Abundance, spatial variance and occupancy: arthropod species distribution in the Azores

KEVIN J. GASTON*, PAULO A. V. BORGES†, FANGLIANG HE‡ and CLARA GASPAR*†

Biodiversity and Macroecology Group, Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK; †University dos Açores, Department de Ciências Agrárias, Terra-Chã, PT-9700-851, Angra do Heroísmo, Terceira, Açores, Portugal; and ‡Department of Renewable Resources, University of Alberta, Edmonton, Alberta, Canada T6G 2H1

Summary

1. The positive abundance–occupancy and abundance–variance relationships are two of the most widely documented patterns in population and community ecology.

2. Recently, a general model has been proposed linking the mean abundance, the spatial variance in abundance, and the occupancy of species. A striking feature of this model is that it consists explicitly of the three variables abundance, variance and occupancy, and no extra parameters are involved. However, little is known about how well the model performs.

3. Here, we show that the abundance–variance–occupancy model fits extremely well to data on the abundance, variance and occupancy of a large number of arthropod species in natural forest patches in the Azores, at three spatial extents, and distinguishing between species of different colonization status. Indeed, virtually all variation about the bivariate abundance–occupancy and abundance–variance relationships is effectively explained by the third missing variable (variance in abundance in the case of the abundance–occupancy relationship, and occupancy in the case of the abundance–variance relationship).

4. Introduced species tend to exhibit lower densities, less spatial variance in these densities, and occupy fewer sites than native and endemic species. None the less, they all lie on the same bivariate abundance–occupancy and abundance–variance, and trivariate abundance–variance–occupancy, relationships.

5. Density, spatial variance in density, and occupancy appear to be all the things one needs to know to describe much of the spatial distribution of species.

Key-words: arthropods, metapopulation dynamics, occupancy, rescue effect, spatial distribution, spatial variance, Taylor's power law.

Introduction

Two of the most general patterns concerning fundamental macroecological variables are the bivariate intraspecific/interspecific abundance–occupancy and abundance–variance relationships. The first is a positive relationship between the average local abundance (μ) of a species and the proportion of available sites (p) at which it occurs (its probability of occurrence; Hanski 1982; Brown 1984; Gaston 1996; Gaston *et al*. 2000). This has been widely documented, both intra- and interspecifically, at a range of spatial scales, and has received particular attention in the contexts of metapopulation dynamics (e.g. Hanski 1991a,b; Gyllenberg & Hanski 1992; Hanski & Gyllenberg 1993; Hanski, Kouki & Halkka 1993), agricultural entomology (Nachman 1981, 1984; Kuno 1986, 1991; Ward *et al*. 1986; Ekbom 1987; Perry 1987; Hepworth & MacFarlane 1992; Feng, Nowierski & Zeng 1993), and conservation biology (e.g. Gaston 1999; Rodrigues, Gaston & Gregory 2000). The second pattern is a positive relationship between the average local abundance of a species across sites and the spatial variance in that abundance (σ^2). While the intraspecific variant of this relationship has attracted by far the majority of attention, it has also been well documented interspecifically, and been found to be manifest at both small and large spatial scales (Taylor 1961, 1984;

Correspondence: Kevin J. Gaston, Biodiversity and Macroecology Group, Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, U.K. Tel. 0114 2220030. Fax: 0114 2220002. E-mail: k.j.gaston@sheffield.ac.uk

Taylor, Woiwod & Perry 1978, 1979, 1980; Perry 1981; Maurer 1994).

In largely distinct literatures, a number of models have been proposed as descriptors of abundance–occupancy and of abundance–variance relationships, respectively, and there has been much discussion of the determinants of the patterns (e.g. Anderson *et al*. 1982; Brown 1984; Binns 1986; Gillis, Kramer & Bell 1986; Perry 1988; Gaston, Blackburn & Lawton 1997; Holt *et al*. 1997; Gaston *et al*. 2000; Harte, Blackburn & Ostling 2001). In both cases it has been observed that explanations rooted in, among others, population demographics, individual behaviour and niche structure are capable of generating empirical patterns similar to those observed. Indeed, given the variety of ways in which they can be explained (often not mutually exclusive), arguably the two relationships capture essential fundamentals of the structuring of the distributions of species.

Although the abundance–occupancy and abundance–variance relationships share abundance as a common currency, until recently there had been little attempt to explore the connection between them. However, He & Gaston (2003) have proposed a general model linking mean abundance, spatial variance in that abundance, and occupancy. It takes as a starting point separate general models for each of the two bivariate relationships. The first one is Taylor's power law for the abundance–variance relationship

$$\sigma^2 = a\mu^b \qquad \text{eqn 1}$$

where *a* and *b* are constants (Taylor 1961). The second model, for the abundance–occupancy relationship, takes the form

$$p = 1 - \left(1 + \frac{\mu}{k}\right)^{-k} \qquad \text{eqn 2}$$

where *k* is a spatial aggregation parameter defined in the domain of $(-\infty, -\mu)$ or $(0, \infty)$ and can be expressed as

$$k = \frac{\mu^2}{\sigma^2 - \mu} \qquad \text{eqn 3}$$

(He & Gaston 2000, 2003; Holt, Gaston & He 2002). Substituting eqn 3 into eqn 2, and recognizing that σ^2 is defined by eqn 1, gives a general model linking mean abundance, spatial variance in that abundance, and occupancy that takes the form

$$p = 1 - \left(\frac{\mu}{\sigma^2}\right)^{\mu^2/\sigma^2 - \mu} \qquad \text{eqn 4}$$

where $\sigma^2 \neq \mu$ but can infinitely approach μ, resulting in $p = 1 - e^{-\mu}$, which is occupancy for the Poisson distribution (He & Gaston 2003; for other derivations of this model see also Wilson & Room 1983; Yamamura 2000).

Using this formulation (henceforth the abundance–variance–occupancy model), He & Gaston (2003) showed

that the abundance–occupancy model (2) and the abundance–variance model (1) can predict each other and are therefore different expressions of the same phenomenon. This suggests that species distributions as described by abundance–occupancy and abundance–variance relationships can be deduced from a single, well-justified model. The abundance–variance model (1) can be equivalently characterized by the abundance–variance–occupancy model (4) by assuming occupancy is a constant, although the predicted abundance–variance relationship does not strictly lead to Taylor's power law (He & Gaston 2003); the predicted abundance–variance relationship exhibits curvilinearity on log-log axes, particularly at low densities, as expected when density is negative binomially distributed (Routledge & Swartz 1991; McArdle & Gaston 1995). The abundance–occupancy model (2) can be characterized by the abundance–variance–occupancy model (4) by assuming that variance σ^2 is a constant. However, while much has been learned about the two bivariate models (1 and 2), little is known about how the trivariate abundance–variance–occupancy model (4) performs.

