

Varying definitions of abundance and incomplete assemblages challenge the generality of the interspecific abundance–distribution relationships

Atte Komonen^{1,*}, Jussi Päivinen² & Janne S. Kotiaho^{3,4}

¹⁾ Department of Biological and Environmental Science, P.O. Box 35, FI-40014 University of Jyväskylä, Finland (*corresponding author's e-mail: atte.komonen@jyu.fi)

²⁾ Metsähallitus, Natural Heritage Services, P.O. Box 36, FI-40101 Jyväskylä, Finland

³⁾ Centre of Excellence in Evolutionary Research, P.O. Box 35, FI-40014 University of Jyväskylä, Finland

⁴⁾ Museum of Natural History, P.O. Box 35, FI-40014 University of Jyväskylä, Finland

Received 17 Dec. 2010, revised version received 5 May 2011, accepted 5 May 2011

Komonen, A., Päivinen, J. & Kotiaho, J. S. 2011: Varying definitions of abundance and incomplete assemblages challenge the generality of the interspecific abundance–distribution relationships. — *Ann. Zool. Fennici* 48: 161–166.

Empirical evidence does not fully support the universal nature of the positive interspecific abundance–distribution relationship. We have earlier documented a negative relationship for butterfly species in Finland, but recently our view was again challenged using a small subset of Finnish butterflies as apparent evidence. Here we scrutinize the critique and identify some general conceptual challenges in analyses of interspecific abundance–distribution relationships. We identify the common problem that the abundance–distribution studies include only a small subset of species, and thus reveal only sample characteristics, not overall patterns in complete assemblages. Small subsets of species are also unlikely to have sufficient power to reveal nonlinear relationships. Second, varying definitions of abundance, especially the practice of using a single point estimate to describe average density, further spur the empirical evidence for the abundance–distribution relationship. To get theoretically relevant results abundance and distribution must be defined and operationalised consistently; otherwise macroecology will reduce to simple documentation of ambiguous patterns and aid little in understanding the biological world.

Introduction

A positive relationship between species abundance and distribution appears to be the prevailing paradigm in contemporary macroecology, and, indeed, it is often believed to be general enough to be considered a universal pattern in ecology (Blackburn *et al.* 2006). However, what

is curious about this universal generalization is that when empirical evidence is scrutinized it reveals that nearly one-third of the studies report non-significant or negative patterns (Blackburn *et al.* 2006, see also Gaston & Lawton 1990, Cowley *et al.* 2001, Päivinen *et al.* 2005). Indeed, the leading authors in the field have concluded that the relationship at hand appears to

vary considerably across assemblages, from significant positive to significant negative and to all shades in between (Blackburn *et al.* 2006). Thus, for us it appears that the empirical evidence for the positive abundance–distribution relationship is slightly short of making it to be as universal as it is often believed to be.

Recently, our view that the relationship between species abundance and distribution in Finnish butterflies may be negative or even nonlinear (Päivinen *et al.* 2005, Komonen *et al.* 2009) was yet again challenged (Selonen & Helos 2010; for the previous challenge see Blackburn and Gaston 2009). In their challenge, Selonen and Helos (2010) first criticize our use of the large-scale butterfly monitoring data (NAFI) and raise several hypothetical problems that these data may have. Second, they aim at disproving the negative abundance–distribution relationship for Finnish butterflies by demonstrating a positive relationship for a set of 15 butterfly species for which mark–recapture based density data was available. Here, we first scrutinize their critique and then point out a few general conceptual pitfalls in the empirical studies of abundance–distribution relationships.

Scrutiny of the critique

We tend to agree with Selonen and Helos (2010), as we did with Blackburn and Gaston (2009), in that one should be careful when using Atlas data, or census data such as NAFI, and to this end, we have repeatedly pointed out the potential problems in the data (Kotiaho *et al.* 2005, Mattila *et al.* 2008, Komonen *et al.* 2009). Nevertheless, Selonen and Helos (2010) make sweeping claims about the behaviour of voluntary lepidopterists biasing the data: “*For example, voluntary lepidopterists often are interested in observing rare species with known occurrence sites, thus making rare species overabundant in the database.*” Then they continue with a number of other hypothetical problems supporting the bias and end up with a conclusion: “... *for rare species the number of observation days is likely to be low, as the known occurrence sites of the [rare species] are often visited only on one or a few days. Consequently, it is questionable*

whether dividing the mean local abundance by the number of observation days [...] treated the density of common and rare species similarly”. The critique is basically that there is a bias in our density estimates caused by the bias in the number of observation days in relation to rare and common species. In effect, this is just a restatement of the issue already raised by Blackburn and Gaston (2009). If Selonen and Helos (2010) had carefully studied our response to the original critique (Kotiaho *et al.* 2009) they should have noticed that there is no evidence of such bias in the data. However, the proof of lack of sampling bias in Kotiaho *et al.* (2009) was in a bit complicated form, and thus below we provide evidence from two further analyses that there is no bias in the observation day data in relation to the rarity of the species.

