1 # Multi-taxa consequences of management for an avian umbrella species

2 ## Abstract

3 Whether management for so-called umbrella species actually benefits co-occurring biota has rarely
4 been tested. Here, we studied consequences for multiple invertebrate taxa of two ground-
5 disturbance treatments designed to support an avian umbrella species (Eurasian stone-curlew,
6 *Burhinus oedicnemus*), and whether analysing ecological requirements across the regional species
7 pool predicted beneficiaries. Responses were assessed for the abundance of five bird species of
8 conservation concern, and the abundance, species richness and composition of carabids,
9 staphylinids, other beetles (non-carabid, non-staphylinid), true bugs and ants, sampling 31258
10 individuals of 402 species in an extensively-replicated experiment across the UK's largest grass-
11 heath. Both treatments provided suitable habitat for the umbrella species, in contrast to controls.
12 Treatment influenced the abundance of only one bird species; but carabid, other beetle and ant
13 richness increased with one or both treatments, while staphylinid richness and abundance increased
14 and true bug richness and abundance decreased with both treatments. Richness of 'priority' (rare,
15 scarce or threatened) invertebrates *a priori* considered to share ecological requirements with the
16 umbrella species (predicted beneficiaries) increased with both treatments. Resampling and
17 rarefaction showed landscapes diversified by treatment supported a greater cumulative species
18 richness of other beetles, ants and true bugs, and importantly priority invertebrates, than a
19 landscape comprising only untreated controls. Such experiments provide strong evidence to assess
20 co-benefits of umbrella species management, but are costly and time consuming. The systematic
21 examination of the autoecological requirements of co-occurring taxa (the 'Biodiversity Audit
22 Approach') successfully predicted likely beneficiaries. Demonstrating wider biodiversity benefits
23 strengthens the case for avian conservation management.

1. Introduction

Major taxonomic biases exist in conservation (Fazey et al. 2005; Troudet et al. 2017). Thus, whilst some groups have received considerable attention (e.g. vertebrates, Clark and May 2002), others have not, with conservation practitioners often lacking information on which species are present within a region, which ought to be prioritized, and the management interventions each requires (Dolman et al. 2012). Single-species surrogates (e.g. 'umbrella' species), especially those that garner public support and funds for conservation ('flagship' species) are often used to drive ecosystem management (Hunter et al. 2016). However, tests of the wider biodiversity consequences of their subsequent conservation have rarely been conducted (Branton and Richardson 2011; Roberge and Angelstam 2004).

Traditional tests of the umbrella species concept examine whether the overall richness or abundance of other species (generally a single taxon or a limited species subset) is higher where the umbrella species is present (e.g. Suter et al. 2002), or whether they occupy similar habitats (e.g. Maslo et al. 2016; Sibarani et al. 2019). Whilst such tests may be appropriate for determining where in the landscape protection should occur (Caro and O'Doherty 1999; Fontaine et al. 2007), they do not establish whether interventions designed for the umbrella species benefit either wider biodiversity, or co-occurring species of conservation concern. Within 'semi-natural' habitats, where conservation requires management interventions that modify vegetation composition and/or structure (Webb 1998; Wright et al. 2012), experiments that evaluate responses to trial prescriptions, across multiple taxa, provide the necessary evidence-base to allow the beneficiaries of umbrella species management to be appraised.

Here, we test the consequences of habitat management for Eurasian stone-curlew (*Burhinus oedicnemus*, hereafter 'stone-curlew'), a UK threatened (Stanbury et al. 2017), high-profile species protected under European legislation (EC 1979) and widely accorded 'flagship' status, not only for a set of other bird species of conservation concern, but also for multiple invertebrate groups. We selected stone-curlew because they breed in managed semi-natural landscapes, where their requirement for bare-open grassland is well-known (Green et al. 2000). Additionally, a systematic bioregional process that classified species with shared autecological requirements into multi-taxa 'management guilds' (the 'Biodiversity Audit Approach'; Dolman et al. 2012), suggests many nationally rare, scarce or threatened species (hereafter 'priority' species) should benefit from similar conditions (see details in Appendix A). However, without validating this experimentally, it is unclear whether management designed for the umbrella species does indeed improve the status of the intended beneficiaries, or whether such gains are offset by reductions in other priority species. We thus devised an extensively replicated landscape-scale experiment across the UK's largest (3850 ha) lowland grass-heath (66 treatment plus 36 control plots, total 204 ha) to examine outcomes of ground-disturbance interventions designed to benefit stone-curlew. Responses were examined for priority bird species and five invertebrate groups, each speciose and taxonomically well-resolved. Although extensive, the assessment of biodiversity is incomplete, as we lacked expertise (fungi, bryophytes, soil macro-fauna), or personnel (vascular plants) to sample other important elements. First, we *a priori* predicted that our treatments would create suitable stone-curlew nesting habitat. Next, we predicted that treatment would increase the abundance of birds and the abundance and species richness (hereafter 'richness') of some invertebrates groups, while diversifying invertebrate composition. We specifically hypothesized treatment would increase the richness of priority invertebrates associated with physically-disturbed and/or heavily-grazed grassland, but decrease the richness of those associated with lightly-grazed undisturbed grassland. We further predicted that the entire landscape with treatments and controls would support greater cumulative richness of each invertebrate group, and of overall priority species, relative to a control-only landscape.

2. Methods

2.1. Study area

The study was carried out in Breckland, a bio-geographic region (1000 km^2) of Eastern England characterised by a semi-continental climate and sandy soils that supports over 2000 priority plant and invertebrate species (26% of all UK priority species, Dolman et al. 2012). Multi-taxa responses to ground-disturbance treatments were examined across grass-heath (lightly-grazed undisturbed grassland) on the Stanford Military Training Area (STANTA) (0°76'E, 52°51'N, 3500 ha), Bridgham Heath (0°83'E, 52°44'N, 150 ha) and Brettenham Heath (0°83'E, 52°43'N, 200 ha) (see Appendix B for study area details, and Fig. A.2 for map).

2.2. Experimental treatments

Between January and early May 2015, 66 treated plots (33 deep- and 33 shallow-cultivated) and 36 uncultivated controls, each 2 ha (a size advocated for stone-curlew, Appendix A), were located at least 100 m apart in open grass-heath mostly excluding, but sometimes near, scattered trees or scrub. Shallow-cultivations were created with a rotovator, which broke up and turned the surface (10 - 13 cm), and deep-cultivations with an agricultural plough, which inverted soil and vegetation (25 - 28 cm) bringing up mineral sub-soil. Following usual conservation practice, shallow-cultivated plots were immediately retreated if the first attempt failed to create sufficient exposed bare substrate (>50%). Four plots (two shallow-cultivated, two deep-cultivated), located within dense heather (*Calluna vulgaris*), were topped with a tractor-mounted flail prior to disturbance. Plots with regenerating thistle (*Cirsium arvense*, *Cirsium vulgare*) or common nettle (*Urtica dioica*) were topped, also with a flail, in late June and July.

The risk posed by unexploded ordnance precluded cultivation treatments in the central STANTA 'impact area', restricting treatments to outer areas of STANTA, plus Bridgham and Brettenham Heaths. Thus constrained, treatments and controls were randomly allocated to plots within four strata, based on soil (NSRI 2014), age since cultivation (Sheail 1979) and plant composition; these were calcareous grass-heath of any age (hereafter 'calcareous grass-heath'), young grass-heath (≤110 years old), intermediate grass-heath (111-167 years old), and ancient-acid grass-heath (≥168 years old) (see Table A.1 for details). Both treatments and controls were distributed similarly with respect to Latitude and Longitude (Kruskal-Wallis, $H = 0.79$, $p = 0.67$; $H = 0.46$, $p = 0.79$, respectively; $n = 102$), but vegetation strata (four groups) were not (Latitude, $H = 19.26$, $p < 0.001$; Longitude, $H = 47.19$, $p < 0.001$; $n = 102$) (Fig. A.2).