A striking feature of the abundance–variance–occupancy model is that it consists explicitly of the three variables (abundance, variance and occupancy), and no extra parameters are involved. This suggests that the spatial distribution of species of any kind (as measured by abundance, variance and occupancy) are uniquely described by the trivariate relationship. In other words, no matter what the species assemblage (e.g. endemic or introduced species, plants or animals) or the spatial scales at which distribution data are sampled, abundances, variances and occupancies should all fall on a common, unique three-dimensional space. This is an extremely general prediction, implying that abundance, spatial variance in abundance, and occupancy are all that one needs to know to describe much of the spatial distribution of species (although, obviously, not the spatial relations of localities). Here we test this prediction using data on the abundance and occurrence of a large number of arthropod species in natural forest patches in the Azores, at three spatial extents, and distinguishing between species of different colonization status. We are particularly interested in testing if the abundance–variance–occupancy model accurately predicts species distribution across different spatial scales [i.e. whether species collected at different scales fall on the same three-dimensional surface of model (4)] and whether endemic, native (nonendemic) and introduced species occupy different parts of abundance–variance–occupancy space.

Methods

AREA OF STUDY

The Azorean archipelago (North Atlantic; 37–40°N, 25–31°W) comprises nine main islands and some small islets. Aligned on a WNW–ESE axis, these extend for

about 615 km across the Mid-Atlantic Ridge, which separates the western group (Flores and Corvo) from the central (Faial, Pico, S. Jorge, Terceira and Graciosa) and the eastern (S. Miguel and S. Maria) groups. All of the islands are of relatively recent volcanic origin, ranging from 8·12 Ma BP (S. Maria) to 300 000 years BP (Pico) (Nunes 1999). The climate is temperate oceanic, with relative atmospheric humidity that can reach 95% in high altitude native semitropical evergreen laurel forest, as well as restricting temperature fluctuations throughout the year. The predominant vegetation form is 'Laurisilva', a humid evergreen broadleaf and microphyllous laurel type forest that originally covered most of western Europe during the Tertiary (Dias 1996). For more details on native vegetation of these islands see Ribeiro *et al*. (2005) and Borges *et al*. 2005, 2006).

EXPERIMENTAL DESIGN

On seven of the Azorean islands (excluding the small and highly modified Graciosa and Corvo) native vegetation was surveyed within defined Natural Forest Reserves and/or NATURA 2000 protected areas (Borges *et al*. 2005). During the summer of 1999 and 2000 randomly placed transects (150-m long, 5-m wide) were established in each (two transects in 10 ha fragments; four transects in 100 ha fragments; eight transects in 1000 ha fragments). On Terceira, this sampling effort was duplicated in 2003. Whenever possible, transects followed a linear direction, although frequently deviations were necessary due to uneven ground and dense vegetation.

Along each transect, 30 pitfall traps were spaced at 5-m intervals and 10 replicates of the three most abundant and common woody plant species (trees and shrubs) were sampled for canopy arthropods. The canopy sampling followed a block design in which, at each 15 m, one branch of each of the most common species was sampled (Ribeiro *et al*. 2005). The endemic tree *Juniperus brevifolia* (Seub.) Antoine was the commonest species, occurring on most transects, and only samples from this source are considered in the present study.

ARTHROPOD SAMPLING AND IDENTIFICATION

We sampled the epigaeic arthropod fauna using pitfall traps set in the ground for at least a 2-week period during the summer. For the canopy arthropod sampling a modified beating tray was used, which consisted of a cloth-inverted pyramid 1 m wide and 60 cm deep (after Basset 1999). For details on sampling procedures see Ribeiro *et al*. (2005) and Borges *et al*. 2005, 2006).

All Araneae, Opiliones, Pseudoscorpiones and insects (excluding Collembola, Diplura, Diptera and Hymenoptera) were first sorted into morphospecies by students under supervision of a trained taxonomist (P.B.) (see Oliver & Beattie 1996). Later, the morphospecies

were identified by one of us (P.B.) using vouchered specimens already available *in situ*, and all unknowns were subsequently sent to several taxonomists for species identification (see Acknowledgements).

To test the abundance–variance–occupancy model we used soil and arboreal arthropod distribution and abundance data at three different scales: (1) a small within reserve scale, including 12 sampling sites (74 species) for soil fauna and 14 sampling sites (73 species) for arboreal fauna (the selected reserve, Serra de St Bárbara e Mistérios Negros on the island of Terceira, is one of the largest and best preserved in the Azores); (2) an intermediate scale, including data from the island of Terceira as a whole, and a total of 39 sampling sites (152 species) for soil fauna and 28 sampling sites (104 species) for arboreal fauna distributed across eight native forest fragments; and (3) a large scale, including 84 sampling transects (291 species) for soil fauna and 57 transects (159 species) for arboreal fauna distributed across seven islands in the Azorean archipelago (see also Ribeiro *et al*. 2005; Borges *et al*. 2005).

The species were classified into three different colonization categories: native but not endemic, endemic, and introduced. In cases of doubt, a species was assumed to be native (see also Borges *et al*. 2006).

ANALYSES

Bivariate patterns

To test how well the abundance–variance–occupancy model (4) can predict the abundance–occupancy and abundance–variance bivariate relationships and the abundance–variance–occupancy trivariate relationship, we followed He & Gaston (2003) in first fitting models (1) and (2) to the empirical abundance–variance and abundance–occupancy relationships, respectively. The linear least squares method was used to fit Taylor's power model (1) to the log-log transformed observed data (the standard method used to fit the power model), and the nonlinear least squares method was used to fit the abundance–occupancy model (2) to the observed data. We then used each of the two fitted bivariate models to predict the other bivariate relationship. The prediction of the abundance–variance pattern was made by substituting the fitted abundance–occupancy model (2) into the abundance–variance–occupancy model (4) in which the relationship between abundance and variance can then be numerically solved. Similarly, prediction of the abundance–occupancy pattern was made by substituting the fitted abundance–variance power model (1) into the abundance–variance–occupancy model (4) and then numerically solving occupancy for each observed abundance. We then tested the fit of these predicted abundance–variance and abundance–occupancy relationships to the empirical data for the respective relationships.