We hypothesize that if there is a bias such that the voluntary lepidopterists would preferentially visit the sites where rare species occur, we should find a signature of this in the observation date data in relation to the rarity of the species (small distribution or red-list status). The observation data we have used for the 95 resident butterfly species in Finland in the previous publications (Komonen *et al.* 2004, Kotiaho *et al.* 2005, Päivinen *et al.* 2005, Komonen *et al.* 2009, Kotiaho *et al.* 2009) is simply massive: total number of observation days is 1 977 436. For each species this translates to a mean and median of 20 815 and 12 713 observation days, respectively, and the range is from 53 to 50 595 observation days. The abundance data do not fall much short of these figures: total number of individuals observed is 1 437 453 with mean and median number of individual per species being 15 131 and 3443, respectively, and ranging from 47 to 149 426 individuals per species. With such data, any bias how ever minute in terms of the voluntary lepidopterists preferentially focusing on rare species should be easily observable.

To test our hypothesis, we conducted two analyses with different criteria for the “rare” species. First criterion was the distribution, i.e. the number of 10 km × 10 km grid cells occupied by the species. If the bias suggested by Selonen and Helos (2010) exists, rare species (species with small distribution) should face greater sampling effort than common species (species with large

distribution). Second criterion was the red-listed status (Rassi *et al.* 2001). If the bias exists, the rare species (red-listed species; $n = 23$) should face greater sampling effort than the common species (species not in the red list; $n = 72$). For each species, the number of observation days in all grid cells occupied by the species (including observation days when the species was not observed) was divided by the number of grid cells occupied by the species. This measure of sampling effort had no significant relationship with the rarity of species, measured as the distribution (Spearman $r = -0.03$, $n = 95$, $p = 0.777$) or as the red-list status (Mann-Whitney $U = 905$, $n_{\text{tot}} = 95$, $p = 0.504$). Thus, as far as the data goes there is no evidence that voluntary lepidopterists would be preferentially visiting the grid cells where rare species occur, making the criticism of Selonen and Helos (2010) unsubstantiated.

Selonen and Helos (2010) continue their criticism with two further statements. First, they state that “*the high-density locations of common species are likely to be under sampled as compared with those of rare species*”. This statement is entirely relying on the assumption that there actually are such high-density locations for the common species. Although this might be the case, Selonen and Helos (2010) provide no evidence for it. The second statement serves as the grand finale of their paper: “*Our results support the conclusion that a positive distribution–abundance relationship in butterflies (Cowley *et al.* 2001) is as general a trend as it was before the work by Päivinen *et al.* (2005) and Komonen *et al.* (2009).*” However, Selonen and Helos (2010) present only the “positive” side of the coin and dismiss the “negative” one, because already in the abstract Cowley *et al.* (2001) conclude that “*Species’ national densities in Britain were positively correlated with their European distribution size, but significantly negatively correlated with their global range sizes*”.

In general, we welcome the attempts to provide critique of our work, and some of the problems that Selonen and Helos (2010) identify may be real. What we find astounding in their critique, however, is that all their claims about the interests of the voluntary lepidopterists participating in NAFI are based on no data. We insist that to pass as rigorous scientific critique,

the critique must be based on solid foundation of empirical or theoretical evidence, or on argumentation pointing out a fundamental flaw in the logic underlying the criticized study.

Conceptual pitfalls in empirical estimation of abundance–distribution relationships

Apart from criticizing and attempting to dismiss our work, Selonen and Helos (2010) aimed to prove that for Finnish butterflies the abundance–distribution relationship is in fact positive and thus in line with the universal nature of this pattern. Unfortunately, neither the arguments nor the analyses by Selonen and Helos (2010) resolve the underlying problems in empirical abundance–distribution relationships; they merely repeat the problems present in many of the extant studies (*see below*).

Analyses based on incomplete assemblages are misleading

Many abundance–distribution studies include only a small non-random sample of species in a given region and thus reveal only sample characteristics, not general or universal patterns. For example, Selonen and Helos (2010) include only 15 species, representing ca. 16% of the butterfly assemblage in Finland. Thus, the title of their paper “*Positive interspecific abundance–distribution relationship in Finnish butterflies*” is clearly misleading. One simply cannot extrapolate — ecologically or statistically — from a (small) non-random sample of species to entire species assemblage. Even if the sample is random in terms of the sites available, the problem is that the species encountered are not a random sample of all species in the entire assemblage; rather the species are included as a function of their commonness (in other words, a random sample of sites is not the same as a random sample of species). We have specifically pointed out that when species are included into the sample in proportion to their commonness, only after ca. 70% of the species in an assemblage are included, the relationship turns

from positive to non-significant or even negative (Komonen *et al.* 2009; *see also* Cowley *et al.* 2001).

