Ecological Archives **E089-115-A1**

Márcio S. Araújo, Paulo R. Guimarães, Jr., Richard Svanbäck, Aluisio Pinheiro, Paulo Guimarães, Sérgio F. dos Reis, and Daniel I. Bolnick. 2008. Network analysis reveals contrasting effects of intraspecific competition on individual vs. population diets. *Ecology* 89:1981–1993.

The individual niche overlap network

The first step in the development of C_{ws} is to define a conceptual framework linking patterns of resource use at the population level and complex network theory. We did this by defining the individual niche overlap network as follows. A network is a representation of associations among elements in a system, in which nodes represent elements, and ‘edges’ are lines that connect those nodes that are somehow associated. To describe resource use variation within a population, we can represent individuals as nodes, with each pairwise combination of individuals connected by an edge if they use any resources in common, or disconnected if they do not share resources. A given edge can be weighted to represent the degree of pairwise niche overlap w_{ij} between two individuals *i* and *j* (adapted from Schoener 1968):

$$w_{ij} = 1 - 0.5\sum_{k=1}^{K}\left|p_{ik} - p_{jk}\right| \tag{A.1}$$

where p_{ik} is the frequency of category *k* in individual *i*’s diet, and p_{jk} is the frequency of category *k* in individual *j*’s diet. The proportion of the *k*-th resource category in individual *i*’s diet, p_{ik} is calculated as:

$$p_{ik} = \frac{n_{ik}}{\sum_{k} n_{ik}} \tag{A.2}$$

where n_{ik} represent the number (or mass) of diet items in individual i's diet that fall into category k. Numbers of prey consumed may be more appropriate for behavioral ecologists studying prey capture decisions, and lends itself to Monte Carlo resampling schemes to test the null hypothesis of no niche variation. Mass may be more appropriate for studies of energy flux through a community or competition. The pairwise niche overlap ranges from 0 (no overlap) to 1 (total overlap). We therefore use w_{ij} to weight the edge connecting individuals i and j.

The average network density of connections

Having defined the niche overlap network, we needed to define a measure of the density of connections that can be incorporated into the clustering coefficient. This is because the degree of clustering of a network is directly proportional to the average network density of connections. This means that any clustering coefficient needs to be corrected for the average density of connections before it can be used to compare the degree of clustering among different networks.

First, we define O, the summation of the total pairwise overlap in the individual niche overlap network:

$$O = \sum w_{ij} \tag{A.3}$$

The average density of connections can be calculated by dividing O by $n(n-1)/2$, which corresponds to the number of edges of a completely connected network with n individuals. By doing this we are standardizing the total pairwise overlap O by the number of potential edges in the network, yielding:

$$\breve{O} = \frac{O}{n(n-1)/2}, \qquad \text{(A.4)}$$

which is the average network density of connections.

The index of clustering $\mathrm{C_{ws}}$

We first take C_{w_i}, which is a combined measure of the number and weight of the edges around individual i and among the nodes directly connected with i defined as:

$$C_{w_i} = \frac{1}{k_i(k_i-1)} \sum_{j,h} \left(\hat{w}_{ij}\hat{w}_{jh}\hat{w}_{hi}\right)^{1/3} \qquad \text{(A.5)}$$

where k_i is the number of edges between individual i and its neighbors; w_{nm} is the weight of the edge between individual n and m; $\hat{w}_{nm} = w_{nm}/\max(w)$. The summation, therefore, quantifies the weights of all edges between individual i and its neighbors ($a_{ij}\ a_{ih}$) that are also neighbors to each other (a_{jh}). We then define C_w, which is the average value of the individual clustering coefficients, C_{w_i}, for all nodes in the network.

