacts were high (Fig.3). For all species except northern brown bandicoots no metric achieved a population recovery within five percent of baseline (Fig.3). The metric which delivered the greatest population benefits varied across species, and on average the largest gains in abundance were produced by species-specific metrics. For example, offsets identified using rarity weighted richness provided the largest gains for northern brown bandicoots, but tiger quolls benefited most from metapopulation connectivity, squirrel gliders from abundance and yellow-bellied gliders from habitat suitability.

Accepted Article

Time since offsetting

For all species, except northern brown bandicoots, we observed a decline in population size 30 years post-offsetting. This was observed in both development scenarios and for both final population size and EMA (Fig S.16, S.17), however it was often notable in the strict avoidance scenario and when examining the minimum abundance of each population (Fig.4). The average EMA declined by 8.1% in both development scenarios (Fig S.17) and the final population size by 0.6% in S1 and 1.4% in S2 between 15 to 30 years after last offset was implemented (Fig S.16). Values rose again by 100 years (Fig.3), to closely match 15-year population values. Thus, there is an apparent time-lag in accrual of development impacts on species' abundance values and a subsequent lag in accrual of offset benefits.

Accepted Article

Area adjusted percentage change in final population

We used our area adjusted percentage change in final population size values to determine per hectare gains in population persistence delivered by each metric. Area adjusted final population values varied within species and between scenarios. However, patterns across species were similar; when impacts were avoided, vegetation area and condition delivered proportional gains in population sizes for a small restoration area (Fig.5). This is particularly apparent for yellow-bellied gliders, where distribution of high-quality restorable habitat was limited and the per hectare gain in population size was more important. When impacts were high (S2), the benefits of vegetation area and condition were lost for all species, but the ratio for species-specific metrics such as habitat suitability, abundance, metapopulation connectivity, and richness increased (Fig.5).

Accepted Article

Discussion

This study highlights the implications of choosing different biodiversity metrics to quantify development-offset trades. We show that achieving NNL or NG in the chosen offset metrics does not always ensure population declines are prevented (Fig.2,3). Our findings concur with existing literature which demonstrates that achieving long-term NNL of biodiversity is more complicated than simply ensuring NG in metric values, or total habitat amount (Buschke & Sinclair, 2019; Marshall, Valavi, et al., 2020; Sonter et al., 2020). Whilst offset success varied between metrics and species, none of the metrics we tested were able to counterbalance long-term impacts to populations when developments removed high-quality habitat (Fig.3). Our results again emphasise that when applying offsets in practice proponents should demonstrably attempt to avoid impacts where possible and direct developments towards areas of least concern for key species (Phalan et al., 2017; Villarroya, Barros, & Kiesecker, 2014). Despite consistent NLs in our targeted scenarios, species-specific metrics do show some promise in terms of gains in abundance per hectare offset (Fig.5). However, examination of species' ecological considerations is required to understand how metrics function in practice.

Extensive habitat clearing has put increasing pressure on biodiversity and restricts access to restorable habitat which could be used to alleviate development impacts through offsetting (Kangas & Ollikainen, 2019). We observed the effects of limited land availability on achieving NNL in habitat and potential long-term implications for species. Some species' habitat limitations meant that losses could not be offset when developments were targeted. For example, when developments removed core yellow-bellied glider habitat there was insufficient high-quality habitat available to match development impacts (Fig.2). Continued clearance of core habitat ensures NL and perpetuates declines even when impacts are routinely offset (Kujala, Whitehead, Morris, & Wintle, 2015).

Similarly, simulations using vegetation area and condition to assign offsets consistently exhausted restorable habitat since like-for-like rules meant offsets could only utilise equivalent vegetation classes. Therefore, we were unable to achieve NNL in habitat suitability, let alone ensure persistence.

Accepted Article

This is partly a limitation of our approach because our modelling framework only restored habitat within our region. Where our simulations stopped after all available habitat within the region was restored, in practice proponents may seek offset sites further away or meet offset requirements via other means i.e., through payment into conservation banks (Bekessy et al., 2010). However, our results highlight the known issue of shifting of conservation actions further from impacts (Gordon et al., 2011; Maron et al., 2016). Ideally offset sites should remain close to impacts so as to preserve the functional role of developed sites in the landscape (Bull, Suttle, Gordon, Singh, & Milner-Gulland, 2013). We show the consequences of shifting offset benefits to accommodate like-for-like requirements, both in terms habitat losses and local population persistence (Buschke & Sinclair, 2019; Moilanen, Kujala, & Mikkonen, 2020).

When impacts were targeted towards high-suitability habitat, none of the tested metrics could consistently facilitate population recovery for our target species. This result, combined with limitations in land availability, further emphasise the need for avoidance as a first solution to preventing biodiversity decline. This conclusion is not novel and it has been quantitatively demonstrated previously (Phalan et al., 2017; Sonter et al., 2020). However, our results have highlighted again the risks to biodiversity of not complying with avoidance measures (Bull, Strange, Smith, & Gordon, 2020; Theis et al., 2019), and the potential gains in abundance which can occur by avoiding impacts to high-quality habitat (Marshall, Valavi, et al., 2020). Impacts which strategically avoided high-quality habitat resulted in NNL or NGs in most cases for all the metrics tested (Fig.3) and were able to deliver equitable gains in abundance per hectare offset, improving outcomes even when available habitat in which to restore is limited (Fig.5).