2.3. Responses to treatment

In June 2015, we assessed habitat suitability for stone-curlew within a randomly selected half (1 ha subplot) of each 2 ha plot. We recorded vegetation height using a sward stick (diameter 90 mm, weight 250 g; following Green and Griffiths 1994) at 42 points distributed evenly along two parallel 100 m transects (placed 30 - 33 m apart), and also whether bare substrate covered over 50% at each point (25 mm diameter), giving incidence from 0 to 42. Where an individual height measurement (on treatment plots) exceeded 2.5 cm due to uneven topography caused by soil disturbance, a new measurement was taken in a different cardinal direction. Vegetation structure of deep-cultivated, shallow-cultivated and control plots was related to thresholds of stone-curlew nesting habitat suitability informed by *a priori* knowledge (see Appendix A) of the species preferences for grass-heath vegetation height and bare ground cover (here quantified as the proportion of 42 sampled points in each plot); examining whether modelled 95% CI limits were <2 cm and >40%, respectively.

120   Birds and ground active invertebrates were also sampled in 2015. Bird sampling was
121   conducted at the scale of the whole 2 ha plot whilst invertebrate sampling was conducted in the
122   same 1 ha subplot as the vegetation structure sampling. Bird responses were examined across all
123   plots, whereas invertebrate responses were examined across 20 deep-cultivated, 20 shallow-
124   cultivated and 16 control plots located outside the impact area, to limit risks from unexploded
125   ordnance.

126   Bird abundance was sampled for five species of UK conservation concern, defined as 'Red' or
127   'Amber' listed (Eaton et al. 2015); woodlark (*Lullula arborea*), yellowhammer (*Emberiza citronella*),
128   common linnet (*Linaria cannabina*), Northern lapwing (*Vanellus vanellus*), and Eurasian curlew
129   (*Numenius arquata*). Abundant, ubiquitous breeding species (e.g. skylark, *Alauda arvensis*), were not
130   considered. Analyses considered the maximum number of individuals detected on any one of three,
131   40-minute visits, made to each plot between 1 April and 31 May (mean visit interval: 13.8 days ± 6.5
132   SD) between dawn and 11:00 during dry, still mornings (Beaufort wind force <4), with at least one
133   week between plot establishment and the first survey to allow territory settlement. During visits, we
134   recorded individuals on, or singing directly above, the plot; initially from a vehicle, followed by
135   walking through the plot edge and center. Observations were restricted to vehicles for three of 306
136   plot-visits (affecting three deep-cultivated treatments) to minimize disturbance to breeding stone-
137   curlew; for these vehicles were repositioned to maximize coverage.

138   Invertebrates on each plot were sampled using six pitfall traps (11 cm depth, 8 cm diameter,
139   covered with 12 mm wire mesh, with 50 ml of a saturated salt solution, NaCL), set 15 m apart in a 15
140   m x 30 m grid, opened for seven consecutive days, separately in each of June, August and October.
141   Samples were pooled across pitfalls and months giving one composite sample per plot. Subsequent
142   analyses controlled for cumulative 'trap-days' arising from partially-successful (considered inactive
143   for half the exposure period) or failed traps (considered inactive for the whole exposure period), and
144   unavoidable, though slight, variation in exposure (traps on two deep-cultivated, one shallow-
145   cultivated, and one control plot were deployed two days longer in August due to military
146   restrictions, affecting 2.4% of all seven-day plot deployments).

147   Invertebrates within five groups were identified to species: carabid, staphylinid, all other
148   beetles, true bugs (but excluding aphids), and ants. Following Dolman *et al.* (2012), priority species
149   were identified as those designated as: (i) threatened (critically endangered, endangered and
150   vulnerable) or near-threatened based upon red list guidelines developed by the International Union
151   for Conservation of Nature; or (ii) Nationally Rare (NR) and Nationally Scarce (NS), or the older
152   designations of Red Data Book (RDB) and Nationally Notable (Na, Nb). A few beetles (1.2%) and true
153   bugs (1.9%) were only identifiable to genus, or a group of species (Table A.2).
154   Following Dolman *et al*. (2012), priority invertebrate species with similar autecological
155   requirements were grouped into 'management guilds', which allowed us to classify those species
156   associated with dry-open terrestrial habitats (excluding other habitats) as either requiring physical
157   disturbance or heavy grazing to create bare ground or short-open turf (hereafter 'disturbed-
158   grassland' species), or associated with lightly-grazed and physically undisturbed to only lightly-
159   disturbed conditions (hereafter 'undisturbed-grassland' species) (Table A.2).

160

2.4. Data analysis

162   Separate analyses considered: vegetation structure (height and bare ground cover, to examine
163   suitability for the umbrella species); bird abundance (five analyses, one per species); richness and
164   abundance of each invertebrate group and of priority invertebrates; and richness of disturbed-
165   grassland and undisturbed-grassland priority invertebrates. Using Generalised Linear Models (GLMs),
166   we examined fixed effects of treatment (three levels), vegetation strata (initially four levels, see
167   below), and for invertebrates the cumulative number of trap-days per sample. Ants were omitted

168 from abundance analyses as coloniality can locally inflate abundance. All models were run in R (R
169 Core Team 2015).

170   For each analysis the appropriate error term (normal, binomial, Poisson, or negative
171 binomial) was selected by examining the ratio of deviance / residual degrees of freedom of full
172 (global) models; with normal error, response variables were log or square-root transformed where
173 appropriate. For parsimony, vegetation strata were combined if parameter estimates were similar
174 and their merger did not reduce model performance (change in Akalike's Information Criterion
175 corrected for small sample size, ΔAICc, on combination ≤2; Burnham and Anderson 2002); where all
176 vegetation strata were similar then levels were not combined and subsequent model selection
177 removed this variable. Next, candidate models comprising all possible variable combinations were
178 examined using the package 'lme4' (Bates et al. 2017); the top ranked model was considered 'best' if
179 ΔAICc >2 relative to the next-ranked model; where competing models were within 2 ΔAICc the most
180 parsimonious was selected, as additional variables lacked strong support (Burnham and Anderson,
181 2002). Where treatment was retained, category means were compared by Tukey's pairwise
182 comparison. For analyses of vegetation structure and bird abundance, separate models were
183 examined that considered either all control plots, including those within the central impact area
184 (where there are no ground-disturbance plots); or that excluded ten 'impact area controls' located
185 >200 m inside the impact area boundary. Spatial autocorrelation of model residuals was examined
186 by Moran's I in the package 'Ape' (Paradis et al. 2004).

187   For each invertebrate group, sampling efficiency of treatments and controls was assessed by
188 comparing sample-based rarefaction (re-scaled to the number of individuals, using the Mao Tau
189 function) extrapolated to the predicted asymptote, using the package iNEXT (Hsieh et al. 2018).

190   Assemblage composition of treatments and controls was examined separately for each
191 invertebrate group by Non-Metric Multidimensional Scaling (NMDS) performed using a Bray–Curtis
192 dissimilarity-matrix of abundance data (square root transformed with Wisconsin double
193 standardization, following Clarke and Warwick 2001) using the 'Vegan' package (Oksanen et al.
194 2018). Invertebrate records not resolved to species, and plots with fewer than three observations for
195 that group, were omitted. Model performance was assessed by examining stress (<0.05 excellent;
196 <0.1 good; <0.2 potentially usefull; >0.3 close to arbitrary, Clarke and Warwick 2001) and NMDS axis
197 one and two scores were compared between treatment and control (three levels) using Kruskal
198 Wallis tests with Dunn-Bonferroni pairwise comparisons (p <0.05).

