

RESEARCH

Open Access



Exclusion of tourist species from assemblages in ecological studies: a methodological approach using spiders

María Florencia Nadal^{1,2*} , Alda González^{1,3} and Gilberto Avalos²

Abstract

Background: The exclusion of tourist species from samples is important to avoid bias in community analyses. However, in practice, this is a very difficult task. The commonly used methods by researchers, when the habitat of the species is not known, have several shortcomings: first, they exclude not only pseudo-rare species but also genuinely rare species; second, the results obtained with those methods depend on the abundance of the sampling; and third, they follow very subjective rules. The aims of this study were: (i) to propose a methodology to detect and exclude habitat-tourist species from the database used to carry out analyses in community ecology studies, (ii) to evaluate how the presence of habitat-tourist species affects the richness estimates, and (iii) to assess the effect of including juvenile spiders in the detection of tourist species and the effect of removing them from the richness estimates.

Results: When the adult + juvenile dataset was considered, twenty-one habitat-tourist species were detected: 8 in forest foliage, 11 in forest leaf litter, and 2 in grassland. When habitat-tourist species were considered with this dataset, richness overestimation was significant in foliage and in leaf litter, and the final slopes of the richness estimation curves were significantly steeper in leaf litter. When only the adult dataset was considered, eight habitat-tourist species were detected: 3 in forest foliage, 4 in forest leaf litter, and just one in grassland. The inclusion of habitat-tourist species in this dataset showed an overestimation of richness, but this was not significant.

Conclusions: The proposed methodology contributes to solving the problem of tourist species, which was recognized as one of the great problems in biodiversity studies. This study showed that common estimators overestimate species richness when habitat-tourist species are included, leading to erroneous conclusions. Besides, this research showed that the inclusion of juveniles (e.g. spiders) could improve the analysis outputs because it allowed the detection of more habitat-tourist species.

Keywords: Pseudo-rare species, Richness estimation, Indicator value, Habitat

Background

Ecosystems are characterized by an exchange of matter and energy (Willis 1997). Since these systems are open, communities of species from a particular biogeographic region or habitat patch within this region can be

expected to interact with communities from surrounding regions or habitats. A kind of interaction occurs due to dispersion, which is one of the factors that regulate the characteristics of a metacommunity or a particular set of communities (Chase et al. 2020). Individuals of species need to colonize new places to ensure the long-term persistence of metapopulations (Hanski 1998). For this, it is expected that some individuals will arrive in habitats or regions that are not suitable for them (Novotný and Bas-set 2000; Archer 2003; Hadidian 2015).

*Correspondence: florencia.nadal@conicet.gov.ar

¹ Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, Argentina
Full list of author information is available at the end of the article

Several species adapted to a particular biogeographic region or habitat, where they have their normal distribution, may occur in other habitats or regions in a very low abundance. These species are known as “tourist”, “transitory”, or “occasional” species. Tourist species (TS) are classified within the pseudo-rare species. The term pseudo-rare species was coined by Rabinowitz (1981) to refer to species that are rare in only one or a few parts of their distribution range. Pseudo-rare species also include those that are assumed to be the product of sampling bias. It is believed that pseudo-rare species appear due to a continuous flow of immigrants, a phenomenon known as the “mass” or “source-sink” effect (Novotný and Basset 2000; Coddington et al. 2009). Furthermore, it is believed that they are not capable of sustaining populations long-term when isolated from the source populations in the primary habitat (Barlow et al. 2010).

Ribeiro and Borges (2010) coined the term “habitat-tourist species” (HTS) to refer to species that are rare in one habitat but abundant and specialists in another. Examples of HTS are certain herbivorous insects that appear on the foliage of a particular plant species only by accident but do not feed on it (Novotný and Basset 2000). So, the HTS is a particular case of TS.

According to Novotný and Basset (2000), excluding TS from the samples is the first logical step of any community analysis. However, no methodology to do it properly has been found. They did not mention the reason for removing these species, but they seem to be the first to make this suggestion. According to Barlow et al. (2010), the main methodological obstacle to excluding TS is to differentiate them from genuinely rare species (i.e. those that have a small geographic range, small habitat range, or low local density) because it is highly context-dependent and requires detailed biological information that remains unavailable for the vast majority of species. This difficulty in separating TS from genuinely rare species was also noted by Archer (2003).

One of the methods used to exclude TS or pseudo-rare species from the data, or their possible effect on the analyses, is to use richness and diversity indices that are not very sensitive to rare species in abundance or frequency of occurrence in the samples (Magurran 2004). Some of these indices are the Simpson diversity index (Simpson 1949), the Abundance-based Coverage Estimator (ACE), and the Incidence-based Coverage Estimator (ICE) (Chao and Yang 1993; Lee and Chao 1994). Another method, which is often applied, consists in depurating the data and removing rare species before carrying out the analyses. For example, Edinger and Risk (2000) exclude species with fewer than three occurrences; Duffy-Anderson et al.

(2006) exclude species that are not present in at least 5% of the samples; Arnold and Lutzoni (2007), Zhan et al. (2014), and Pinto et al. (2021) remove singletons (species represented by a single individual); and Barlow et al. (2010) exclude species with fewer than 10 records. These methods have several shortcomings. First and foremost, removing rare species in abundance (or not considering them in the case of the indexes mentioned above) excludes not only pseudo-rare species but genuinely rare species as well. This implies the risk of losing valuable information on conservation values (Barlow et al. 2010). Second, the results obtained from these methods depend on the abundance of the sampling, and in particular, on the abundance obtained for each species, which can fluctuate from year to year or from season to season (Bonte et al. 2002; Nadal et al. 2018). Third, as noted by Barlow et al. (2010), the methods follow very subjective rules about how many species to remove or where to remove them from (e.g. rare species are removed from the samples or from the entire sample). Archer (2003) and Eggleton et al. (2021) used different techniques to remove TS and HTS, respectively, based on empirical knowledge of the natural history of each species. Although these techniques are more appropriate than those methods mentioned before, they require prior knowledge of the natural history or habitat preferences of all species, and these authors have not suggested a specific technique to assess habitat or region preference when this information is not available.

Some of the ecological indices that may be most affected by the inclusion of HTS are the estimators of richness since several of these are sensitive to rare species in abundance (Magurran 2004). Eggleton et al. (2021) have confirmed this with litter-dwelling beetles and Archer (2003) with wasps and bees. Spiders are a taxon where estimation curves almost never reach an asymptote (Toti et al. 2000; Sørensen 2004; Rico-G et al. 2005). This shortcoming has been attributed to the high species richness that exists and the abundance of sampling that is never enough to complete inventories (Toti et al. 2000). This drawback has also been attributed to the discard of juveniles from the analyses, which leads to the waste of a large amount of information (Jiménez-Valverde and Lobo 2006). However, the effect that TS or HTS could have on richness estimates in this taxon has not been assessed.

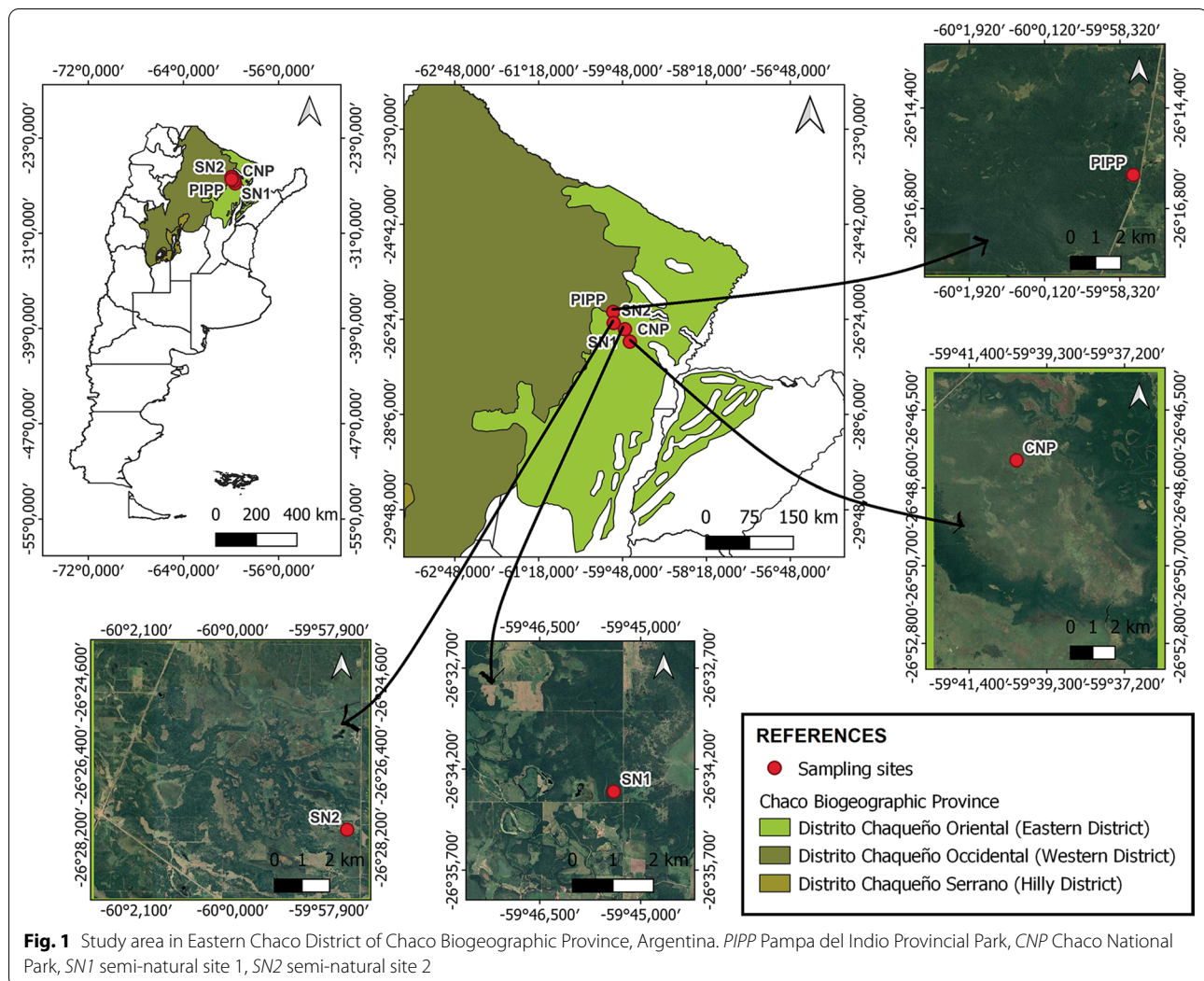
The spiders of the Salticidae family are the most representative family of the Araneae order, with more than 6000 described species (World Spider Catalog 2022). They hunt during the day by jumping on their prey (e.g. without using webs) and occupy a wide variety of habitats

(Cumming and Wesolowska 2004; Foelix 2011; Nadal et al. 2018). Herein, spider assemblage is presented as a case study in order to: (i) propose a methodology to detect and eliminate HTS from the database used to carry out analyses in community ecology studies, particularly analyses of assemblages of a given taxon; (ii) evaluate how the inclusion/exclusion of the HTS affects richness estimates; and (iii) assess the effect of including juveniles of spiders in the detection of HTS and the effect of removing them from the richness estimates. Regarding the second aim, it is expected to find an overestimation of the richness by including HTS, while regarding the third aim, it is expected to find a higher number of HTS, and richness estimates close to the asymptote when juveniles are included.