Exact methods for testing goodness-of-fit for non-linear models do not exist. We therefore evaluated

Table 1. Goodness-of-fit (R^2) for the individual bivariate models ('fit') and for the general model (4) ('predict') for soil and canopy species groups (introduced, endemic and native) at scales of reserve, island and archipelago. *P*-values are inappropriate for nonlinear regression and are therefore omitted (see Methods)

	Scale	Status	Abundance–variance		Abundance–occupancy	
			fit	predict	fit	predict
Soil	Reserve	Introduced	0·965	0·940	0·901	0·941
		Endemic	0·950	0·950	0·863	0·856
		Native	0·960	0·931	0·851	0·893
Soil	Island	Introduced	0·967	0·957	0·684	0·727
		Endemic	0·948	0·959	0·707	0·649
		Native	0·968	0·954	0·867	0·905
Soil	Archipelago	Introduced	0·973	0·955	0·657	0·745
		Endemic	0·956	0·952	0·691	0·727
		Native	0·971	0·955	0·797	0·803
Canopy	Reserve	Introduced	0·940	0·947	0·651	0·517
		Endemic	0·963	0·955	0·919	0·913
		Native	0·975	0·972	0·880	0·889
Canopy	Island	Introduced	0·942	0·941	0·804	0·762
		Endemic	0·955	0·949	0·970	0·983
		Native	0·976	0·967	0·864	0·881
Canopy	Archipelago	Introduced	0·929	0·901	0·709	0·715
		Endemic	0·966	0·952	0·769	0·791
		Native	0·978	0·969	0·871	0·900

goodness-of-fit by $R^2 = 1 - [\sum(y_i - \hat{y}_i)^2]/[\sum(y_i - \bar{y}_i)^2]$, where y_i is the observed value (log-transformed in the case of variance), $\hat{y}_i$ is the fitted value, and $\bar{y}$ is the mean of the observed values (Ryan 1997). The alternative approach of a lack-of-fit test, using an asymptotic *F* distribution, requires a large sample size that is not met by our data. Although formal methods are available to test whether two linear regression lines are statistically indistinguishable (see Graybill 1976), such methods are not available for comparing nonlinear models. A simple and effective approach is to compare the 95% confidence intervals of the estimates of the parameters. If the intervals overlap, the two parameters are not significantly different.

Trivariate patterns

The abundance–variance–occupancy model (4) uniquely defines a three-dimensional surface relating abundance, variance in abundance and occupancy regardless of species status and spatial scales of the data. If this model describes the data well, the observed data points should all lie on this surface. Because the model does not involve any free parameters, we simply substituted the observed abundance, variance and occupancy of each species to see how well the species lay on the three-dimensional space (there are no formal tests available to assess the fit but, as will become apparent, this is not a major issue in the present case).

Results

Both for canopy and soil habitats, at all three spatial scales, the species assemblages exhibit positive interspecific abundance–occupancy and abundance–variance relationships for native, endemic and introduced species (Table 1; Figs 1 and 2). The goodness-of-fits of the empirical data to the bivariate models (1) and (2) and to the relationships predicted using the abundance–variance–occupancy model (4) (by substituting the fitted abundance–variance and abundance–occupancy models) are compared in Table 1. As indicated by the R^2 values, both the fitted models and the predicted relationships describe the empirical data very well in almost all cases. This suggests that the abundance–variance–occupancy model (4) can be equally useful for deducing bivariate abundance–occupancy and abundance–variance patterns. Note, standard probabilities on the R^2 values are inappropriate for nonlinear regression of model (2) (see Ryan 1997), they are therefore omitted from Table 1. In other cases, all the R^2 are highly significantly different from zero at $P < 0.001$.

The confidence intervals for parameter estimates for the majority of the subsets of the data (different spatial scales and species of different status) overlap both for abundance–variance and abundance–occupancy models, suggesting that they fall on common relationships (Table 2).

The abundance–variance–occupancy model (4) describes the trivariate relationship between these variables for all the empirical data almost perfectly, regardless of spatial scales and taxonomic groups (Figs 3a and 4a). Departures of observed from predicted occupancy using the abundance–variance–occupancy (4) are negligible (Figs 3b and 4b).

None the less, there is some suggestion in the above patterns that at different spatial scales, species of different status tend disproportionately to occupy particular regions of abundance–variance–occupancy space. This is indeed the case (Table 3). Using a Kruskal–Wallis
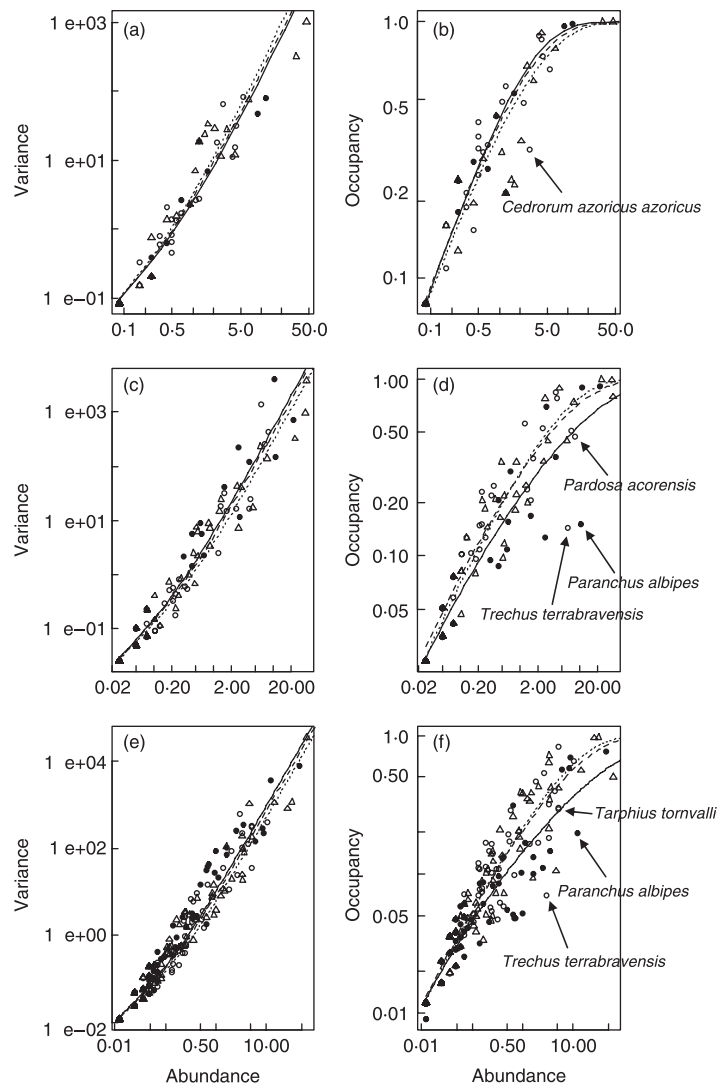
Fig. 1. (a,c,e) The abundance–variance relationship, and (b,d,f) the abundance–occupancy relationship, for the soil data at different spatial extents (a,b) reserve, (c,d) island, and (e,f) archipelago. The smooth curves are the two patterns predicted from the abundance–variance–occupancy model (4) for each species group, denoted by different symbols (to avoid clutter, fitted curves are not shown). Filled circle: introduced; open circle: endemic; open triangle: native. Solid curve: introduced; dashed curve: endemic; dotted curve: native.

