

DOI: 10.4081/nhs.2025.894
Submitted: 20 March 2025
Accepted: 14 September 2025

Combined use of redundancy and indicator species analyses for the selection of optimal pressure-specific bioindicators: a case study on Carabidae

Stefano Macchio, Michela Gori*, Luisa Nazzini, Susanna D'Antoni

Italian Institute for Environmental Protection and Research (ISPRA), Rome, Italy

*Corresponding author: michela.gori@isprambiente.it

Abstract - This article presents a method for identifying highly effective bioindicator species within a taxonomic group, even in the absence of full knowledge of their ecology or tolerance to specific stresses. This method was tested using trapping data of ground beetle species (Coleoptera: Carabidae) in organic and conventional hazelnut *Corylus avellana* orchards, in order to select the species most sensitive to plant protection products. The dataset includes 10,565 individuals belonging to 57 different ground beetle species, collected in the Piedmont and Latium regions (Italy), where hazelnut cultivation is widespread. A comparison was made between the results obtained through Redundancy Analysis conducted on the entire sample and the same analysis conducted on a portion of the sample obtained by selecting, through the Indicator Value Index, the species with the highest bioindication capacity for the considered pressure. The species were ordered according to the difference of the Indicator Value Index of each, respectively for conventional and organic cultivation, and a battery of Redundancy Analyses was conducted by progressively excluding the species with lower bioindication capacity. The results with the best combination of explained inertia and statistical significance (permutation test) were then selected, defining this approach as Backward Redundancy Analysis Sequence (BRS). Partial Redundancy Analyses were subsequently conducted to isolate the impact of plant protection products from that produced by other factors. Regarding the bioindication capacity of ground beetles on the effects of Plant Protection Products on hazelnut orchards, the results obtained confirm a significantly clearer and more consistent response by applying the BRS compared to the classical approach.

Key words: Backward redundancy analysis sequence, biomonitoring, carabidae, indicator value index, plant protection products impact.

Riassunto - Uso combinato delle analisi di Ridondanza e delle specie indicatrici per la selezione di bioindicatori pressione-specifici: un caso di studio sui Carabidae.

L'articolo presenta un metodo per identificare specie bioindicatrici altamente efficaci all'interno di un gruppo tassonomico, anche in assenza di una piena conoscenza della loro ecologia o tolleranza a stress specifici. Questo metodo è stato testato utilizzando dati di trappolaggio di specie di coleotteri carabidi (Coleoptera: Carabidae) in nocciuleti (*Corylus avellana*) biologici e convenzionali, al fine di selezionare le specie maggiormente sensibili ai prodotti fitosanitari. Il set di dati include 10.565 individui appartenenti a 57 diverse specie di carabidi, in Piemonte e Lazio (Italia) dove la coltivazione del nocciolo è ampiamente diffusa. È stato effettuato il confronto tra i risultati ottenuti tramite l'Analisi di Ridondanza condotta sull'intero campione e la stessa analisi condotta su una parte del

campione ottenuta selezionando tramite l'Indice del Valore Indicatore (Dufrene & Lagrange, 1997) le specie con maggiore capacità di bioindicazione per la pressione considerata. Le specie sono state ordinate in funzione della differenza dell'Indice del Valore Indicatore di ciascuna rispettivamente per la coltivazione convenzionale e quella biologica, ed è stata condotta una batteria di Analisi di Ridondanza andando ad escludere progressivamente le specie con capacità di bioindicazione inferiori. È stato quindi selezionato il risultato con la migliore combinazione di inerzia spiegata e significatività statistica (permutation test); definendo tale approccio Analisi di Ridondanza in Sequenza Inversa (BRS). Sono state successivamente condotte Analisi di Ridondanza parziale per isolare l'impatto dei prodotti fitosanitari da quello prodotto da altri fattori. Relativamente alla capacità di bioindicazione dei Carabidi sugli effetti dei fitofarmaci sui nocchioletti, i risultati ottenuti confermano un responso nettamente più chiaro e consistente applicando la BRS rispetto all'approccio classico.

Parole chiave: Analisi di ridondanza in sequenza inversa, biomonitoraggio, carabidi, *indicator value index*, impatto dei prodotti fitosanitari.

Introduction

In the field of bioindication, variables can be classified into compositional, structural, and functional categories (Noss, 1990). Species, as compositional variables, are recognized for their strong bioindication potential, as indicators derived from them can be universally applied across ecosystems and they are: (i) key factors of the biodiversity; (ii) the basic elements of ecosystems; (iii) sensitive to changes; (iv) defined and measurable in an unambiguous way (Marchetti, 2004). Since it is neither possible nor necessary to consider all species, it is essential to focus on an effective sample able to represent the ecosystem as a whole and be as sensitive as possible to human pressures (Marchetti, 2004). Assessing the degree to which ecological systems are impacted by anthropogenic disturbances and alterations in structure and function is crucial for the long-term conservation of biotic diversity. The capacity to gauge the state and trends of ecological system conditions facilitates the early identification of existing or emerging issues before they develop into major crises. However, the complex and varied nature of ecological systems requires the use and proper validation of a select set of biological condition indicators to enable effective monitoring (Canterbury *et al.*, 2000). Proposed subcategories of bioindicators include environmental, ecological, and biodiversity indicators. Ecological bioindicators are employed to detect and monitor the impacts of stressors on biota, while environmental indicators facilitate the detection and monitoring of changes in specific environmental states, and biodiversity indicators aid in identifying and monitoring species diversity within particular regions (Russo *et al.*, 2021). The use of specific taxocenosis for ecological bioindication purposes has become a well-established practice. Given that many organisms respond to environmental characteristics and their alterations, the potential pool of bioindicator candidates is extensive, complicating the selection process. Selecting appropriate species is crucial as it influences political and management decisions, and an excessive number of bioindicators can yield conflicting results and induce confusion (Russo *et al.*, 2021). Within a given taxocenosis, it is reasonable to expect varying degrees of sensitivity among species to specific issues or parameters, ranging from highly sensitive to unaffected. It is imperative to exclude species with inadequate bioindication capacities from analyses, as they may obscure or distort outcomes, hindering a clear understanding of the impacts of treatments, pressures, or environmental conditions on a community. According to Andreasen (Marchetti, 2004), species selection for use as indicators should be guided by expert judgment grounded in comprehensive knowledge of the natural history of candidate organisms, their habitat affiliations, interactions with other organisms, and roles within the ecosystem. Marchetti

(2004) proposes the utilization of ecological profiles as a tool for evaluating the suitability of selected species based on the criteria mentioned above. These profiles should encompass a range of ecological, policy, and management characteristics for each species, including their habitat utilization, preferences for abiotic conditions, trophic levels and guild associations, spatial distributions, sensitivity to specific human pressures, threat status, endemism, policy implications, and other relevant attributes. Multivariate analyses are commonly employed in biological community assessments due to their accuracy and sensitivity in quantifying human impacts and biological recoveries. However, the treatment of rare species in these analyses remains controversial. Some researchers advocate for excluding rare species, arguing that they may introduce noise and provide minimal additional information compared to more common species, while others maintain that rare species are better indicators of ecosystem stress than common ones (Poos & Jackson, 2012). Nevertheless, the qualitative and quantitative responses of each component of a taxocenosis to ecological factors of interest are often inadequately understood.