If only a small subset of species in the assemblage is included, the selection criteria of the species becomes of paramount importance. For example, in Selonen and Helos' (2010) data 60% of the species were threatened, whereas the true percentage of threatened Finnish butterflies is around 24%; Hughes (2000) included only the butterfly Family Lycaenidae; whereas Gutiérrez and Menéndez (1997) classified carabid beetle species in three groups according to their wing size and distribution, and then analysed the abundance–distribution relationships separately for these groups. To be fair, also we have done demarcations by studying butterflies and excluding other Lepidoptera. All such subjective demarcations, although often justifiable for practical reasons, may have little biological rationale and may lead to publication bias supporting the prevailing paradigm (Kuhn 1996, Simmons *et al.* 1999).

Furthermore, with the near complete assemblage of the Finnish butterflies, we have pinpointed an intriguing possibility that rather than being universally positive or negative, the true relationship between abundance and distribution may in fact be nonlinear comprising both negative and positive zones across the rarity of the species (Komonen *et al.* 2009). Unfortunately, small subsets of species, such as Selonen and Helos' (2010), are unlikely to have the power to reveal such nonlinear relationships. Thus, we urge that studies attempting to address macroecological questions should clearly distinguish the inferences based on a small sample of species and inferences based on the near-complete assemblages.

Point estimates do not reveal average density

Abundance and distribution are concepts, which can be defined in many ways and in many scales (Blackburn *et al.* 2006). The most natural way to define abundance is population density, i.e. the number of individuals per unit of area. In macroecological context, the question is: do wide-

spread species have a higher population density *on average* than species with more restricted distribution. Although we welcome the approach by Selonen and Helos (2010) in that they aim at obtaining scientifically rigorous data on species abundances based on mark–recapture studies, their data do not represent average abundance, but rather a point estimate of the abundance at only one site or in one metapopulation (*see also* Gaston & Lawton 1990). Given the high profile of the abundance–distribution relationship in macroecology, it is perhaps surprising that there are very few studies (with the notable exception of birds; Gaston & Blackburn 2000), that have reliably documented *average abundance* in this context. If there is variance in the abundance of the species across the sites where they occur (which almost inevitable is the case), a single point estimate of the abundance may very well be largely misleading.

The problem with a single point estimate can be neatly illustrated with the data of Selonen and Helos (2010: table 1). They provide two density estimates for a common species *Aphantopus hyperanthus* (574 and 1925 individuals per ha). Assuming that also many other species are likely to have similar over three-fold differences in point estimates of density (and probably even much more if more estimates are available), but only one site happened to be studied using mark–recapture, it becomes obvious that the positive relationship obtained by Selonen and Helos (2010) is likely to be very unstable. In their table 1, Selonen and Helos (2010) provide two density estimates also for the rare *Parnassius mnemosyne* (143 and 274 individuals per ha). What is noteworthy is that these estimates come from the exactly same metapopulation and as such are not independent density estimates — yet they are very different. Closer look of the studies reveals that the lower estimate was based on a few-day pilot study (Välimäki *et al.* 2000), while the higher density estimate was obtained over the entire flight period (Välimäki & Itämies 2003). Density estimates based on markedly different sampling efforts should not be directly averaged at all.

Given that point estimates tend to be highly variable, it is perhaps no longer surprising that the density estimates in our data (uncorrected

or corrected for observation days) and those provided by Selonen and Helos (2010) do not correspond to each other (Spearman $r = 0.12$, $n = 15$, $p = 0.666$ and $r = -0.23$, $n = 15$, $p = 0.405$, respectively). Because Selonen and Helos (2010) only speculate regarding the average abundance of species over larger areas based on a sample size of one (or two at best) for each species, their contribution to the abundance–distribution debate is limited. We do agree that in order to make progress in studies of interspecific abundance–distribution relationships one needs reliable abundance estimates. Mark–release–recapture approach is one potential solution, although the method is challenging to conduct over large areas and for many species.

Conclusions

The positive relationship between species abundance and distribution is a dogma. It is relatively easy to obtain positive relationships by using varying definitions and operationalisations of abundance and distribution (of course, the logic applies to the negative relationships as well). However, to obtain theoretically relevant results these concepts must be defined and operationalised consistently, or the results based on different definitions and operationalisations must be investigated and reported. The requirement for consistent definitions is not semantics, but should be the norm in all science. This is because all concepts inherently influence our understanding of biotic phenomena and ecological processes (Mayr 1997, Pigliucci 2009). One could argue that because the positive relationships do exist, despite varying definitions and operationalisations, it is robust and truly universal in nature. However, we contend that exactly because of the ambiguity in the underlying concepts, the true nature of the abundance–distribution relationship remains unresolved. Our concern is that if we want to advance the *Science* of macroecology, we must pay more attention to the definitions and operationalisations of abundance and distribution. If we do not, macroecology will reduce to simple documentation of ambiguous patterns and aid little our understanding of the ecological processes taking place in the biological world.

