Testing the heterospecific attraction hypothesis with time-series data on species co-occurrence

Esther Sebastián-González^{1,2,*}, José Antonio Sánchez-Zapata²,
Francisco Botella² and Otso Ovaskainen³

¹Doñana Biological Station-CSIC, Américo Vespucio s/n 41092, Sevilla, Spain

²Department of Applied Biology, Miguel Hernández University, Cra. Beniel Km 3.2, 03312 Orihuela, Alicante, Spain

³Department of Biosciences, University of Helsinki, Vukinkaari 1, 00014 Helsinki, Finland

The distributional patterns of actively moving animals are influenced by the cues that the individuals use for choosing sites into which they settle. Individuals may gather information about habitat quality using two types of strategies, either directly assessing the relevant environmental factors, or using the presence of conspecifics or heterospecifics as an indirect measure of habitat quality. We examined patterns of heterospecific attraction with observational time-series data on a community of seven waterbird species breeding in artificial irrigation ponds. We fitted to the data a multivariate logistic regression model, which attributes the presence–absence of each species to a set of environmental and spatial covariates, to the presence of con- and heterospecifics in the previous year and to the presence of heterospecifics in the same year. All species showed a clear tendency to continue breeding in the same sites where they were observed in the previous year. Additionally, the presence of heterospecifics, both in the previous year and in the same year, generally increased the probability that the focal species was found breeding on a given pond. Our data thus give support for the heterospecific attraction hypothesis, though causal inference should be confirmed with manipulative experiments.

Keywords: species distribution; multivariate logistic regression; waterbirds; artificial wetlands; social information; heterospecific attraction hypothesis

1. INTRODUCTION

A key challenge for better understanding distributional patterns of species with active habitat selection is to identify the cues that the individuals use to choose their breeding sites. The individuals may assess sites directly based on their environmental characteristics, such as availability of food, refuge from predators and a safe environment for nesting (Westphal *et al.* 2003). As an alternative, the individuals may also use social information, the presence of con- and heterospecifics indicating that the site is likely to be suitable for breeding (e.g. Stamps 1988; Leibold & Mikkelsen 2002; Ahlring & Faaborg 2006).

Heterospecific attraction is expected to lead to positive co-occurrence patterns among species, and such patterns may thus signal the presence of direct or indirect species interactions. Non-random associations in species distributions have been studied by comparing species co-occurrence patterns with those generated by appropriate null models (Gotelli 2000; Peres-Neto *et al.* 2001; Gotelli & McCabe 2002; Sfenthourakis *et al.* 2005). As species assemblages are affected by environmental conditions and a multitude of direct and indirect interactions (Diamond 1975), it is however difficult to

conclusively infer the causal links from observational data. Two species may co-occur especially often if they facilitate each other directly, but also if they simply share their habitat requirements. Similarly, the species may appear to avoid each other either if they show competitive exclusion, but also if they simply have dissimilar habitat requirements.

In the case of shared habitat requirements, individuals may use heterospecific presence as an indicator of patch quality (Mönkkönen & Forsman 2002; Forsman & Thomson 2008; Hromada *et al.* 2008). Other factors behind the heterospecific attraction hypothesis (HAH, Mönkkönen *et al.* 1990) include e.g. reduced predation risk in multi-species communities, increased social stimulus and improved mating opportunities. The HAH has been experimentally confirmed for a few bird and newt species (Mönkkönen & Forsman 2002; Pupin *et al.* 2007). While manipulative experiments provide the most conclusive evidence, they can be difficult to perform for many taxa, and they may fail to quantify the effect size under natural conditions. Thus, it would also be valuable to be able to address the HAH hypothesis with observational data.

In this paper, we consider as a case study spatio-temporal data on waterbirds breeding in a network of irrigation ponds. We attribute patterns of habitat occupancy to species-specific habitat requirements, to the presence of the same and of the other species in the previous year and to the presence of other species in the same

* Author for correspondence (esebastian@umh.es).

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2010.0244> or via <http://rspb.royalsocietypublishing.org>.

year. Technically, our modelling approach is based on Bayesian multivariate logistic regression (O'Brien & Dunson 2004), which has earlier been applied in the ecological context to snap-shot data on co-occurrence of wood-decaying fungi (Ovaskainen *et al.* in press). Here, we add to the set of explanatory variables the presence of the species in the previous year, thus extending the applicability of the model to time-series data.