Materials and methods

Study area

The study area was located in the Chaco Biogeographic Province of Argentina, specifically in the Eastern Chaco District (Arana et al. 2021). Four sites were selected in this area: two protected natural parks and two not protected and semi-natural sites located between the first (Fig. 1). The two protected areas selected were the Chaco National Park (CNP) ($26^{\circ} 50' S$, $59^{\circ} 48' W$), which extends through the Presidencia de la Plaza and the Sargento Cabral departments, and the Pampa del Indio Provincial Park (PIPP) ($26^{\circ} 16' S$, $59^{\circ} 58' W$), located in the Libertador General San Martín department. The two not protected areas were the semi-natural site 1 (SN1) ($26^{\circ} 34' S$, $59^{\circ} 46' W$), located in the Sargento Cabral



department, and the semi-natural site 2 (SN2) (26° 27' S, 59° 59' W), located in the Veinticinco de Mayo department. The map of Fig. 1 was made with QGIS version 3.16.7 (QGIS Development Team 2021) and the satellite images were downloaded from SAS Planet (SAS.Planet Development Team 2019).

The main plant communities of the Eastern Chaco District consisted of xerophytic forests of *Schinopsis balansae*, *Aspidosperma quebracho-blanco*, *Bulnesia* spp., and other species (Cabrera 1971, 1976). However, these have lost quality due to the selective logging of hardwood trees, mainly *S. balansae*, conducted since the nineteenth century (Cabrera 1971; Morello and Rodrigues 2009). Additionally, in the last four decades, the forests have lost coverage due to clearcutting (Morello and Rodrigues 2009). The remaining communities are represented by riparian forests, shrublands, palm groves, grasslands, savannahs, and other plant communities associated with wetlands (Cabrera 1971, 1976; Morello and Rodrigues 2009; Arana et al. 2021). There are pieces of evidence that grasslands have reduced coverage in Chaco Province as a consequence of the introduction of cattle. This has favored shrub encroachment via selective grazing (Grau et al. 2015).

The climate of the Eastern Chaco District according to the Köppen-Geiger climate classification (Peel et al. 2007) is humid subtropical without dry seasons and with hot summers. The mean annual temperature follows a N-S gradient, ranging from approximately 23 °C to 24 °C in the north and 18 °C to 19 °C in the south (Morello et al. 2012). The climate, in general, is more humid than in the other districts, increasing precipitation from west to east. Thus, the mean annual rainfall in the west (the border region with the Western Chaqueño District or Chaco Seco) is around 750 mm yr⁻¹, while in the east it is approximately 1300 mm yr⁻¹ (Cabrera 1971; Morello et al. 2012; Arana et al. 2021).

Fieldwork

The samplings were conducted seasonally: from March 6 to 9 of 2017 (summer), from August 7 to 10 of 2017 (winter), from December 4 to 7 (spring), and from May 28 to 31 of 2018 (autumn). The total sampling comprised 720 samples (4 dates × 4 sites × 3 transects × 5 samples per transect × 3 techniques). The samples were not joint in any case.

The samplings were conducted systematically, covering three types of habitats to achieve the greatest representativeness of the Salticidae species: forest foliage, forest leaf litter, and grassland. In forests and grasslands, three 200-m transects were drawn, separated by a distance of no less than 2 km to increase the representativeness of

spider species. A sample was taken every 50 m from the grassland transects. Two samples were taken every 50 m from the forest transects, one for leaf litter and one for foliage. Each leaf litter sample was taken before the foliage sample to avoid contamination of the leaf litter samples with spiders that might have fallen from foliage. So, the presence of HTS by sampling artifacts is unlikely in this study. The distance of 50 m between samples was established to avoid problems of non-independence of the samples and follows the sampling standards for spiders in northeastern Argentina (Avalos et al. 2018; Nadal et al. 2018; Achitte-Schmutzler et al. 2022).

Three collection techniques were applied in this study. In the forest, epigeal spiders were obtained from leaf litter samples from an area of 1 × 1 m each. The leaf litter samples were passed through a 1 cm mesh opening sieve on a white canvas. Foliage spiders were collected by beating the vegetation, with 15 beatings on the shrubby vegetation and the lower portion of the tree stratum up to a height of 2.5 m, and collecting the material on a 2 × 2 m white canvas. In grasslands, samples were taken by vacuuming, using a G-Vac garden vacuum cleaner (NIWA SNW260, China). For each sample, the vegetation was sucked into the vacuum in an area of 2 × 2 m for a designated time of 1 min. The material collected with the vacuum cleaner was scattered on a white canvas. In the three sampling types, once the spiders were detected on the canvas, they were captured with entomological forceps and placed in bottles with 70% ethyl alcohol. All samples were taken during daytime hours (from 8 AM until 5 PM), when Salticidae spiders are more active. The techniques applied herein are similar to those employed in Avalos et al. (2018); Nadal et al. (2018), and Achitte-Schmutzler et al. (2022).

Laboratory work

Adult and juvenile spiders were identified using the information provided by the genitalia or the habitus of each specimen, respectively. The information provided by the habitus is the general appearance of the specimen, including the shape and color of the body parts. Adult specimens were identified to the genus or species level using the original descriptions available in the World Spider Catalog (2022) and the Metzner Catalog of Salticidae Spiders (Metzner 2022). Juvenile specimens were identified to the genus or species level based on the habitus of adults and a set of supporting evidence or information, as follows: (i) geographical distribution of the potential species, (ii) occurrence of juvenile and adult individuals in the same sample, and (iii) comparison with specimens previously collected in the Eastern Chaco District. The specimens that could not be identified, not even as

morphospecies, were excluded from the analyses, e.g. (i) juveniles at a very early stage of development and juveniles of two *Gastromicans* species that had very similar habitus; and (ii) adult males of the genus *Gastromicans*, which could not be differentiated since their palps (male genitalia) and chelicerae were very similar. Unidentified specimens were presented in the species list as “not determinable” and “*Gastromicans* spp.”, and they represented 8% of the total.

For the identification of the material, different stereoscopic microscopes available at the Biología de los Artrópodos Laboratory of the Facultad de Ciencias Exactas y Naturales y Agrimensura (FACENA), Universidad Nacional del Nordeste (UNNE) were used: Olympus SZ51, Olympus SZ40, Leica ES2 and Leica EZ4. The identified material was deposited in the CARTROUNNE collection (Facultad de Ciencias Exactas y Naturales y Agrimensura, Corrientes, Argentina) and the IBSI-Ara collection (Instituto de Biología Subtropical, Misiones, Argentina).

Data analysis

The following analyses were performed separately with adult and adult + juvenile datasets:

(i) In order to detect HTS, the habitat (or habitats) to which the species should be adapted or be resident (sensu Archer 2003 and Eggleton et al. 2021) was assessed by analyzing the habitat specificity of the species. For this purpose, the IndVal index (Dufrêne and Legendre 1997) was used with the addition of the square root by De Cáceres and Legendre (2009) and the combinations of groups proposed by De Cáceres et al. (2010), through the following formula:

$$\sqrt{\text{IndVal}}^g = \sqrt{A^g \times B} = \sqrt{\frac{a_C^g}{a^g} \times \frac{n_C}{N_C}}$$

$$a_C^g = \frac{N}{K} \sum_{j \in C} (a_i / N_i)$$

$$a^g = \frac{N}{K} \sum_{j \in K} (a_i / N_i)$$

where $\sqrt{\text{IndVal}}^g$ is the indicator value of a species in a group (in this paper habitat) of sites (in this paper samples), or indicator value of a species in a combination of groups; A^g is the calculation of specificity with combination of groups, that is, calculation of the average abundance of a species in the sites of a group or a combination of groups compared to all groups in the study; B is the calculation of fidelity, that is, the relative frequency of occurrence of a species in the sites of a group; a is the sum of abundances of the target species over all sites; a_i is

the sum of abundances of the target species over all sites in site group i ; C is the a set of c site groups conforming a particular group of sites; K is the set of all k site groups; n is the number of sites where the target species occurs; N is the number of sites in the data set; and N_i is the number of sites belonging to site group i . A^g is maximum (1 or 100%) when a species is only present in a group or only in a group association. B is maximum when a species is present at all sites in a group. The IndVal index is maximum when individuals of a species are observed at all sites of a single group or at all sites of a group association.

The IndVal index was evaluated with the R version 4.0.2 program using the *indicspecies* and *stats* packages, with a significance level of 0.05 and 999 permutations (R Core Team 2020). The IndVal index is normally used to evaluate the preference of a species for a characteristic (Dufrêne and Legendre 1997; Cáceres and Legendre 2009), so that species can serve as an indicator of that condition if they have a high value of IndVal (e.g. > 0.70). For this purpose, one characteristic often considered is habitat disturbance (Larrivée et al. 2008). Thus, if a disturbance indicator species is found at a place, that place is probably disturbed. The combination of groups is an extension of the original index but its use is optional.

In this study, we have not looked for the species to be indicators, but only to know the specificity they have for habitats. Therefore, here species adapted to a certain habitat (or combination of habitats) were those that showed a significant IndVal value ($p < 0.05$) (hereafter termed IndVal-significant species) and a specificity value (component A of the IndVal index) greater than 0.75. Only the specificity component was considered for two reasons: (1) The use of the fidelity component in the IndVal Index only makes sense when species will be used as indicators in the field. This is because species have to be easily found in the group in which they are indicators. In this sense, if the only aim is to know the specificity of the habitat of a particular species, the usefulness of this fidelity component does not make sense. (2) It is difficult for all species related to a particular habitat to have high fidelity values (be homogeneously distributed and occur in most samples), which is a requirement to achieve high fidelity (see formula above). Therefore, excluding the fidelity component allows the proposed methodology to be used with species without a homogeneous distribution (e.g. grouped distribution). The specificity component of the IndVal index is similar to the Species Specialization Index (SSI) (Julliard et al. 2006) as they account for the degree of habitat specialization. However, the SSI differs from the specificity component of IndVal in that it computes the variation in average densities across habitat classes via standard deviation. Although SSI is a powerful index, the disadvantage of using SSI instead

of the specificity component of IndVal is that it does not vary from zero to one. Therefore, a standard umbral value of high specialization, as herein is established with a value of 0.75, cannot be established with SSI.

(ii) Those species that were adapted to another habitat (e.g. species considered residents of another habitat) were removed from the target habitat data, on the assumption that the species that occurred in the habitat to which they are not adapted were HTS.

(iii) Once these data were depurated, richness estimation analyses were carried out with the Estimates software version 9.1.0 (Colwell 2013) in order to show how the presence of HTS affected the richness estimates. The estimation of species richness was made for each of the four sites, discriminating by type of habitat: foliage, leaf litter, and grassland. To assess the effect of HTS, this analysis was performed both by including and excluding HTS. The following non-parametric estimators were used: ACE (Chao and Yang 1993; Lee and Chao 1994; Chazdon et al. 1988; Magurran 2004; Gotelli and Colwell 2011), ICE (Chao and Yang 1993; Lee and Chao 1994; Chazdon et al. 1988; Gotelli and Colwell 2011), Jackknife1 (Burnham and Overton 1978, 1979; Heltshe and Forrester 1983), Jackknife2 (Burnham and Overton 1978; Smith and Belle 1984; Gotelli and Colwell 2011), Chao1 (Chao 1984, 1987; Magurran 2004; Gotelli and Colwell 2011), and Chao2 (Colwell and Coddington 1994; Gotelli and Colwell 2011). The last two estimators were used in their bias-corrected versions. ACE and Chao1 are based on abundance, while ICE, Jackknife1, Jackknife2, and Chao2 are based on presence-absence. The difference between observed and estimated richness was also calculated and the values obtained were compared with a parametric (*t*-test) or a nonparametric (Kruskal–Wallis) test depending on whether the distribution of the data subset had a normal or nonparametric distribution, respectively. The distribution of the data was evaluated with the Shapiro–Wilk test. In addition to the estimated and observed richness values, estimation curves were plotted. These curves were expressed as a function of the number of samples. The values of the indices and the observed richness used to plot the curves were obtained from the mean of 100 randomizations. To compare the curves more rigorously, the value of the final slope (between the last 10 samples of the abscissa) of each curve was calculated. This calculation was used in other works (e.g. Hortal and Lobo 2005). The values of the slopes were compared using the same tests mentioned above.