C_w is approximately equal to the average network density of connections (measured by $\breve{O}$) of a totally random network, so that $C_w \sim \breve{O}$. In our case, a totally random network consists of individuals that sample randomly from the population niche. This means that two random networks will differ in their measures of C_w simply if they differ in the average density of connections $\breve{O}$. As a consequence, directly using C_w as a measure of the degree of clustering may be misleading, especially if one wants to compare different networks. As a way to circumvent this problem, we define C_{ws}, which is a correction of C_w that controls for the effect of $\breve{O}$:

$$C_{ws} = \frac{(C_w - \breve{O})}{(C_w + \breve{O})} \tag{A.6}$$

Now, the degree of clustering is measured relative to $\breve{O}$, and in a totally random network $C_{ws} \sim 0$. An interesting feature of C_{ws} is that it can assume both positive and negative values. C_{ws} will be positive ($C_w > \breve{O}$) if the local density of connections is higher than the overall density of connections, indicating that the population is characterized by clusters of individuals sharing common resources. In contrast, C_{ws} will be negative ($C_w < \breve{O}$) when the local density of connections is lower than the overall density of connections, indicating that individuals usually use a very particular combination of resources that differs from that of other individuals (i.e. individuals' diets are overdispersed).

The index E *of among-individual diet variation*

The average network density of connections $\breve{O}$ (A.4) can also be interpreted as a measure of among-individual diet variation as follows. In a network composed of individuals whose diets are identical (no among-individual niche variation), all individuals are connected and $w_{ij} = 1$ for all pairs of individuals. In such a network, the summation of the total pairwise overlap O (A.3) equals $n(n-1)/2$, the number of edges of a completely connected network with n individuals. However, if there is diet variation, at least part of the w_{ij} values will be less than 1 and, as a consequence, O will be less than $n(n-1)/2$. The higher the degree of diet variation, the smaller the value of O. Therefore, the value of $\breve{O}$, which divides O by $n(n-1)/2$, can be interpreted as a measure of the degree of among-individual diet variation. $\breve{O}$ will range from 1 when there is no diet variation, towards 0 as variation increases. This opens the possibility of creating a new index based on complex network theory to measure the degree of diet variation. We therefore define the index:

$$E = 1 - \breve{O} \tag{A.7}$$

, which is positively related to niche variation and ranges from zero when all individuals have identical diets and there is no diet variation, to 1 as diet variation increases. The sampling variance of E is known (see below), allowing parametric comparisons between populations or tests of null hypotheses of no diet variation, making E preferable to previous indices of diet variation (Bolnick et al. 2002).

We found significant among-individual diet variation within all enclosures, and in the wild-caught control fish (all $E > 0.60$; all $P < 0.001$, Monte Carlo simulations). The *t*-tests using the calculated variances for E indicated that the observed E-values were larger than zero in all samples ($P < 0.0005$), in agreement with the non-parametric tests. Consistent with the previous analysis by Svanbäck and Bolnick (2007), our new index of diet variation indicates higher average among-individual diet variation at high (HD) than low density (LD) treatments (HD, average E =0.786; LD, average E = 0.687; paired *t*-test, $t_4 = -4.171$, $P = 0.014$). Diet variation was significantly higher in the HD treatment than the in the control sample (one-sample paired *t*-test, $t_4 = 6.073$; $P = 0.004$), whereas there was no difference between the LD treatment and the control ($t_4 = 0.393$; $P = 0.714$).

The sampling variance of the index E

A Jackknife estimation of the variance of the average density of connections $\breve{O}$ can be derived using the formalism of *U*-statistics (Arversen 1969). We first note that

$$\breve{O} = \binom{n}{2}^{-1} \sum_{1 \le i \le j \le n} w_{ij} \tag{A.8}$$

i.e. $\breve{O}$ is a *U*-statistic of degree 2 and kernel given by w_{ij} (A.1). It is asymptotically normal with mean θ and variance $4\xi_1^2$ for

$$\theta = Ew_{12} = 1 - 0.5\sum_{k=1}^{K} E\left|p_{1k} - p_{2k}\right| \tag{A.9}$$