Our study system consists of a waterbird community that uses irrigation ponds as their breeding habitat. These ponds were constructed to store the water coming from an inter-river water transfer. The network of irrigation ponds is surrounded by an unsuitable habitat matrix for waterbirds except for the presence of some natural wetlands, and the environmental context is thus conceptually similar to an archipelago of islands (De Meester *et al.* 2005; Céréghino *et al.* 2008). The system is highly dynamic partly because the irrigation ponds are regularly perturbed owing to management, suffering from abrupt changes in the water level, and being treated chemically with algacides added to the water. The colonization–extinction patterns of many species are known to be influenced by variation in environmental covariates, such as vegetation type, and by spatial covariates, such as patch location and size (Sánchez-Zapata *et al.* 2005; Sebastián-González *et al.* 2010a,b). The objective of this study was to extend the earlier species-specific results to a description of metacommunity dynamics. Our main focus is on the role of social information and heterospecific attraction in structuring the species community. As facilitative processes are expected to dominate over competitive processes in continuously perturbed environments (Bertness & Callaway 1994; Villarreal-Barajas & Martorell 2009), we predicted the predominance of positive co-occurrence patterns in this system. However, as some of the species have very similar niches and thus are likely to show strong competitive interactions, we predicted that the co-occurrence pattern could be negative for some pairs of species. The null hypothesis against which these predictions are tested is that community structure would be determined by the environmental conditions influencing each species independently of the other species, and so the co-occurrence patterns would be non-significant after accounting for the environmental and spatial patch characteristics.

2. MATERIAL AND METHODS

(a) *The empirical study system*

The study was conducted in the Vega Baja Valley in southeast Spain, an agricultural area of 95 840 ha dominated by citrus fruit trees, vegetables, palm trees and housing developments (Sánchez-Zapata *et al.* 2005). The climate is semiarid Mediterranean, with low annual rainfall (300 mm) and warm mean annual temperature (18°C).

In the study area, there are several seminatural wetlands. Some of these (Pantano de El Hondo, Lagunas de la Mata y Torrevieja, Salinas de San Pedro and Salinas de Santa Pola) are international SPAs or RAMSAR sites and have a regional protection status because of their importance for waterbirds (Ramsar Convention Secretariat 2006). An additional type of habitat for waterbirds is provided by a network of ca 2627 irrigation ponds (0.01–6.61 ha in area), which have been constructed since the 1980s. The ponds

and the wetlands form a well-connected system for mobile waterbirds (Cramp 1998).

(b) *Waterbird surveys*

Waterbirds were counted during the breeding season (between May and June, depending on the year) annually from 2002 to 2008 in a set of 221 irrigation ponds that were selected for the survey. The selected ponds are a random sample among the ponds that were available for the study, i.e. we excluded those that were located in fenced private areas or that we could not find in the field. The total census lasted three to four consecutive weeks, and we surveyed waterbirds during daylight. As the ponds are privately owned and fenced, access to all ponds was not possible during all years. Moreover, some ponds were destroyed or abandoned and dried during the study period. The number of yearly surveyed ponds varied between the minimum of 176 and the maximum of 221, being on average 194 ponds. Each pond was visited once a year. We stayed at the pond the time needed to assure a total bird count.

We focused on seven species of waterbirds that were known to breed commonly on the irrigation ponds (Sánchez-Zapata *et al.* 2005): little grebe (*Tachybaptus ruficollis*), black-winged stilt (*Himantopus himantopus*), little-ringed plover (*Charadrius dubius*), moorhen (*Gallinula chloropus*), common coot (*Fulica atra*), common shelduck (*Tadorna tadorna*) and mallard (*Anas platyrhynchos*). Black-winged stilt and little-ringed plover are migrants, whereas the remaining five species are residents in the system of irrigation ponds and the adjacent wetland areas. Little grebe and the common shelduck are the earliest species to show signs of breeding activity (the rest of the resident species start the breeding process more or less at the same time but are less conspicuous), whereas the black-winged stilt is the last species to arrive to breed in the pond system.

Presence–absence of these species was recorded using binoculars and scopes. Given the small size of the ponds (ranging between 0.01 and 6.61 ha, average size = 0.63 ha) and their poor vegetation cover (less than the 30% of the ponds presented vegetation at their shore), the survey error is likely to be small.

(c) *Pond characteristics*

We characterized the ponds using seven environmental variables that were known to affect the occurrence of this species community (Sánchez-Zapata *et al.* 2005; Sebastián-González *et al.* 2010a). As covariates, we used (i) the log-transformed pond area (A), (ii) the log-transformed distance to closest wetland (D), and (iii) the log-transformed connectivity to other ponds (S). As fixed factors, we used (iv) the pond construction design, (v) the presence of submerged vegetation, (vi) the presence of shore vegetation, and (vii) the presence of reed (*Phragmites australis*).

We used digitalized aerial photographs and a geographical information system (GRASS 5.0) to calculate pond areas and the distance of each pond to the closest wetland. The connectivity of pond i to the remaining ponds was computed as

$$S_i = \sum_{j \neq i} \sqrt{A_j} \left(\frac{d_{ij}}{1000} \right), \quad (2.1)$$

where the index j goes over all the known ponds ($n = 2627$) except the focal pond i , A_j is the area of pond j (in hectares)

Table 1. Components included (x) in the four versions of the multivariate logistic regression models for predicting the occurrence of the focal species in year t .

model	environmental and spatial covariates	survey year as a random effect	presence of all species in year $t - 1$	presence of heterospecifics in year t
M1(I)	x	x		
M1(R)	x	x		x
M2(I)	x	x	x	
M2(R)	x	x	x	x

and d_{ij} is the distance between ponds i and j (in metres). The mean value of log-transformed connectivity was 7.1, with range in 4.9–7.7.