Utilising avoidance measures also reduces reliance on high multipliers, maximises restorable habitat, and allows gains to accrue within required time periods (Gibbons et al., 2016; Kiesecker et al., 2009). Here we observed a time-lag in development impacts on population size and subsequent recovery even when development impacts avoided high-quality habitat and when we assumed that restoration

Accepted Article

was immediately effective (Fig.4; S.16; S.17; Dorrough et al. 2019). Time-lags are a widely discussed topic in offsetting (Maron et al., 2012) but are likely to go unnoticed in practice because most offset-evaluations are conducted using different biodiversity metrics to those used during planning (Marshall, Wintle, et al., 2020). So even if offsets manage to meet their designated requirements (e.g. total vegetation area restored), impacts on populations and species may still go on unnoticed for decades (Maron et al., 2012).

Given that our simulations assume restoration is immediately and entirely effective there should be no risks of delayed benefits to habitat or species, other than those caused by the metrics themselves. Here, we did not start to observe some development impacts on species until 30 years after development and offsets were implemented (Fig.4). Moreover, for some species these lags were more pronounced in the avoidance scenario, suggesting the impacts of these development actions may be unclear for some time after their implementation even when offset gains appear to be successful (Fig.4). In practice this means that if populations are evaluated within 15 or 20 years, which is a common time-frame within which offsets must achieve benefits (Gibbons & Lindenmayer, 2007), there may be a false presumption that offsets are delivering true long-term benefits. Consequences of such a presumption are further exacerbated in practice because in reality habitat creation and restoration have a high likelihood of failure (Dorrough et al., 2019; Maron et al., 2012).

When impacts are not avoided long-term outcomes of offsets are nuanced and species-specific, requiring some understanding of how species' demographic variables and population parameters effect outcomes (Buschke & Sinclair, 2019; Reside et al., 2019; Vanderduys, Reside, Grice, & Rechetelo, 2016). Northern brown bandicoots for example seemed to be better equipped to recover through offsetting even under intense development pressure (Fig.3). This is probably because they have high population growth rates (*Rmax* = 1.5), occur in high densities (*d* = 0.985), and produce multiple generations per year (Kemper et al., 1990), allowing populations to recover when using metrics which account for these processes (Fig.3). Conversely, tiger quolls have large dispersal

Accepted Article

distances, occur at low densities (d = 0.05), and require large home-territories (Körtner, Holznagel, Fleming, & Ballard, 2015). Consequently, stochasticity in placement of development impacts and restoration efforts could drastically alter their population structures, resulting in a decline in abundance even when habitat is being restored (Fig.3). Similar source-sink dynamics have been observed when simulating offsets for other highly dispersive species (e.g. birds), making it even more difficult to determine long-term impacts of developments for wide ranging, locally rare species (Devictor et al., 2010; Marshall, Valavi, et al., 2020). We would expect that offsets are likely to have better long-term outcomes if metrics accurately describe the processes impacting species they intend to protect (Cristescu et al., 2013; Marshall, Wintle, et al., 2020). We observed some evidence to support using such metrics, particularly when considering the total area required to implement beneficial offsets (Fig.5).

This research illustrates that long-term benefits delivered through offsets are dependent on adherence to the mitigation hierarchy, availability of land for restoration and understanding species-specific processes underlying persistence. Designing offset metrics that capture species-specific processes in practice will make accounting for multiple species in offset programs quite complex (Kujala et al., 2015). However, considering the dynamics of species we are aiming to protect should improve our understanding of how offsets are likely to perform for all species (Bekessy et al., 2010; Schmeller et al., 2017). Using species distributions and demographic parameters to help strategically plan development offset trades could mean avoiding impacts on species which cannot be accounted for through offsetting (Moilanen & Kotiaho, 2018; Villarroya et al., 2014), both in terms of extent of their remaining habitat (e.g. yellow-bellied gliders) and their ability to recover (e.g. squirrel gliders and tiger quolls). Simulation approaches, such as those used here, could also be useful for evaluating potential impacts of a proposed development on habitat and species prior to implementation and help estimate long-term risks and benefits of offsets (Bedward, Ellis, & Simpson, 2009).

Accepted Article

There is no one size fits all approach for measuring biodiversity and future policy solutions will likely be nuanced and context dependent. Regardless of the approach, ecological objectives of an offset need to be explicitly stated and metrics should reflect those objectives. A key aspect of this is ensuring offset targets remain achievable within habitat available, and ideally offset gains should be accrued prior to impacts. If development impacts cannot be demonstrably offset with the metrics and habitat available in an adequate time they should not be approved.

Acknowledgements: This work was supported by the Australian Government's National Environment Science Program Threatened Species Recovery Hub. We thank the Hunter and Central Coast Regional Environmental Management Strategy (HCCREMS), and A. Whitehead for compiling the species data and environmental variables. HK also acknowledges support from Academy of Finland Strategic Research Council grant #312559.

Conflict of interest: None declared

Authors' contributions: EM, DS, BAW, and HK conceived the research ideas. EM and RV developed the simulation methodology and tested the codes. CV and EM built the population models. DW and EM implemented the simulations and analysed the results. EM led the writing of the manuscript. All authors contributed critically to the manuscript.

Data availability: Data and code required to repeat these simulations and population models are available via the OSF repository http://doi.org/10.17605/OSF.IO/NG4ZT (Marshall et al., 2021).