Results

Association of the species with the assessed habitats

A total of 1,697 spiders from the Salticidae family were collected, 27% of them in the adult stage. Of the 69

species identified, 59 were adults (Additional file 1: Table S1). Considering adults + juveniles, IndVal index showed the three groups tested and a combination: foliage, leaf litter, grassland, and leaf litter + grassland. Within these groups, the analysis showed 41 IndVal-significant species: 17 for foliage, two for leaf litter, 21 for grassland, and one for leaf litter + grassland. Considering adults, IndVal index showed the three groups tested: foliage, leaf litter, and grassland groups. Within these groups, the analysis showed 27 IndVal-significant species: 11 for foliage, two for leaf litter, and 14 for grassland (Table 1).

Detection of HTS

When adults + juveniles were considered, eight HTS were detected in foliage, 11 in leaf litter, and two in grassland. In contrast, when only adults were considered, three HTS were detected in foliage, four in leaf litter, and one in grassland (Table 2).

Estimation of species richness with and without HTS

In all cases where HTS were present, there was an overestimation of species richness. Analysis of the adult + juvenile dataset revealed an overestimation of species richness, mainly in leaf litter, which was maximum in SN1 (67%), followed by SN2 (52%), CNP (50%), and finally PIPP (45%). In foliage, the overestimation of richness was maximum in SN2 (28%), followed by PIPP (24%) and SN1 (22%), while in grassland the maximum overestimation occurred in PIPP (13%), followed by CNP (6%). Analysis of the adult dataset resulted in an overestimation of species richness, especially in leaf litter with an overestimation of up to 52% in SN2, followed by 41% in SN1. In foliage, richness overestimation was the highest in PIPP and SN1 (both with 36% overestimation), followed by CNP (19%) and finally SN2 (11%). In grassland, the overestimation was up to 10% in the PIPP (Table 3 and 4, Appendix Figs. 2, 3, 4, 5, 6, 7, 8 and 9). Richness overestimation due to the inclusion of HTS was significant in foliage and leaf litter when the adult + juvenile dataset was considered (Additional file 1: Table S2). The final slopes of the curves that included HTS were steeper than those that excluded HTS for all cases (Appendix Figs. 2, 3, 4, 5, 6, 7, 8 and 9, Additional file 2: Table S4 and S5). These differences were significant in leaf litter when the adult + juvenile dataset was considered (Additional file 1: Table S3).

Estimation of species richness considering adults + juveniles or adults

With few exceptions, richness estimates were higher when adults + juveniles were considered than when only adults were considered (Tables 3 and 4). However, the difference between richness overestimation between both

Table 1 IndVal-significant species in three habitats of the studied area, Eastern Chaco District, Argentina

Adults + juveniles					Adults				
Foliage group (17 species)					Foliage group (11 species)				
Species	Indval	p	A	B	Species	Indval	p	A	B
<i>Philira micans</i>	0.530	0.001	0.99	0.28	<i>Philira micans</i>	0.337	0.001	0.97	0.12
<i>Scopocira histrio</i>	0.444	0.001	0.99	0.20	<i>Scopocira histrio</i>	0.250	0.001	1.00	0.06
<i>Colonus germaini</i>	0.393	0.001	1.00	0.15	<i>Colonus germaini</i>	0.129	0.031	1.00	0.02
<i>Maeota dichrura</i>	0.371	0.001	1.00	0.14	<i>Maeota dichrura</i>	0.129	0.036	1.00	0.02
<i>Chira spinosa</i>	0.350	0.001	0.89	0.14	<i>Chira spinosa</i>	0.196	0.004	0.76	0.05
<i>Arachnomura querandi</i>	0.303	0.001	1.00	0.09	<i>Arachnomura querandi</i>	0.258	0.001	1.00	0.07
<i>Gypogyna forceps</i>	0.303	0.001	1.00	0.09	<i>Gypogyna forceps</i>	0.194	0.001	1.00	0.04
<i>Chira gounellei</i>	0.252	0.001	0.89	0.07	<i>Chira gounellei</i>	0.144	0.013	1.00	0.02
<i>Bryantella smaragda</i>	0.242	0.001	1.00	0.06	–				
<i>Titanattus parvus</i>	0.233	0.001	1.00	0.05	<i>Titanattus parvus</i>	0.183	0.001	1.00	0.03
<i>Wedoquella macrothecata</i>	0.224	0.001	1.00	0.05	–				
<i>"Coryphasia" sp.n</i>	0.197	0.001	0.93	0.04	–				
<i>Cotinusa vittata</i>	0.171	0.001	1.00	0.03	<i>Cotinusa vittata</i>	0.158	0.008	1.00	0.02
<i>Cyllodania zoobotanica</i>	0.171	0.004	1.00	0.03					
<i>Gastromicans sp.n.1</i>	0.158	0.006	1.00	0.02	<i>Gastromicans sp.n.1</i>	0.158	0.006	1.00	0.02
<i>Scoturius tigris</i>	0.158	0.009	1.00	0.02	–				
<i>Frigga quintensis</i>	0.149	0.019	0.89	0.02	–				
Leaf litter group (2 species)					Leaf litter group (2 species)				
Species	Indval	p	A	B	Species	Indval	p	A	B
<i>Semiopyla viperina</i>	0.570	0.001	1.00	0.32	<i>Semiopyla viperina</i>	0.173	0.005	0.90	0.03
<i>Hisukattus transversalis</i>	0.171	0.002	1.00	0.03	<i>Hisukattus transversalis</i>	0.158	0.006	1.00	0.02
Grassland group (21 species)					Grassland group (14 species)				
Species	Indval	p	A	B	Species	Indval	p	A	B
<i>Maeota dorsalis</i>	0.620	0.001	0.99	0.39	<i>Maeota dorsalis</i>	0.399	0.001	0.98	0.16
<i>Pseudofluda capandegui</i>	0.505	0.001	0.97	0.26	<i>Pseudofluda capandegui</i>	0.413	0.001	1.00	0.17
<i>"Euophrys" melanoleuca</i>	0.456	0.001	1.00	0.21	<i>"Euophrys" melanoleuca</i>	0.428	0.001	1.00	0.18
<i>"Dendryphantes" aff. mordax</i>	0.386	0.001	0.97	0.15	<i>"Dendryphantes" aff. mordax</i>	0.242	0.001	1.00	0.06
<i>Akela ruricola</i>	0.334	0.001	0.92	0.12	<i>Akela ruricola</i>	0.281	0.001	1.00	0.08
<i>Pachomius rubrogastrus</i>	0.321	0.001	0.95	0.11	–				
<i>Pseudofluda palachiyaxa</i>	0.312	0.001	0.97	0.10	<i>Pseudofluda palachiyaxa</i>	0.144	0.008	1.00	0.02
<i>Tartamura adfectuosa</i>	0.311	0.001	0.97	0.10	<i>Tartamura adfectuosa</i>	0.259	0.001	0.95	0.07
<i>Neonella sp.n.1</i>	0.274	0.001	1.00	0.07	<i>Neonella sp.n.1</i>	0.250	0.001	1.00	0.06
<i>Zygoballus sp.n</i>	0.251	0.001	0.94	0.07	<i>Zygoballus sp.n</i>	0.144	0.014	1.00	0.02
<i>Neonella sp.n.2</i>	0.242	0.001	1.00	0.06	<i>Neonella sp.n.2</i>	0.214	0.001	1.00	0.05
<i>Jollas sp.</i>	0.227	0.001	0.83	0.06	<i>Jollas sp.</i>	0.189	0.005	0.77	0.05
<i>Tapsatella albocastanea</i>	0.224	0.001	1.00	0.05	–				
<i>Pachomius areteguazu</i>	0.216	0.001	0.93	0.05	<i>Pachomius areteguazu</i>	0.144	0.011	1.00	0.02
<i>Ailluticus aff. rotundus</i>	0.208	0.001	0.95	0.05	–				
<i>Aphirape gamas</i>	0.204	0.001	1.00	0.04	<i>Aphirape gamas</i>	0.129	0.040	1.00	0.02
<i>Aphirape riparia</i>	0.194	0.001	1.00	0.04	–				
<i>Cotinusa cf. melanura</i>	0.171	0.001	1.00	0.03	–				
<i>Pachomius sp.n</i>	0.171	0.002	1.00	0.03	–				
<i>Semyopila? sp.n</i>	0.171	0.003	1.00	0.03	<i>Semyopila? sp. n</i>	0.171	0.002	1.00	0.03
<i>Semiopyla cathaphracta</i>	0.134	0.050	0.86	0.86	–				

Table 1 (continued)

Leaf litter + Grassland group (1 species)									
Species	Indval	p	A	B	Species	Indval	p	A	B
<i>Corythalia conferta</i>	0.196	0.027	0.87	0.04	–				

A: specificity component of the Indval index, B: fidelity component of the Indval index, Indval: Indicator Value Index, p: significance value, question mark indicates that the identification of the species in the genus is doubtful (i.e. it may be a new genus), quoted words indicate that the species should be included in another genus

Table 2 HTS in the studied area, Eastern Chaco District, Argentina

HTS	Primary habitat	Ad + J	Ad
Foliage (8)			
<i>"Dendryphantes" aff. mordax</i>	Grassland	•	
<i>Jollas sp.</i>	Grassland	•	•
<i>Maeota dorsalis</i>	Grassland	•	•
<i>Pachomius areteguazu</i>	Grassland	•	
<i>Pseudofluda capandegui</i>	Grassland	•	
<i>Tartamura adfectuosa</i>	Grassland	•	•
<i>Zygoballus sp.n</i>	Grassland	•	
<i>Semiopyla cathaphracta</i>	Grassland	•	
Leaf litter (11)			
<i>Ailluticus aff. rotundus</i>	Grassland	•	
<i>Akela ruricola</i>	Grassland	•	
<i>Chira gounellei</i>	Foliage	•	
<i>Chira spinosa</i>	Foliage	•	•
<i>"Coryphasia" sp.n</i>	Foliage	•	•
<i>Frigga quintensis</i>	Foliage	•	
<i>Jollas sp.</i>	Grassland	•	•
<i>Phillira micans</i>	Foliage	•	•
<i>Pseudofluda palachiyaxa</i>	Grassland	•	
<i>Scopocira histrio</i>	Foliage	•	
<i>Pachomius rubrogastrus</i>	Grassland	•	
Grassland (2)			
<i>Chira gounellei</i>	Foliage	•	
<i>Chira spinosa</i>	Foliage	•	•

Ad adults, J juveniles, HTS habitat-tourist species, primary habitat: a term used by Barlow et al. (2010) and can be understood as habitat where the species are expected to be resident (sensu Archer 2003 and Eggleton et al. 2021), quoted words indicate that the species should be included in another genus

datasets was significant only in leaf litter when HTS were included (Additional file 1: Table S4). The richness estimation curves did not reach an asymptote in most cases. However, there were several exceptions, such as the leaf litter of CNP with adult and adult + juvenile datasets without HTS, and the grassland of SN2, where asymptotes were obtained from the adult dataset (Appendix Figs. 2D, 3C – D and 7E – F).

In grassland and in foliage with HTS, and in leaf litter without HTS, the curves made with the adult

dataset were further from reaching the asymptote than the curves made with the adult + juvenile dataset (Appendix Figs. 2, 3, 4, 5, 6, 7, 8 and 9, Additional file 1: Table S5). However, these differences were not significant (Table S6). On the contrary, in leaf litter with HTS, and in foliage without HTS, the final slopes of the curves were steeper when adults + juveniles were considered (Figs. 2, 3, 4, 5, 6, 7, 8 and 9, Additional file 1: Table S5). In leaf litter with HTS, these differences were significant (Additional file 1: Table S6).