We distinguished two types of pond according to their construction design: low-density polyethylene (LDP) and high-density polyethylene (HDP). LDP ponds are covered by a layer of gravel to protect the plastic from solar radiation, providing the pond with a more natural appearance than in the case of HDP ponds.

We considered three vegetation types: shore vegetation (vegetation without a direct contact with the water), submerged vegetation (vegetation inside the water) and reed (*P. australis*) vegetation. Reed vegetation was considered separately because it has been considered important for providing breeding and refuge sites for the moorhen and the common coot (Cramp 1998). Each of the three vegetation types was classified as present or absent, using a minimum surface cover of 10 per cent as a threshold.

(d) Multivariate logistic regression model

We analysed the data using Bayesian multivariate logistic regression (O'Brien & Dunson 2004; Ovaskainen *et al.* in press) that models the probability of a given combination of species found from a given sampling unit. Using the index i for the n sampling units and the index j for the m species, the model can be formulated as

$$\left. \begin{aligned} y_{ij} &= 1(z_{ij} > 0), \\ z_{ij} &= \mu_{ij} + \text{logit}[F(e_{ij})], \\ \mu &= \mathbf{X}\boldsymbol{\beta}, \\ \text{and } \mathbf{e}_i &\sim N(\mathbf{0}, \mathbf{R}) \end{aligned} \right\}, \quad (2.2)$$

where the data $\mathbf{y}$ are arranged as an $n \times m$ matrix, $\mathbf{X}$ is the $n \times k$ design matrix (where k is the number of covariates used as explanatory variables), $\boldsymbol{\beta}$ is the $k \times m$ matrix of regression coefficients and the $m \times m$ correlation matrix $\mathbf{R}$ describes whether each species pair co-occurs on the same sampling unit more or less often than expected by random. F is the cumulative density function of univariate normal distribution with mean 0 and variance 1, and the link function is defined by $\text{logit}(x) = \log(x/(1-x))$. The function $1(x > 0)$ is the indicator function with value 1 if $x > 0$ and with value 0 otherwise.

The technical description of the model (equation (2.2)) may look somewhat dissimilar from the usual notation of logistic regression,

$$\text{logit}[\text{Pr}(y_{ij} = 1)] = \mu_{ij}. \quad (2.3)$$

However, for the univariate case, equation (2.2) is identical to equation (2.3). Further, equation (2.2) is a natural multivariate generalization of equation (2.3) in the sense that the marginal distribution for each species follows the

corresponding univariate logistic regression (O'Brien & Dunson 2004).

The explanatory variables were categorized into three types, providing information (i) on the environmental and spatial covariates, (ii) the presence of conspecifics and heterospecifics in the previous year, and (iii) the presence of heterospecifics in the same year. We fitted to the data four versions of the model, with differing sets of explanatory variables included (table 1). All of the models included the environmental and spatial covariates. All models also included the survey year as a random factor, with covariance structure specified by a species-to-species variance–covariance matrix $\mathbf{V}$. Model M1(I) includes only these components, and it assumes independence among species by replacing the correlation matrix $\mathbf{R}$ by the identity matrix $\mathbf{I}$ (matrix with ones at the diagonal and zeros elsewhere). Model M1(R) is a multivariate model in which the correlation matrix $\mathbf{R}$ was fitted to the data. Models M2(I) and M2(R) are otherwise similar to models M1(I) and M1(R), but they use information on the occurrence pattern in the previous year to predict the occurrence pattern for the following year. More precisely, these models include as additional predictors (fixed factors) the presence of each species in that pond in the previous year.

Models M1(I) and M2(I) can be considered as environmentally constrained null models (Peres-Neto *et al.* 2001), which do not account for non-random co-occurrence patterns unlike models M1(R) and M2(R). Models M1(I) and M1(R) view the data from each year as dynamically independent snapshots. Models M2(I) and M2(R) are time-series models that include the presences of all species in the previous year as explanatory variables.

We assumed uniform priors for the regression coefficients $\boldsymbol{\beta}$, a marginally uniform prior (Barnard *et al.* 2000) for the correlation matrix $\mathbf{R}$ and the Wishart prior $W(\mathbf{I}_m, m+1)$ for the $m \times m$ variance–covariance matrix $\mathbf{V}$. The posterior distributions of the four models were sampled using the MCMC procedure described in Ovaskainen *et al.* (in press).

(e) Model fit

To assess whether the structural assumptions of the model were in line with the data, we examined the fit of the four model variants by comparing model predictions (posterior predictive data) against the real data. As our main interest is in patterns of species co-occurrence, we evaluated the model fit in terms of predicted number of ponds with no birds, with a single bird species and with two or more bird species. As the data come from an irregular patch network covering a large region, we also examined the possibility of spatially correlated residuals. To do so, we manually split the entire patch network into 10 sub-networks with the aim

Figure 1. An example of (a) observed versus (b) predicted distribution. The intensity of the colour refers to the number of years during which the pond was recorded occupied, white corresponding to no observations, black to an observation in all years. Figure shown for little grebe, prediction by model M1(I). For a similar comparison of the other species, see electronic supplementary material, figures S3–S8. The numbers refer to the 10 sub-networks that were used to compare the observed and predicted prevalence (table 5 and figure 2).