References

Albuquerque, F., & Beier, P. (2015). Rarity-weighted richness: A simple and reliable alternative to integer programming and heuristic algorithms for minimum set and maximum coverage problems in conservation planning. *PLoS ONE*, *10*(3), 1–7. doi: 10.1371/journal.pone.0119905

Accepted Article

Bedward, M., Ellis, M. V, & Simpson, C. C. (2009). Simple modelling to assess if offsets schemes can prevent biodiversity loss , using examples from Australian woodlands. *Biological Conservation*, *142*(11), 2732–2742. doi: 10.1016/j.biocon.2009.06.026

Bekessy, S. A., Wintle, B. A., Lindenmayer, D. B., McCarthy, M. A., Colyvan, M., Burgman, M. A., & Possingham, H. P. (2010). The biodiversity bank cannot be a lending bank. *Conservation Letters*, *3*, 151–158.

Boykin, K. G., Kepner, W. G., Bradford, D. F., Guy, R. K., Kopp, D. A., Leimer, A. K., … Gergely, K. J. (2012). A national approach for mapping and quantifying habitat-based biodiversity metrics across multiple spatial scales. *Ecological Indicators*.

Bull, J. W., Strange, N., Smith, R. J., & Gordon, A. (2020). Reconciling multiple counterfactuals when evaluating biodiversity conservation impact in social-ecological systems. *Conservation Biology*, *00*(0), 1–12. doi: 10.1111/cobi.13570

Bull, J. W., Suttle, K. B., Gordon, A., Singh, N. J., & Milner-Gulland, E. J. (2013). Biodiversity offsets in theory and practice. *Oryx*, *47*(3), 369–380.

Buschke, F. T., & Sinclair, S. P. (2019). Adding ecological and evolutionary processes to restoration biodiversity offset models using neutral theory. *Diversity and Distributions*, *25*(9), 1351–1361. doi: 10.1111/ddi.12943

Cristescu, R. H., Rhodes, J., Frére, C., & Banks, P. B. (2013). Is restoring flora the same as restoring fauna? Lessons learned from koalas and mining rehabilitation. *Journal of Applied Ecology*, *50*(2), 423–431. doi: 10.1111/1365-2664.12046

Devictor, V., Clavel, J., Julliard, R., Lavergne, S., Mouillot, D., Thuiller, W., … Mouquet, N. (2010). Defining and measuring ecological specialization. *Journal of Applied Ecology*, *47*, 15–25. doi: 10.1111/j.1365-2664.2009.01744.x

Dorrough, J., Sinclair, S. J., Oliver, I., Management, E., Branch, S., Division, S., … Wales, N. S.

Accepted Article

(2019). Expert predictions of changes in vegetation condition reveal perceived risks in biodiversity offsetting. *PLOS One*, 1–21.

Gibbons, P., Evans, M. C., Maron, M., Gordon, A., Le Roux, D. S., von Hase, A., … Possingham, H. P. (2016). A loss-gain calculator for biodiversity offsets and the circumstances in which No Net Loss is feasible. *Conservation Letters*, *9*(4), 252–259. doi: 10.1111/conl.12206

Gibbons, P., & Lindenmayer, D. B. (2007). Offsets for land clearing: No net loss or the tail wagging the dog? *Ecological Management and Restoration*, *8*(1), 26–31. doi: 10.1111/j.1442-8903.2007.00328.x

Gordon, A., Langford, W. T., Todd, J. A., White, M. D., Mullerworth, D. W., & Bekessy, S. A. (2011). Assessing the impacts of biodiversity offset policies. *Environmental Modelling and Software*, *26*(12), 1481–1488. doi: 10.1016/j.envsoft.2011.07.021

Hanford, J. K., Crowther, M. S., & Hochuli, D. F. (2016). Effectiveness of vegetation-based biodiversity offset metrics as surrogates for ants. *Conservation Biology*, *31*(1), 161–171. doi: 10.1111/cobi.12794

Hanski, I. (1994). A Practical Model of Metapopulation Dynamics. *Journal of Animal Ecology*, *63*(1), 151–162.

Harwood, T. D., Donohue, R. J., Williams, K. J., Ferrier, S., McVicar, T. R., Newell, G., … Anderson, B. (2016). Habitat Condition Assessment System: a new way to assess the condition of natural habitats for terrestrial biodiversity across whole regions using remote sensing data. *Methods in Ecology and Evolution*, *7*(9), 1050–1059. doi: 10.1111/2041-210X.12579

Hodgson, J. A., Thomas, C. D., Wintle, B. A., & Moilanen, A. (2009). Climate change, connectivity and conservation decision making: Back to basics. *Journal of Applied Ecology*, *46*(5), 964–969. doi: 10.1111/j.1365-2664.2009.01695.x

Accepted Article

Kangas, J., & Ollikainen, M. (2019). Economic Insights in Ecological Compensations: Market Analysis With an Empirical Application to the Finnish Economy. *Ecological Economics*, *159*(November 2018), 54–67. doi: 10.1016/j.ecolecon.2019.01.003

Kemper, C., Kitchener, D. J., Humphreys, W. F., How, R. A., Schmitt, L. H., & Bradley, A. (1990). The biology of the northern brown bandicoot, isoodon macrourus (Marsupialia: Peramelidae) at mitchell plateau, Western Australia. *Australian Journal of Zoology*, *37*(6), 609–616. doi: 10.1071/ZO9890627

Kiesecker, J. M., Copeland, H., Pocewicz, A. M. Y., Nibbelink, N., Kenney, B. M. C., Dahlke, J., … Stroud, D. A. N. (2009). A Framework for Implementing Biodiversity Offsets : Selecting Sites and Determining Scale. *BioScience*, *59*(1), 77–84.