Table 2. Estimates and their 95% confidence intervals for the parameters of the abundance–variance model (a, b) and the abundance–occupancy model (k) for soil and canopy species groups (introduced, endemic and native) at scales of reserve, island and archipelago

	Scale	Status	a	b	k
Soil	Reserve	Introduced	1·119 (0·841, 1·397)	1·446 (1·301, 1·591)	0·655 (0·332, 0·978)
		Endemic	1·205 (0·984, 1·426)	1·507 (1·374, 1·640)	0·554 (0·381, 0·727)
		Native	1·376 (1·143, 1·609)	1·532 (1·416, 1·648)	0·429 (0·288, 0·570)
Soil	Island	Introduced	2·379 (2·119, 2·640)	1·666 (1·579, 1·753)	0·196 (0·131, 0·261)
		Endemic	1·859 (1·555, 2·164)	1·606 (1·481, 1·731)	0·241 (0·162, 0·320)
		Native	1·832 (1·648, 2·016)	1·521 (1·452, 1·589)	0·350 (0·273, 0·427)
Soil	Archipelago	Introduced	3·015 (2·821, 3·208)	1·712 (1·654, 1·770)	0·106 (0·083, 0·128)
		Endemic	2·472 (2·250, 2·695)	1·602 (1·526, 1·678)	0·134 (0·106, 0·163)
		Native	2·394 (2·233, 2·554)	1·557 (1·508, 1·606)	0·194 (0·162, 0·226)
Canopy	Reserve	Introduced	1·555 (1·132, 1·979)	1·641 (1·416, 1·866)	0·251 (0·127, 0·374)
		Endemic	1·215 (1·000, 1·430)	1·461 (1·355, 1·567)	0·796 (0·524, 1·069)
		Native	1·745 (1·511, 1·978)	1·698 (1·593, 1·802)	0·349 (0·246, 0·452)
Canopy	Island	Introduced	1·652 (1·291, 2·014)	1·489 (1·341, 1·637)	0·275 (0·180, 0·370)
		Endemic	1·576 (1·422, 1·730)	1·472 (1·409, 1·535)	0·596 (0·464, 0·728)
		Native	1·938 (1·732, 2·144)	1·602 (1·520, 1·683)	0·324 (0·236, 0·411)
Canopy	Archipelago	Introduced	2·171 (1·807, 2·536)	1·516 (1·388, 1·644)	0·182 (0·123, 0·240)
		Endemic	2·261 (2·056, 2·467)	1·595 (1·515, 1·674)	0·246 (0·187, 0·305)
		Native	2·264 (2·075, 2·452)	1·589 (1·527, 1·651)	0·241 (0·194, 0·287)

© 2006 The Authors.
Journal compilation
© 2006 British
Ecological Society,
*Journal of Animal
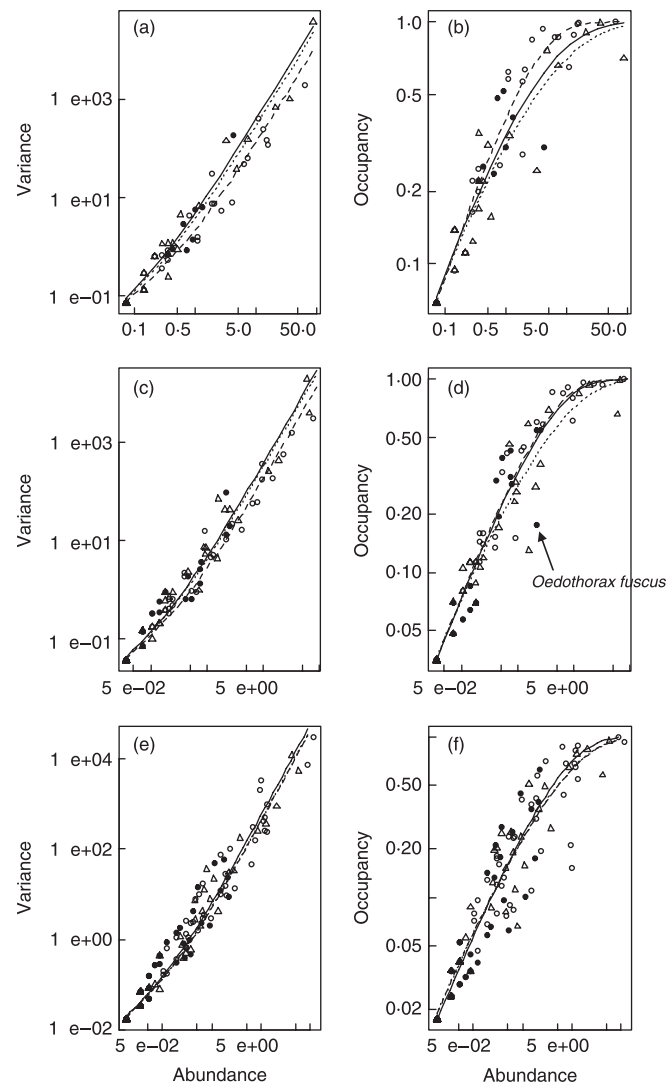Ecology*, **75**,
646–656

Fig. 2. (a,c,e) The abundance–variance relationship, and (b,d,f) the abundance–occupancy relationship, for the canopy data at different spatial extents (a,b) reserve, (c,d) island, and (e,f) archipelago. The smooth curves are the two patterns predicted from the abundance–variance–occupancy model (4) for each species group, denoted by different symbols (to avoid clutter, fitted curves are not shown). Filled circle: introduced; open circle: endemic; open triangle: native. Solid curve: introduced; dashed curve: endemic; dotted curve: native.

one-way analysis of variance by ranks, the different groups of species in the soil habitat have significantly different density, variance and occupancy at the island scale, and occupancy at the archipelago scale, and different density, variance and occupancy at the archipelago scale in the canopy, with introduced species always having lower abundance, occupancy and variance in both soil and canopy at all spatial scales, whether this is statistically significant or not (Table 3a). Likewise, almost invariably, the density, occupancy and variance of species was higher at the scale of the reserve than at that of the island, which in turn was higher than that of the archipelago, both in soil and canopy habitats (Table 3b).