Table 2. The effect of the spatial and environmental covariates in model M1(R). The values show the median estimates of the regression coefficients. Cases for which the posterior probability for a positive (or negative) response was at least 95% are shown bold. The coefficient for material corresponds to the effect of LDP (the effect of HDP set to zero), and the coefficients for submerged, shore and reed vegetation correspond to the presence of that vegetation type. Log-transformed variables are indicated with an asterisk.

species	A* (ha)	D* (m)	S* (m)	material	submerged	shore	reed
little grebe	0.270	−0.626	−0.407	0.969	1.331	0.205	0.497
black-winged stilt	0.492	−0.706	−0.390	1.320	0.706	−0.132	−0.697
little-ringed plover	0.424	−0.204	−0.360	0.608	0.687	−0.506	−0.919
moorhen	0.068	−0.583	0.155	0.680	1.270	0.348	2.474
mallard	0.545	−0.669	0.104	1.178	0.987	0.448	−0.00
common shelduck	0.944	−0.661	−0.806	0.371	0.057	−0.051	−0.404
common coot	0.179	−0.276	−0.052	1.579	1.277	0.518	1.538

that they would be semi-independent (figure 1), and compared the observed prevalence with the predicted prevalence for each sub-network.

3. RESULTS

(a) Effect of the spatial and environmental covariates

As the marginal distributions of the multivariate model are identical to those of univariate models (see above), the results relating to effects of the spatial and environmental covariates are identical for models M1(I) and M1(R). In these models, the occurrence probabilities of all species increased with increasing pond size and increasing amount of submerged vegetation, and they were higher for ponds constructed with the LDP material than the HDP material (table 2). The distance to the closest wetland had a negative effect, i.e. ponds close to wetlands had birds with a higher frequency than ponds far from the wetlands. Somewhat surprisingly,

connectivity had generally a negative effect, meaning that birds were most often recorded in those parts of the network in which pond density was low. The responses to the presence of shore vegetation and to the presence of reed varied among the species.

The random effect survey year models any variation (e.g. owing to changing environmental conditions) on top of the fluctuations explained by the model. We did not find a high statistical support for species occurrence covarying with time for any of the species pairs (electronic supplementary material, table S1). Thus, residual fluctuations were not in synchrony between the species.

(b) Effect of the species community in the previous year

The diagonal values of table 3 are positive for all species, thus a species found from a given pond in one year was likely to be found there also the next year. Also the presence of other species in the previous year generally

Table 3. The effect of the presence of the species in the previous year to the presence of the species in the focal year, based on model M2(**R**). Symbols as in table 2.

species	presence at $t - 1$						
	lg	bws	lrp	mh	m	cs	cc
presence at t							
little grebe (lg)	1.477	0.439	-0.076	0.315	0.217	0.289	0.440
black-winged stilt (bws)	0.468	1.763	0.405	0.347	0.092	1.195	0.574
little-ringed plover (lrp)	0.221	0.654	1.359	-0.500	-1.147	1.082	-0.141
moorhen (mh)	0.985	0.508	-0.048	1.182	0.062	0.418	0.569
mallard (m)	0.487	0.539	0.472	0.978	0.862	0.044	1.348
common shelduck (cs)	0.624	1.448	0.172	0.088	-0.162	1.375	0.654
common coot (cc)	0.899	0.476	0.185	1.164	0.262	-0.613	2.491

Table 4. Species co-occurrence as measured by the correlation matrix **R**. The lower diagonal represents the correlation coefficients (elements of the matrix **R**) for model M1(**R**), which includes as explanatory variables the environmental and spatial covariates and survey year. Upper diagonal represents model M2(**R**), which additionally includes the presence of all species in the previous year. A positive value indicates that the two species co-occur more often than expected by random, a negative value the opposite case. Symbols as in table 2.

species	lg	bws	lrp	mh	m	cs	cc
little grebe (lg)	1	0.33	0.07	0.28	0.19	0.18	0.46
black-winged stilt (bws)	0.43	1	0.42	-0.05	0.31	0.29	0.37
little-ringed plover (lrp)	0.17	0.46	1	-0.01	0.06	0.06	-0.15
moorhen (mh)	0.32	0.08	-0.04	1	0.09	-0.01	0.25
mallard (m)	0.33	0.41	0.05	0.23	1	0.13	0.17
common shelduck (cs)	0.32	0.50	0.24	0.10	0.25	1	0.15
common coot (cc)	0.50	0.34	-0.15	0.46	0.34	0.21	1

increased the occurrence probability of the focal species in the next year. This was especially the case with the little grebe and the black-winged stilt, the presence of which in year $t - 1$ had a positive effect on the occurrence probability of most of the other species in the year t . As an exception to the generally positive heterospecific effects, the presence of mallard in the previous year made it less likely that little-ringed plover would appear in the same pond the next year. This effect was asymmetric, as the presence of the little-ringed plover in the previous year did not influence the occurrence probability of the mallard in the following year.