Körtner, G., Holznagel, N., Fleming, P. J. S., & Ballard, G. (2015). Home range and activity patterns measured with GPS collars in spotted-tailed quolls. *Australian Journal of Zoology*, *63*(6), 424–431. doi: 10.1071/ZO16002

Kujala, H., Whitehead, A. L., Morris, W. K., & Wintle, B. A. (2015). Towards strategic offsetting of biodiversity loss using spatial prioritization concepts and tools: A case study on mining impacts in Australia. *Biological Conservation*, *192*, 513–521. doi: 10.1016/j.biocon.2015.08.017

Liu, C., White, M., & Newell, G. (2013). Selecting thresholds for the prediction of species occurrence with presence-only data. *Journal of Biogeography*, *40*(4), 778–789. doi: 10.1111/jbi.12058

Maron, M., Hobbs, R. J., Moilanen, A., Matthews, J. W., Christie, K., Gardner, T. A., … Mcalpine, C. A. (2012). Faustian bargains ? Restoration realities in the context of biodiversity offset policies. *Biological Conservation*, *155*, 141–148.

Maron, M., Ives, C. D., Kujala, H., Bull, J. W., Maseyk, F. J. F., Bekessy, S. A., … Evans, M. C.

Accepted Article

(2016). Taming a Wicked Problem : Resolving Controversies in Biodiversity Offsetting. *BioScience*, *66*(6), 1–10. doi: 10.1093/biosci/biw038

Marshall, E., Valavi, R., O'Connor, L., Cadenhead, N., Southwell, D., Wintle, B. A., & Kujala, H. (2020). Quantifying the impact of vegetation-based metrics on species persistence when choosing offsets for habitat destruction. *Conservation Biology*, cobi.13600. doi: 10.1111/cobi.13600

Marshall, E., Visintin, C., Valavi, R., Wilkinson, D. P., Southwell, D., Wintle, B. A., & Kujala, H. (2021) Data from: Integrating species metrics into biodiversity offsetting calculations to improve long-term persistence OSF Digital Repository. http://doi.org/10.17605/OSF.IO/NG4ZT.

Marshall, E., Wintle, B. A., Southwell, D., & Kujala, H. (2020). What are we measuring? A review of metrics used to describe biodiversity in offsets exchanges. *Biological Conservation*, *241*(May 2019), 108250. doi: 10.1016/j.biocon.2019.108250

Martin, N., Evans, M., Rice, J., Lodhia, S., & Gibbons, P. (2016). Using offsets to mitigate environmental impacts of major projects : A stakeholder analysis. *Journal of Environmental Management*, *179*, 58–65. doi: 10.1016/j.jenvman.2016.04.054

McKenney, B. A., & Kiesecker, J. M. (2010). Policy Development for Biodiversity Offsets : A Review of Offset Frameworks. *Environmental Management*, *45*(1), 165–176. doi: 10.1007/s00267-009-9396-3

Merow, C., Smith, M. J., & Silander, J. A. (2013). A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography*, *36*(10), 1058–1069. doi: 10.1111/j.1600-0587.2013.07872.x

Moilanen, A., & Kotiaho, J. S. (2018). Fifteen operationally important decisions in the planning of biodiversity offsets. *Biological Conservation*, *227*(September), 112–120. doi:

Accepted Article

10.1016/j.biocon.2018.09.002

Moilanen, A., Kujala, H., & Mikkonen, N. (2020). A practical method for evaluating spatial biodiversity offset scenarios based on spatial conservation prioritization outputs. *Methods in Ecology and Evolution*, *1*(1), 1–2. doi: 10.1111/j.2041-210x.2010.00016.x

Office of Environment and Heritage. (2017). *Biodiversity Assessment Method*.

Phalan, B., Hayes, G., Brookes, S., Marsh, D., Howard, P., Costelloe, B., … Whitaker, S. (2017). Avoiding impacts on biodiversity through strengthening the first stage of the mitigation hierarchy. *Oryx*. doi: 10.1017/S0030605316001034

Phillips, S. J., Dudik, M., & Dudík, M. (2008). Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. *Ecography*, *31*(December 2007), 161–175. doi: 10.1111/j.0906-7590.2008.5203.x

R Development Core Team. (2017). *R: A language and environment for statistical computing.* Retrieved from https://www.r-project.org/.

Reside, A. E., Cosgrove, A. J., Pointon, R., Trezise, J., Watson, J. E. M., & Maron, M. (2019). How to send a finch extinct. *Environmental Science and Policy*, *94*(November 2018), 163–173. doi: 10.1016/j.envsci.2019.01.005

Schmeller, D. S., Weatherdon, L. V., Loyau, A., Bondeau, A., Brotons, L., Brummitt, N., … Regan, E. C. (2017). A suite of essential biodiversity variables for detecting critical biodiversity change. *Biological Reviews*. doi: 10.1111/brv.12332

Schooley, R. L., & Branch, L. C. (2011). Habitat quality of source patches and connectivity in fragmented landscapes. *Biodiversity Conservation*, 1611–1623. doi: 10.1007/s10531-011-0049-5

Sonter, L. J., Simmonds, J. S., Watson, J. E. M., Jones, J. P. G., Kiesecker, J. M., Costa, H. M., …

Accepted Article

Maron, M. (2020). Local conditions and policy design determine whether ecological compensation can achieve No Net Loss goals. *Nature Communications*, 1–11. doi: 10.1038/s41467-020-15861-1