199   To determine whether diversifying the landscape through umbrella species management,
200 would support a higher cumulative richness of each invertebrate group and of (multi-taxa) priority
201 species, we used rarefaction (resampling among plots). Five scenarios were examined: resampling
202 control plots only ('control-only' landscape); the existing landscape diversified by management for
203 the umbrella species ('treatment-control' landscapes, resampling either from: control and shallow-
204 cultivated plots; control and deep-cultivated plots; or control, deep-cultivated, and shallow-
205 cultivated plots); and a 'treatment-only' landscape (resampling both deep- and shallow-cultivated
206 plots, but without control plots). For each scenario, sample-based rarefaction was extrapolated to
207 the sampling intensity (number of individuals) of the full set, examining the overlap in 95% CIs.
208

3. Results

3.1. Stone-curlew habitat suitability

Vegetation height decreased progressively from control to shallow- then deep-cultivation (Fig. 1) and was lower on pooled young and calcareous than on pooled intermediate and ancient-acid grass-heath. Conversely, bare ground cover increased from control to shallow then deep cultivation and was not influenced by vegetation strata (Table A.5). Both treatments, but not controls, provided habitat suitable for nesting stone-curlew (Fig. 1).

Residuals from the bare ground cover model (but not the vegetation height model) were weakly spatially autocorrelated (Moran's I = 0.05), suggesting some variation attributable to a spatially correlated factor(s) not considered in the modeling. Nevertheless, we consider inference for treatment effects to be robust, as treatments and controls were distributed randomly and were balanced across vegetation strata (Table A.1). Impact-area controls were retained in both analyses, as their inclusion did not affect model inference.

3.2.1. Management consequences for wider biodiversity

Sampling gave 187 bird registrations, comprising 52 woodlark, 41 Northern lapwing, 49 Eurasian curlew, 12 yellowhammer, and 33 common linnet; and 31258 invertebrates from 402 species (Table A.2), including: 4740 carabids from 57 species; 5580 staphylinids from 116 species; 1617 other beetles from 139 species (largely from 4 families; 498 Curculionidae, 31%; 243 Elateridae, 15%; 216 Chrysomelidae, 13%; 192 Scarabaeidae, 12%); 1874 true bugs from 72 species; and 17447 ants from 18 species. Of the 402 invertebrates, 52 were priority species, including eleven carabids, ten staphylinids, twenty other beetles, nine true bugs, and two ants. For carabids, staphylinids, other beetles, and ants, sample-based rarefaction approached the asymptote in treatments and controls (Fig. A.1) indicating sampling had effectively captured the composition of these assemblages. However, true bugs were uncommon on, and therefore insufficiently sampled from, deep-cultivated treatments (but approached the asymptote on controls and shallow-cultivated treatments).

Although residuals of the carabid, staphylinid, other beetle, true bug and ant richness and abundance models (both all species and priority species models) were not spatially autocorrelated, residuals from two of the five models considering bird abundance (Northern lapwing and Eurasian curlew) and the priority disturbed-grassland species model (but not the priority undisturbed-grassland species model) were significantly, though again weakly, autocorrelated (Moran's I = 0.05, 0.04, and 0.09, respectively). However, for each affected model, we consider inference of treatment effects to be robust for the reasons explained above. For the priority disturbed-grassland species model, where Moran's I was greatest, removing one particularly abundant species (*Philonthus lepidus*, that comprised 48% of the 471 individuals in this group), removed autocorrelation and gave similar interpretation.

3.2.2. Birds

Of the five species considered, treatment only influenced Northern lapwing abundance (Table A.4), with more recorded on shallow-cultivated treatments than controls (with abundance intermediate on deep-cultivated treatments, Fig. 2). Vegetation strata influenced woodlark and yellowhammer abundance (but was not important in other models), with fewer woodlarks on young than pooled intermediate, calcareous and ancient-acid grass-heath, and more yellowhammers on ancient-acid than on pooled young, intermediate and calcareous grass-heath (Table A.3 and Table A.5). Impact-area controls were retained in all analyses, as their inclusion did not affect model inference.

3.2.3. Carabids

For all carabids, richness was greater on shallow-cultivated treatments than deep-cultivated treatments or controls (Fig. 3) and was not influenced by vegetation strata (Table A.4), while abundance was lower on ancient-acid than pooled young, intermediate and calcareous grass-heath, but was not influenced by treatment (Table A.3 and Table A.5). For priority carabids, richness and abundance were greater on the controls than deep-cultivated treatments (while shallow-cultivated treatments did not differ from controls, they held a greater abundance than deep-cultivated treatments), and on pooled young, intermediate and calcareous than ancient-acid grass-heath. Assemblage composition of treatments and controls differed along NMDS axis one (but not axis two, Fig. 4), with composition contrast relative to controls increasing from shallow to deep-cultivation.

3.2.4. Staphylinids

For all staphylinids, richness and abundance were greater on both treatments than controls (Fig. 3), and for abundance only, on young than pooled intermediate, calcareous and ancient-acid grass-heath (Table A.5) (though retaining both young and calcareous grass-heath as distinct gave similar models, Table A.3). For priority staphylinids, neither treatment nor vegetation strata influenced richness (Table A.4), but abundance was again greater on young grass-heath (again retaining young and calcareous grass-heath as distinct gave similar models). Composition of treatments and controls differed along NMDS axis one (but not axis two, Fig. 4), with contrast relative to controls again increasing from shallow- to deep-cultivation.

3.2.5. Other beetles

For all other beetles, richness was greater on shallow-cultivated than deep-cultivated treatments or controls (Fig. 3), and on pooled young and intermediate than pooled ancient-acid or calcareous grass-heath (Table A.3 and Table A.5). Abundance was greater on shallow-cultivated than deep-cultivated treatments, but with controls similar to both and no effect of vegetation strata (Table A.4). For priority other beetles, richness and abundance were greater on both treatments than controls (vegetation strata were not important). Composition of shallow- and deep-cultivated treatments were similar, but differed from controls along both axes (Fig. 4).

3.2.6. True bugs

Richness and abundance of true bugs declined from control to shallow- then deep-cultivation (Fig. 3) and was greater on intermediate than pooled young, calcareous, and ancient-acid grass-heath (Table A.3 and Table A.5). Neither treatment nor vegetation strata influenced the richness or abundance of priority true bugs (Table A.4). Composition of treatments differed from controls along NMDS axis one (Fig. 4), while on axis two controls and shallow-cultivation differed from deep-cultivation.

3.2.7. Ants

For all ants, richness was greater on both treatments than on controls (Fig. 3, Table A.5) (vegetation strata were not important, Table A.4). Neither treatment nor vegetation strata influenced the richness of priority ants. Composition of shallow- and deep-cultivated treatments were similar, but differed from controls along NMDS axis two (but not axis one, Fig. 4).

3.2.8. Disturbed-grassland and undisturbed-grassland priority invertebrates

299 Of the 52 priority invertebrate species, 22 (471 individuals) and 19 (287 individuals) were classified
300 as disturbed-grassland and undisturbed-grassland species, respectively (ten were not principally
301 associated with dry-open terrestrial habitats, and for one, insufficient autecological information
302 prohibited classification, see Table A.2). Richness of disturbed-grassland species was greater on both
303 treatments than controls, whilst the richness of undisturbed-grassland species was not influenced by
304 treatment (Fig. 5). Pooled young, intermediate and calcareous grass-heath held a greater richness of
305 disturbed-grassland and undisturbed-grassland species than ancient-acid grass-heath (Table A.5);
306 though disturbed-grassland models which retained intermediate and ancient-acid grass-heath as
307 distinct, or all vegetation strata categories as distinct, were similar (Table A.3).

308

309 **3.3. Landscape-scale management consequences**

310 Extrapolated rarefaction curves showed that the cumulative richness of other beetles, true bugs,
311 ants, and priority invertebrates (but not of carabids or staphylinids) increased when both treatments
312 were combined with controls (treatment-control landscape), relative to the control-only landscape
313 (Fig. 6). Scenarios that re-sampled controls plus only one treatment, or both treatments but not
314 controls (the treatment-only landscape), gave greater cumulative richness of other beetles, ants and
315 priority invertebrates, compared to the control-only landscape, but did not differ from resampling
316 controls plus both treatments. For true bugs, resampling either deep-cultivated and control plots, or
317 a treatment-only landscape, increased cumulative richness relative to the control-only landscape;
318 however resampling shallow-cultivated plus control plots did not. For staphylinids, a treatment-only
319 landscape gave lower cumulative richness than any treatment-control landscape, but not than the
320 control-only landscape.