Including the species information at year $t - 1$ in the model somewhat changed the estimated effects of the environmental and spatial covariates (electronic supplementary material, table S2). Models M1 and M2 largely agreed about the direction of a given effect, but model M2 generally predicted a smaller effect size. This was expected, as information on the presence of the species in the previous year takes away some of the explanatory power from the environmental and spatial covariates.

(c) Within-year co-occurrence among the species

Different bird species generally occurred together with a higher frequency than expected by random (table 4). Out of the 21 species pairs, we found in model M1(**R**) strong statistical support (posterior probability at least 95%) for a positive co-occurrence for 16 species pairs. The only case for which there was strong statistical support for a negative association consists of the species pair little-ringed plover and common coot, which

occurred together less often than expected by random. The inclusion of the presence of the species in the previous year as explanatory factors (model M2(**R**)) explained away some of the within-year co-occurrence patterns. However, strong statistical support for additional positive (negative) within-year co-occurrence was still found for 11 (1) species pairs.

(d) Model fit

Models M1(**I**) and M2(**I**) that do not account for within-year co-occurrence underestimated the number of ponds with no birds, overestimated the number of ponds with one bird species and somewhat underestimated the number of ponds with two or more species (figure 2). This reflects the finding that most species pairs tended to co-occur more often than expected by random (table 4). As expected, the bias was largely removed in models M1(**R**) and M2(**R**), as these models include a measure of species-to-species co-occurrence. However, even these models somewhat overestimated the number of ponds with a single bird and underestimated the number of ponds with two or more species, suggesting the presence of some higher order associations.

To examine whether the data were affected by spatial factors not included in the model, we compared the observed and predicted prevalence within sub-networks (figure 1). For all the species, there was much spatial variation in the raw data (table 5). Model M1 captures more than half of this variation, and thus region-specific variation in prevalence can be to a large extent explained by the environmental and spatial covariates. The exception is the little-ringed plover, for which the prediction

Figure 2. Fraction of ponds in which zero, one or more than one species were present. The open symbols depict the real data, the filled ones with error bars the median and 95% credibility intervals in posterior predictive data. The four cases (from left to right) correspond to models M1(I), M1(R), M2(I) and M2(R), respectively.

of model M1 does not explain a significant part of variation among the sub-networks. However, for most species, some systematic deviations remain (figure 3). The sub-network-specific residuals were positive with 95 per cent credibility for one or more species in three sub-networks, while the opposite was true in five sub-networks (figure 3). Model M2 performs much better than model M1 in predicting the spatial patterns (table 5 and figure 3). This is not surprising, as it uses knowledge about the distribution of the species in the previous year.

4. DISCUSSION

The aim of our work was to use observational data to assess the HAH. Based on our results, the community of waterbirds has a very strong tendency to aggregate at the very same ponds, both within a year and between years. Such a result could follow simply from overall variation in habitat quality, the birds aggregating at the highest quality ponds. However, our results are based on a model that accounts for all environmental and spatial covariates that we expected to be of major importance (Cramp 1998; Sánchez-Zapata *et al.* 2005; Abellán *et al.* 2006; Arocena 2007; Sebastián-González *et al.* 2010a). Further, the within-year aggregation pattern was evident also in model M2(R), which uses information on the species presence in the previous year, and thus includes an accurate assessment of habitat quality from the species point of view. Thus, we conclude that our results give rather strong support for the HAH hypothesis.

Waterbirds may obtain several benefits from settling in sites already occupied by heterospecifics (see review in Mönkkönen & Forsman 2002). First, the presence of other species might provide protection against predators, by means of communal defence or alarming behaviour. For example, the black-winged stilt displays noisy alarm calls in the presence of predators that probably act as warning signals also for the other species. Second, heterospecific attraction can lead to an aggregated distribution also within species, increasing the number of encounters and thus facilitating the mating process. Third, the presence of residents can be a signal of low risk of predation, as the residents may already have been predated in patches with a higher predation risk. Heterospecifics can also act as indicators of habitat

Figure 3. Residual prevalence (predicted minus observed) in each of the 10 sub-networks shown in figure 1. Each dot corresponds to one species. Cases in which the real value was outside the 95% credibility interval of posterior predictive data are shown by black. (a) Model M1(I); (b) model M2(I).

Table 5. Variation in prevalence among the 10 sub-networks of figure 1. The first two columns show the mean and standard deviation (among the sub-networks) of species prevalence (fraction of occupied ponds). The latter two columns show the amount of variation explained in linear regression in which sub-network-specific prevalence is explained by the median of posterior predictive data. Cases in which the predictor was significant ($p < 0.05$) are shown in bold. Based on M1(I) and M2(I), but species-specific predictions for M1(R) and M2(R) are identical (§2).

species	mean	sdv	R^2 (M1)	R^2 (M2)
little grebe	0.23	0.05	0.62	0.83
black-winged stilt	0.24	0.05	0.72	0.88
little-ringed plover	0.14	0.02	0.08	0.56
moorhen	0.07	0.02	0.68	0.89
mallard	0.09	0.02	0.55	0.77
common shelduck	0.06	0.02	0.53	0.88
common coot	0.05	0.02	0.63	0.87

quality, especially for migrant species arriving to breeding sites (Thomson *et al.* 2003; Forsman & Thomson 2008; Hromada *et al.* 2008; Forsman *et al.* 2009). Migrants need to make fast but reliable decisions about where to breed, whereas residents have had much more time to evaluate patch quality.