Discussion

Population and community ecology has recently entered an exciting phase of pattern unification (e.g. Hanski & Gyllenberg 1997; Bell 2001; Hubbell 2001; Pachepsky

et al. 2001; Borda-de-Água, Hubbell & McAllister 2002; He & Gaston 2003; McGill & Collins 2003). A variety of models are being derived that predict the forms of two or more patterns based on the same sets of core assumptions. For example, several other ecological patterns in addition to those investigated in this study have been modelled using the same core statistical model (2) – the species–area relationship (He & Legendre 2002), species composition across landscapes (Plotkin & Muller-Landau 2002), the endemics–area relationship (Green & Ostling 2003), and percolation patterns (He & Hubbell 2003). Empirical analyses to test these predictions have, however, thus far lagged significantly behind the development of theory.

Here, we have shown that the abundance–variance–occupancy model of He & Gaston (2003) fits extremely well to data for diverse species assemblages at different spatial scales, irrespective of the status of those species. There is obvious variation about the interspecific
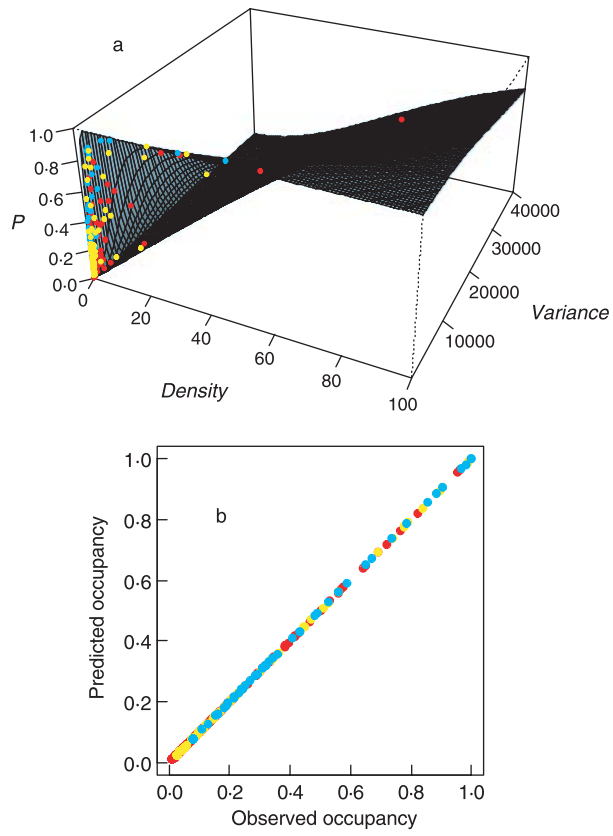
Fig. 3. (a) Three-dimensional surface describing the trivariate pattern for soil data. The dots are the observed data. Cyan: reserve; yellow: island; red: archipelago. (b) Comparison of predicted and observed occupancy.
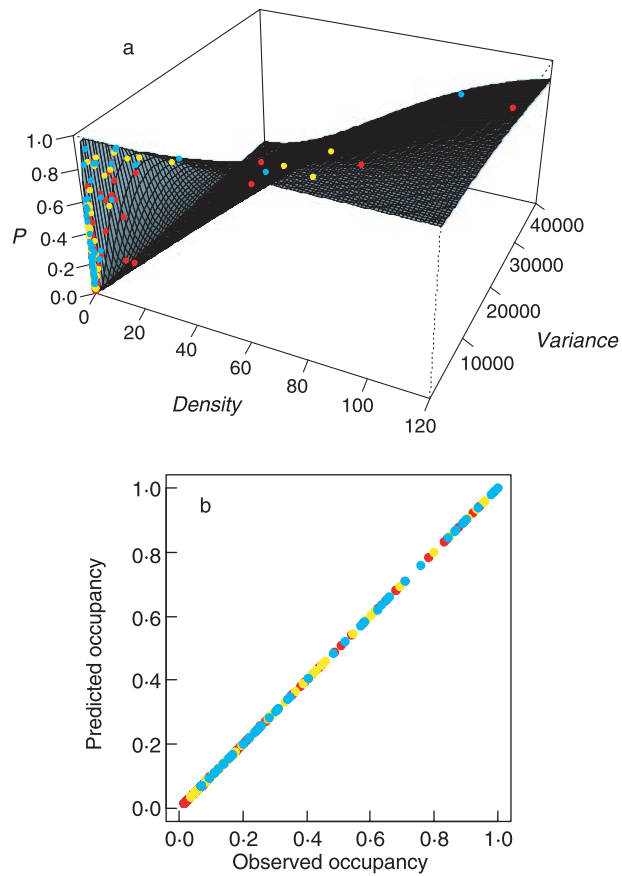

Fig. 4. (a) Three-dimensional surface describing the trivariate pattern for canopy data. The dots are the observed data. Cyan: reserve; yellow: island; red: archipelago. (b) Comparison of predicted and observed occupancy.