Theis, S., Ruppert, J. L. W., Roberts, K. N., Minns, C. K., Koops, M., & Poesch, M. S. (2019). Compliance with and ecosystem function of biodiversity offsets in North American and European freshwaters. *Conservation Biology*, *0*(0), 1–12. doi: 10.1111/cobi.13343

Thuiller, W., Münkemüller, T., Schiffers, K. H., Georges, D., Dullinger, S., Eckhart, V. M., … Schurr, F. M. (2014). Does probability of occurrence relate to population dynamics? *Ecography*, *37*(12), 1155–1166. doi: 10.1111/ecog.00836

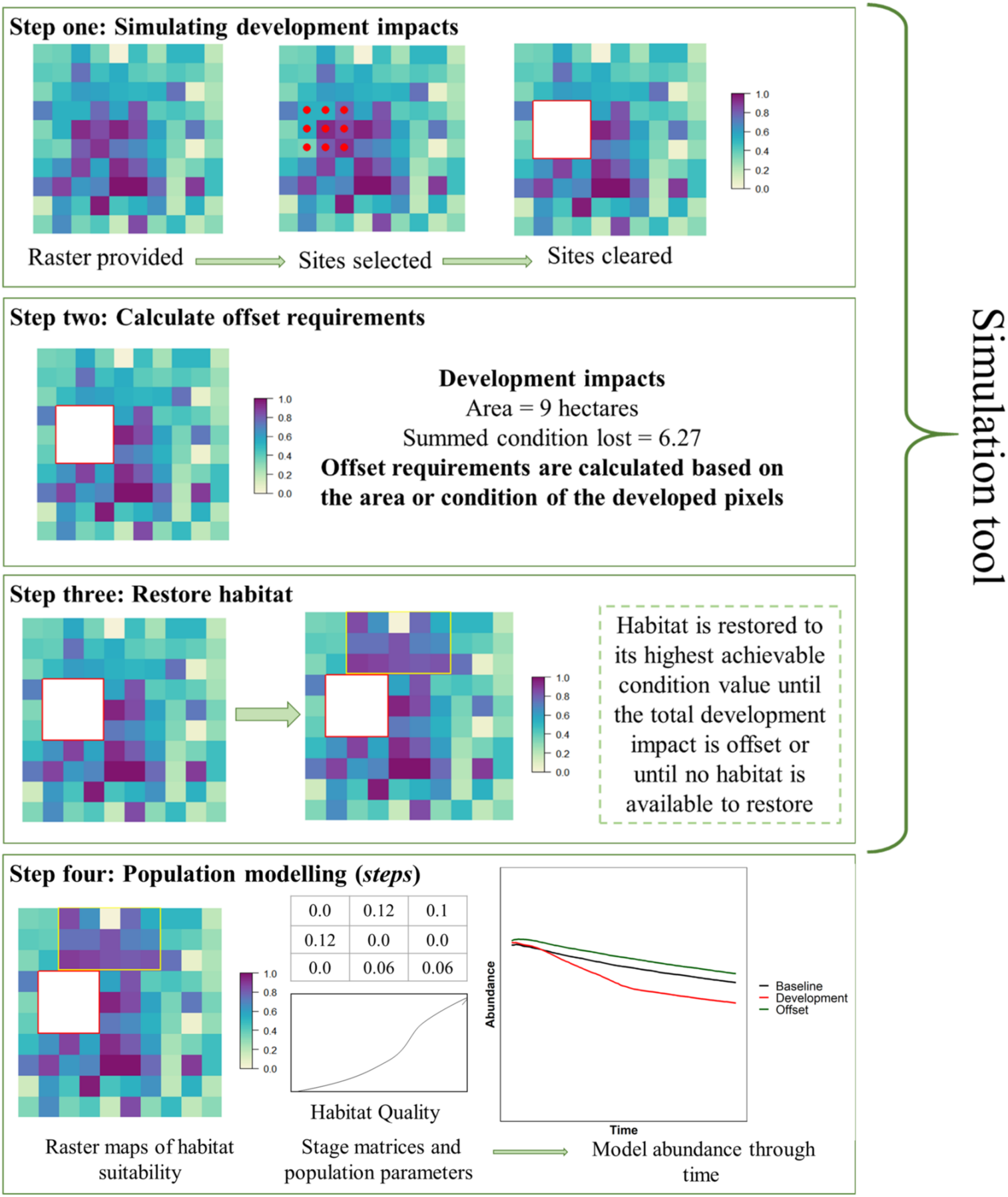
Vanderduys, E. P., Reside, A. E., Grice, A., & Rechetelo, J. (2016). Addressing Potential Cumulative Impacts of Development on Threatened Species : The Case of the Endangered Black-Throated Finch. *PLOS One*, 1–19. doi: 10.1371/journal.pone.0148485

Villarroya, A., Barros, A. C., & Kiesecker, J. M. (2014). Policy development for environmental licensing and biodiversity offsets in Latin America. *PLoS ONE*, *9*(9). doi: 10.1371/journal.pone.0107144