Among the waterbird community studied here, the little grebe is the most widespread resident species in the irrigation ponds (Sánchez-Zapata *et al.* 2005; Sebastián-González *et al.* 2010a,b). Consistent with the

above considerations, the little grebe showed a positive co-occurrence with all the other species. The common shelduck prefers the natural wetlands for wintering, and returns to the ponds for breeding before the migrant black-winged stilt [E. Sebastián-González, T. A. Sánchez-Zapata & F. Botella, 2002–2009, unpublished data]. The high positive co-occurrence between these two species suggests that the black-winged stilt may use the common shelduck as an indicator of patch quality.

Heterospecific attraction is likely to be an advantage only if competition between the species is sufficiently low (Mönkkönen *et al.* 1999). Classical niche theory (for a review, see Chase & Leibold 2003) predicts that coexistence is possible if intraspecific competition is stronger than interspecific competition, requiring a partitioning of the niche space (MacArthur & Levins 1967; Tilman 1982). Among the waterbirds included in this study, the little grebe is the only diving bird, feeding mainly in the middle of the water column (Cramp 1998). Therefore, it does not compete for the resources with the other species, giving an additional explanation why its high co-occurrence with all other species is feasible. However, contrary to the expectation that coexistence would be possible mainly for species using different prey types or prey sizes (e.g. Nudds *et al.* 1994, 2000; Pöysä *et al.* 1996), also species with similar foraging and breeding niches frequently appeared together at the ponds. For example, both the black-winged stilt and the little-ringed plover feed on invertebrates and locate their nests at the shore of the ponds (Cramp 1998), but their co-occurrence was still much higher than what would be expected by random. This may indicate that resource (prey and nest sites) availability is sufficiently high, as in such a situation a species' presence is not limited by the presence of its competitors (Chesson 2000).

Facilitative processes are likely to be more important than competitive processes, in continuously perturbed environments (Bertness & Callaway 1994; Villarreal-Barajas & Martorell 2009), as is the case for the majority of the irrigation ponds. Variation in disturbance level may also explain that the somewhat surprising result that the prevalence of most species was highest in the sparse parts of the network. Ponds in areas of high pond density are likely to be visited, and thus disturbed, by a larger number of pond owners than in the case of isolated ponds. As distances between all irrigation ponds are small compared with the dispersal capacity of the waterbirds, low pond density may not be a problem from the dispersal point of view.

Hierarchical modelling approaches provide a flexible framework for partitioning variation in complex and structured datasets, and they are becoming increasingly popular in the ecological literature (Cressie *et al.* 2009). In the present case, the application of the Bayesian multivariate logistic regression model allowed us to assess the dynamics of the entire community in a single model, giving a compact description of the species-to-species relationships both between and within the years. While our results strongly support the HAH hypothesis, we recall that they are based solely on observational data, and must thus be interpreted with caution. As the next step, the results of this study can be used to design manipulative experiments to

examine the processes behind the patterns reported here. For example, our results suggest that the migrant black-winged stilt uses the resident common shelduck as an indicator of habitat quality. This could be tested by placing artificial ducks (imitating the shape and colours of the common shelduck) at a randomized subset of ponds before the breeding season, and by monitoring whether these ponds are favoured by the black-winged stilt over the control ponds.

The co-occurrence patterns in the data are generated by similar (or dissimilar) responses to the shared environment, and by other factors such as direct and indirect interactions among the species. It is clear that our model includes only some of the relevant environmental and spatial covariates, the effects of other covariates remaining in the model residual. For example, we did not include any characteristics of the habitats surrounding the ponds, such as the presence of rabbit burrows, which can be used for breeding for the common shelduck (Urios *et al.* 1991). Further, as the data come from a spatial context, unsurprisingly the model residuals showed a level of spatial correlation. Such correlations can be caused by the dispersal process or by unmeasured spatially correlated environmental variation not included in our model. Nevertheless, we argue that the presence of the focal species in the previous year is an accurate and integrative species-specific measure of habitat quality. Thus, our strongest argument for the results supporting the HAH hypothesis is that higher than random co-occurrence of many species pairs was found also in the model that contained as an explanatory variable the presence of the species in the previous year.

This study was supported by the Generalitat Valenciana (the I + D project CTIDIB/2002/142), the Conselleria de Territori i Habitatge, the Spanish Ministry of Education (FPU grant to E.S.-G.), the Academy of Finland (grant no. 124242 to O.O.) and the European Research Council (ERC Starting grant no. 205905 to O.O.). R. A. Sempere, J. D. Anadón, M. Carrete, C. Villacorta, C. Díaz, J. Navarro, M. A. Richarte, M. Blázquez, M. D. Antón, I. González, J. Aliaga, R. Ballestar and S. Bordonado helped with the fieldwork.