4. Discussion

We quantified the multi-taxa consequences of stone-curlew management through a well-replicated landscape-scale experiment making over 200 bird observations and obtaining 30000 invertebrate records. Both treatments, unlike controls, supported suitable stone-curlew nesting habitat, and five treatment plots (four deep-cultivated, one shallow-cultivated, but not controls) were colonized by breeding pairs during this study (2015) (confirmed during avian surveys or additional follow up visits). Whilst staphylinid richness / abundance and ant richness increased with both treatments, and carabid richness, other beetle richness, and Northern lapwing abundance increased with shallow-cultivated treatments (but not deep-cultivated), the richness / abundance of one group (true bugs) declined with both treatments. However, the richness of disturbed-grassland priority invertebrate species (predicted beneficiaries) increased with both treatments, invertebrate assemblage composition differed between treatments and controls for every group, and diversified landscapes with treatments and control plots supported a higher cumulative richness of ants, other beetles, true bugs, and priority invertebrates, than a control-only landscape.

4.1. Contrasting taxonomic responses

Treatment only influenced Northern lapwing abundance from the five bird species studied (increased with shallow-cultivation); though woodlark subsequently increased when management was re-applied annually, and was then greater on both treatments than controls (Hawkes et al. 2018). Either shallow-cultivation, or both shallow- and deep-cultivation, were associated with increased staphylinid, carabid, other beetle and ant richness, as well as staphylinid abundance, but lower true bug richness and abundance. Consistent with evidence from Swedish semi-natural grasslands, where different taxa have contrasting habitat-process requirements (Vessby et al. 2002), our results demonstrated that whilst stone-curlew management benefitted some groups, others were disadvantaged. When only priority species were considered, only the richness and abundance of priority carabids (that declined with deep- but not shallow-cultivation) and other beetles (that increased with both treatments) were influenced by treatment. Importantly, just four of the 52 priority species were found exclusively on controls, of which two, *Harpalus pumilus* and *Leiodes longipes*, are regarded as NR / RDB. Thus, although the richness / abundance of priority species was similar among treatment and control plots for most groups, few species were lost entirely though treatment.

Invertebrate assemblage composition differed between treatments and controls for all five groups, probably because the open sparsely-vegetated structure created by treatments promoted the warmer micro-climate (Krämer et al. 2012) required by thermophilous species (Cameron and Leather 2011; Pedley et al. 2013), whilst controls retained the taller vegetation and plant assemblage required by many herbivores (Woodcock and Pywell 2010). Interestingly, this difference increased progressively from control to shallow- then deep-cultivation for carabids, staphylinids and true bugs, consistent with the observed increase in extent of bare ground from controls (largely closed swards), through shallow-cultivation then greatest in deep-cultivation (almost entirely bare) (Fig. 1).

Our knowledge that stone-curlew and large numbers of priority species share similar management requirements (informed by Biodiversity Auditing, Dolman et al. 2012) provides us with an important justification for regarding the bird as an umbrella species, but establishing whether these species respond as expected is important. In agreement with our *a priori* prediction, the richness of disturbed-grassland priority invertebrates, the predicted beneficiaries, increased with both treatments. It is noteworthy that sampling revealed 22 priority disturbed-grassland species within eight months of treatment establishment, as many of the most range-restricted species within this region are often absent from isolated suitable habitat (Bertoncelj and Dolman 2013; Lin et al. 2006). We anticipate more disturbed-grassland associated priority species will gradually

370 accrue, as reported for specialist carabids and plants in response to similar ground-disturbance
371 management (Pedley et al. 2013).

372

373 **4.2. Consequences of landscape diversification and transformation**

374 Thus far we have examined how birds and invertebrates differed in richness (alpha diversity),
375 abundance and composition between treatments and controls. Whilst this demonstrates the
376 consequences of management, focal-species conservation is rarely implemented across entire
377 landscapes. We were therefore particularly interested in the effects management might have on
378 total species-richness (beta diversity) by diversifying the landscape.

379 Resampling a hypothetical landscape comprising treatments plus controls, gave greater
380 cumulative richness of ants, other beetles, true bugs and priority invertebrates compared to a
381 control-only landscape. However, the decision to implement either one, or both, treatments did
382 little to influence this outcome in all but one instance (true bug richness did not increase with the
383 addition of shallow-cultivations). This reflects that, although invertebrate assemblage composition
384 tended to be distinctly different between treatments and controls (Fig. 4), the two treatments
385 tended to show at least some (carabids) or considerable (other beetles, ants) overlap in
386 composition. Differences between our hypothetical intervention (treatment-only or treatment-
387 control) and non-intervention (control-only) landscapes are consistent with other studies, which
388 show a positive relationship between richness or abundance and landscape heterogeneity (Smith et
389 al. 2010; Weibull et al. 2000). Considering a hypothetical treatment-only landscape lacking any
390 undisturbed grassland gave a greater cumulative richness of ants, other beetles, true bugs and
391 priority invertebrates, than a control-only landscape, but fewer staphylinids than any control-
392 treatment landscape. This suggests that conservation scenarios that diversify but do not entirely
393 replace grasslands through stone-curlew management, would support the most species.

394

395 **4.3. Consequences for Biodiversity conservation of grassland and heath**

396 Lowland heath and dry (chalk, acid or dune) grasslands have distinct assemblages but similar
397 ecological processes (Rodwell 1991, 1992); and biodiversity response to management (Dolman et al.
398 2012). In this mosaic of soils (podsol to rendzina) and grass-heath age, vegetation strata influenced
399 woodlark and yellowhammer abundance; carabids, staphylinids, other beetles, and true bugs
400 richness / abundance; and cross-taxa richness of both disturbed and undisturbed-grassland priority
401 species. Crucially, no stratum was consistently better than others, which is unsurprising given
402 grassland invertebrates differ in their habitat preference (Woodcock and Pywell 2010).

403 Invertebrate assemblage composition consistently differed between treatments and
404 controls, though richness / abundance (alpha diversity, per plot) did not increase for all groups.
405 Focusing on priority invertebrates, just two groups (carabids and other beetles) were influenced by
406 treatment, while for those priority species associated with disturbed-grassland (i.e. predicted
407 beneficiaries) cross-taxa richness increased with both treatments, consistent with our *a priori*
408 prediction. Combining ground-disturbance management with controls increased the cumulative
409 landscape-wide richness (beta-diversity) of most invertebrate groups, and importantly of priority
410 invertebrates across groups. It is on this basis that we conclude that stone-curlew represent a
411 suitable management surrogate within this landscape, and recommend strategies that promote
412 heterogeneity by implementing stone-curlew management, across vegetation strata, whilst also
413 maintaining undisturbed grassland. Further work will investigate whether the immediate benefits
414 observed by this study are retained with follow-up management (e.g. repeat disturbance to maintain
415 open habitats), or whether retaining some plots fallow beyond their suitability for stone-curlew
416 offers different outcomes to broader biodiversity

5. Conclusion

Because the requirements of single species rarely embrace the requirements of all co-occurring biota, the appropriateness of the umbrella species concept has been questioned (e.g. Roberge and Angelstam 2004). However, most tests of efficacy fail to consider the consequences of umbrella species management on other biota; the appropriate appraisal within semi-natural landscapes, where adoption of umbrella species drives interventions beyond site protection (e.g. Thornton et al. 2016). Our study demonstrated that appropriate experiments can provide strong evidence to assess both co-beneficiaries and the disadvantaged from umbrella species management. Importantly, responses differed strongly between taxa suggesting that experimental tests of surrogacy must be broad in taxonomic scope. A precautionary approach, to retain areas untreated, is also advised given uncertain responses of unsampled taxonomic groups.