This paper aims to establish an effective method for identifying the best bioindicators within a taxocenosis, even in the absence of sufficient knowledge regarding the ecology and tolerance levels of each taxon to specific pressures. To evaluate the proposed method, a dataset concerning the presence and abundance of ground beetle species (Coleoptera: Carabidae) sampled in both organic and conventional hazelnut orchards (*Corylus avellana*) was analyzed, with the use of Plant Protection Products (PPP) serving as the pressure against which the sensitivity of each Carabidae taxon must be assessed.

Most European agricultural landscapes, particularly those characterized by a fine-grained mosaic and low-intensity production systems, were once biodiversity-rich. However, in recent decades, many previously common species have decreased in number or disappeared due to intensified agricultural production and the accompanying decline of semi-natural landscape elements, primarily attributable to reduced habitat heterogeneity and increased land-use practices, particularly fertilizer and PPP application rates (Billeter *et al.*, 2008).

Among the various taxa constituting soil fauna, ground beetles are one of the most prevalent families, both in terms of species diversity and total biomass. Widely distributed worldwide and present in all environments, with over 40,000 species globally and approximately 1,350 in Italy (Vigna Taglianti, 2005), ground beetles are frequently utilized as bioindicators due to their apparent sensitivity to habitat changes, well-understood ecology and systematics, rapid response to habitat alterations, and ease of collection through classic pitfall traps (Eyre *et al.*, 1990; Niemelä, 1990; Luff *et al.*, 1992; Niemelä *et al.*, 1993; Loreau, 1994; Lövei & Sunderland, 1996).

Consequently, this group of ground-dwelling arthropods is increasingly employed in studies assessing both the environmental impacts of human activities on terrestrial ecosystems and the effects of different management practices (Rushton *et al.*, 1990; Magura *et al.*, 2000; Avgin & Luff, 2010).

Although many comparisons of farming systems have involved monitoring Carabidae, often only total carabid numbers or species are compared between systems, potentially masking effects on individual species and yielding misleading conclusions (Büchs *et al.*, 1997; Holland & Luff, 2000).

Data on the presence and abundance of ground beetles in organic and conventional hazelnut orchards were collected during a 5-year Italian project (2015-2019) funded by the Italian Ministry of the Environment and Energy Security (MASE), coordinated by the Italian Institute for Environmental Protection and Research (ISPRA), and conducted in collaboration with the Regional Agency for the Protection of the Environment (ARPA) of Latium and Piedmont Regions, the University of Turin, and the University of Rome Tor Vergata. The project aimed to ascertain whether organic farming and

agroecological best practices are more compatible with biodiversity conservation than conventional farming employing PPP. Additionally, the study aimed to identify bioindicators useful for evaluating the effects of PPP on biodiversity (D'Antoni *et al.*, 2020).

Conducted in the Piedmont and Latium regions during the 2015-16 and 2018-19 campaigns, the project focused on various crops, including rice fields, vineyards, arable crops, and hazelnut orchards. To compare organic and conventional farming and observe the effects of PPP on biodiversity at different spatial and temporal scales, a wide range of bioindicators were selected and tested.

Materials and Methods

Covariates and the selection of fields in pairs

To accomplish the aforementioned objectives, six hazelnut orchards under organic cultivation (referred to as OH) and six hazelnut orchards from conventional farms (labeled as CH) were selected in the Latium region. These orchards were compared for the presence and abundance of ground beetle species as bioindicators. To minimize the effects of covariates, study areas were paired (organic vs. conventional) based on similar area, location, and environmental characteristics. The selection process was guided by covariates organized into thematic groups based on the coherence of their information, such as data pertaining to agronomic practices, chemical treatments, presence of biodiversity attractors, and soil characteristics, as outlined in Macchio *et al.* (2024).

Multivariate methods for the reduction of covariates

More than 100 different covariates were identified (Macchio *et al.*, 2024). Given the statistical objectives, the extensive number of original covariates and their substantial multicollinearity necessitate reduction and removal of correlations among them. Such correlations can induce significant distortions, potentially amplifying the influence of certain variables while obscuring others. To address this, the original covariates were grouped using Principal Coordinates Analysis (PCoA), a method suited for handling diverse datasets simultaneously (Legendre & Legendre, 1998). The resulting derived covariates include X10UP, primarily linked to soil granulometry, and X0107UP, associated with mechanical agricultural practices, as well as the presence of natural vegetation and elements of the rural landscape. The Plant Protection Products Index (PPI), indicative of in-field treatments, and the Plant Protection Products input in Neighboring Fields (PNF), reflecting chemical treatments in adjacent fields, were identified as the two variables correlated with the use of PPP (refer to Macchio *et al.*, 2024). In multivariate analyses, PPI and PNF were employed as target explanatory variables, alongside the PCoA derived covariates (X10UP and X0107UP).

Sampling method for Carabidae

Carabidae have been sampled using pitfall trapping, a widely employed method in ecology owing to its undeniable advantages, which can be succinctly outlined as follows: (i) passive collection, rendering results independent of collector skill; (ii) high catching efficiency, facilitating statistical analysis due to abundance; (iii) user-friendly, even for non-specialists; (iv) quick setup and low operational costs; (v) extensive reference literature availability (Adis, 1979; Brown & Matthews, 2016; Southwood, 1978; Woodcock, 2005).

A total 108 plastic cups (400 cc capacity, 13 cm height, 8.5 cm diameter mouth) were utilized as pitfall-traps each one labeled with an alphanumeric code denoting sampling site and trap location. Each cup, pre-drilled near the upper rim to prevent overflow during rain, was filled three-quarters with white wine vinegar and buried flush with the ground. A ceramic tile (20 cm square, white) was

then placed atop each cup, positioned approximately 2 cm above ground level to permit invertebrate passage while preventing flooding, leaf/debris accumulation, and to avoid bycatch of non-target animals.

Within each of the 12 hazelnut orchards, nine pitfall traps were spaced approximately 10 m apart, forming a 20 m square. Sampling involved 10 annual sessions from May to October. Data collection spanned the 2015-16 and 2018-19 campaigns. Ground beetles from each trap were preserved in 70% ethanol, later sorted and identified to a specific taxonomic level in the laboratory.

Because of unexpected constraints, it was not possible to maintain a consistent number of valid pitfall traps across orchards and sessions. Consequently, the average number of individuals per 100 pitfall traps was calculated for each species and orchard group (organic and conventional). Furthermore, given the substantial variations in population sizes among different taxa within natural populations (Mateos, 2016; Krebs, 1994), the count averages were transformed using the natural logarithm [$\ln(n + 1)$] to mitigate the potential influence of this disparity on the results.