$$\xi_1^2 = Var\varphi_1(\mathbf{X}_1) \tag{A.10}$$

where $\mathbf{X}_1 = (p_{11},\ldots,p_{1K})'$ and $\varphi_1(\mathbf{X}_1) = E\{1 - 0.5\sum_{k=1}^{K}\left|p_{1k} - p_{2k}\right| \mid \mathbf{X}_1\}$. As long as $\xi_1^2 > 0$, the *U*-statistic is non-degenerate. In our case, given $n_{i\bullet}$ (the total number of food items of the *i*-th individual), and taking the distribution of the food items $\mathbf{X}_i = (x_{i1},\ldots,x_{iK})'$ to be $M(n_{i\bullet},\pi_1,\ldots,\pi_K)$, we can then write

$$\varphi_1(\mathbf{X}_1) = 1 - 0.5\sum_{v_2=1}^{+\infty} P(n_{2\bullet} = v_2)\sum_{k=1}^{K}\sum_{x_{2k}=0}^{v_2}\left|p_{1k} - \frac{x_{2k}}{v_2}\right|\binom{v_2}{x_{2k}}\pi_k^{x_{2k}}(1-\pi_k)^{v_2 - x_{2k}} \tag{A.11}$$

If, and only if, there is one single food category, i.e. $\pi_1 = 1$ and $\pi_2 = \cdots = \pi_K = 0$, so that $\varphi_1(\mathbf{X}_1) = 0$ almost surely, will the *U*-statistic be degenerate. This will happen when, and only when, all individuals are specialized on the same single resource. Otherwise, this *U*-statistic will behave in a reasonable fashion and asymptotic normality is attained. Under asymptotic normality, one can employ the standard deviation for building asymptotic intervals and to perform asymptotically powerful tests. Moreover, the variance of $\breve{O}$ can be obtained by Jackknifing the *U*-statistics by the following formula (Sen 1960, Arversen 1969):

$$Var(\breve{O}) = Kn^2(n-1)\binom{n-1}{2}^{-2}\sum_{c=0}^{2}(cn-4)S_c \qquad \text{(A.12)}$$

where $S_c = \sum w_{i_1,i_2} w_{i_3,i_4}$, for any resample $\{i_1, i_2, i_3, i_4\}$ from $\{1, \ldots, n\}$, $c = 0, 1, 2$ being the number of coincident indices, and the sum in S_c being taken for all such quadruples. Note that resampling is performed among the individuals and not among food items for a single individual. This is done to preserve the underlying stochastic dependency structure within individual resource distributions and, therefore, produce a more robust estimate, without the need and the associated shortcomings of assuming some specific dependency setup. The variance of the index E in turn is given by

$$Var(E) = Var(1-\breve{O}) = (-1)^2 Var(\breve{O}) = Var(\breve{O}) \qquad \text{(A.13)}$$

so that the variance of $\breve{O}$ holds for E.

LITERATURE CITED

Arversen, J. N. 1969. Jackknifing *U*-statistics. Annals of Mathematical Statistics **40**:2076-2100.

Bolnick, D. I., L. H. Yang, J. A. Fordyce, J. M. Davis, and R. Svanbäck. 2002. Measuring individual-level resource specialization. Ecology **83**:2936-2941.

Saramäki, J., M. Kivelä, J. P. Onnela, K. Kaski, and J. Kertész. 2007. 2002. Generalizations of the clustering coefficient to weighted complex networks. Physical Review E **75**: 027105.

Schoener, T. W. 1968. The *Anolis* lizards of Bimini: resource partitioning in a complex fauna. Ecology **49**:704-726.

Sen, P. K. 1960. On some convergence properties of *U*-statistics. Calcutta Statistical Association Bulletin **10**:1-18.

Svanbäck, R., and D. I. Bolnick. 2007. Intraspecific competition drives increased resource use diversity within a natural population. Proceedings of the Royal Society of London. Series B: Biological Sciences **274**:839-844.