Whilst experiments such as ours provide the best means of assessing umbrella species management efficacy, they are costly and time consuming. Systematically examining autoecological requirements across multiple co-occurring taxa (the 'Biodiversity Audit Approach') successfully predicted the beneficiaries of umbrella management in our study and may offer a feasible alternative. Here, we were able to test these *a priori* expectations with an experiment, with both our results and existing autoecological knowledge providing important justification for the widespread adoption of avian conservation management. Similar approaches would be valuable in regions with high concentrations of priority species, especially where management interventions currently focus on a limited, unrepresentative and biased subset of species.

439 **Supplementary data**

440 Supplementary data to this article can be found online at:

Literature cited

Bates, D., Maechler, M., Bolker, B., Walker, S., 2017. Linear Mixed-Effects Models using 'Eigen' and S4, the Comprehensive R Archive Network. https://cran.r-project.org/web/packages/lme4/ (accessed May 2018). In R package Version 1.1-1.5. The Comprehensive R Archive Network.

Bertoncelj, I., Dolman, P.M., 2013. Conservation potential for heathland carabid beetle fauna of linear trackways within a plantation forest. Insect. Conserv. Diver. 6, 300-308.

Branton, M., Richardson, J.S., 2011. Assessing the value of the umbrella-species concept for conservation planning with meta-analysis. Conserv. Biol. 25, 9-20.

Burnham, K.P., Anderson, D.R., 2002. Model selection and multi-model inference: a practical information-theoretic approach. Springer, New York.

Cameron, K.H., Leather, S.R., 2011. Heathland management effects on carabid beetle communities: the relationship between bare ground patch size and carabid biodiversity. J. Insect. Conserv. 16, 523-535.

Caro, T.M., O'Doherty, G., 1999. On the use of surrogate species in conservation biology. Conserv. Biol. 13, 805-814.

Clark, J.A., May, R.M., 2002. Taxonomic bias in conservation research. Science 297, 191-192.

Clarke, K.R., Warwick, R.M., 2001. Change in marine communities: an approach to statistical analysis and interpretation, 2nd edn edn. Plymouth Marine Laboratory, Plymouth.

Dolman, P.M., Panter, C.J., Mossman, H.L., 2012. The biodiversity audit approach challenges regional priorities and identifies a mismatch in conservation. J. Appl. Ecol. 49, 986-997.

Eaton, M., Aebischer, N., Brown, A., Hearn, R., Lock, L., Musgrove, A., Noble, D., Stroud, D., Gregory, R., 2015. Birds of conservation concern 4: the population status of birds in the UK, Channel Islands and Isle of Man. Brit. Birds. 108, 708-746.

Fazey, I., Fischer, J., Lindenmayer, D.B., 2005. What do conservation biologists publish? Biol. Conserv. 124, 63-73.

Fontaine, B., Gargominy, O., Neubert, E., 2007. Priority sites for conservation of land snails in Gabon: testing the umbrella species concept. Divers. Distrib. 13, 725-734.

Green, R.E., Griffiths, G.H., 1994. Use of preferred nesting habitat by stone curlews Burhinus oedicnemus in relation to vegetation structure. J. Zool. 233, 457-471.

Green, R.E., Tyler, G.A., Bowden, C.G.R., 2000. Habitat selection, ranging behaviour and diet of the stone curlew (*Burhinus oedicnemus*) in southern England. J. Zool. 250, 161-183.

Hawkes, R.W., Smart, J., Brown, A., Jones, H., Dolman, P.M., 2018. Experimental evidence that ground-disturbance benefits Woodlark *Lullula arborea*. Ibis 161, 447-452.

Hsieh, T., C, Ma, K., H, Chao, A., 2018. Interpolation and extrapolation for species diversity. http://chao.stat.nthu.edu.tw/blog/software-download (accessed July 2018).

Hunter, M., Westgate, M., Barton, P., Calhoun, A., Pierson, J., Tulloch, A., Beger, M., Branquinho, C., Caro, T., Gross, J., Heino, J., Lane, P., Longo, C., Martin, K., McDowell, W.H., Mellin, C., Salo, H., Lindenmayer, D., 2016. Two roles for ecological surrogacy: indicator surrogates and management surrogates. Ecol. Indic. 63, 121-125.

Krämer, B., Kämpf, I., Enderle, J., Poniatowski, D., Fartmann, T., 2012. Microhabitat selection in a grassland butterfly: a trade-off between microclimate and food availability. J. Insect. Conserv. 16, 857-865.

488   Lin, Y.-C., James, R., Dolman, P.M., 2006. Conservation of heathland ground beetles
489   (Coleoptera, Carabidae): the value of lowland coniferous plantations. Biodivers.
490   Conserv. 16, 1337.
491   Maslo, B., Leu, K., Faillace, C., Weston, M.A., Pover, T., Schlacher, T.A., 2016. Selecting
492   umbrella species for conservation: a test of habitat models and niche overlap for
493   beach-nesting birds. Biol. Conserv. 203, 233-242.
494   NSRI, 2014. National soil map of England and Wales - NATMAP, 1:250,000 Scale, Cranfield
495   University, Bedfordshire. http://www.landis.org.uk (accessed July 2014).
496   Oksanen, J., Blanchet, F., G, Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P., R,
497   O'Hara, R., B, Simpson, G., L, Solymos, P., Henry, M., Stevens, H., Szoecs, E., Wagner,
498   H., 2018. Community Ecology Package Version 2.5-2. https://cran.r-
499   project.org/web/packages/vegan/vegan.pdf (accessed July 2018).
500   Paradis, E., Claude, J., Strimmer, K., 2004. APE: Analyses of Phylogenetics and Evolution in R
501   language. Bioinformatics 20, 289-290.
502   Pedley, S.M., Franco, A.M.A., Pankhurst, T., Dolman, P.M., 2013. Physical disturbance
503   enhances ecological networks for heathland biota: A multiple taxa experiment. Biol.
504   Conserv. 160, 173-182.
505   R Core Team, 2015. *R: a language and environment for statistical computing. Version 3.2.2*.
506   R Foundation for Statistical Computing, Vienna.
507   Roberge, J.-M., Angelstam, P.E.R., 2004. Usefulness of the umbrella species concept as a
508   conservation tool. Conserv. Biol. 18, 76-85.
509   Rodwell, 1991. British Plant Communities. Volume 2: Mires and Heaths. Cambridge
510   University Press, Cambridge.
511   Rodwell, 1992. British Plant Communities. Volume 3: Grasslands and Montane Vegetation.
512   Cambridge University Press, Cambridge.
513   Sheail, J., 1979. Documentary evidence of the changes in the use, management and
514   appreciation of the grass-heaths of Breckland. J. Biogeogr. 6, 277-292.
515   Sibarani, M.C., Di Marco, M., Rondinini, C., Kark, S., 2019. Measuring the surrogacy potential
516   of charismatic megafauna species across taxonomic, phylogenetic and functional
517   diversity on a megadiverse island. J. Appl. Ecol. 56, 1220-1231.
518   Smith, H.G., Dänhardt, J., Lindström, Å., Rundlöf, M., 2010. Consequences of organic farming
519   and landscape heterogeneity for species richness and abundance of farmland birds.
520   Oecologia 162, 1071-1079.
521   Stanbury, A., Brown, A., Eaton, M., Aebischer, N., Gillings, S., Hearn, R., Noble, D., Stroud, D.,
522   Gregory, R., 2017. The risk of extinction for birds in Great Britain. Brit. Birds. 110,
523   502-517.
524   Suter, W., Graf, R.F., Hess, R., 2002. Capercaillie (*Tetrao urogallus*) and avian biodiversity:
525   testing the umbrella-species concept. Conserv. Biol. 16, 778–788.
526   Thornton, D., Zeller, K., Rondinini, C., Boitani, L., Crooks, K., Burdett, C., Rabinowitz, A.,
527   Quigley, H., 2016. Assessing the umbrella value of a range-wide conservation
528   network for jaguars (*Panthera onca*). Ecol. Appl. 26, 1112-1124.
529   Troudet, J., Grandcolas, P., Blin, A., Vignes-Lebbe, R., Legendre, F., 2017. Taxonomic bias in
530   biodiversity data and societal preferences. Sci. Rep. 7, 9132.
531   Vessby, K., Söderström, B., Glimskär, A., Svensson, B., 2002. Species-richness correlations of
532   six different taxa in Swedish seminatural grasslands. Conserv. Biol. 16, 430-439.
533   Webb, N.R., 1998. The traditional management of European heathlands. J. Appl. Ecol. 35,
534   987-990.