Discussion

This work proposes a simple methodology to detect and eliminate HTS from the database before performing any kind of analysis of assemblages of a given taxon. The proposed methodology included the use of a known index, the Indicator Value Index or IndVal. As an example, it was shown how HTS affect the richness estimates in spider assemblages of the Salticidae family in one study area of Eastern Chaco, Argentina. Knowledge of the primary habitat of these species made it possible to detect them and remove them from the habitats in which they appeared as HTS. Analyses of richness estimates with and without HTS showed that HTS overestimated richness, although not always significantly.

According to Eggleton et al. (2021) HTS are expected to find among highly dispersive taxa. Spiders of the Salticidae family have the ability to disperse by various mechanisms. One of the most important ones includes ballooning (Horner 1975), which consists of dispersal by air (Richman and Jackson 1992). The presence of HTS in the studied habitats could be related to dispersal using this mechanism. The higher number of HTS in leaf litter could be due to the fact that the probability of spiders falling from the foliage was added to this type of dispersal. Although there seem to be no reports on the process of falling of the foliage-dwelling spiders, it can be deduced that this process does occur since one of the widely used collection techniques for foliage-dwelling spiders is foliage beating, which makes the spiders fall to the ground (Nadal et al. 2018). Therefore, this could occur naturally due to the effect of gravity or the movement of foliage by the wind.

Table 3 Estimation of richness with HTS in the studied area, Eastern Chaco District, Argentina

		Habitat	S	ACE	ICE	Chao1	Chao2	Jack1	Jack2
CNP	Ad + J	Foliage	16	27.75	28.28	25.17	33.70	24.85	32.60
		Leaf litter	5	5.44	5.40	5.00	5.00	5.98	6.00
		Grassland	28	39.68	41.22	36.96	38.82	38.82	45.65
	Ad	Foliage	11	26.65	26.87	15.97	16.16	17.88	21.80
		Leaf litter	2	2.00	2.00	2.00	2.00	2.00	2.00
		Grassland	20	33.96	32.41	28.90	28.85	28.85	34.70
SN1	Ad + J	Foliage	20	26.70	27.83	23.72	25.16	26.88	30.80
		Leaf litter	9	23.82	23.21	23.48	23.75	14.90	20.70
		Grassland	23	25.66	25.75	23.85	23.66	26.93	23.20
	Ad	Foliage	13	38.78	27.52	40.43	40.53	20.87	28.60
		Leaf litter	5	15.00	15.00	7.50	7.95	8.93	11.85
		Grassland	15	19.25	20.00	19.93	22.38	20.90	25.75
SN2	Ad + J	Foliage	27	35.97	37.07	38.95	35.85	36.83	42.70
		Leaf litter	8	18.45	17.97	17.76	17.83	12.92	17.75
		Grassland	19	20.02	19.93	19.33	19.33	20.97	21.00
	Ad	Foliage	18	22.32	21.87	21.69	21.69	23.90	26.85
		Leaf litter	6	21.00	21.00	10.29	10.92	10.92	14.80
		Grassland	16	16.00	16.00	16.00	16.00	16.00	10.30
PIPP	Ad + J	Foliage	24	35.56	32.50	30.96	30.88	31.87	36.75
		Leaf litter	8	29.00	29.00	15.29	15.38	13.90	18.75
		Grassland	19	20.35	20.55	20.49	19.98	21.95	22.95
	Ad	Foliage	13	23.10	22.16	19.77	19.88	19.88	24.75
		Leaf litter	3	6.67	7.00	3.80	3.98	4.97	6.90
		Grassland	17	23.96	23.77	27.35	27.33	23.88	29.70

ACE Abundance-based Coverage Estimator, Ad adults, CNP Chaco National Park, HTS habitat-tourist species, ICE Incidence-based Coverage Estimator, Jack1 Jackknife1, Jack2 Jackknife2, J juveniles, PIPP Pampa del Indio Provincial Park, S species richness, SN1 semi-natural site 1, SN2 semi-natural site 2

The richness overestimates obtained by including HTS herein are in agreement with what was found by Archer (2003) studying wasps and bees, and with Eggleton et al. (2021) with beetles. These results have important implications for conservation policies because it is very common to use species richness and other diversity indices as an indicator of habitat quality (Avalos et al. 2018; Nadal et al. 2018; Nether et al. 2019; Pinto et al. 2021). In this sense, HTS could increase species richness in habitats that are not necessarily of good quality and lead to wrong conclusions. The steepest asymptotes obtained by Jiménez-Valverde and Lobo (2006) when considering adults of spiders instead of adults + juveniles do not match with what was found herein as the only significant difference indicated the opposite of what was obtained by those authors. However, they worked with the family of crab spiders (Thomisidae), which could partly explain the differences.

One of the limitations of the proposed methodology is that the detection of IndVal-significant species by

the Indval index requires the species to be abundant enough (Dufrêne and Legendre 1997). Julliard et al. (2006) detected the same problem with the SSI index. However, this drawback is offset by the low probability that a rare species in one habitat will appear as HTS in another, due to its lower abundance. Despite the aforementioned limitation, this is the first study to propose an unbiased way of detecting tourist-habitat species, by discriminating them from the rest of the data. To our knowledge, Barlow et al. (2010) were the only authors who proposed a numerical way to treat TS. However, their methodology continues to have the drawbacks inherent to other methods, which they themselves criticize and which are also criticized herein. Thus, these authors propose to exclude species with less than 10 records in the sample but they do not distinguish genuine rare species from TS properly. Archer (2003) and Eggleton et al. (2021) used more appropriate procedures than that of previous authors, based on the knowledge of the habitat preference of the species, but not in a very objective way.

Table 4 Estimation of richness without HTS in the studied area, Eastern Chaco District, Argentina

		Habitat	S	ACE	ICE	Chao1	Chao2	Jack1	Jack2
CNP	Ad + J	Foliage	16	27.75	28.28	25.17	33.70	24.85	32.60
		Leaf litter	3	3.00	3.00	3.00	3.00	3.00	3.00
		Grassland	27	36.76	38.25	34.17	35.85	36.83	42.70
	Ad	Foliage	11	26.65	26.87	15.97	16.16	17.88	21.80
		Leaf litter	2	2.00	2.00	2.00	2.00	2.00	2.00
		Grassland	20	33.96	32.41	28.90	28.85	28.85	34.70
SN1	Ad + J	Foliage	16	18.22	19.42	16.74	17.48	19.93	20.95
		Leaf litter	4	7.90	7.32	4.95	4.98	5.97	7.90
		Grassland	23	25.66	25.75	23.85	23.66	26.93	23.20
	Ad	Foliage	11	26.01	19.44	25.68	25.75	16.90	22.70
		Leaf litter	3	5.00	5.95	5.00	5.95	5.95	8.85
		Grassland	15	19.25	20.00	19.93	22.38	20.90	25.75
SN2	Ad + J	Foliage	23	27.29	27.27	27.98	25.95	28.90	30.90
		Leaf litter	5	6.88	6.69	5.97	5.98	6.97	8.90
		Grassland	19	20.02	19.93	19.33	19.33	20.97	21.00
	Ad	Foliage	17	20.15	19.83	19.46	19.46	21.92	23.90
		Leaf litter	4	10.00	10.00	5.20	5.48	6.95	8.90
		Grassland	16	16.00	16.00	16.00	16.00	16.00	10.30
PIPP	Ad + J	Foliage	20	25.47	23.87	23.32	23.28	24.92	27.85
		Leaf litter	4	7.00	7.00	4.48	4.49	5.97	6.95
		Grassland	18	18.72	18.86	18.50	18.33	19.97	20.00
	Ad	Foliage	10	13.77	13.39	11.93	11.97	13.93	15.90
		Leaf litter	3	6.67	7.00	3.80	3.98	4.97	6.90
		Grassland	16	21.27	21.12	23.39	23.38	21.90	26.75

ACE Abundance-based Coverage Estimator, Ad adults, CNP Chaco National Park, HTS habitat-tourist species, ICE Incidence-based Coverage Estimator, Jack1 Jackknife1, Jack2 Jackknife2, J juveniles, PIPP Pampa del Indio Provincial Park, S species richness, SN1 semi-natural site 1, SN2 semi-natural site 2

Conclusions

The methodology proposed in this work contributes to solving the problem of HTS that was recognized a long time ago, but for which an effective solution has not been found. This method corrects the bias of methods of other authors where species less abundant are eliminated with subjective rules. However, it can be used in a complementary way with the methods that use the empirical knowledge of the natural history or habitat preference of species. In this sense, empirical knowledge of the habitat preference of species could support this method to have a double confirmation on whether a species is HTS or TS.

This study demonstrates that estimators overestimate species richness when HTS are included which can lead to erroneous conclusions in ecological studies. The overestimation of richness by including HTS reinforces the importance of removing these species before performing ecological analyses. Contrary to expectations, the inclusion of juveniles did not contribute to richness estimation curves being more

asymptotic than excluding them. However, this study showed that the with adult + juvenile dataset can be detected more HTS than with adult dataset. So, the inclusion of juveniles could improve the analyses, at least in spiders. Although the proposed methodology focused on HTS, and how they affect richness estimation, it can be applied to TS at broader scales (e.g. at the level of biogeographic regions), other community ecology analyses (e.g. alpha and beta diversity indices), or using other multiple or single taxa. In this sense, it is recommended to analyze species preference for regions, with the same procedure with which the preference for habitat was evaluated herein. In the same way, it is recommended to remove HTS or TS before analyzing diversity indices in a particular habitat or region, respectively. Regarding future prospects, knowing the identity of HTS can be very useful in upcoming studies to assess the effect of habitat fragmentation, since more HTS is expected to be obtained the more fragmented the habitats are.

Appendix

See the Figs. 2, 3, 4, 5, 6, 7, 8 and 9 below.