Multivariate methods

A Detrended Correspondence Analysis (DCA) was conducted followed by a Redundancy Analysis (RDA) performed using the entire set of sampled species (referred to as RDA_{TOT}). Subsequently, the Indicator Value Index (IndVal; Dufrene & Legendre, 1997) was employed to assess the bioindication capacity of each species for both organic and conventional hazelnut orchards and to rank the species accordingly. Following this, a Backward RDA Sequence (BRS) was executed, gradually reducing the number of species by excluding those with lower bioindication capabilities in each iteration. Finally, partial RDAs were conducted to differentiate the impact of PPP treatment on the Carabidae community from that attributed to the covariates.

These analyses were conducted using all species (pRDA_{TOT}) and using only those selected through BRS (pRDA_{BRS}) to evaluate the efficacy of the proposed method.

The Detrended Correspondence Analysis and the Redundancy Analysis

A DCA was conducted on the collected biological data to assess whether the dataset exhibited homogeneity or heterogeneity, indicating suitability for linear RDA or unimodal (Canonical Correspondence Analysis, CCA) sorting methods, respectively. Notably, the length of the first axis derived from DCA using ln-transformed biological data was found to be less than 3 standard deviations, prompting the application of linear RDA over unimodal CCA (Lepš & Šmilauer, 2003). RDA is a multivariate method employed to elucidate the relationships between biological communities (comprising taxa composition and abundance) and their environment. This method aims to derive synthetic environmental gradients from ecological datasets. Gradients serve as the foundation for synthetically describing and visualizing the various habitat preferences (niches) of taxa through an ordering diagram. RDA operates under the assumption of a linear relationship in the habitat preferences of taxa.

The indicator value index

As previously mentioned, IndVal was utilized to evaluate the indicator value of all species across both organic and conventional hazelnut orchards. Unlike SIMPER (Clarke, 1993), another method used for assessing bioindication capacity in species assemblages, IndVal solely relies on comparisons of within-species abundance and occurrence, without considering interspecies comparisons. This feature provides an advantage as the value assigned to a species by IndVal remains unaffected by the

abundance of other species (Dufrene & Legendre, 1997; Legendre & Legendre, 1998). The significance of each species' indicator value is determined through a site randomization procedure. IndVal is a straightforward method to find bioindicator species or species assemblages that characterize groups of sites, without relying on hierarchical or non-hierarchical site classification methods.

Indicator species, for the computation of IndVal, are defined as the most characteristic species of each group, predominantly found in a single group of the typology and present in the majority of sites belonging to that group. For each species i in each site group j , the product of A_{ij} , representing the mean abundance of species i in the sites of group j compared to all groups in the study, and B_{ij} , the relative frequency of occurrence of species i in the sites of group j , is computed as follows:

$$A_{ij} = N_{\text{individuals}_{ij}} / N_{\text{individuals}_i}$$

$$B_{ij} = N_{\text{sites}_{ij}} / N_{\text{sites}_j}$$

$$\text{IndVal}_{ij} = A_{ij} * B_{ij}, \times 100$$

Given that study areas were selected in pairs of fields (organic vs. conventional) with similar environmental characteristics and considering the relatively uniform environment of cultivated hazelnut groves, it would be theoretically expected the same species composition and IndVal value for each group of hazelnut orchards, with minor random variations. Therefore, any difference in IndVal exhibited by a species between the two groups of hazelnut orchards could be attributed to chance or to PPP treatment and/or differences in certain covariates. It is reasonable to infer that the greater this difference, the lower the probability that it is attributable to chance and the higher the probability that the species is a reliable bioindicator.

In contrast to the definition of indicator species for IndVal computation, defined as the most characteristic species of each group, found mostly in a single group and present in the majority of the sites belonging to that group (Dufrene & Legendre, 1997), in this study any species that clearly responds to the ecological factor of interest by demonstrating an increase or reduction in abundance and/or distribution is considered a reliable bioindicator. Evaluation of a species' bioindication capacity relies not on its IndVal value within an individual group, but rather on the magnitude of the difference between IndVal values across the two treatment groups, here denoted as DIFF%.

Subsequently, species were sorted based on their DIFF% value, assigning them an ID from 1 (highest DIFF%) to 57 (lowest DIFF%).

The Backward RDA Sequence

A BRS was conducted, gradually reducing the number of species by excluding those with lower DIFF% at each iteration. Concurrently, as the BRS progressed, the percentage of inertia explained by the explanatory environmental variables was computed. Additionally, the permutation test was employed to determine the probability that the explained inertia was due to chance rather than the effect of the environmental variables themselves.

For each step of the BRS, a rank was assigned to both the percentage of inertia explained (with an increasing rank reflecting the progressive increase in inertia explained by the environmental variables) and the permutation probability (with a decreasing rank reflecting the progressive reduction of statistical significance). The sum of these ranks (referred to as ranks sum or RS) served as an indicator of the quality of the result obtained in the corresponding RDA step. The species with the highest RS was deemed the first with minimum but sufficient bioindication capacity, and species with DIFF% greater than or equal to it were considered optimal bioindicators for the treatment and sample under examination. The RDA conducted using only the bioindicator species selected through BRS

was denoted as RDA_{BRS} . Subsequently, the results of RDA_{TOT} and RDA_{BRS} were compared based on inertia explained, overall test, and Procrustes analysis (Jackson, 1995). The PROTEST test (PROcrustean Randomization TEST: Jackson, 1995) was then applied to the result of the Procrustes analysis on the two different ordinations (Andreella *et al.*, 2023; Poos & Jackson, 2012; Digby & Kempton, 1987; Forcino *et al.*, 2015).

The partial RDA

Finally, to validate the effectiveness of the proposed method, partial RDAs were conducted both on the entire set of species ($pRDA_{TOT}$) and solely on those selected using the BRS ($pRDA_{BRS}$). Subsequently, the results were compared to assess the impact on the carabid community attributable to the treatment (PPI and PNF) versus that attributable to the covariates (X10UP and X0107UP).

Results

Biological sample

The biological sample comprises 10,565 individuals from 57 identified species of ground beetles. Only one trapped individual identified at the gender level was excluded from the sample and further analyses. A total of 215 valid sampling sessions were conducted, with an overall sampling effort of 1,855 trap days over a four-year period (2015-2016 and 2018-2019). On average, 1.73 sampling days with 8.63 operational pitfall traps per month were conducted for each cultivated hazelnut grove between April and October (D'Antoni *et al.*, 2020) (Fig. 1).

Three species (*Calathus fuscipes*, *Nebria brevicollis*, *Pterostichus melas*) collectively account for over 62% of the entire standardized catch on 100 pitfall traps, while 43 out of 58 species each represent less than 1% of the total catch.

Results of RDA_{TOT}

The results of RDA_{TOT} applied to all sampled species are illustrated in Fig. 2. Both this graph and Fig. 3 are structured with "type 2" scaling, wherein the angles between variables signify their correlation, and incorporate fitted site scores, whereby site scores are represented as linear combinations of the environmental variables. Points representing sites with the same type of treatment are visually enclosed within their respective colored convex hulls (the smallest convex polygon containing a given set), facilitating the assessment of their dispersion in the two-dimensional space defined by the first two principal axes. The centers of gravity of the hulls corresponding to the two treatments are denoted by the acronyms CON and ORG in red. Closer proximity of sites (and smaller convex hulls) indicates greater similarity in the structure and composition of the sampled Carabidae communities therein.