535 Weibull, A.-C., Bengtsson, J., Nohlgren, E., 2000. Diversity of butterflies in the agricultural
536   landscape: the role of farming system and landscape heterogeneity. Ecography 23,
537   743-750.
538 Woodcock, B.A., Pywell, R.F., 2010. Effects of vegetation structure and floristic diversity on
539   detritivore, herbivore and predatory invertebrates within calcareous grasslands.
540   Biodivers. Conserv. 19, 81-95.
541 Wright, H.L., Lake, I.R., Dolman, P.M., 2012. Agriculture—a key element for conservation in
542   the developing world. Conserv. Lett. 5, 11-19.
543

Figures
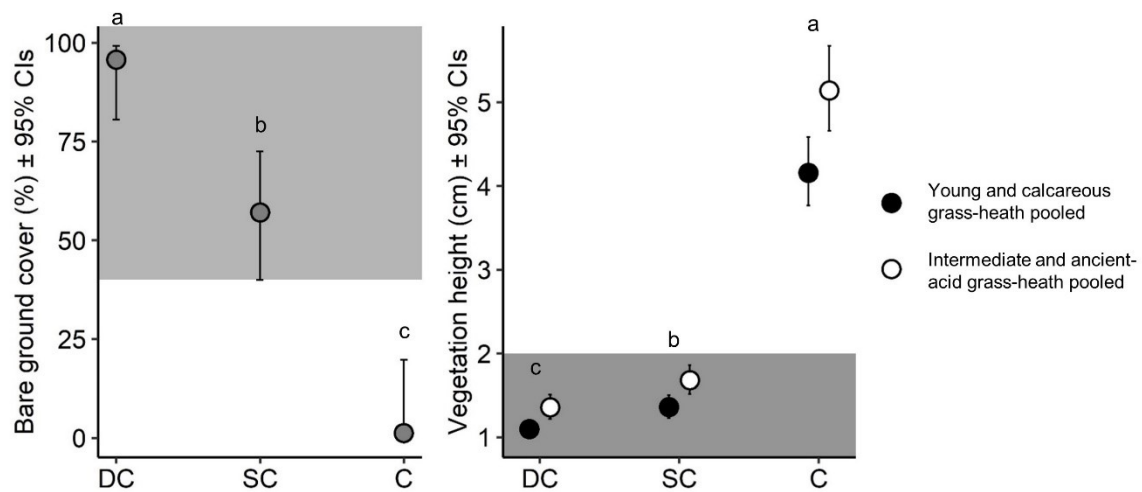
Figure 1. Bare ground cover and vegetation height across deep-cultivated (DC), shallow-cultivated (SC) and control (C) grass-heath plots (*n* = 102). Showing means and 95% CI (bars) from Generalized Linear Models that included the fixed effect of treatment (three levels, both models), and vegetation strata in the bare ground cover model (two categories, see legend); means that share a superscript (homogenous sub-sets, a – c) do not differ significantly (Tukey pairwise comparisons *p* >0.05). Shading denotes limits of suitable Eurasian stone-curlew *Burhinus oedicnemus* nesting habitat. See Table A.5 for selected models.
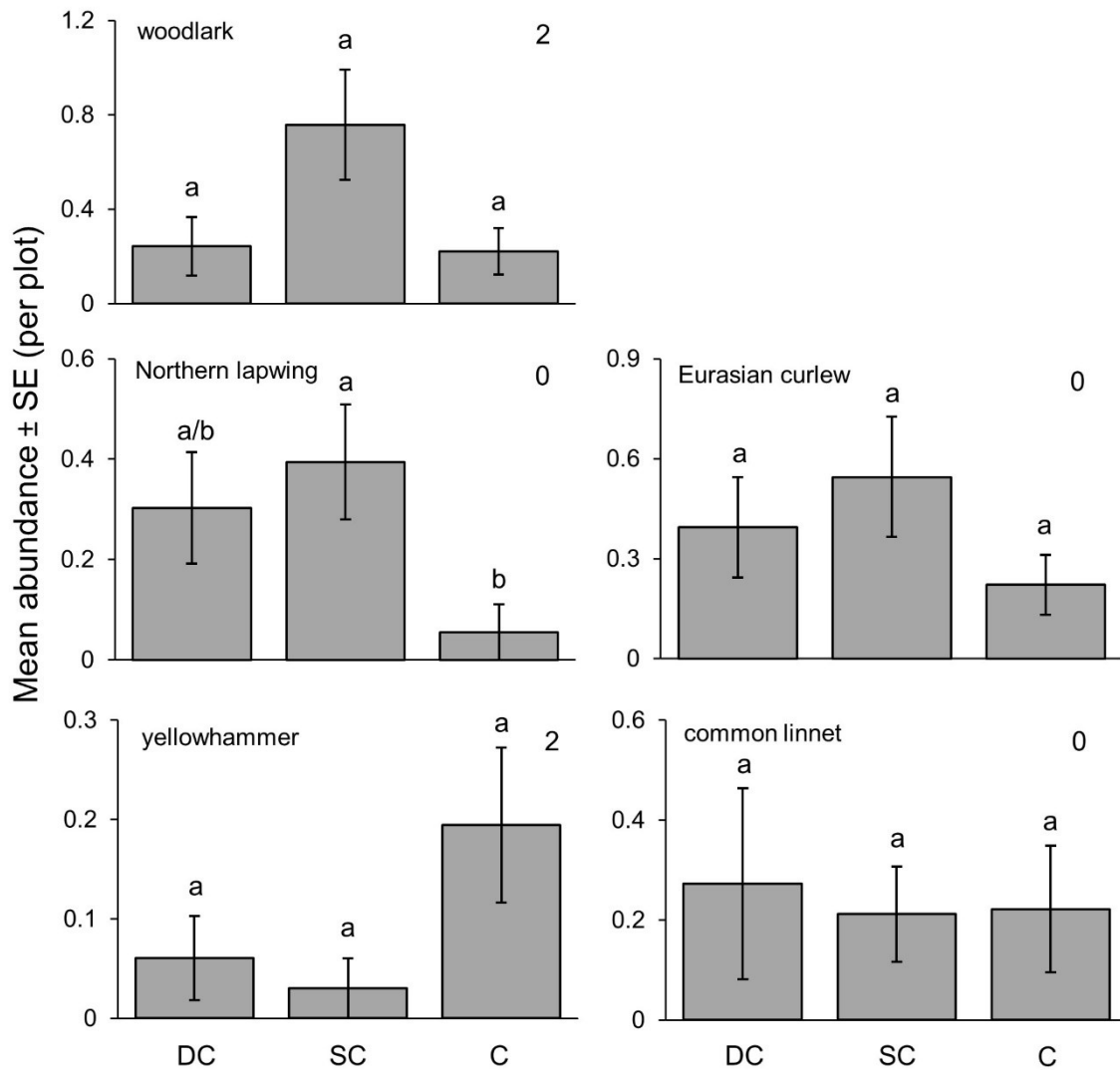
Figure 2. Observed mean (± S.E.) abundance of five bird species of conservation concern across deep-cultivated (DC), shallow-cultivated (SC), and control (C) grass-heath plots (n = 102). Treatments and controls were compared by Generalized Linear Models, controlling for vegetation