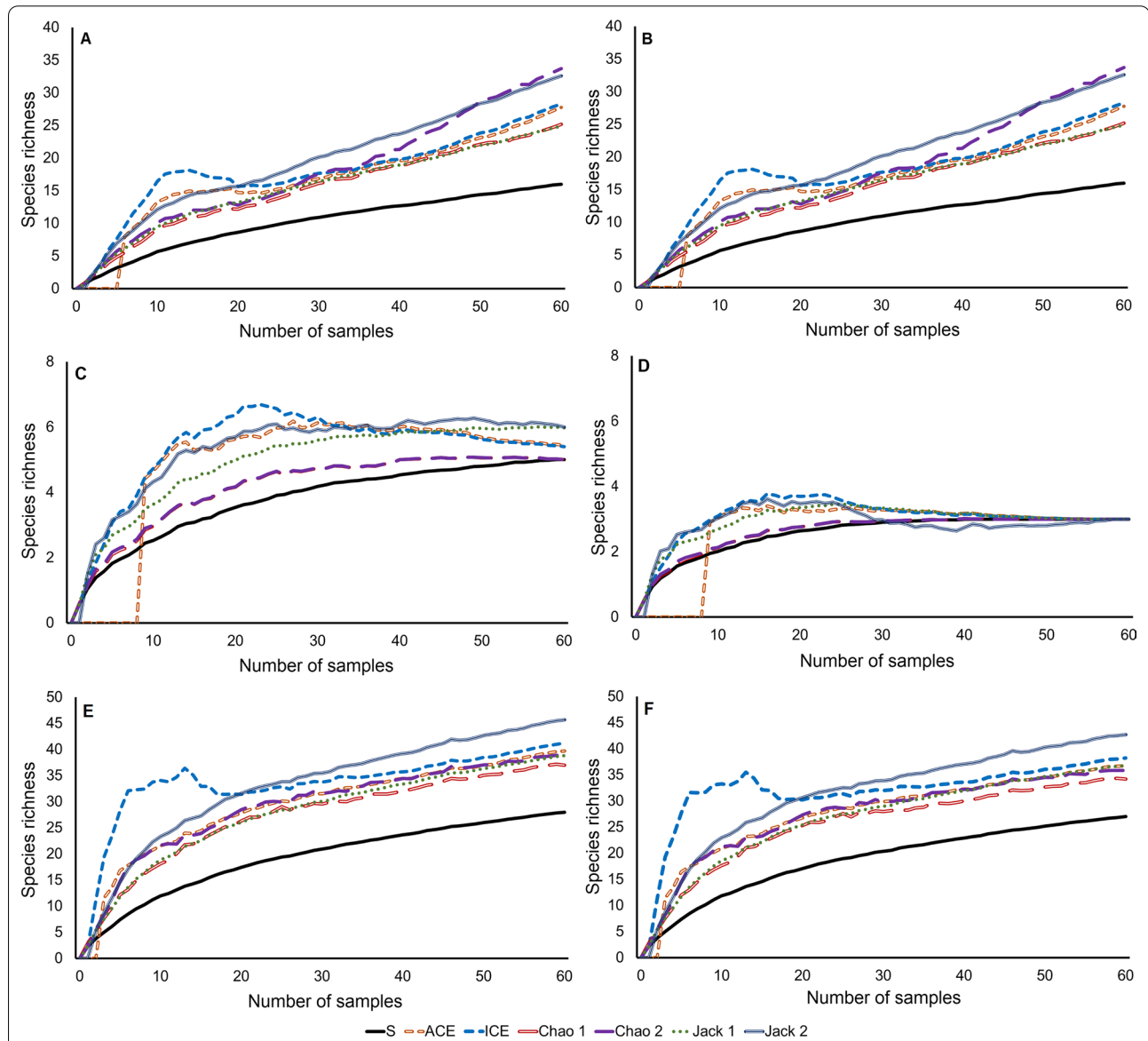
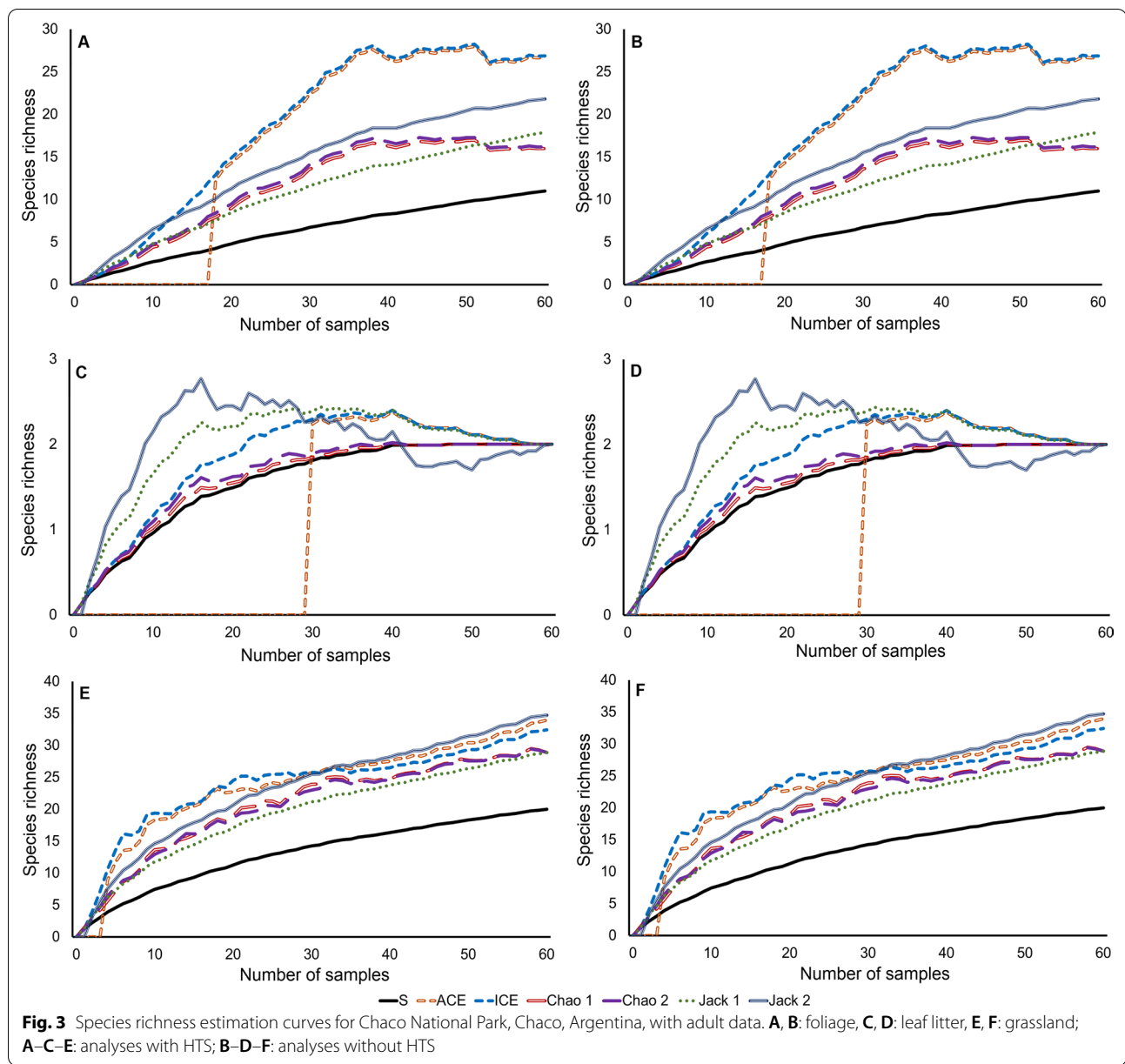
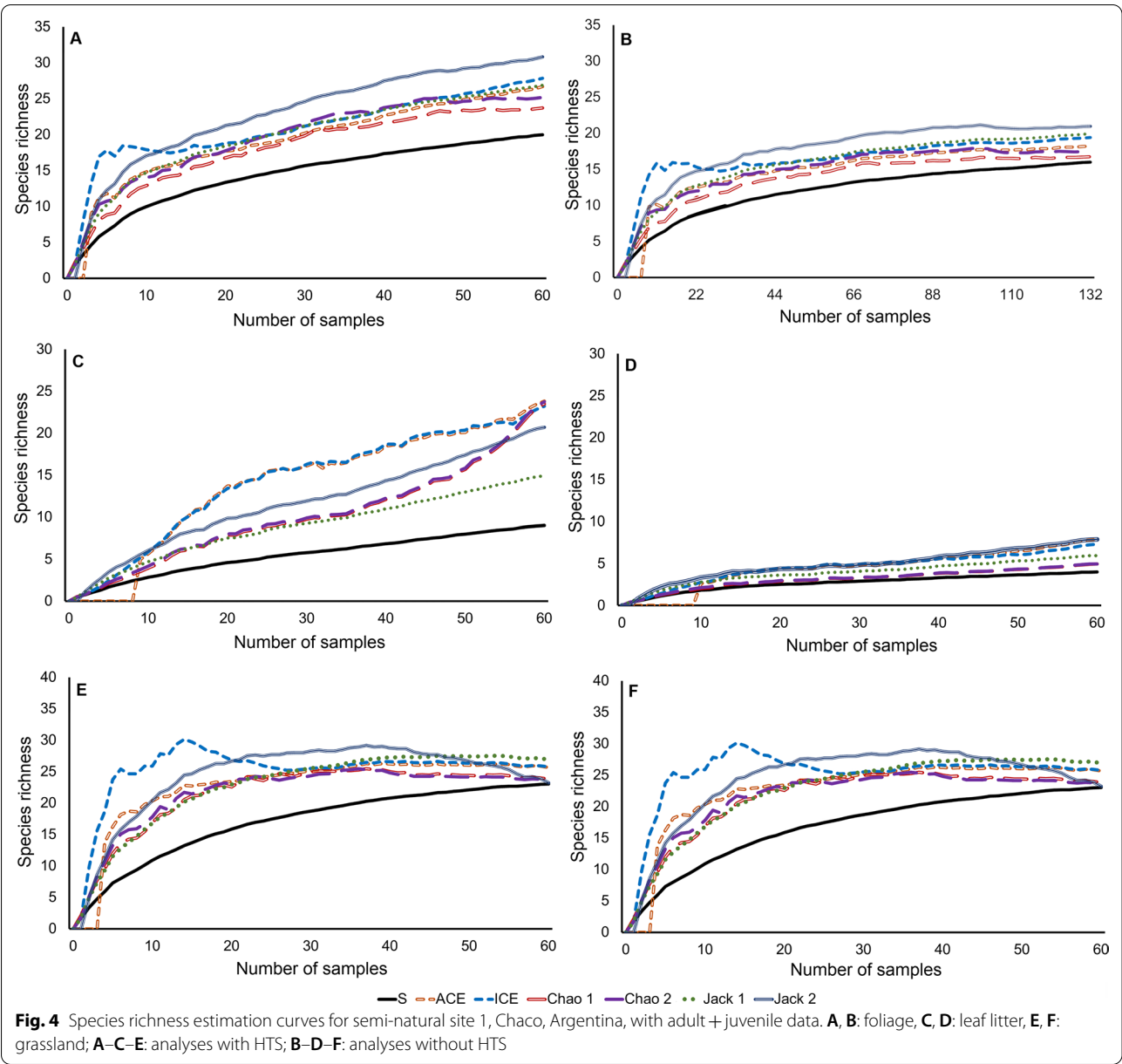


Fig. 2 Species richness estimation curves for Chaco National Park, Chaco, Argentina, with adult + juvenile data. **A, B:** foliage, **C, D:** leaf litter, **E, F:** grassland; **A–C–E:** analyses with HTS; **B–D–F:** analyses without HTS