The plot generated from the RDA_{TOT} using all species sampled demonstrates a slight distinction between the convex hulls of organic and conventional hazelnut orchards, suggesting a moderate divergence in Carabidae communities between the two sets of sites with distinct treatments. However, the canonical axes account for the majority of the variability (51.72%). The validity of the analysis was confirmed by a permutation test, revealing a statistically significant fit, thereby affirming the efficacy of the selected environmental variables in elucidating the variability in the bioindicator community (Tab. 1).

The two target explanatory variables linked to the use of PPP (PPI and PNF) exhibit stronger correlations with the first of the two main axes, while the two derived covariates (X10UP and X0107UP) show stronger correlations with the second axis. Thus, in comparison to the derived

covariates, PPI and PNF exert a greater influence in delineating the environmental gradient along which species are primarily distributed. Conversely, they make a comparatively minor contribution to the second gradient. None of the four RDA canonical axes achieves statistical significance in the permutation test, as indicated in Tab. 1; this suggests that none of the four gradients derived from the linear combinations of the explanatory variables precisely fit the abundance distribution of all observed carabid species.

Results of IndVal application and BRS

All species have been arranged in ascending order based on the increasing DIFF% value (refer to Tab. 2) to facilitate the execution of a BRS. During each iteration of the BRS, the species with the lowest DIFF% were systematically excluded. As delineated in § 2.4.3, at each step of the BRS, both the percentage of inertia explained by the environmental variables (INERTIA rank) and the permutation probability (statistical significance) (P-PERM rank) were assigned rank values. The sum of these ranks (RS) was regarded as an indicator of the quality of the outcome obtained in the corresponding RDA step.

BRS statistics are presented in Tab. 2. The optimal combination (RS=92) of inertia explained by the explanatory variables and statistical significance of the overall test (lower permutation probability) has been identified, utilizing the RS value, corresponding with the species *Notiophilus rufipes* (the tenth species ranked by decreasing DIFF% values). As detailed in the Materials and methods section (§4.3), the species with the highest RS value was designated as the initial one possessing sufficient bioindication capacity for subsequent analyses (RDA_{BRS}). Furthermore, species with a DIFF% greater than that were deemed effective bioindicators for the treatment and the sample examined (*Notiophilus rufipes*, *Amara aenea*, *Leistus fulvibarbis*, *Notiophilus substriatus*, *Brachinus sclopeta*, *Harpalus distinguendus*, *Nebria brevicollis*, *Pterostichus melas*, *Carabus rossii*, *Carabus convexus*). These ten species (highlighted in bold in Tab. 2) were selected for RDA_{BRS}.

The graph resulting from RDA_{BRS} (Fig. 3) exhibits a notable separation of the convex hulls representing organic and conventional hazelnut orchards. This indicates a significant differentiation in the Carabidae communities between the two groups of sampled sites subjected to different treatments. The canonical axes explain a substantial portion of the variance (57.19%).

The statistical significance of the solution generated by the analysis, assessed by the permutation test, confirms the efficacy of the selected environmental variables in elucidating the variability within the bioindicator community, even when considering a reduced number of species involved (Tab. 3). The two explanatory variables associated with the first axis are X0107UP and PPI, with the latter exhibiting notably higher correlation with the second axis as well. Regarding the two environmental gradients along which species are distributed, it is noteworthy that PPI significantly influences both the first and, to a greater extent, the second gradient.

It is noteworthy that in the RDA_{BRS} results, the percentage of inertia explained by the first canonical axis surpasses that explained by the sum of the other three axes (29.10 vs 28.09), whereas in RDA_{TOT} results, it is notably lower (19.09% vs. 32.64%). The first canonical axis demonstrates statistical significance in the permutation test ($p < 0.05$), indicating a strong alignment between the corresponding environmental gradient and the abundance distribution of the carabid species selected via BRS (Tab. 3). The PROTEST test reveals significance in the Procrustes analysis between the two RDA variants, with a p-value of 0.072 (*) and a correlation of 0.563 in a symmetric Procrustes rotation.

The outcome of the PROTEST test underscores a significant disparity between the ordination derived from the entire dataset of carabids compared to that obtained solely from the subset comprising the top 10 indicator species. Although the correlation between the two ordinations remains consistent as expected, the difference between them approaches statistical significance.

Results of partial RDA

Tab. 4 presents the results of the partial RDAs conducted on both the entire species dataset and the subset selected by BRS. A comparative analysis of the outcomes was performed to assess the effectiveness of the selected subset of species in evaluating the impact of treatment (PPI and PNF) and covariates (X10UP and X0107UP).

The total species sample exhibits a considerably higher total variance compared to the reduced sample (40.012 vs. 15.210). Notably, the percentage of variance explained solely by the covariates (Conditional Inertia) is nearly identical in both approaches (32.04% vs. 31.42%). However, the portion of variability attributed to unknown factors and unexplainable by the considered variables (Unconstrained Inertia) diminishes with the BRS approach (48.28% vs. 42.81%), while the portion of inertia solely explained by the explanatory variables linked to PPP usage (Constrained Inertia) markedly increases (19.68% vs. 25.77%).

The RDA obtained by removing the effect of the covariates is not statistically significant when using all the species in the sample, while it is statistically significant when using only the species selected via BRS ($P < 0.05$ *).

Discussion

Effective bioindicators play a crucial role in assessing the impacts of anthropogenic disturbances and guiding conservation efforts (Heink & Kowarik, 2010; Müller *et al.*, 2000). By accurately monitoring changes in biodiversity, policymakers and land managers can make informed decisions to mitigate environmental degradation and promote sustainable practices. This applies to both natural environments and contexts already heavily modified by humans and subject to constant pressures, such as agroecosystems.

However, as observed by Büchs *et al.* (1997) and Holland & Luff (2000), the utilization of carabid species assemblages as bioindicators of different farming systems, employing all encountered carabid species indiscriminately yields weak evidence of the impact of PPP, resulting in a relatively low bioindication capacity attributed to the taxocenosis regarding this pressure. Conversely, evidence provided through the BRS species selection process can offer clearer and more significant indications of the effects caused by PPP, attributing a high bioindication capacity to the 10 species selected.

The use of this subset of species, compared to the utilization of all sampled species, brings about several enhancements that, from a statistical point of view, can be described as follows:

- increase in the percentage of variability explained by the effect of PPP (Constrained Inertia), while the percentage of variability induced by covariates (Conditional Inertia) remains nearly unchanged;
- reduction in variability due to unknown factors (Unconstrained Inertia);
- optimization of the environmental gradients represented by the main axes, with the first axis capable of elucidating most of the canonical variability;
- enhancement in the statistical significance of the permutation tests;
- improvement in the outcome of the PROTEST test.

Therefore, the statistical results show that the proposed method reduces both the influence of anthropogenic and environmental covariates and unknown factors, better highlighting the effect of the selected pressure.