Table 3. Kruskal–Wallis one-way analysis of variance by ranks, corrected for ties. Shown in the table is the mean rank for each group. *P*-values are the χ^2 probabilities for the Kruskal–Wallis statistic (statistically significant values in bold). (a) Test for differences in density, spatial variance and occupancy between species groups (introduced, endemic and native) within reserve, island and archipelago. (b) Test for differences in density, spatial variance and occupancy between scales (reserve, island and archipelago) within introduced, endemic and native species

(a)		Soil				Canopy			
Scale	Variable	Introd.	Endemic	Native	P-value	Introd.	Endemic	Native	P-value
Reserve	density	35·31	38·36	37·87	0·89	31·7	41·83	34·66	0·23
	variance	35·09	37·84	38·47	0·87	31·67	41·20	35·36	0·30
	occupancy	35·59	39·54	36·62	0·80	31·53	42·90	33·61	0·12
Island	density	64·16	86·85	80·10	**0·035**	44·62	58·37	51·88	0·18
	variance	65·40	85·08	80·15	0·073	44·87	57·60	52·49	0·24
	occupancy	62·12	89·84	79·97	**0·0087**	44·75	59·22	50·95	0·14
Archip.	density	136·21	160·61	143·94	0·14	69·63	95·85	72·24	**0·0044**
	variance	137·65	159·09	143·81	0·22	70·56	95·32	72·08	**0·006**
	occupancy	131·51	163·66	145·69	**0·038**	68·51	95·39	73·50	**0·005**

(b)		Soil				Canopy			
Scale	Variable	Reserve	Island	Archip.	P-value	Reserve	Island	Archip.	P-value
Introduced	density	124·00	87·28	70·45	**< 0·0001**	55·64	45·56	36·07	**0·02**
	variance	110·12	84·90	74·04	**0·012**	52·07	42·96	38·88	0·19
	occupancy	135·38	90·44	66·87	**< 0·0001**	60·80	48·12	32·72	**0·0002**
Endemic	density	96·38	77·77	62·61	**0·0009**	72·37	61·82	59·98	0·30
	variance	88·64	74·49	66·84	0·059	67·70	60·46	63·37	0·72
	occupancy	106·57	81·30	57·41	**< 0·0001**	78·93	64·31	54·82	**0·014**
Native	density	150·13	114·77	89·67	**< 0·0001**	79·67	71·17	50·76	**0·001**
	variance	138·67	112·52	93·90	**0·001**	76·68	69·49	53·29	**0·01**
	occupancy	160·67	118·51	84·85	**< 0·0001**	83·00	71·67	48·85	**< 0·0001**

bivariate abundance–variance relationship (Figs 1 and 2) and, yet more so, about the abundance–occupancy relationship (Figs 1 and 2), as has frequently been observed before (e.g. Gaston 1996; Blackburn, Gaston & Gregory 1997; Gaston *et al.* 2000; He & Gaston 2003). Thus, for example, outliers to the abundance–occupancy relationship include (1) restricted specialized endemic species (e.g. *Trechus terrabravensis* Borges, Serrano & Amorim – Carabidae; Fig. 1d,f; *Cedrorum azoricus azoricus* Borges & Serrano – Carabidae, Fig. 1b) that occupy only pristine sites but are quite abundant there (see also Borges *et al.* 2006); (2) introduced soil habitat specialists such as *Paranchus albipes* (Fabricius) (Carabidae; Fig. 1d,f); (3) restricted endemics particularly abundant in some pristine sites (e.g. *Tarphius tornvalli* Gillerfors – Zopheridae; Fig. 1f); (4) specialized grassland spiders (e.g. *Pardosa acorensis* Simon – Lycosidae; Fig. 1d) that tend to occupy marginal sites in native forest, and consequently being abundant there occur in fewer sites than otherwise expected; (5) specialized grassland spiders (e.g. *Oedothorax fuscus* Blackwall – Linyphiidae; Fig. 2d) that also occur in canopies infrequently; and (6) some species that probably were not sampled adequately with the methods applied or that have aggregated behaviour (e.g. some coccids in canopies). This variation is, however, effectively explained by the third missing variable (variance in abundance in the case of the abundance–occupancy relationship, and occupancy in the case of the abundance–variance

relationship). Indeed, the overall fit of the abundance–variance–occupancy model is extraordinarily good (Figs 3 and 4). As postulated, the spatial distribution of species (as measured by abundance, variance and occupancy) does seem to be uniquely described by the trivariate relationship.

Although introduced species tend to exhibit lower densities, less spatial variance in their abundances and to occupy fewer sites than native and endemic species (Table 3a), they none the less lie on the same bivariate abundance–variance and abundance–occupancy, and trivariate abundance–variance–occupancy, relationships (for other evidence see also Holt & Gaston 2003). This implies that the fundamental population dynamics of all three groups are broadly similar, whatever other biological differences they may display (for review see Williamson 1996; for a recent case study in insular systems see Borges *et al.* 2006), or the same statistical assumptions underlie the distribution of the species. Likewise, although for all three groups of species spatial variance in abundance and occupancy decline toward larger spatial scales, regardless of scale data points fall on the same trivariate abundance–variance–occupancy relationship (Figs 3 and 4).

He & Gaston (2003) suggested that one simple, but potentially widely generalizable, mechanistic interpretation of the abundance–variance–occupancy model (4) derives from metapopulation dynamics. A positive abundance–occupancy relationship can result from (1)

the carrying capacity hypothesis – different species in an assemblage have different local carrying capacities, and those that attain higher local population sizes have a lower extinction rate and/or a higher colonization rate than those that attain small local population sizes, and therefore occupy more patches at equilibrium (Nee, Gregory & May 1991); or (2) the rescue effect hypothesis – immigration decreases the probability of a local population going extinct (the rescue effect; Brown & Kodric-Brown 1977), and the rate of immigration per patch increases as the proportion of patches that are occupied increases, leading often to species with higher population sizes occupying more patches (Hanski 1991a,b; Gyllenberg & Hanski 1992; Hanski & Gyllenberg 1993; Hanski *et al.* 1993). Such an effect is captured in the neutral theory of biodiversity (Bell 2001; Hubbell 2001). A likely process that can also result theoretically in the positive abundance–variance relationship is the immigration of individuals from high-density sites to low-density and vacant ones (Bell 2003; He & Gaston 2003). While very general, and potentially describable in a variety of terms (e.g. niche theory, individual behaviour), the potential applicability of such mechanisms to the empirical data analysed here is variable, and presumably declines moving from within to between-island patterns, as the role of dispersal between areas lessens.