strata (panels show the number of categories in each analysis, see Table A.3 for identities; '0' indicates vegetation strata is excluded from the model); means that share a superscript (homogenous sub-sets, a – b) did not differ significantly (Tukey pairwise comparisons p >0.05). See Table A.5 for selected models.
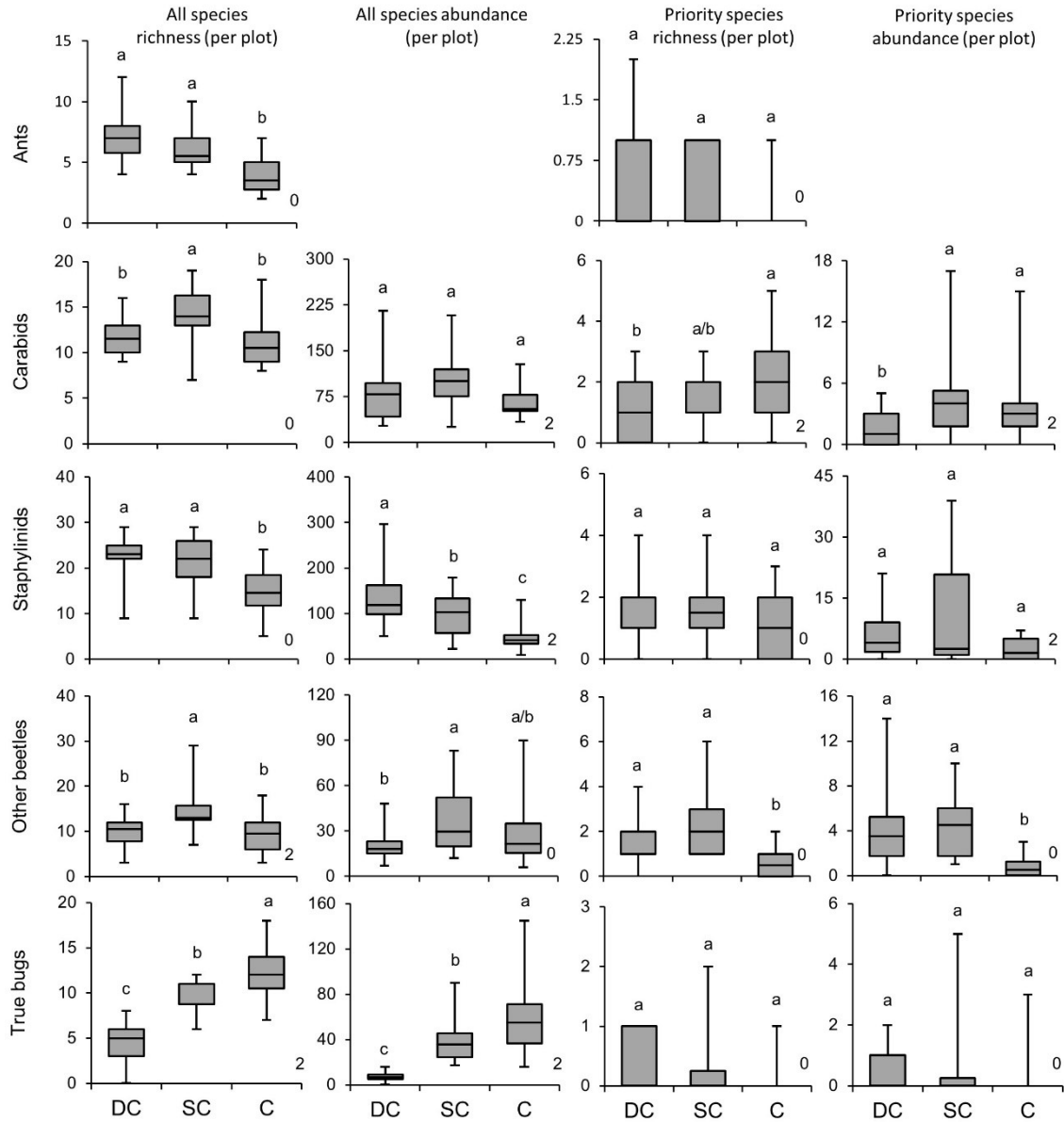
Figure 3. Observed richness and abundance of all species or just priority (rare, scarce or threatened) species, separately for each of five invertebrate groups, across deep-cultivated (DC), shallow-cultivated (SC) and control (C) grass-heath plots (n = 56). Treatments and controls were compared by Generalized Linear Models controlling for vegetation strata (panels show the number of categories in each analysis, see Table A.3 for identities; '0' indicates vegetation strata is excluded from the model); means that share a superscript (homogenous sub-sets, a – c) did not differ significantly (Tukey pairwise comparisons p >0.05). Box plots show the median (central line), 25th and 75th percentiles (box) and range (whiskers).
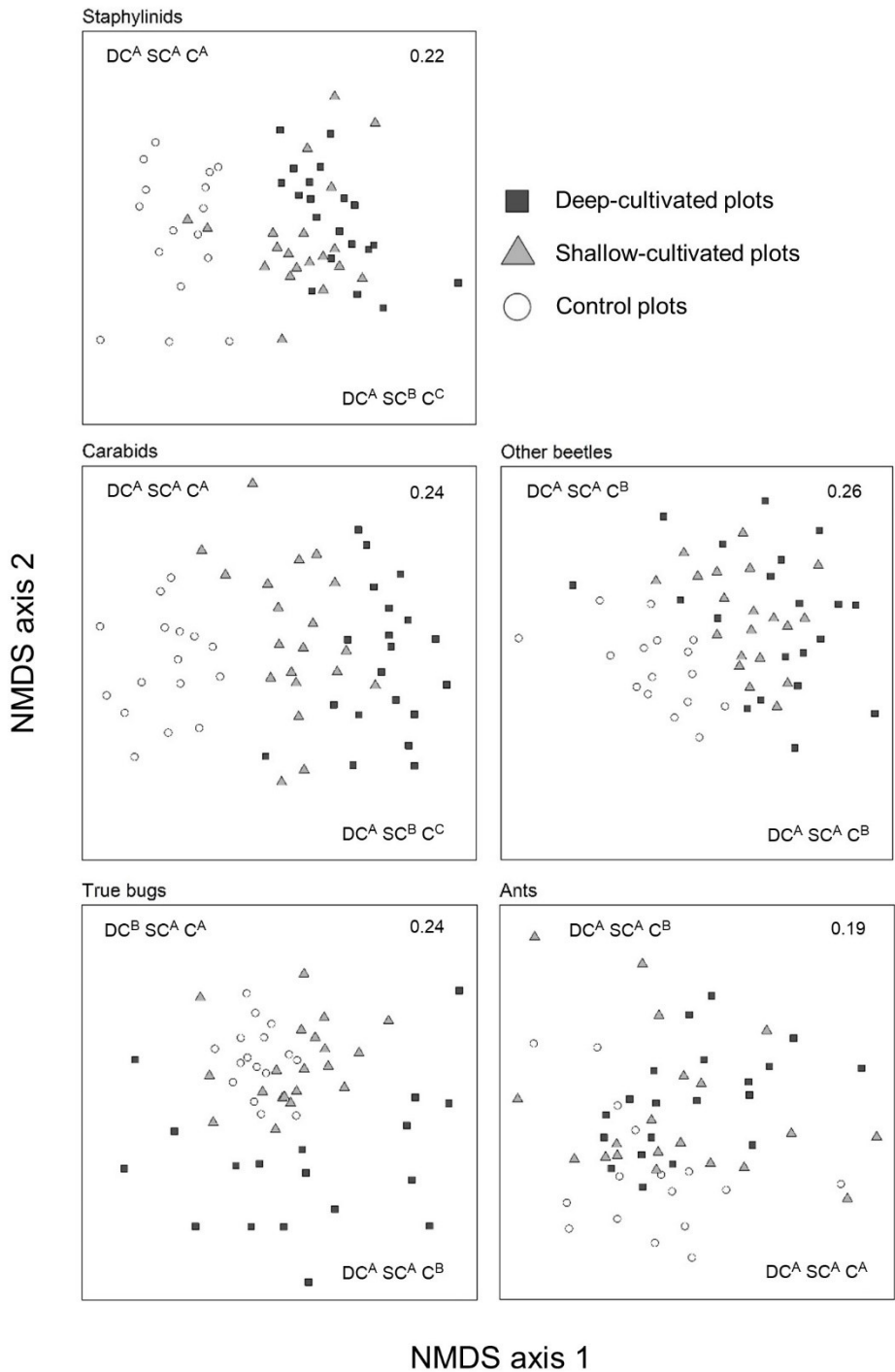
Figure 4. Non-metric Multidimensional Scaling (NMDS) ordination for each of five invertebrate groups across deep-cultivated (DC), shallow-cultivated (SC), and control (C) plots (n = 56; except for true bugs n = 54 as two deep-cultivated plots with fewer than three observations were omitted). Differences between axis scores of treatment and control plots were compared through Kruskal Wallis tests; categories that share a superscript (homogeneous subset, a-c ranked highest to lowest) do not differ significantly (Dunn-Bonferroni pairwise comparisons $p < 0.05$), reported separately for axis 1 (bottom right) and axis 2 (top left) on each. Stress values are shown on each (top right).
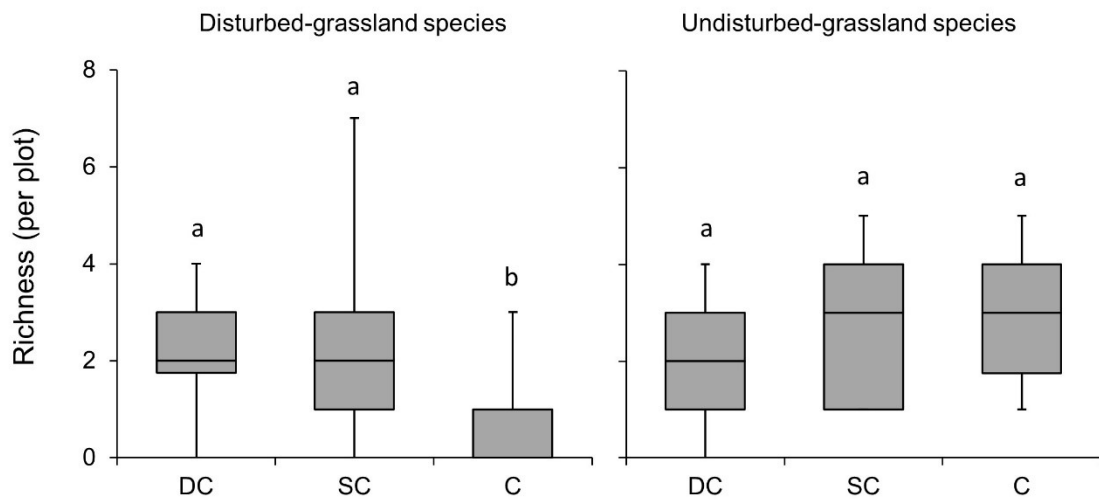
Figure 5. Observed cross-taxa richness of priority (rare, scarce or threatened) invertebrate species, shown separately for those associated with disturbed (i.e. physically-disturbed and/or heavily-grazed) grassland or with undisturbed grassland, across deep-cultivated (DC), shallow-cultivated (SC) and control (C) grass-heath plots (n = 56). Treatments and controls were compared by Generalized Linear Models controlling for vegetation strata (ancient-acid grass-heath vs pooled young, intermediate and calcareous grassland); means that share a superscript (homogenous sub-sets, a – b) do not differ significantly (Tukey pairwise comparisons p >0.05). Box plots show the median (central line), 25th and 75th percentiles (box) and range (whiskers).
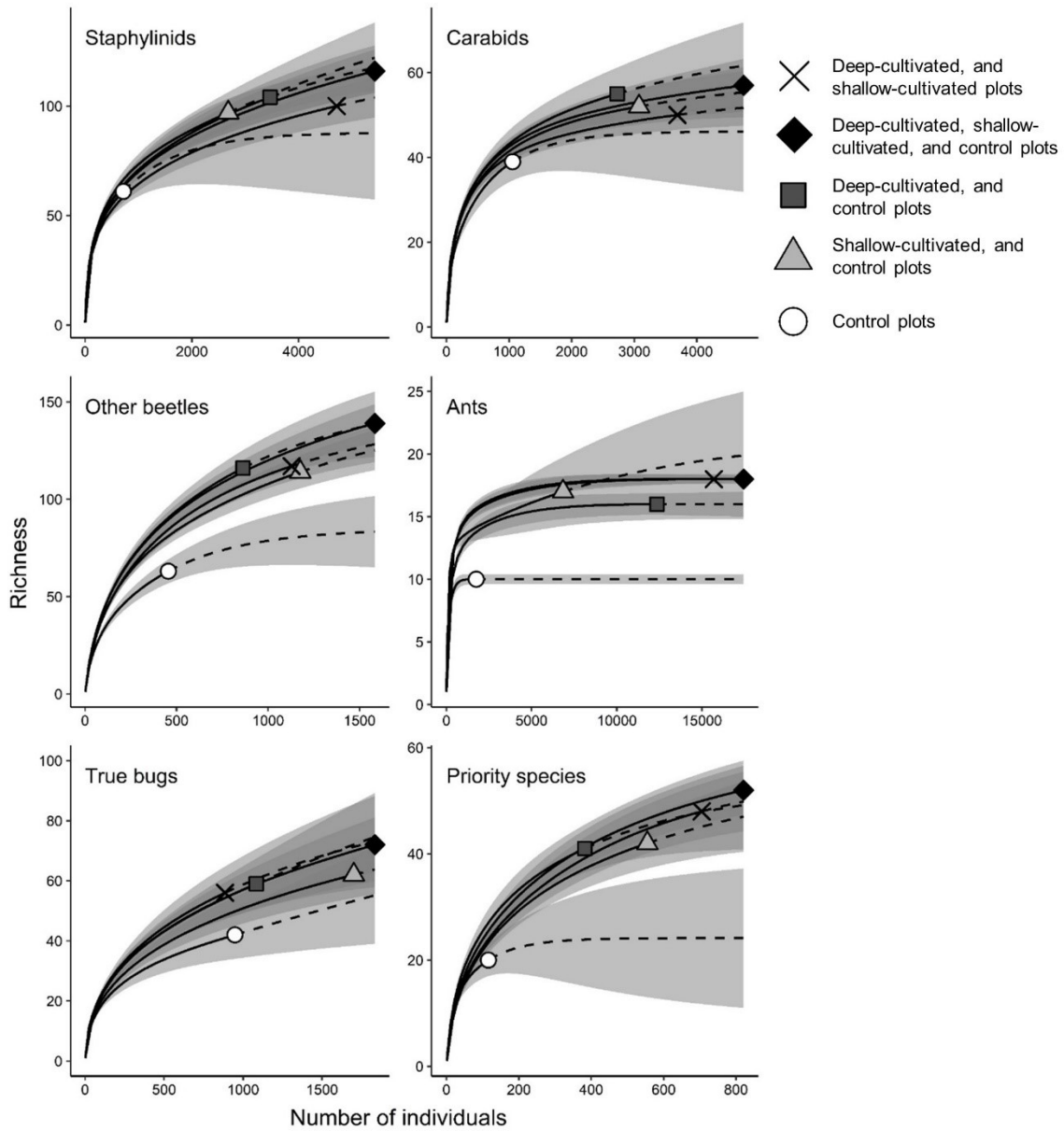
Figure 6. Extrapolated rarefaction curves and 95% CI limits for each of five invertebrate groups, and for all priority (rare, scarce or threatened) invertebrate species, for five hypothetical landscapes comprising different combinations of control and treatment plots (see key). Symbols denote observed richness; solid and dashed lines interpolated and extrapolated richness respectively; shading represents 95% CI bounds. Sample-based rarefactions are rescaled to numbers of sampled individuals.