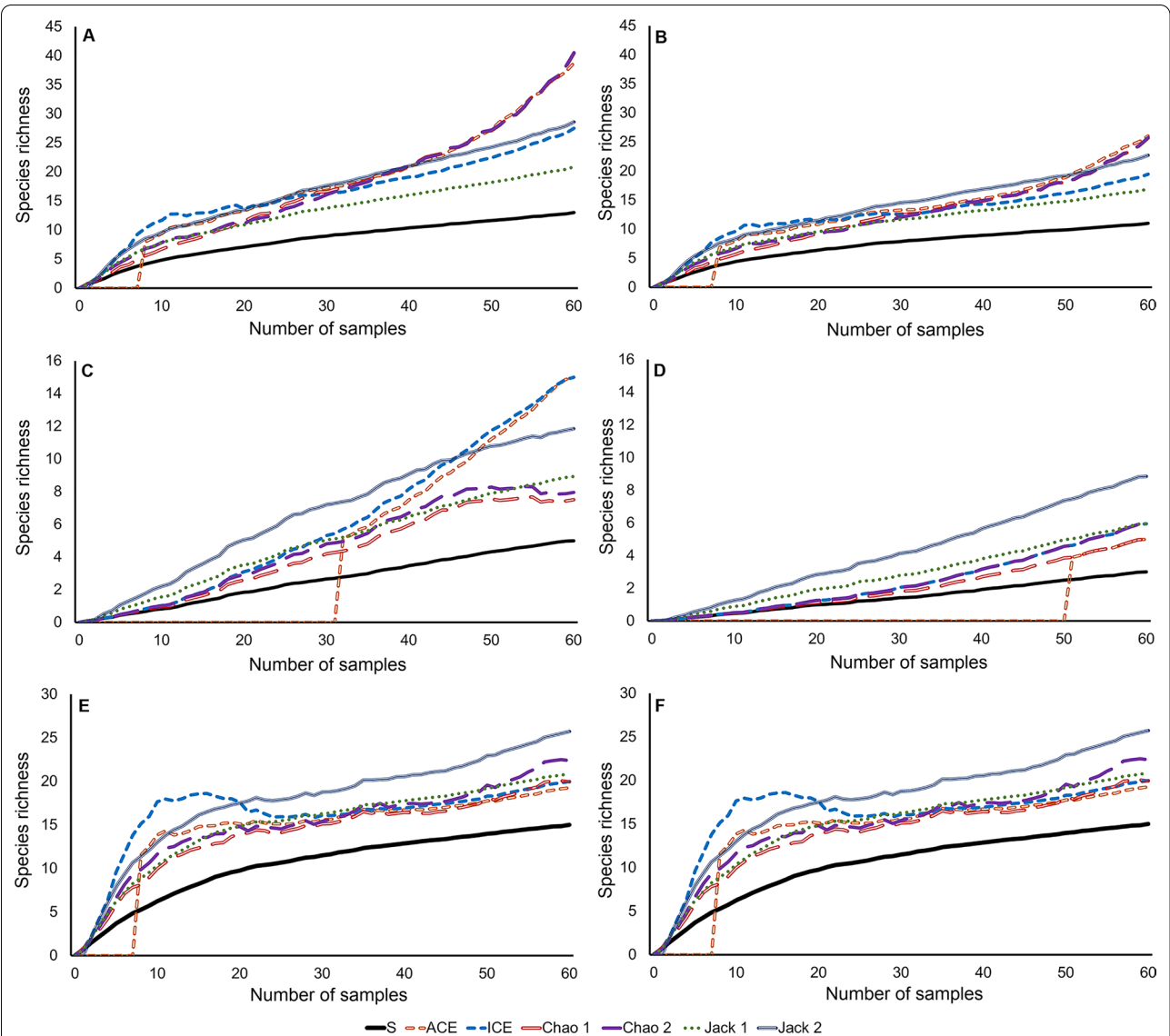


Fig. 5 Species richness estimation curves for semi-natural site 1, Chaco, Argentina, with adult data. **A, B:** foliage, **C, D:** leaf litter, **E, F:** grassland; **A–C–E:** analyses with HTS; **B–D–F:** analyses without HTS

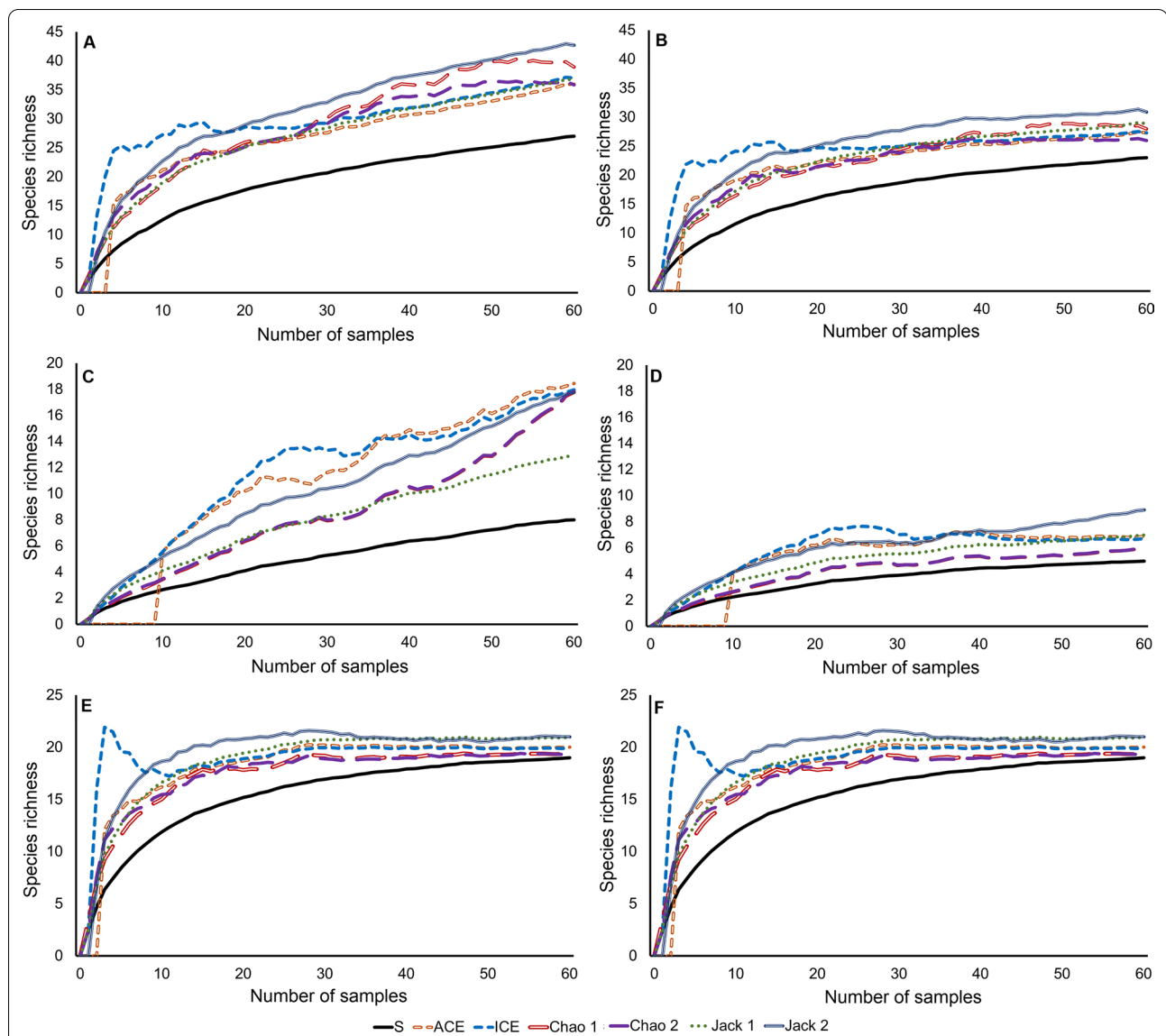
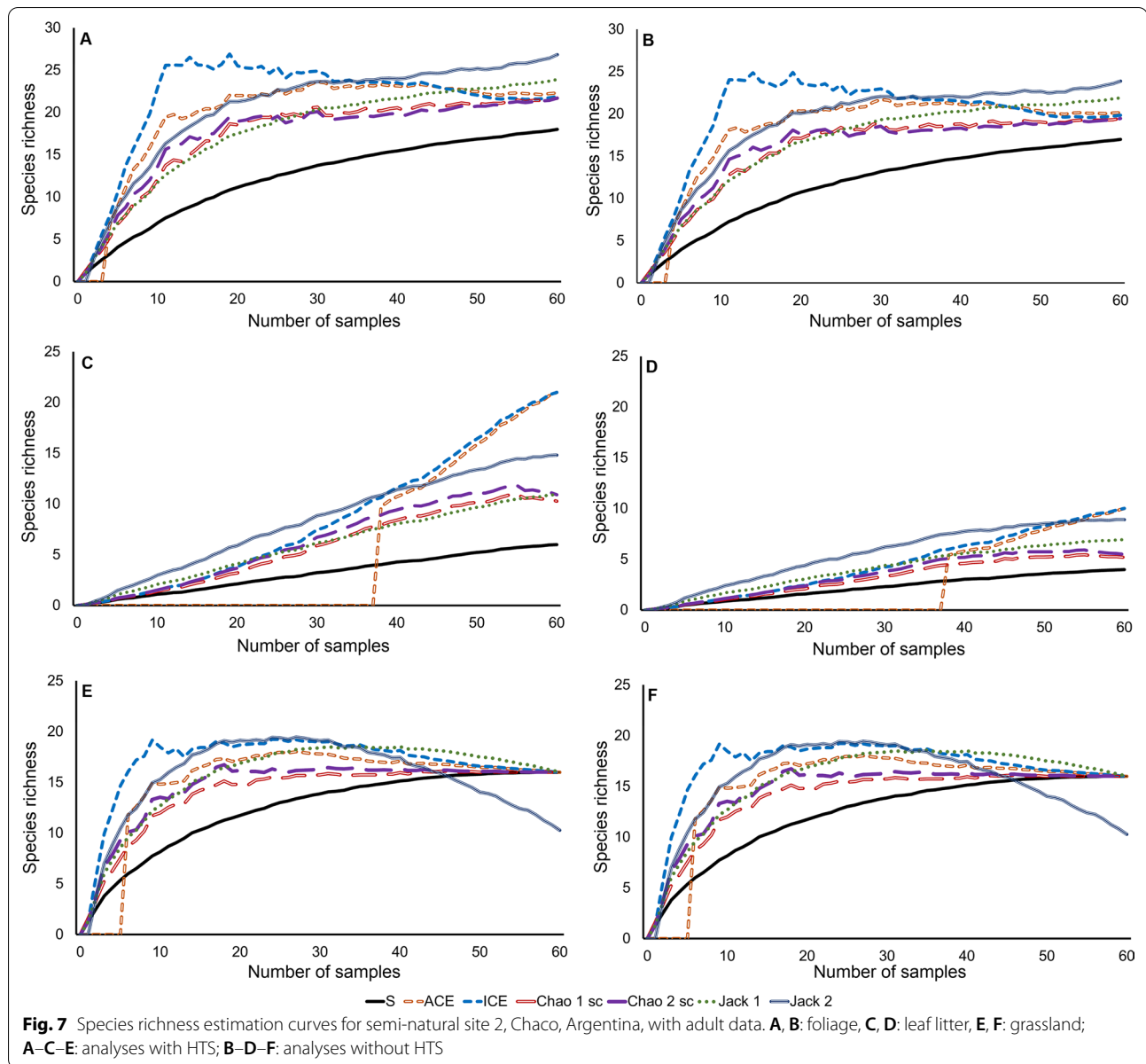
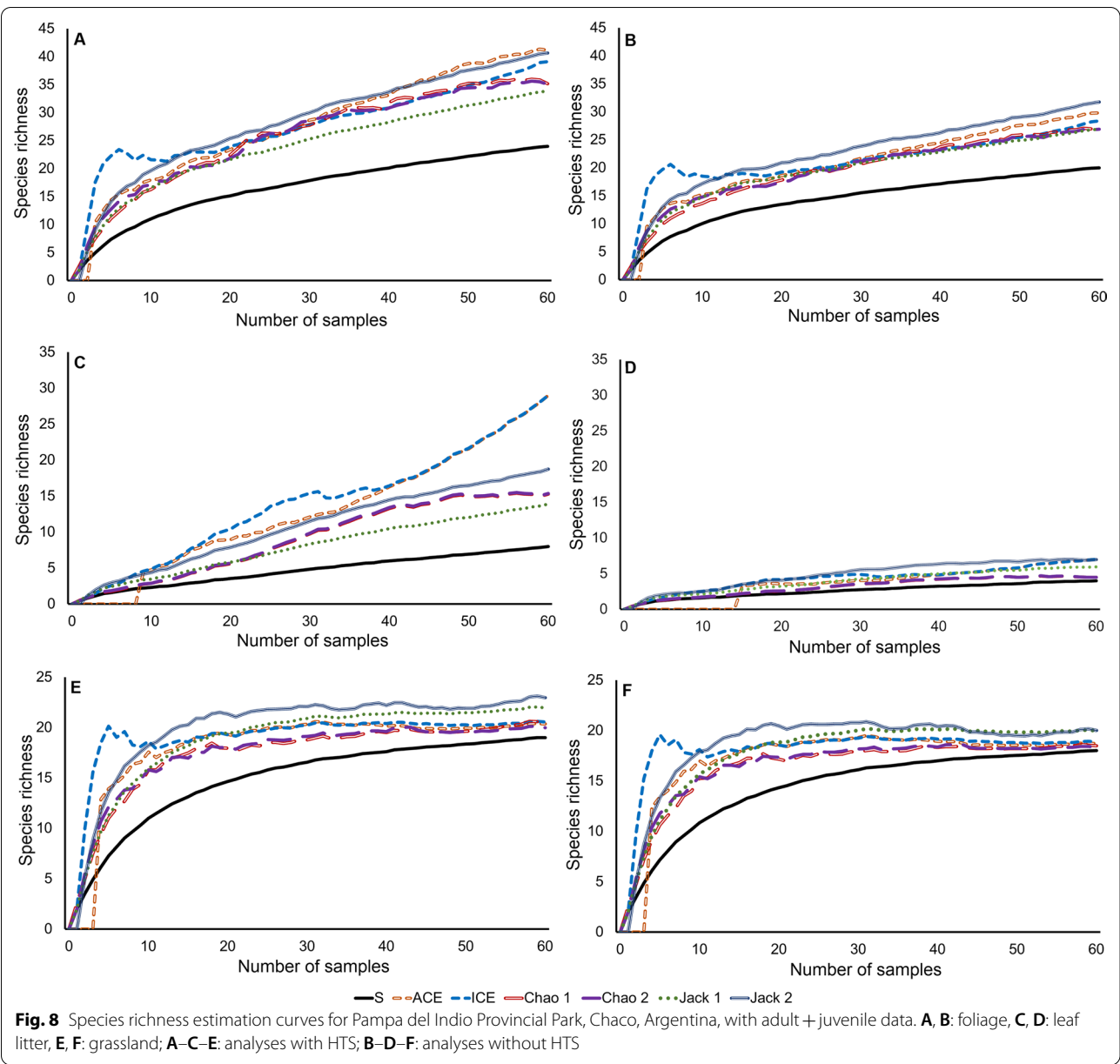
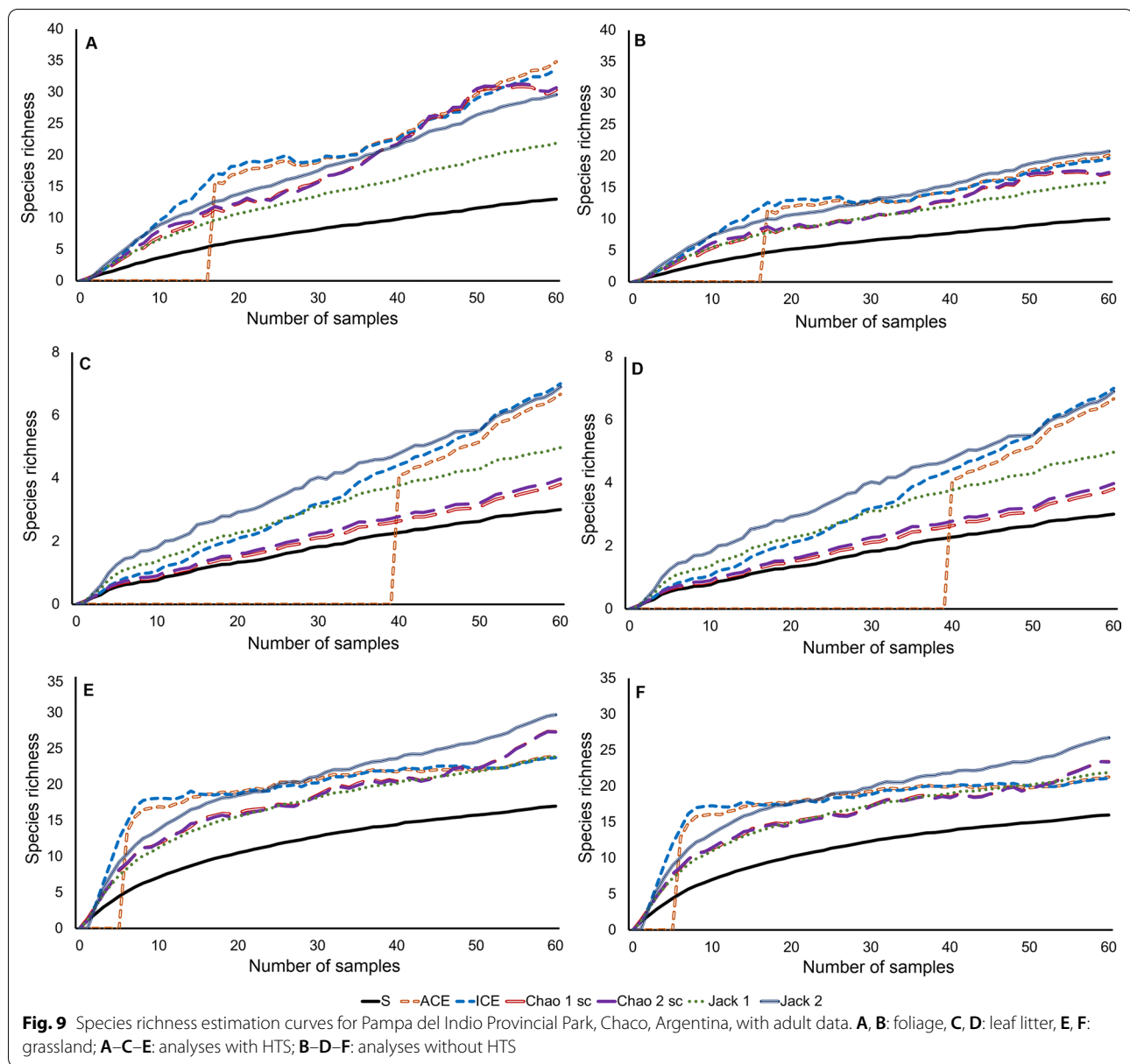