References

- Blackburn, T. M. & Gaston, K. J. 2009: Sometimes the obvious answer is the right one: a response to 'Missing the rarest: is the positive interspecific abundance–distribution relationship truly general macroecological pattern?' — *Biology Letters* 5: 777–778.
- Blackburn, T. M., Cassey, P. & Gaston, K. J. 2006: Variations on a theme: sources of heterogeneity in the form of the interspecific relationship between abundance and distribution. — *Journal of Animal Ecology* 75: 1426–1439.
- Cowley, M. J. R., Thomas, C. D., Roy, D. B., Wilson, R. J., León-Cortés, J. L., Gutiérrez D., Bulman, C. R., Quinn, R. M., Moss, D. & Gaston, K. J. 2001: Density–distribution relationships in British butterflies. I. The effect of mobility and spatial scale. — *Journal of Animal Ecology* 70: 410–425.
- Gaston, K. J. & Blackburn, T. M. 2000: *Pattern and process in macroecology*. — Blackwell Science, Oxford.
- Gaston, K. J. & Lawton, J. H. 1990: Effects of scale and habitat on the relationship between regional distribution and local abundance. — *Oikos* 58: 329–335.
- Gutiérrez, D. & Menéndez, R. 1997: Patterns in the distribution, abundance and body size of carabid beetles (Coleoptera: Caraboidea) in relation to dispersal ability. — *Journal of Biogeography* 24: 903–914.
- Hughes, J. B. 2000: The scale of resource specialization and the distribution and abundance of lycaenid butterflies. — *Oecologia* 123: 375–383.
- Komonen, A., Päivinen, J. & Kotiaho, J. S. 2009: Missing the rarest: is the positive interspecific abundance–distribution relationship truly general macroecological pattern? — *Biology Letters* 5: 492–494.
- Komonen, A., Grapputo, A., Kaitala, V., Kotiaho, J. S. & Päivinen, J. 2004: The role of niche breadth, resource availability and range position on the life history of butterflies. — *Oikos* 105: 41–54.
- Kotiaho, J. S., Komonen, A. & Päivinen, J. 2009: On the obvious positive interspecific relationship between abundance and distribution: a reply to Blackburn and Gaston. — *Biology Letters* 5: 779–780.
- Kotiaho, J. S., Kaitala, V., Komonen, A. & Päivinen, J. 2005: Predicting the risk of extinction from shared ecological characteristics. — *Proceedings of the National Academy of Sciences of the USA* 102: 1963–1967.
- Kuhn, T. S. 1996: *The structure of scientific revolutions*. — University of Chicago Press, Chicago.
- Mattila, N., Kotiaho, J. S., Kaitala, V., Komonen, A. & Päivinen, J. 2008: Interactions between ecological traits and host plant type explain distribution change in noctuid moths. — *Conservation Biology* 23: 703–709.
- Mayr, E. 1997: *This is biology, the science of the living world*. — Harvard University Press, Cambridge.
- Pigliucci, M. 2010: *Nonsense on stilts: How to tell science from bunk*. — University of Chicago Press, Chicago.
- Päivinen, J., Grapputo, A., Kaitala, V., Komonen, A., Kotiaho, J. S., Saarinen, K. & Wahlberg, N. 2005: Negative density–distribution relationship in butterflies. — *BMC Biology* 3: 5.

- Rassi, P., Alanen, A., Kanerva, T. & Mannerkoski, I. (eds.) 2001: *Suomen lajien uhanalaisuus 2000 [The 2000 red list of Finnish species]*. — Ministry of the Environment and Finnish Environment Institute, Helsinki [In Finnish with English Summary].
- Selonen, V. & Helos, I. 2010: Positive density-distribution relationship in Finnish butterflies. — *Annales Zoologici Fennici* 47: 347–350.
- Simmons, L. W., Tomkins, J. L., Kotiaho, J. S. & Hunt, J. 1999: Fluctuating paradigm. — *Proceedings of the Royal Society of London B* 266: 593–595.
- Välimäki, P. & Itämies, J. 2003: Migration of the clouded Apollo butterfly *Parnassius mnemosyne* in a network of suitable habitats — effects of patch characteristics. — *Ecography* 26: 679–691.
- Välimäki, P., Itämies, J. & Helminen, O. 2000: Pikkuapollon (*Parnassius mnemosyne*) esiintyminen Rauman ympäristössä vuonna 1999 [The occurrence of clouded apollo (*Parnassius mnemosyne*) around Rauma in 1999]. — *Baptia* 25: 61–69. [In Finnish with English Summary].