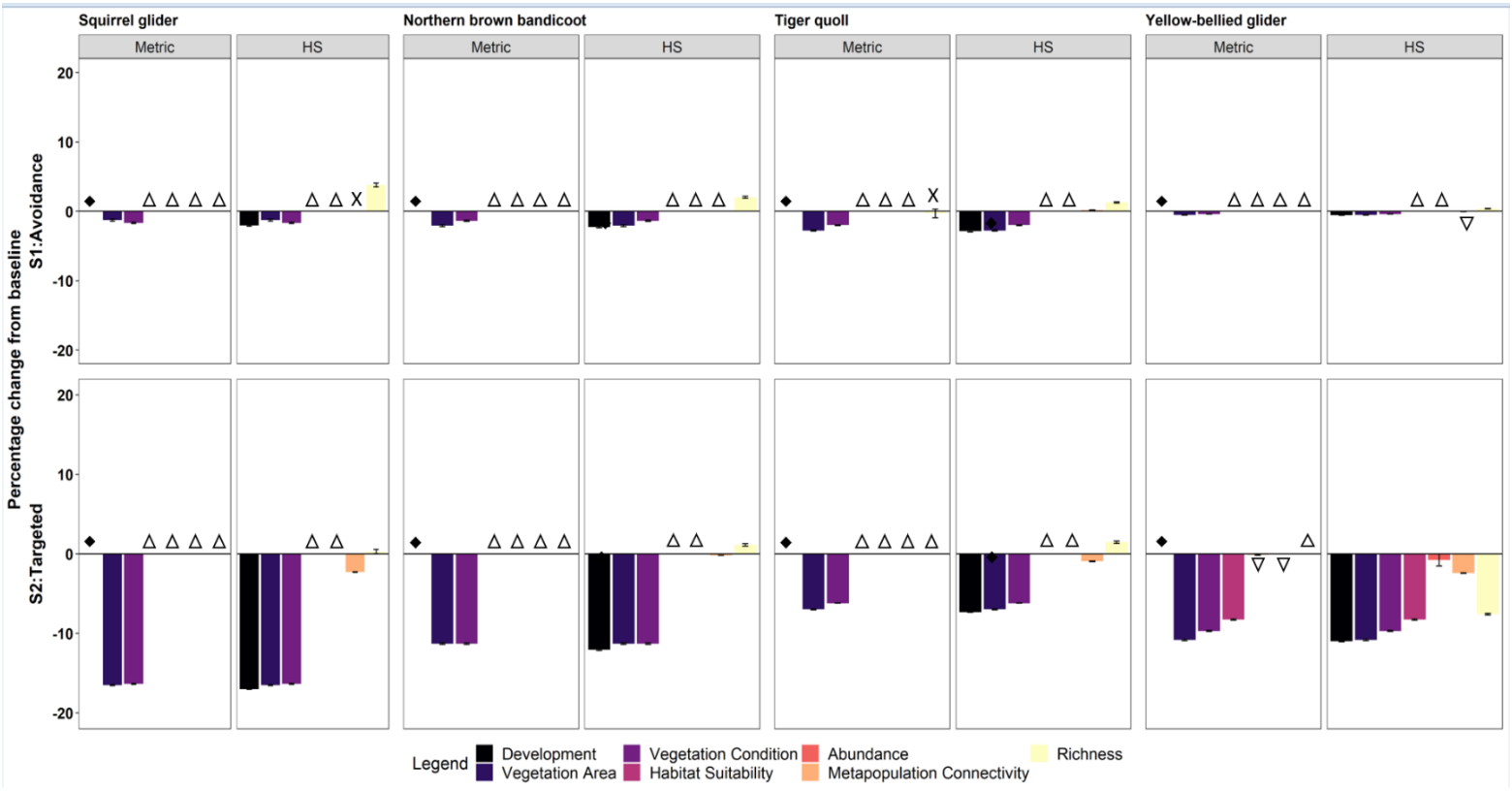
Visintin, C., Briscoe, N. J., Woolley, S. N. C., Lentini, P. E., Tingley, R., Wintle, B. A., & Golding, N. (2020). Steps : Software for Spatially□ and Temporally□ Explicit Population Simulations . *Methods in Ecology and Evolution*, *2020*(July 2019), 1–8. doi: 10.1111/2041-210x.13354

Weissgerber, M., Roturier, S., Julliard, R., & Guillet, F. (2019). Biodiversity offsetting: Certainty of the net loss but uncertainty of the net gain. *Biological Conservation*, *237*(June), 200–208. doi: 10.1016/j.biocon.2019.06.036