The species selected to form the subset considered for the analyses show the characteristics required to be considered good bioindicators (Noss, 1990; Rüdiger *et al.*, 2012; Battisti & Zullo, 2019), as: (i) they are taxonomically stable; (ii) most of them are common, widespread, relatively easy to detect and to sample with a minimum field and economic effort; (iii) they are sensitive to a specific variable (in our case PPP); (iv) allow changes to be monitored at a small-scale spatial resolution; (v) universally applicable and spatially comparable, and (vi) the interpretation and communication of the results obtained from the indicators can be easily carried out by means of simple, quantitative, and reproducible uni-variate metrics.

The method described in this paper, when applied within a DPSIR framework (Smeets & Weterings, 1999), allows Impact indicators to be identified, since the pressures considered alter the presence and abundance of certain species and therefore the structure of their taxocenosis. From an operational point of view, this is an approach that could be applied in contexts where conservation managers deal with specific target-based and/or threat-based projects (e.g. Battisti *et al.*, 2024).

In conclusion, this methodological approach has proven effective in selecting a subset of species with strong bioindication capabilities. It can be applied in all contexts where there is limited knowledge of individual species' ecology and their response to the pressures under consideration. However, to validate the effectiveness of the method, it would be advisable to increase the number of sampling sites, as the BRS requires processing a substantial amount of data to yield more robust results. Future research directions will focus primarily on expanding the number of sampling sites and, subsequently, on studying the biological and ecological aspects that render these 10 species effective indicators of PPP pressure in agro-ecosystems.

Acknowledgments

The Authors would like to acknowledge Marco Mattoccia (University of Rome Tor Vergata), Lorenzo Talarico and Jacopo Lorusso for their contributions to the sampling and data collection process of ground beetles. We wish to thank the anonymous reviewer as well, who provided useful suggestions and comments which largely improved the first draft of the manuscript.

Funding

This work stems from a project supported by the Italian Ministry of the Environment and Energy Security. The Ministry played a significant role in the project design, defining objectives and execution modalities in coordination with ISPRA. However, the analysis and interpretation of data, the writing of the paper and the decision to submit the paper for publication are fully managed by the authors.

Part of the funds received by the Ministry were used by ISPRA to enter into agreements with other Research Entities. Specifically, the data on Carabidae, used for drafting this article, were collected under the agreement signed on 31/05/2018, between ISPRA and the University of Rome Tor Vergata.

Availability of data and materials

Data supporting the results reported in the article can be found at the following hyperlinks:

<https://www.geonode.nnb.isprambiente.it/catalogue/#!/document/974>

<https://www.geonode.nnb.isprambiente.it/catalogue/#!/document/976>

<https://www.geonode.nnb.isprambiente.it/catalogue/#/document/975>

<https://www.geonode.nnb.isprambiente.it/catalogue/#/document/973>

References

- Adis J., 1979 - Problems of interpreting arthropod sampling with pitfall traps. *Zoologischer Anzeiger Jena*, Vol.202, pp. 177–84.
- Andreella A. & Finos L., 2022 - Procrustes analysis for high-dimensional data. *Psychometrika*, Vol.87 No.4, pp. 1422–1438.
- Avgin S.S. & Luff M.L., 2010 - Ground beetles (Coleoptera: Carabidae) as bioindicators of human impact. *Munis Entomology and Zoology*, Vol.5 No.1, pp. 209-215.
- Battisti C., Testi A., Fanelli G., Rastrelli M., Giovacchini P. & Marsili L., 2024 - A Synthetic Framework to Match Concepts and Approaches When Managing Anthropogenic Threats. *Conservation*, Vol.4(3), pp. 395–401. <https://doi.org/10.3390/conservation4030026>
- Battisti C. & Zullo F., 2019 - A recent colonizer bird as indicator of human-induced landscape change: Eurasian collared dove (*Streptopelia decaocto*) in a small Mediterranean island. *Reg Environ Change*, Vol.19, pp. 2113–2121. <https://doi.org/10.1007/s10113-019-01547-y>
- Billeter R., Liira J., Bailey D., Bugter R., Arens P., Augenstein I., Aviron S., Baudry J., Bukacek R., Burel F., Cerny M., De Blust G., De Cock R., Diekötter T., Dietz H., Dirksen J., Dormann C., Durka W., Frenzel M., Hamersky R., Hendrickx F., Herzog F., Klotz S., Koolstra B., Lausch A., Le Coeur D., Maelfait J.P., Opdam P., Roubalova M., Schermann A., Schermann N., Schmidt T., Schweiger O., Smulders M.J.M., Speelmans M., Simova P., Verboom J., Van Wingerden W.K.R.E., Zobel M. & Edwards P.J., 2008 - Indicators for biodiversity in agricultural landscapes: a pan-European study. *Journal of Applied Ecology*, Vol.45, pp. 141-150. <https://doi.org/10.1111/j.1365-2664.2007.01393.x>
- Brown G.R. & Matthews I.M., 2016 - A review of extensive variation in the design of pitfall traps and a proposal for a standard pitfall trap design for monitoring ground-active arthropod biodiversity. *Ecological Evolution*, Vol. 6 No.12, pp. 3953-64. doi: 10.1002/ece3.2176.
- Buchs W., Harenberg A. & Zimmermann J., 1997 - The invertebrate ecology of farmland as a mirror of the intensity of the impact of man? – an approach to interpreting results of field experiments carried out in different crop management intensities of a sugar beet and an oil seed rape rotation including set-aside. *Biological Agriculture and Horticulture*, Vol.15, pp. 83–108.
- Canterbury G.E., Martin T.E., Petit D.R., Petit L. & Bradford D.E., 2000 - Bird Communities and Habitat as Ecological Indicators of Forest Condition in Regional Monitoring. *Conservation Biology*, Vol.14 No. 2, pp. 544-558.
- Clarke K.R., 1993 - Non-parametric multivariate analyses of changes in community structure. *Austral Ecology*, Vol.18 No.1, pp. 117-143. doi: 10.1111/j.1442-9993.1993.tb00438.x
- D’Antoni S., Bonelli S., Gori M., Macchio S., Maggi C., Nazzini L., Onorati F., Rivella E. & Vercelli M., 2020 - La sperimentazione dell’efficacia delle Misure del Piano d’Azione Nazionale per l’uso sostenibile dei prodotti fitosanitari (PAN) per la tutela della biodiversità. *ISPRA*, Rapporto 330/2020.
- Digby P.G.N. & Kempton R.A., 1987 - Multivariate Analysis of Ecological Communities. *Chapman and Hall*, London.
- Dufrene M. & Legendre P., 1997 - Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecological Monographs*, Vol. 67, pp. 345-366.
- Eyre M.D., Lu M.L. & Rushton S.P., 1990 - The ground beetle (Coleoptera, Carabidae) fauna of intensively managed agricultural grasslands in northern England and southern Scotland. *Pedobiologia*, Vol. 34, pp. 11-18.