More generally, the spatial distribution of individuals is fundamental to understanding macroecological patterns. The fact that many patterns can arise from model (2) (He & Gaston 2000, 2003; He & Legendre 2002; Plotkin & Muller-Landau 2002; Green & Ostling 2003; He & Hubbell 2003) suggests that the spatial distributions of species is likely to be widely approximated by the negative (or positive if $k < 0$) binomial distribution from which model (2), and therefore the abundance–variance–occupancy model (4), is derived. The inter-relationships of abundance, variance and occupancy are inherently constrained by the negative/positive binomial models (Royle, Nichols & Kéry 2005), so that the third variable (occupancy or variance) is an approximation of the other two (i.e. for a given abundance, occupancy and variance are negatively correlated). Therefore, the abundance–variance–occupancy model (4) allows explanation of the residual variation left by the individual bivariate models.

In conclusion, regardless of the mechanisms that may generate the bivariate patterns of abundance–occupancy and abundance–variance, our results show that the distribution of species should be entirely determined by abundance, occupancy and variance. In other words, the abundance of a species, its spatial variation and the area of occupancy on landscapes are uniquely constrained, involving no further parameters.

Acknowledgements

References

Anderson, R.M., Gordon, D.M., Crawley, M.J. & Hassell, M.P. (1982) Variability in the abundance of animal and plant species. *Nature*, **296**, 245–248.

Basset, Y. (1999) Diversity and abundance of insect herbivores collected on *Castanopsis acuminatissima* (Fagaceae) in New Guinea: relationships with leaf production and surrounding vegetation. *European Journal of Entomology*, **96**, 381–391.

Bell, G. (2001) Neutral macroecology. *Science*, **293**, 2413–2418.

Bell, G. (2003) The interpretation of biological surveys. *Proceedings of the Royal Society, London B*, **270**, 2531–2542.

Binns, M.R. (1986) Behavioural dynamics and the negative binomial distribution. *Oikos*, **47**, 315–318.

Blackburn, T.M., Gaston, K.J. & Gregory, R.D. (1997) Abundance-range size relationships in British birds: is unexplained variation a product of life history? *Ecography*, **20**, 466–474.

Borda-de-Água, L., Hubbell, S.P. & McAllister, M. (2002) Species-area curves, diversity indices, and species abundance distributions: a multifractal analysis. *American Naturalist*, **159**, 138–155.

Borges, P.A.V., Aguiar, C., Amaral, J., Amorim, I.R., André, G., Arraiol, A., Baz, A., Dinis, F., Enghoff, H., Gaspar, C., Ilharco, F., Mahnert, V., Melo, C., Pereira, F., Quartau, J.A., Ribeiro, S., Ribes, J., Serrano, A.R.M., Sousa, A.B., Strassen, R.Z., Vieira, L., Vieira, V., Vitorino, A. & Wunderlich, J. (2005) Ranking protected areas in the Azores using standardized sampling of soil epigean arthropods. *Biodiversity and Conservation*, **14**, 2029–2060.

Borges, P.A.V., Lobo, J.M., Azevedo, E.B., Gaspar, C.S., Melo, C. & Nunes, L.V. (2006) Invasibility and species richness of island endemic arthropods: a general model of endemic vs. exotic species. *Journal of Biogeography*, **33**, 169–187.

Brown, J.H. (1984) On the relationship between abundance and distribution of species. *American Naturalist*, **124**, 255–279.

Brown, J.H. & Kodric-Brown, A. (1977) Turnover rates in insular biogeography: effect of immigration on extinction. *Ecology*, **58**, 445–449.

Dias, E. (1996) *Vegetação Natural dos Açores: Ecologia e Sintaxonomia das Florestas Naturais*. PhD Thesis, Universidade dos Açores, Angra do Heroísmo.

Ekbom, B.S. (1987) Incidence counts for estimating densities of *Rhopalosiphum padi* (Homoptera: Aphididae). *Journal of Economic Entomology*, **80**, 933–935.

Feng, M.C., Nowierski, R.M. & Zeng, Z. (1993) Populations of *Sitobion avenae* and *Aphidius ervi* on sprint wheat in the northwestern United States. *Entomologia Experimentalis et Applicata*, **67**, 109–117.

Gaston, K.J. (1996) The multiple forms of the interspecific abundance-distribution relationship. *Oikos*, **75**, 211–220.

Gaston, K.J. (1999) Implications of interspecific and intraspecific abundance–occupancy relationships. *Oikos*, **86**, 195–207.

Gaston, K.J., Blackburn, T.M., Greenwood, J.J.D., Gregory, R.D., Quinn, R.M. & Lawton, J.H. (2000) Abundance–occupancy relationships. *Journal of Applied Ecology*, **37** (Suppl. 1), 39–59.

Gaston, K.J., Blackburn, T.M. & Lawton, J.H. (1997) Interspecific abundance-range size relationships: an appraisal of mechanisms. *Journal of Animal Ecology*, **66**, 579–601.

Gillis, D.M., Kramer, D.L. & Bell, G. (1986) Taylor's power law as a consequence of Fretwell's Ideal Free Distribution. *Journal of Theoretical Biology*, **123**, 281–287.

Graybill, F.A. (1976) *Theory and Application of the Linear Model*. Duxbury, North Scituate, MA.

Green, J.L. & Ostling, A. (2003) Endemics-area relationships: the influence of species dominance and spatial aggregation. *Ecology*, **84**, 3090–3097.

Gyllenberg, M. & Hanski, I. (1992) Single-species metapopulation dynamics; a structured model. *Theoretical Population Biology*, **42**, 35–61.

Hanski, I. (1982) Dynamics of regional distribution: the core and satellite species hypothesis. *Oikos*, **38**, 210–221.

Hanski, I. (1991a) Single-species metapopulation dynamics: concepts, models and observations. *Biological Journal of the Linnean Society*, **42**, 17–38.

Hanski, I. (1991b) Reply to Nee, Gregory and May. *Oikos*, **62**, 88–89.

Hanski, I. & Gyllenberg, M. (1993) Two general metapopulation models and the core-satellite species hypothesis. *American Naturalist*, **142**, 17–41.

Hanski, I. & Gyllenberg, M. (1997) Uniting two general patterns in the distribution of species. *Science*, **275**, 397–400.

Hanski, I., Kouki, J. & Halkka, A. (1993) Three explanations of the positive relationship between distribution and abundance of species. *Species Diversity in Ecological Communities: Historical and Geographical Perspectives* (eds R.E. Ricklefs & D. Schluter), pp. 108–116. University of Chicago Press, Chicago, IL.

Harte, J., Blackburn, T. & Ostling, A. (2001) Self-similarity and the relationship between abundance and range size. *American Naturalist*, **157**, 374–386.