Fig. 6 Species richness estimation curves for semi-natural site 2, Chaco, Argentina, with adult + juvenile data. **A, B:** foliage, **C, D:** leaf litter, **E, F:** grassland; **A–C–E:** analyses with HTS; **B–D–F:** analyses without HTS







Abbreviations

A: Specificity component of the Indicator Value Index; ACE: Abundance-based Coverage Estimator; B: Fidelity component of the Indicator Value Index; CAR-TROUNNE: Colección de Artrópodos de la Universidad Nacional del Nordeste; CNP: Chaco National Park; FACENA: Facultad de Ciencias Exactas y Naturales y Agrimensura; HTS: Habitat-tourist species; ICE: Incidence-based Coverage Estimator; IBSI-Ara collection: The arachnological collection of the Instituto de Biología Subtropical, Misiones, Argentina; IndVal: Valor Indicador; *p*: Statistical significance value; PPIP: Pampa del Indio Provincial Park; S: Species richness; SN1: Semi-natural site 1; SN2: Semi-natural site 2; SSI: Species Specialization Index; TS: Tourist-species; UNNE: Universidad Nacional del Nordeste.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13717-022-00398-6>.

Additional file 1: Table S1. Species of the family Salticidae and their abundance in Pampa del Indio Provincial Park (PIPP), Chaco National Park (CNP) and semi-natural sites (SN1 and SN2) in Chaco, Argentina. **Table S2.** Significance assessment of the differences between the richness overestimation data, calculated as the difference between observed and estimated richness, when Habitat-tourist species were considered or not. **Table S3.** Significance assessment of the differences between the data of final slopes of richness estimation curves, when Habitat-tourist

species were considered or not. **Table S4.** Significance assessment of the differences between the richness overestimation data, calculated as the difference between observed and estimated richness, when adults were considered or not. **Table S5.** Average values of the final slopes of the richness estimation curves, when adults were considered or not. **Table S6.** Significance assessment of the differences between the data of final slopes of richness estimation, when adults were considered or not.

Additional file 2: Table S1. Difference between observed and estimated richness when habitat-tourist species were considered. **Table S2.** Difference between observed and estimated richness when habitat-tourist species were not considered. **Table S3.** Verification of the normality assumption of the richness overestimation data, calculated as the difference between observed and estimated richness. **Table S4.** Values of the final slopes of the richness estimation curves when habitat-tourist species were considered. **Table S5.** Values of the final slopes of the richness estimation curves when habitat-tourist species were not considered. **Table S6.** Verification of the normality assumption of the final slope data of the richness estimation curves.

Acknowledgements

First, we thank the reviewers and the editor for their valuable comments that improved this article. Second, we thank colleagues I. Zanone, P. Gonzalez, C. Achitte, G. Rubio, P. Cuaranta, E. Toledo, R. Aguirre, M. De los Santos, D. Larrea and A. Raimundo, who collaborated with the fieldwork. Third, we thank G. D. Rubio for collaborating with the identification of several species in the early stages of the study (e.g. *Mburuvicha* sp., *Neonella* spp., *S. cathaphracta*, *T. yungae*, *W. punctata*). Fourth, we thank Alejandra Scotti for the language editing. Finally, we thank the administrators and park rangers of Pampa del Indio Provincial Park and Chaco National Park, as well as the owners and foremen of the fields for their permission to carry out the sampling in those places.

Author contributions

MFN and GA performed the field sampling, MFN determined the species, performed the data analysis, and wrote the manuscript. GA and AG edited the manuscript. All authors gave the final approval for publication. All authors read and approved the final manuscript.

Authors' information

MFN is concluding her doctoral thesis financed by a doctoral fellowship from CONICET and supervised by AG and GA. Her research topic is focused on the ecology and taxonomy of spiders of the Salticidae family, AG is an Ad-Honorem Independent Researcher of CONICET with experience in spider ecology, and GA is a professor and researcher from FACENA-UNNE with experience in spider ecology.

Funding

This research was financed by project PI F 003/2015 of the SGCyT—UNNE (Argentina) and by a CONICET doctoral fellowship awarded to the first author.

Availability of data and materials

There is relevant information available in the supporting information. The datasets used in this paper will become public at Figshare repository (<https://figshare.com>) on May 9 2024. The DOIs are as follow: Data to run in Estimates: 10.6084/m9.figshare.19733683. Data to run in R with Indicspecies package: 10.6084/m9.figshare.19733716.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, Argentina. ²Laboratorio de Biología de los Artrópodos, Facultad de Ciencias Exactas y Naturales y Agrimensura, Universidad Nacional del Nordeste, Avenida Libertad 5470, 3400 Corrientes, Argentina. ³Centro de Estudios Parasitológicos y de Vectores (CEPAVE-UNLP), Boulevard 120 s/n (e/60-64), 1900 La Plata, Argentina.