REFERENCES

- Abellán, P., Sánchez-Fernández, D., Millán, A., Botella, F., Sánchez-Zapata, J. A. & Giménez, A. 2006 Irrigation ponds as macroinvertebrate habitat in a semi-arid agricultural landscape (SE Spain). *J. Arid Environ.* **67**, 255–269. (doi:10.1016/j.jaridenv.2006.02.009)
- Ahlering, M. A. & Faaborg, J. 2006 Avian habitat management meets conspecific attraction: if you build it, will they come? *Auk* **123**, 301–312. (doi:10.1642/0004-8038(2006)123[301:AHMMCA]2.0.CO;2)
- Arocena, R. 2007 Effects of submerged aquatic vegetation on macrozoobenthos in a coastal lagoon of the southwestern Atlantic. *Int. Rev. Hydrobiol.* **92**, 33–47. (doi:10.1002/iroh.200610881)
- Barnard, J., McCulloch, R. & Meng, X. L. 2000 Modeling covariance matrices in terms of standard deviations and correlations, with application to shrinkage. *Stat. Sin.* **10**, 1281–1311.
- Bertness, M. D. & Callaway, R. 1994 Positive interactions in communities. *Trends Ecol. Evol.* **9**, 191–193. (doi:10.1016/0169-5347(94)90088-4)