- Forcino F.L., Leighton L.R., Twerdy P. & Cahill J.F., 2015 - Reexamining Sample Size Requirements for Multivariate, Abundance-Based Community Research: When Resources are Limited, the Research Does Not Have to Be. *PLoS ONE* Vol.10 No.6, e0128379.
<https://doi.org/10.1371/journal.pone.0128379>
- Heink U. & Kowarik I., 2010 - What are indicators? On the definition of indicators in ecology and environmental planning. *Ecol Indic* 10: 584–593. doi.org/10.1016/j.ecolind.2009.09.009
- Holland J.M. & Luff M.L., 2000 - The effects of agricultural practices on Carabidae in temperate agroecosystems. *Integrated Pest Management Reviews*, Vol.5, pp.109–129.
- Jackson D.A., 1995 - PROTEST: a Procrustean randomization test of community environment concordance. *Ecoscience*, Vol. 2, pp. 297–303.
- Krebs C.J., 1994 - Ecology: The Experimental Analysis of Distribution and Abundance. *Harper Collins College Publishers*, New York.
- Legendre P. & Legendre L., 1998 - Numerical Ecology (2nd ed). *Elsevier Science*, Amsterdam.
- Lepš J. & Šmilauer P., 2003 - Multivariate Analysis of Ecological Data using CANOCO. *Cambridge University Press*, New York.
- Loreau M., 1994 - Ground beetles in a changing environment: determinants of species diversity and community assembly. In: Biodiversity, Temperate Ecosystems and Global Change. Boyle T.J. & Boyle C.E.B. (Eds.). *Springer*, Berlin, Heidelberg, pp. 77-98.
- Lövei G. & Sunderland K.D., 1996 - Ecology and behavior of ground beetles (Coleoptera: Carabidae). *Annual Review of Entomology*, Vol.41, pp.231-256.
- Luff M.L., Eyre M.D. & Rushton S.P., 1992 - Classification and prediction of grassland habitats using ground beetles (Coleoptera, Carabidae). *Journal of Environmental Management*, Vol.35, pp. 301-315.
- Macchio S., Vercelli M., Gori M., Nazzini L., Bellucci V., Bianco P.M., Jacomini C., Rivella E. & D'Antoni S., 2024 - A standardized method for estimating environmental and agronomic covariates to discriminate the explanatory variables effects on bioindicators: a case study on soil fauna. *International Journal of Sustainable Agricultural Management and Informatics*, Vol.10 No.1, pp. 1-26.
- Magura T., Tothmeresz B. & Bordan Zs., 2000 - Effects of nature management practice on carabid assemblages (Coleoptera: Carabidae) in a non-native plantation. *Biological Conservation*, Vol.93, pp. 95-102.
- Marchetti M. (ed), 2004 - Monitoring and indicators of forest biodiversity in Europe - From ideas to operationality. EFI Proceedings No 51. *European Forest Institute Series*, pp.526.
- Mateos E., 2016 - La fauna del sòl i les xarxes tròfiques edàfiques. *L'Atzavara*, Vol.26, pp. 15-24.
- Müller F., Hoffmann-Kroll R. & Wiggering H., 2000 - Indicating ecosystem integrity-theoretical concepts and environmental requirements. *Ecol Model* 130(1–3):13–23. [doi.org/10.1016/S0304-3800\(00\)00210-6](https://doi.org/10.1016/S0304-3800(00)00210-6)
- Niemelä J., 1990 - Effect of changes in the habitat on carabid assemblages in a wooded meadow on the Åland Islands. *Notulae Entomologicae*, Vol.69, pp. 169-174.
- Niemelä J., Langor D. & Spence J.R., 1993 - Effects of clear-cut harvesting on boreal ground-beetle assemblages (Coleoptera: Carabidae) in western Canada. *Conservation Biology*, Vol.7, pp. 55-561.
- Noss R.F., 1990 - Indicators for Monitoring Biodiversity: A Hierarchical Approach. *Conservation Biology*, Vol.4, pp. 355-364. <https://doi.org/10.1111/j.1523-1739.1990.tb00309.x>
- Poos M.S. & Jackson D.A., 2012 - Addressing the removal of rare species in multivariate bioassessments: The impact of methodological choices. *Ecological Indicators*, Vol.18, pp. 82–90.

- Rüdisser J., Tasser E. & Tappeiner U., 2012 - *Distance to nature* - A new biodiversity relevant environmental indicator set at the landscape level. *Ecological Indicators*, Vol.15(1), pp. 208-216. <https://doi.org/10.1016/j.ecolind.2011.09.027>
- Rushton S.P., Eyre M.D. & Lu M.L., 1990 - The effects of management on the occurrence of some carabid species in grassland. In: The Role of Ground Beetles in Ecological and Environmental Studies. Stork N.E. (Ed.), *Intercept Publications*, Andover.
- Russo D., Salinas-Ramos V., Cistrone L., Smeraldo S., Bosso L. & Ancillotto L., 2021 - Do We Need to Use Bats as Bioindicators? *Biology*, Vol.10, No.8, p. 693. doi:[10.3390/biology10080693](https://doi.org/10.3390/biology10080693)
- Smeets E. & Weterings R., 1999 - Environmental indicators: Typology and overview. Technical report No 25, *European Environment Agency*, Copenhagen.
- Southwood T.R.E., 1978 - Ecological Methods with particular reference to insect populations, 2nd Ed. *Chapman and Hall*, Cambridge.
- Vigna Taglianti A., 2005 - Checklist e corotipi delle specie di Carabidae della fauna italiana. In: I Coleotteri Carabidi per la valutazione ambientale e la conservazione della biodiversità. Brandmayr P., Zetto T., Pizzolotto R. (Eds.). *APAT, Manuali e Linee Guida* 34/2005, pp. 186–225.
- Woodcock B.A., 2005 - Pitfall trapping in ecological studies. In: Insect sampling in forest ecosystems. Leather S. (ed.). *Blackwell Publishing*, Oxford, pp. 37–57. <https://doi.org/10.1002/9780470750513.ch3>