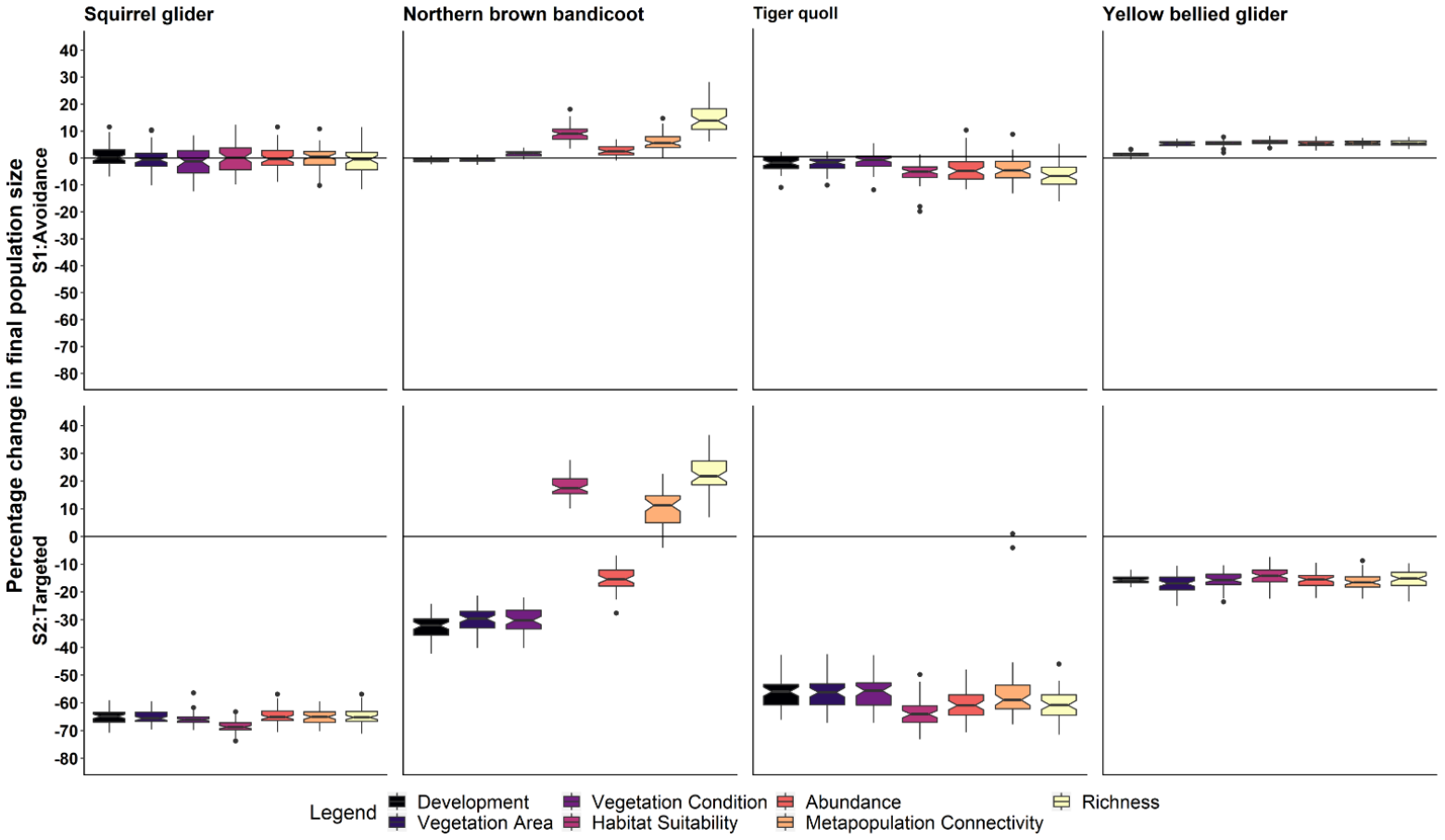
Williams, P., Gibbons, D., Margules, C., Rebelo, A., Humphries, C., & Pressey, R. (1996). A comparison of richness hotspots, rarity hotspots, and complementary areas for conserving diversity of British birds. *Conservation Biology*, *10*(1), 155–174. doi: 10.1046/j.1523-1739.1996.10010155.x