- Cressie, N., Calder, C. A., Clark, J. S., Hoef, J. M. V. & Wikle, C. K. 2009 Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. *Ecol. Appl.* **19**, 553–570. (doi:10.1890/07-0744.1)
- Cérèghino, R., Biggs, J., Oerti, B. & Declerck, S. 2008 The ecology of European ponds: defining the characteristics of a neglected freshwater habitat. *Hydrobiologia* **597**, 1–6. (doi:10.1007/s10750-007-9225-8)
- Chase, J. M. & Leibold, M. A. 2003 *Ecological niches. Linking classical and contemporary approaches*. Chicago, IL: University of Chicago Press.
- Chesson, P. 2000 General theory of competitive coexistence in spatially-varying environments. *Theor. Popul. Biol.* **58**, 211–237. (doi:10.1006/tpbi.2000.1486)
- Cramp, S. 1998 *The complete birds of the western Palearctic in CD-room*. London, UK: Oxford University Press.
- De Meester, L., Declerck, S., Stoks, R., Louette, G., van de Meutter, F., de Brie, T., Michels, E. & Brendonck, L. 2005 Ponds and pools as model systems in conservation biology, ecology and evolutionary biology. *Aquat. Conserv. Mar. Freshw. Ecosyst.* **15**, 715–725. (doi:10.1002/aqc.748)
- Diamond, J. M. 1975 Assembly of species communities. In *Ecology and evolution of communities* (eds M. L. Cody & J. M. Diamond), pp. 342–444. Cambridge, MA: Harvard University Press.
- Forsman, J. T. & Thomson, R. L. 2008 Evidence of information collection from heterospecifics in cavity-nesting birds. *Ibis* **150**, 409–412. (doi:10.1111/j.1474-919X.2007.00773.x)
- Forsman, J. T., Hjernquist, M. B. & Gustafsson, L. 2009 Experimental evidence for the use of density based interspecific social information in forest birds. *Ecography* **32**, 539–545. (doi:10.1111/j.1600-0587.2008.05635.x)
- Gotelli, N. J. 2000 Null model analysis of species co-occurrence patterns. *Ecology* **81**, 2606–2621. (doi:10.1890/0012-9658(2000)081[2606:NMAOSC]2.0.CO;2)
- Gotelli, N. J. & McCabe, D. L. 2002 Species co-occurrence: a metaanalysis of J. M. Diamond's assembly rules. *Ecology* **83**, 2091–2096. (doi:10.1890/0012-9658(2002)083[2091:SCOAMA]2.0.CO;2)
- Hromada, M., Antczak, M., Valone, T. J. & Tryjanowski, P. 2008 Settling decisions and heterospecific social information use in shrikes. *PLoS ONE* **3**, e3930. (doi:10.1371/journal.pone.0003930)
- Leibold, M. A. & Mikkelsen, G. M. 2002 Coherence, species turnover, and boundary clumping: elements of meta-community structure. *Oikos* **97**, 237–250. (doi:10.1034/j.1600-0706.2002.970210.x)
- Mac Arthur, R. & Levins, R. 1967 The limiting similarity, convergence, and divergence of coexisting species. *Am. Nat.* **101**, 377–385. (doi:10.1086/282505)
- Mönkkönen, M. & Forsman, J. T. 2002 Heterospecific attraction among forest birds: a review. *Ornithol. Sci.* **1**, 41–51. (doi:10.2326/osj.1.41)
- Mönkkönen, M., Helle, P. & Soppela, K. 1990 Numerical and behavioural responses of migrant passerines to experimental manipulation of resident tits (*Parus* spp.): heterospecific attraction in northern breeding bird communities? *Oecologia* **85**, 218–255. (doi:10.1007/BF00319404)
- Mönkkönen, M., Härdling, R., Forsman, J. T. & Tuomi, J. 1999 Evolution of heterospecific attraction: using other species as cues in habitat selection. *Evol. Ecol.* **13**, 91–104. (doi:10.1023/A:1006590215306)
- Nudds, T. D., Sjöberg, K. & Lundberg, P. 1994 Ecomorphological relationships among Palearctic dabbling ducks on Baltic coastal wetlands and a comparison with the Nearctic. *Oikos* **69**, 295–303. (doi:10.2307/3546150)
- Nudds, T. D., Elmberg, J., Sjöberg, K., Pöysä, H. & Nummi, P. 2000 Ecomorphology in breeding Holarctic dabbling ducks: the importance of lamellar density and body length varies with habitat type. *Oikos* **91**, 583–588. (doi:10.1034/j.1600-0706.2000.910321.x)
- O'Brien, S. M. & Dunson, D. B. 2004 Bayesian multivariate logistic regression. *Biometrics* **60**, 739–746. (doi:10.1111/j.0006-341X.2004.00224.x)
- Ovaskainen, O., Hottola, J. & Siitonen, J. In press. Modelling species co-occurrence by multivariate logistic regression generates new hypotheses on fungal interactions. *Ecology*. (doi:10.1890/10-0173.1)
- Peres-Neto, P. R., Olden, J. D. & Jackson, D. A. 2001 Environmentally constrained null models: site suitability as occupancy criterion. *Oikos* **93**, 110–120. (doi:10.1034/j.1600-0706.2001.930112.x)
- Pöysä, H., Elmberg, J., Nummi, P. & Sjöberg, K. 1996 Are ecomorphological associations among dabbling ducks consistent at different spatial scales? *Oikos* **76**, 608–612. (doi:10.2307/3546358)
- Pupin, F., Sacchi, R., Gentili, A., Galeotti, P. & Fasola, M. 2007 Discrimination of toad calls by smooth newts: support for the heterospecific attraction hypothesis. *Anim. Behav.* **74**, 1683–1690. (doi:10.1016/j.anbehav.2007.03.020)
- Ramsar Convention Secretariat. 2006 *The Ramsar Convention manual: a guide to the Convention on Wetlands*, 4th edn. Gland, Switzerland: Ramsar Convention Secretariat.
- Sánchez-Zapata, J. A., Anadón, J. D., Carrete, M., Giménez, A., Navarro, J., Villacorta, C. & Botella, F. 2005 Breeding waterbirds in relation to artificial pond attributes: implications for the design of irrigation facilities. *Biodivers. Conserv.* **14**, 1627–1639. (doi:10.1007/s10531-004-0534-1)
- Sebastián-González, E., Sánchez-Zapata, J. A. & Botella, F. 2010a Agricultural ponds as alternative habitat for waterbirds: spatial and temporal patterns of abundance and management strategies. *Eur. J. Wildl. Res.* **56**, 11–20. (doi:10.1007/s10344-009-0288-x)
- Sebastián-González, E., Botella, F., Paracuellos, M. & Sánchez-Zapata, J. A. 2010b Processes driving temporal dynamics in the nested pattern of waterbird communities. *Acta Oecol.* **36**, 160–165. (doi:10.1016/j.actao.2009.11.006)
- Sfenthourakis, S., Tzanatos, E. & Giokas, S. 2005 Species co-occurrence: the case of congeneric species and a causal approach to patterns of species association. *Global Ecol. Biogeogr.* **15**, 39–49. (doi:10.1111/j.1466-822X.2005.00192.x)
- Stamps, J. A. 1988 Conspecific attraction and aggregation in territorial species. *Am. Nat.* **131**, 329–347. (doi:10.1086/284793)
- Thomson, R. L., Forsman, J. T. & Mönkkönen, M. 2003 Positive interactions between migrant and resident birds: testing the heterospecific attraction hypothesis. *Oecologia* **134**, 431–438. (doi:10.1007/s00442-002-1140-0)
- Tilman, D. 1982 *Resource competition and community structure*. Princeton, NJ: Princeton University Press.
- Uriós, V., Escobar, J. V., Pardo, R. & Goómez, J. A. 1991 *Atlas de aves nidificantes de la Comunidad Valenciana*. Valencia, Spain: Conselleria de Agricultura y Pesca.
- Villarreal-Barajas, T. & Martorell, C. 2009 Species-specific disturbance tolerance, competition and positive interactions along an anthropogenic disturbance gradient. *J. Veg. Sci.* **20**, 1027–1040. (doi:10.1111/j.1654-1103.2009.01101.x)
- Westphal, M. I., Field, S. A., Tyre, A. J., Paton, D. & Possingham, H. P. 2003 Effects of landscape pattern on bird species distribution in the Mt. Lofty Ranges, South Australia. *Landsc. Ecol.* **18**, 413–426. (doi:10.1023/A:1026115807529)