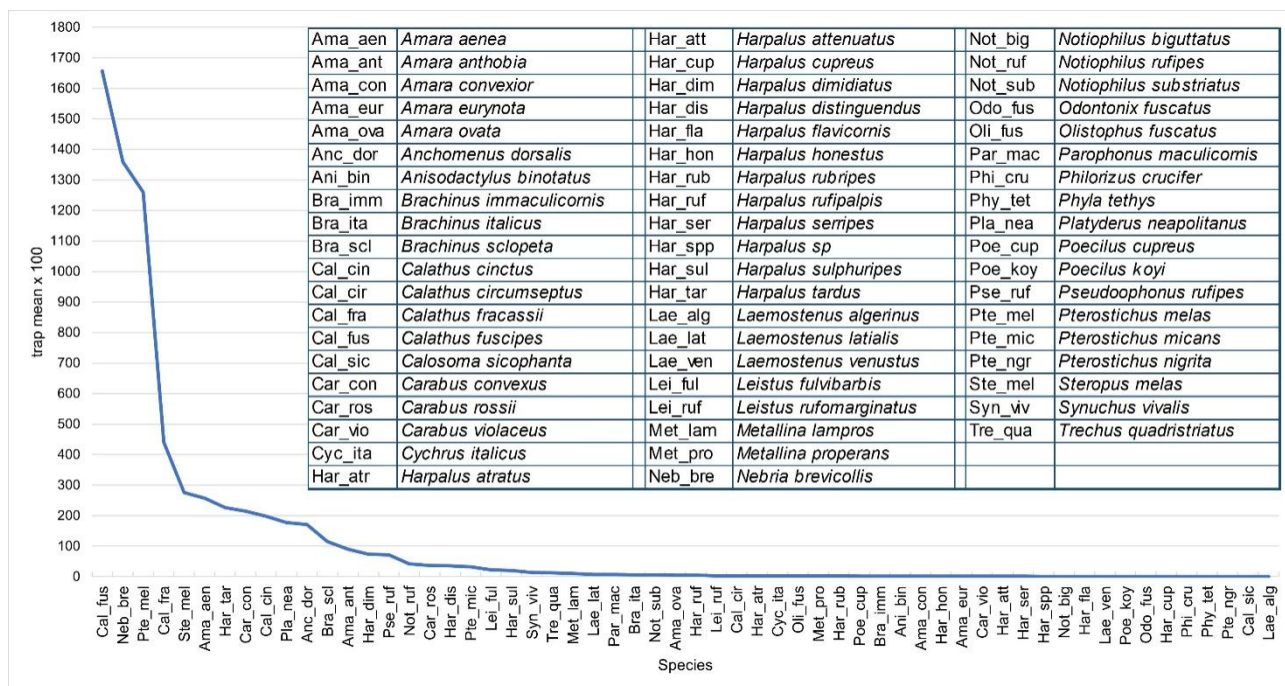


Fig. 1 – Structure of the overall biological sample with the abbreviations of the scientific names of the species. / Struttura dell'intero campione biologico con indicate le abbreviazioni dei nomi scientifici delle specie.

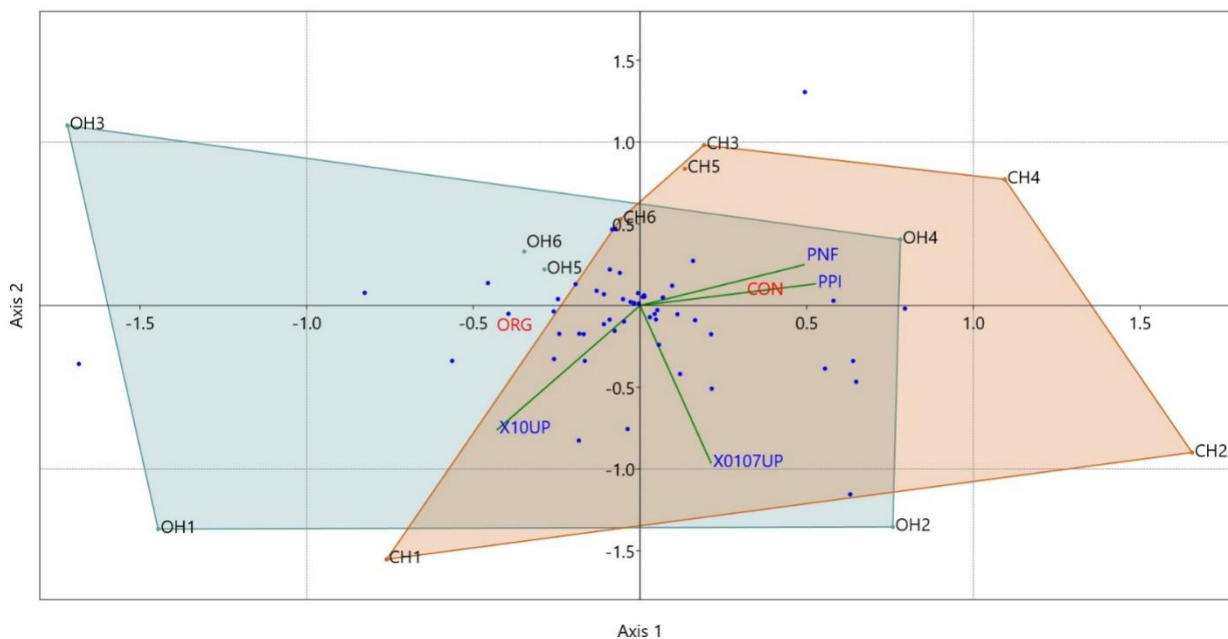
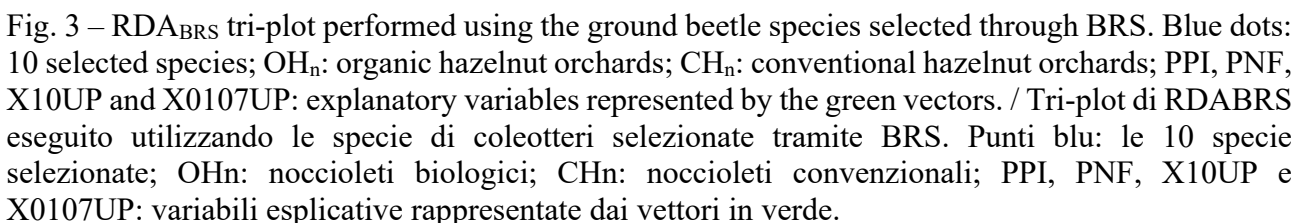


Fig. 2 – RDA_{TOT} tri-plot using all the sampled species of ground beetles. Blue dots: 57 sampled species; OH_n: organic hazelnut orchards; CH_n: conventional hazelnut orchards; PPI, PNF, X10UP and X0107UP: explanatory variables represented by the green vectors. / Tri-plot della RDA_{TOT} effettuato utilizzando tutte le specie di carabidi campionate. Punti blu: 57 specie campionate; OH_n: noccioleti biologici; CH_n: noccioleti convenzionali; PPI, PNF, X10UP e X0107UP: variabili esplicative rappresentate dai vettori in verde.



Sum eigenvalue canonical axes = 20.715						Overall test	
Axis	Eigenvalue	Axis inertia explained %	Cumulative inertia explained %	Taxa - environment correlations (R)	Axis p-permutation test significance	R ²	0.517
RDA1	7.645	19.09	19.09	0.848	0.107 N.S	R ² _{adj}	0.241
RDA2	6.636	16.57	35.66	0.903	0.084 (*)	F	1.875
RDA3	4.517	11.28	46.93	0.887	0.307 N.S	p-permutation (n=999)	0.006 **
RDA4	1.917	4.79	51.72	0.824	0.747 N.S		
Sum eigenvalue residual axes = 19.338							

Tab. 2 – BRS statistics. IndVal values for each species and treatment; performance indicators ranks of the BRS steps: inertia explained and permutation test probability. All species have been listed based on the increasing DIFF% value. The 10 species of carabids selected by BRS are in bold. / Statistiche della BRS. Valori di IndVal per ciascuna specie e per ciascun trattamento; ranghi di valore per gli indicatori di prestazione per ogni step della BRS: inerzia spiegata e probabilità del test di permutazione. Tutte le specie sono state elencate in base al valore DIFF% crescente. Le 10 specie di carabidi selezionate dalla BRS sono riportate in grassetto.