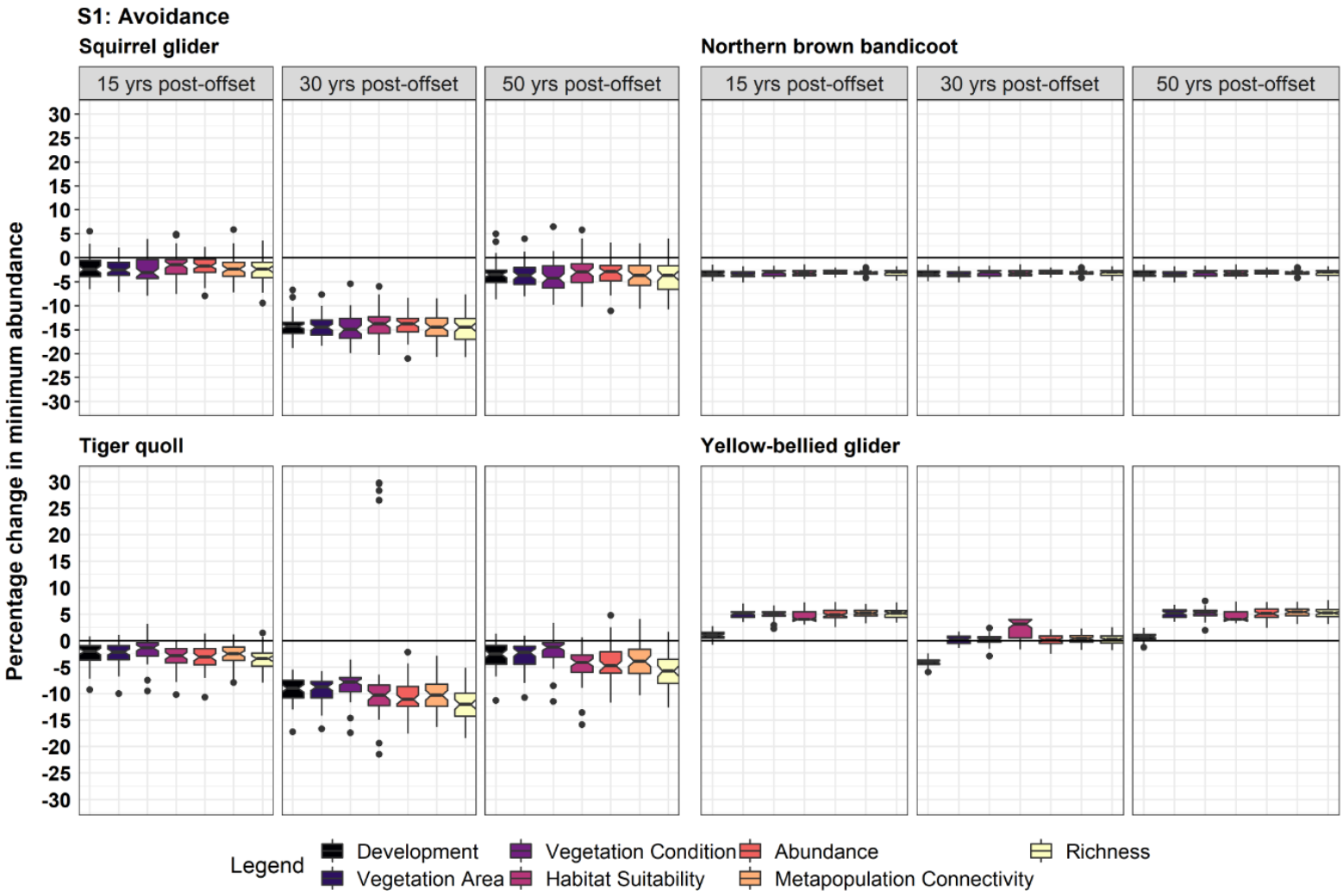
JPE_14117_Figure1 TP.png



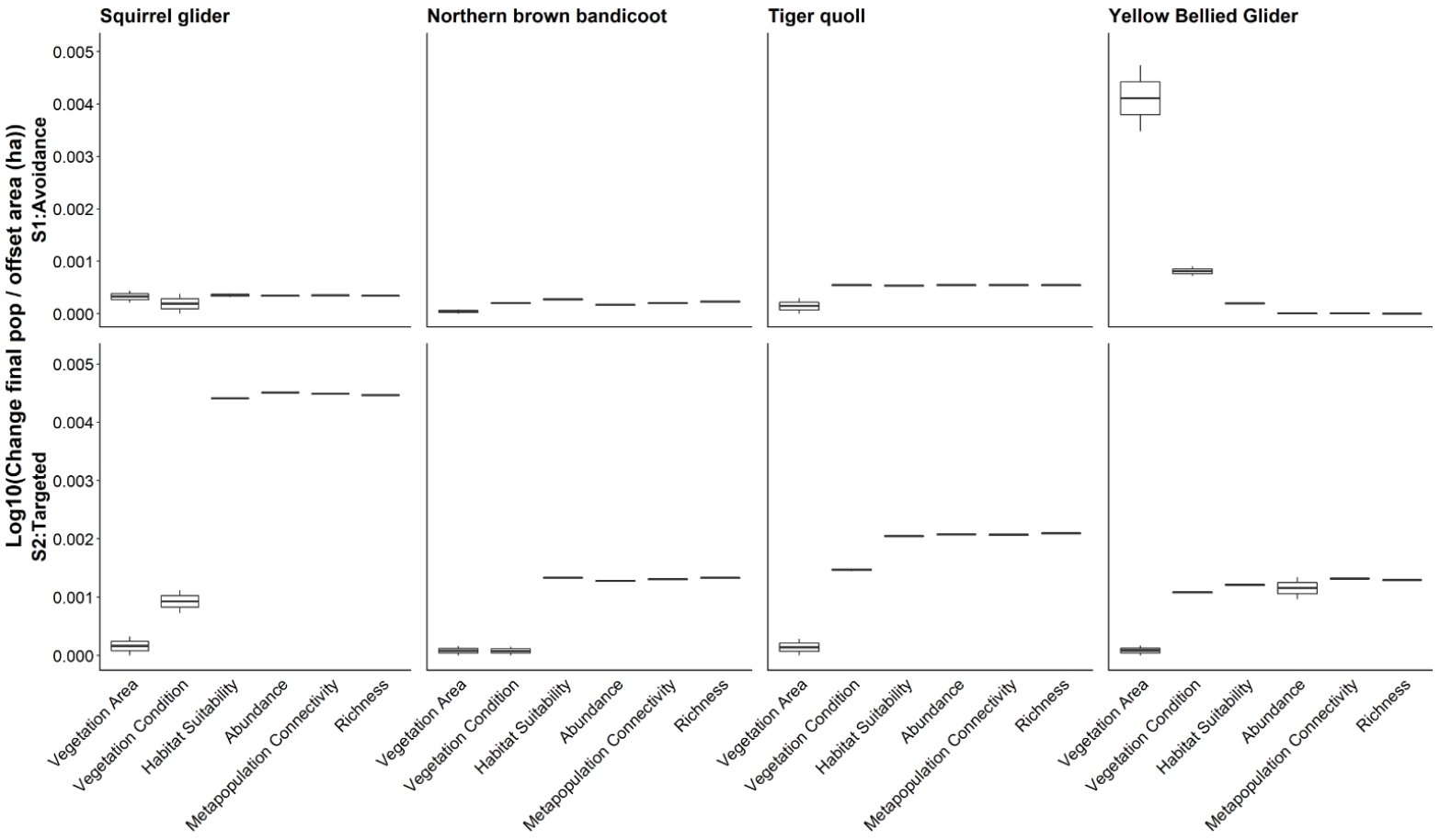
JPE_14117_Figure 2 TP.png



JPE_14117_Figure 3 TP.png



JPE_14117_Figure 4 TP.png



JPE_14117_Figure 5.png