Received: 9 May 2022 Accepted: 19 August 2022

Published online: 15 September 2022

References

- Achitte-Schmutzler HC, Avalos G, Oscherov EB (2022) Diversidad taxonómica de Thomisidae (Araneae) en ambientes heterogéneos del sitio Ramsar Humedales Chaco, Argentina. *Caldasia* 44(1):119–129. <https://doi.org/10.15446/caldasia.v44n1.83581>
- Arana MD, Natale E, Ferretti N, Romano G, Oggero A, Martínez G, Posadas P, Morrone JJ (2021) Esquema biogeográfico de la República Argentina. *Opera lilloana* 56, Fundación Miguel Lillo, Tucumán
- Archer ME (2003) The wasps and bees (Hymenoptera: Aculeata) of Messingham Sand Quarry in Watsonian Lincolnshire with special reference to resident and tourist species. *Naturalist* 128:93–102
- Arnold AE, Lutzoni F (2007) Diversity and host range of foliar fungal endophytes: are tropical leaves biodiversity hotspots? *Ecology* 88(3):541–549. <https://doi.org/10.1890/05-1459>
- Avalos G, Achitte-Schmutzler HC, de los Santos ME (2018) Caracterización de la fauna de arañas en monocultivos de *Eucalyptus* y *Pinus* de la Reserva del Iberá, Corrientes, Argentina. *Rev Mex Biodivers* 89(1):134–148. <https://doi.org/10.22201/ib.20078706e.2018.1.1910>
- Barlow J, Gardner TA, Louzada J, Peres CA (2010) Measuring the conservation value of tropical primary forests: the effect of occasional species on estimates of biodiversity uniqueness. *PLoS One* 5(3):e9609. <https://doi.org/10.1371/journal.pone.0009609>
- Bonte D, Baert L, Maelfait JP (2002) Spider assemblage structure and stability in a heterogeneous coastal dune system (Belgium). *J Arachnol* 30(2):331–343. [https://doi.org/10.1636/0161-8202\(2002\)030\[0331:SASASJ\]20CO;2](https://doi.org/10.1636/0161-8202(2002)030[0331:SASASJ]20CO;2)
- Burnham KP, Overton WS (1978) Estimation of the size of a closed population when capture probabilities vary among animals. *Biometrika* 65(3):625–633. <https://doi.org/10.1093/biomet/653625>
- Burnham KP, Overton WS (1979) Robust estimation of population size when capture probabilities vary among animals. *Ecology* 60(5):927–936. <https://doi.org/10.2307/1936861>
- Cabrera AL (1971) Fitogeografía de la República Argentina. *Bol Soc Argent Bot* 14:1–42
- Cabrera AL (1976) Regiones fitogeográficas argentinas. Editorial ACME, Buenos Aires
- Chao A (1984) Nonparametric estimation of the number of classes in a population. *Scand J Stat* 11(4):265–270
- Chao A (1987) Estimating the population size for capture-recapture data with unequal catchability. *Biometrics* 43(4):783–791
- Chao A, Yang MC (1993) Stopping rules and estimation for recapture debugging with unequal failure rates. *Biometrika* 80(1):193–201. <https://doi.org/10.1093/biomet/801193>
- Chase JM, Jeliaskov A, Ladouceur E, Viana DS (2020) Biodiversity conservation through the lens of metacommunity ecology. *Ann NY Acad Sci* 1469(1):86–104. <https://doi.org/10.1111/nyas14378>
- Chazdon RL, Colwell RK, Denslow JS, Guariguata MR (1988) Statistical methods for estimating species richness of woody regeneration in primary and secondary rain forest of northeastern Costa Rica. In: Dallmeier F, Comiskey JA (eds) *Forest biodiversity research monitoring and modeling: conceptual background and Old World case studies*. Parthenon Publishing, Paris, pp 285–309
- Coddington JA, Agnarsson I, Miller JA, Kuntner M, Hormiga G (2009) Under-sampling bias: the null hypothesis for singleton species in tropical arthropod surveys. *J Anim Ecol* 78(3):573–584
- Colwell RK, Coddington JA (1994) Estimating terrestrial biodiversity through extrapolation. *Philos Trans R Soc London Ser B* 345(1311):101–118. <https://doi.org/10.1098/rstb19940091>

- Colwell RK (2013) EstimateS: Statistical estimation of species richness and shared species from samples. <https://viceroyeeb.uconn.edu/estimates>. Accessed 14 Mar 2022.
- Cumming MS, Wesolowska W (2004) Habitat separation in a species-rich assemblage of jumping spiders (Araneae: Salticidae) in a suburban study site in Zimbabwe. *J Zool* 262(1):1–10. <https://doi.org/10.1017/S0952836903004461>
- De Cáceres M, Legendre P (2009) Associations between species and groups of sites: indices and statistical inference. *Ecology* 90(12):3566–3574. <https://doi.org/10.1890/08-1823.1>
- De Cáceres M, Legendre P, Moretti M (2010) Improving indicator species analysis by combining groups of sites. *Oikos* 119(10):1674–1684. <https://doi.org/10.1111/j.1600-0706.2010.18334x>
- Duffy-Anderson JT, Busby MS, Mier KL, Deliyandis CM, Stabeno PJ (2006) Spatial and temporal patterns in summer ichthyoplankton assemblages on the eastern Bering Sea shelf 1996–2000. *Fish Oceanogr* 15:80–94. <https://doi.org/10.1111/j.1365-2641.2005.00348x>
- Dufrène M, Legendre P (1997) Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecol Monogr* 67(3):345–366
- Edinger EN, Risk MJ (2000) Reef classification by coral morphology predicts coral reef conservation value. *Biol Conserv* 92:1–13. [https://doi.org/10.1016/S0006-3207\(99\)00067-1](https://doi.org/10.1016/S0006-3207(99)00067-1)
- Eggleton P, Smith J, Holdsworth S, Inward K, Carpenter D (2021) Tourist species bias estimates of extrapolated species density in dispersive taxa: a case study from a litter beetle assemblage in temperate woodland. *Community Ecol* 22(1):107–112. <https://doi.org/10.1007/s42974-021-00040-z>
- Foelix R (2011) Biology of spiders. Oxford University Press, New York
- Gotelli NJ, Colwell RK (2011) Estimating species richness. In: Magurran AE, McGill BJ (eds) *Frontiers in measuring biodiversity*. Oxford University Press, Oxford, pp 39–54
- Grau HR, Torres R, Gasparri NI, Blendinger PG, Marinaro S, Macchi L (2015) Natural grasslands in the Chaco. A neglected ecosystem under threat by agriculture expansion and forest-oriented conservation policies. *J Arid Environ* 123:40–46. <https://doi.org/10.1016/j.jaridenv.2014.12.006>
- Hadidian J (2015) Wildlife in US cities: managing unwanted animals. *Animals* 5(4):1092–1113. <https://doi.org/10.3390/ani5040401>
- Hanski I (1998) Metapopulation dynamics. *Nature* 396(6706):41–49. <https://doi.org/10.1038/23876>
- Heltshe JF, Forrester NE (1983) Estimating species richness using the jackknife procedure. *Biometrics* 39(1):1–11
- Horner NV (1975) Annual aerial dispersal of jumping spiders in Oklahoma (Araneae Salticidae). *J Arachnol* 2(2):101–105
- Hortal J, Lobo JM (2005) An ED-based protocol for optimal sampling of biodiversity. *Biodivers Conserv* 14(12):2913–2947. <https://doi.org/10.1007/s10531-004-0224-z>
- Jiménez-Valverde A, Lobo JM (2006) The ghost of unbalanced species distribution data in geographical model predictions. *Divers Distrib* 12(5):521–524. <https://doi.org/10.1111/j.1366-9516.2006.00267x>
- Julliard R, Clavel J, Devictor V, Jiguet F, Couvet D (2006) Spatial segregation of specialists and generalists in bird communities. *Ecol Lett* 9(11):1237–1244. <https://doi.org/10.1111/j.1461-0248.2006.00977.x>
- Larivée M, Drapeau P, Fahrig L (2008) Edge effects created by wildfire and clear-cutting on boreal forest ground-dwelling spiders. *For Ecol Manage* 255(5–6):1434–1445. <https://doi.org/10.1016/j.foreco.2007.10.062>
- Lee SM, Chao A (1994) Estimating population size via sample coverage for closed capture-recapture models. *Biometrics* 50(1):88–97
- Magurran AE (2004) *Measuring biological diversity*. Blackwell Science Ltd, Oxford
- Metzner H (2022) Jumping spiders (Arachnida: Araneae: Salticidae) of the world. <https://www.jumping-spiders.com>. Accessed 15 Mar 2022.
- Morello JH, Rodríguez AF (eds) (2009) *El Chaco sin bosques: La Pampa o el desierto del futuro*. Orientación Gráfica Editora, Buenos Aires
- Morello J, Matteucci SD, Rodríguez AF, Silva ME, Mesopotámica P, Llana P (2012) *Ecorregiones y complejos Ecosistémicos de Argentina*. Orientación Gráfica Editora, Buenos Aires
- Nadal MF, Achitte-Schmutzler HC, Zanone I, Gonzalez PY, Avalos G (2018) Diversidad estacional de arañas en una reserva natural del Espinal en Corrientes, Argentina. *Caldasia* 40(1):129–143. <https://doi.org/10.15446/caldasiav40n167362>
- Nether MC, Dudek J, Buschini MLT (2019) Trophic interaction and diversity of cavity-nesting bees and wasps (Hymenoptera: Aculeata) in Atlantic forest fragments and in adjacent matrices. *Apidologie* 50(1):104–115. <https://doi.org/10.1007/s13592-018-0623-x>
- Novotný V, Basset Y (2000) Rare species in communities of tropical insect herbivores: pondering the mystery of singletons. *Oikos* 89(3):564–572. <https://doi.org/10.1034/j.1600-0706.2000.890316x>
- Peel MC, Finlayson BL, McMahon TA (2007) Updated world map of the Köppen-Geiger climate classification. *Hydrol Earth Syst Sci* 11(5):1633–1644. <https://doi.org/10.5194/hess-11-1633-2007>
- Pinto CM, Pairo PE, Bellocq MI, Filloy J (2021) Different land-use types equally impoverish but differentially preserve grassland species and functional traits of spider assemblages. *Sci Rep* 11:10316. <https://doi.org/10.1038/s41598-021-89658-7>
- QGIS Development Team (2021) QGIS Geographic Information System Open Source Geospatial Foundation Project. <https://qgis.org>. Accessed 17 Jul 2021.
- R Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing [version 4.0.2] Vienna, Austria. <https://www.R-project.org/>. Accessed 23 Jun 2020
- Rabinowitz D (1981) Seven forms of rarity. In: Synge H (ed) *Biological aspects of rare plant conservation*. John Wiley and Sons Ltd, New York, pp 205–217
- Ribeiro SP, Borges PAV (2010) Canopy habitat area effect on the arthropod species densities in the Azores: pondering the contribution of tourist species and other life histories. In: Serrano ARM, Borges PAV, Boieiro M, Oromí P (Eds). *Terrestrial arthropods of Macaronesia—Biodiversity Ecology and Evolution*. Security Print—Sociedade de Indústria Gráfica Lda, Loures, pp 89–114
- Richman DB, Jackson RR (1992) A review of the ethology of jumping spiders (Araneae Salticidae). *Bull Br Arachnol Soc* 9(2):33–37
- Rico-G A, Beltrán-A JP, Álvarez-D A, Flórez-D E (2005) Diversidad de arañas (Arachnida: Araneae) en el Parque Nacional Natural Isla Gorgona Pacífico Colombiano. *Biota Neotrop* 5:99–110. <https://doi.org/10.1590/S1676-06032005000200009>
- SAS.Planet Development Team (2019) SAS.Planet [version 190707]. <http://www.sasgis.org/download>. Accessed 30 March 2020.
- Simpson EH (1949) Measurement of diversity. *Nature* 163(4148):688–688
- Smith EP, van Belle G (1984) Nonparametric estimation of species richness. *Biometrics* 40(1):119–129
- Sørensen LL (2004) Composition and diversity of the spider fauna in the canopy of a montane forest in Tanzania. *Biodivers Conserv* 13(2):437–452. <https://doi.org/10.1023/B:BIOC.0000006510.49496.1e>
- Toti DS, Coyle FA, Miller JA (2000) A structured inventory of Appalachian grass bald and heath bald spider assemblages and a test of species richness estimator performance. *J Arachnol* 28(3):329–345. [https://doi.org/10.1636/0161-8202\(2000\)028\[0329:ASIOAG\]2.0.CO;2](https://doi.org/10.1636/0161-8202(2000)028[0329:ASIOAG]2.0.CO;2)
- Willis AJ (1997) The ecosystem: an evolving concept viewed historically. *Funct Ecol* 11:268–271
- World Spider Catalog (2022) World Spider Catalog. Version 23.5. Natural History Museum Bern. <http://wsc.nmbe.ch>. Accessed 1 May 2022
- Zhan A, He S, Brown EA, Chain FJ, Theriault TW, Abbott CL, Heath DD, Cristescu ME, MacIsaac HJ (2014) Reproducibility of pyrosequencing data for biodiversity assessment in complex communities. *Methods Ecol Evol* 5(9):881–890. <https://doi.org/10.1111/2041-210X.12230>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.