<i>species</i>	INDVAL				Backward Rda Sequence				
	<i>ORG</i>	<i>CON</i>	<i>DIFF%</i>	Species_ID	INERTIA explained by environmental variables (%)	P-PERM probability by permutation	INERTIA rank	P- PERM rank	RS ranks sum
Car_vio	8.333	8.333	0.00	57	51.72	0.005	3	46	49
Har_att	8.333	8.333	0.00	56	51.79	0.009	4	14	18
Ama_eur	8.333	8.333	0.00	55	51.88	0.007	5	33	38
Oli_fus	16.670	16.670	0.00	54	51.92	0.006	6	37	43
Har_dim	33.780	32.890	0.89	53	51.99	0.009	7	14	21
Syn_viv	11.900	9.524	2.38	52	52.42	0.010	11	11	22
Har_sul	23.810	26.190	2.38	51	52.38	0.012	8	7	15
Har_ruf	10.000	6.667	3.33	50	52.67	0.009	13	14	27
Har_atr	5.556	11.110	5.55	49	52.42	0.009	10	14	24
Lei_ruf	8.333	16.670	8.34	48	52.38	0.005	9	46	55
Pte_mic	22.920	32.810	9.89	47	52.42	0.005	12	46	58
Cal_cin	31.310	45.790	14.48	46	53.45	0.002	14	55	69
Pse_ruf	38.810	53.420	14.61	45	55.59	0.004	40	51	91
Pla_nea	41.670	58.330	16.66	44	54.88	0.001	19	56	75
Har_rub	22.220	5.556	16.66	43	54.88	0.006	20	37	57
Tre_qua	38.100	21.430	16.67	42	54.89	0.009	21	14	35
Ama_con	0.000	16.670	16.67	41	54.93	0.008	24	28	52
Ani_bin	0.000	16.670	16.67	40	54.99	0.004	27	51	78
Bra_imm	16.670	0.000	16.67	39	54.90	0.009	23	14	37
Bra_ita	16.670	0.000	16.67	38	55.00	0.006	29	37	66
Cal_sic	0.000	16.670	16.67	37	55.34	0.008	32	28	60
Har_cup	16.670	0.000	16.67	36	55.36	0.007	33	33	66
Har fla	16.670	0.000	16.67	35	55.41	0.003	36	54	90
Har_ser	16.670	0.000	16.67	34	55.41	0.009	37	14	51
Lae_alg	0.000	16.670	16.67	33	55.38	0.006	34	37	71
Lae_ven	0.000	16.670	16.67	32	55.41	0.004	35	51	86
Met_lam	0.000	16.670	16.67	31	55.46	0.007	38	33	71
Met_pro	16.670	0.000	16.67	30	55.00	0.009	28	14	42
Not_big	0.000	16.670	16.67	29	54.85	0.008	17	28	45
Odo_fus	0.000	16.670	16.67	28	54.81	0.012	16	7	23
Phi_cru	16.670	0.000	16.67	27	54.85	0.006	18	37	55
Phy_tet	16.670	0.000	16.67	26	54.89	0.009	22	14	36
Poe_koy	0.000	16.670	16.67	25	54.94	0.010	25	11	36
Pte_ngr	16.670	0.000	16.67	24	54.99	0.009	26	14	40
Har_tar	59.110	40.890	18.22	23	55.03	0.011	30	9	39
Ama_ant	38.890	61.110	22.22	22	55.58	0.009	39	14	53
Lae_lat	29.170	2.083	27.09	21	56.38	0.006	47	37	84
Par_mac	37.500	8.333	29.17	20	56.22	0.005	44	46	90
Cal_fra	32.120	0.606	31.51	19	56.28	0.009	45	14	59
Anc_dor	46.780	14.910	31.87	18	53.78	0.010	15	11	26
Cyc_ita	33.330	0.000	33.33	17	55.91	0.005	41	46	87
Cal_cir	33.330	0.000	33.33	16	55.91	0.008	42	28	70
Ama_ova	11.110	44.440	33.33	15	56.14	0.011	43	9	52

Har_hon	33.330	0.000	33.33	14	56.47	0.006	48	37	85
Poe_cup	33.330	0.000	33.33	13	56.50	0.009	49	14	63
Ste_mel	8.333	41.670	33.34	12	56.55	0.007	50	33	83
Cal_fus	67.830	32.170	35.66	11	55.21	0.006	31	37	68
Not_ruf	17.440	54.260	36.82	10	57.19	0.006	55	37	92
Ama_aen	69.530	30.470	39.06	9	56.86	0.008	53	28	81
Lei_ful	21.740	61.590	39.85	8	56.31	0.015	46	5	51
Not_sub	4.762	47.620	42.86	7	56.83	0.013	52	6	58
Bra_scl	54.600	9.052	45.55	6	57.66	0.009	56	14	70
Har_dis	13.510	59.460	45.95	5	57.06	0.042	54	4	58
Neb_bre	25.390	74.610	49.22	4	56.59	0.045	51	3	54
Pte_mel	23.990	76.010	52.02	3	45.87	0.253	2	2	4
Car_ros	71.790	18.800	52.99	2	40.43	0.366	1	1	2
Car_con	65.730	0.467	65.26	1	*	*	*	*	*

(*) Values are not calculable, as the RDA cannot be performed on one species.

Tab. 3 – RDA_{BRS} statistics and overall permutation test to assess the statistical significance of the environmental variables effect. / Statistiche della RDABRS e test di permutazione complessiva per valutare la significatività statistica dell'effetto delle variabili ambientali.

Sum eigenvalue canonical axes = 8.698						Overall test	
Axis	Eigenvalue	Axis inertia explained %	Cumulative inertia explained %	taxa - environment correlations (R)	Axis p-permutation test significance	R ²	0.572
RDA1	4.425	29.10	29.10	0.924	0.011 *	R ² _{adj}	0.327
RDA2	2.429	15.97	45.06	0.863	0.145 N.S	F	2.338
RDA3	1.249	8.21	53.28	0.661	0.510 N.S	p-permutation (n=999)	0.006 **
RDA4	0.595	3.91	57.19	0.715	0.688 N.S		
Sum eigenvalue residual axes = 6.512							

Tab. 4 – Statistics and overall permutation tests comparison between partial RDA performed with all species and with the 10 species selected by BRS. / Confronto tra statistiche e test di permutazione complessiva ottenuti eseguendo le RDA parziali con tutte le specie e RDA parziali con le 10 specie selezionate da BRS.

PARTIAL RDA							
	all species			only species selected by BRS			
Inertia	(n)	(%)		(n)	(%)		
Total	40.012	100.00		15.210	100.00		
Conditional	12.821	32.04		4.779	31.42		
Constrained	7.873	19.68		3.919	25.77		
Unconstrained	19.318	48.28		6.512	42.81		
<i>p-permutation test significance</i>	0.157 N.S.			0.022*			
<i>Eigenvalues for constrained axes</i>	RDA1	RDA2		RDA1		RDA2	
	5.112	2.761		2.882		1.038	
	PC1	PC2	PC3	PC4	PC5	PC6	PC7
<i>Eigenvalues for unconstrained axes for all species</i>	6.912	4.466	2.903	1.820	1.539	1.066	0.612
<i>Eigenvalues for unconstrained axes only for species selected by BRS</i>	2.525	1.624	1.092	0.532	0.433	0.276	0.030