He, F. & Gaston, K.J. (2000) Estimating species abundance from occurrence. *American Naturalist*, **156**, 553–559.

He, F. & Gaston, K.J. (2003) Occupancy, spatial variance, and the abundance of species. *American Naturalist*, **162**, 366–375.

He, F. & Hubbell, S.P. (2003) Percolation theory for the distribution and abundance of species. *Physical Review Letters*, **91**, 198103.

He, F. & Legendre, P. (2002) Species diversity patterns derived from species-area models. *Ecology*, **83**, 1185–1198.

Hepworth, G. & MacFarlane, J.R. (1992) Systematic presence-absence sampling method applied to two-spotted spider mite (Acari: Tetranychidae) on strawberries in Victoria, Australia. *Journal of Economic Entomology*, **85**, 2234–2239.

Holt, A.R. & Gaston, K.J. (2003) Interspecific abundance–occupancy relationships of British mammals and birds: is it possible to explain the residual variation? *Global Ecology and Biogeography*, **12**, 37–46.

Holt, A.R., Gaston, K.J. & He, F. (2002) Occupancy-abundance relationships and spatial distribution. *Basic and Applied Ecology*, **3**, 1–13.

Holt, R.D., Lawton, J.H., Gaston, K.J. & Blackburn, T.M. (1997) On the relationship between range size and local abundance: back to basics. *Oikos*, **78**, 183–190.

Hubbell, S.P. (2001) *A Unified Theory of Biodiversity and Biogeography*. Princeton University Press, Princeton, NJ.

Kuno, E. (1986) Evaluation of statistical precision and design of efficient sampling for the population estimates based on frequency of sampling. *Researches in Population Ecology*, **28**, 305–319.

Kuno, E. (1991) Sampling and analysis of insect populations. *Annual Review of Entomology*, **36**, 285–304.

Maurer, B.A. (1994) *Geographical Population Analysis: Tools for the Analysis of Biodiversity*. Blackwell Scientific, Oxford.

McArdle, B.H. & Gaston, K.J. (1995) The temporal variability of densities: back to basics. *Oikos*, **74**, 165–171.

McGill, B. & Collins, C. (2003) A unified theory for macroecology based on spatial patterns of abundance. *Evolutionary Ecology Research*, **5**, 469–492.

Nachman, G. (1981) A mathematical model of the functional relationship between density and spatial distribution of a population. *Journal of Animal Ecology*, **50**, 453–460.

Nachman, G. (1984) Estimates of mean population density and spatial distribution of *Tetranychus urticae* (Acarina: Tetranychidae) and *Phytoseiulus persimilis* (Acarina: Phytoseiidae) based upon the proportion of empty sampling units. *Journal of Applied Ecology*, **21**, 903–913.

Nee, S., Gregory, R.D. & May, R.M. (1991) Core and satellite species: theory and artefacts. *Oikos*, **62**, 83–87.

Nunes, J.C.C. (1999) A Actividade Vulcânica na ilha do Pico do Plistocénico Superior ao Holocénico: Mecanismo Eruptivo e Hazard Vulcânico. PhD Thesis, Universidade dos Açores.

Oliver, I. & Beattie, A.J. (1996) Invertebrate morphospecies as surrogates for species: a case study. *Conservation Biology*, **10**, 99–109.

Pachepsky, E., Crawford, J.W., Bown, J.L. & Squire, G. (2001) Towards a general theory of biodiversity. *Nature*, **410**, 923–926.

Perry, J.N. (1981) Taylor's power law for dependence of variance on mean in animal populations. *Applied Statistics*, **30**, 254–263.

Perry, J.N. (1987) Host-parasitoid models of intermediate complexity. *American Naturalist*, **130**, 955–957.

Perry, J.N. (1988) Some models for spatial variability of some animal species. *Oikos*, **51**, 124–130.

Plotkin, J.B. & Muller-Landau, H.C. (2002) Sampling the species composition of a landscape. *Ecology*, **83**, 3344–3356.

Ribeiro, S.P., Borges, P.A.V., Gaspar, C., Melo, C., Serrano, A.R.M., Amaral, J., Aguiar, C., André, G. & Quartau, J.A. (2005) Canopy insect herbivores in the Azorean Laurisilva forests: key host plant species in a highly generalist insect community. *Ecography*, **28**, 315–330.

Rodrigues, A.S.L., Gaston, K.J. & Gregory, R.D. (2000) Using presence-absence data to establish reserve selection procedures that are robust to temporal species turnover. *Proceedings of the Royal Society, London B*, **267**, 897–902.

Routledge, R.D. & Swartz, T.B. (1991) Taylor's power law re-examined. *Oikos*, **60**, 107–112.

Royle, J.A., Nichols, J.D. & Kéry, M. (2005) Modelling occurrence and abundance of species when detection is imperfect. *Oikos*, **110**, 353–359.

Ryan, T.P. (1997) *Modern Regression Methods*. Wiley & Sons, New York.

Taylor, L.R. (1961) Aggregation, variance and the mean. *Nature*, **189**, 732–735.

Taylor, L.R. (1984) Assessing and interpreting the spatial distributions of insect populations. *Annual Review of Entomology*, **29**, 321–357.

Taylor, L.R., Woiwod, I.P. & Perry, J.N. (1978) The density-dependence of spatial behaviour and the rarity of randomness. *Journal of Animal Ecology*, **47**, 383–406.

Taylor, L.R., Woiwod, I.P. & Perry, J.N. (1979) The negative binomial as a dynamic ecological model for aggregation and the density dependence of. *K. Journal of Animal Ecology*, **48**, 289–304.

Taylor, L.R., Woiwod, I.P. & Perry, J.N. (1980) Variance and the large scale spatial stability of aphids, moths and birds. *Journal of Animal Ecology*, **49**, 831–854.

Ward, S.A., Sunderland, K.D., Chambers, R.J. & Dixon, A.F.G. (1986) The use of incidence counts for estimation of cereal aphid populations. 3. Population development and the incidence-density relation. *Netherlands Journal of Plant Pathology*, **92**, 175–183.

Williamson, M. (1996) *Biological Invasions*. Chapman & Hall, London.

Wilson, L.T. & Room, P.M. (1983) Clumping patterns of fruit and arthropods in cotton, with implications for binomial sampling. *Environmental Entomology*, **12**, 50–54.

Yamamura, K. (2000) Colony expansion model for describing the spatial distribution of populations. *Researches on Population Ecology*, **42**, 161–169.