Exploring different turnover metrics

Cape vs SWA publication

Ruan van Mazijk

2018-12-28

Definitions of turnover &/or similarity

Table 1: A glossary

Symbol	Description	Derivation
A	No. species in site 1	
B	No. species in site 2	
J ($\equiv a$)	No. species shared by sites 1 & 2	$\lvert A \cap B \rvert$
b	No. species exclusively in site 1	$\lvert A - J \rvert$
c	No. species exclusively in site 2	$\lvert B - J \rvert$

Jaccard distance (`vegan::vegdist(x, method = "Jaccard")`)

$$\beta_{d_J} = \frac{A + B - 2J}{A + B - J}$$

Jaccard similarity (Koleff et al. 2003. *J. Anim. Ecol.* 72(3))

$$\beta_{s_J} = \frac{a}{a + b + c}$$

(Note: $\beta_{d_J} = 1 - \beta_{s_J}$)

β_g distance (Gaston et al. 2001 in Koleff et al. 2003. *J. Anim. Ecol.* 72(3))

$$\beta_{d_g} = \frac{b + c}{a + b + c}$$

β_{gl} distance (Lennon et al. 2001 *J. Anim. Ecol.* 70(6))

$$\beta_{d_{gl}} = \frac{2|b - c|}{2a + b + c}$$

Expressing turnover in common terms

As above,

$$\beta_{d_J} = \frac{A + B - 2J}{A + B - J}$$

Thus, given the indentities in Table 1,

$$
\begin{aligned}
\beta_{d_g} &= \frac{b+c}{a+b+c} \\
&= \frac{(A-J)+(B-J)}{J+(A-J)+(B-J)} \\
&= \frac{A+B-2J}{A+B-J} \\
&\therefore \beta_{d_g} \equiv \beta_{d_J}
\end{aligned}
$$

And,

$$
\begin{aligned}
\beta_{d_{gl}} &= \frac{2|b-c|}{2a+b+c} \\
&= \frac{2|(A-J)-(B-J)|}{2J+(A-J)+(B-J)} \\
&= \frac{2|A-B|}{A+B}
\end{aligned}
$$

Also, Bray-Curtis distance (`vegan::vegdist(x, method = "bray")`) can be derived from Jaccard distance, as, given by `?vegan::vegdist`

$$
\beta_{d_J} = \frac{2\beta_{d_{BC}}}{1+\beta_{d_{BC}}}
$$

Thus,

$$
\beta_{d_{BC}} = \frac{-\beta_{d_J}}{\beta_{d_J}-2}
$$

Also, as in Table 1, let $b = |A-J|$ and $c = |B-J|$. Thus,

$$
\begin{aligned}
|A \cup B| &= a+b+c \\
&= J+|A-J|+|B-J| \\
&= A+B-J
\end{aligned}
$$

Is “the no. species not shared” the same as $\gamma - \alpha$?

Table 2: Another glossary

Symbol	Description	Derivation
$\mathbf{N}$	A neighbourhood of $2 \leq n \leq 4$ cells	
δ	No. species not shared by sites	$\lvert A + B - 2J \rvert$
γ	No. species across sites	$\lvert A + B - J \rvert$
$\beta_{d_J ij}$	Jaccard distance between cells i & j	See above
$\overline{\beta_{d_J ij}}$		$\frac{\delta}{\gamma}$
$\overline{\alpha}$	Mean no. species in sites	

We wish to quantify the notion of the absolute number of species gained or lost when moving from one cell to another: δ.

The average no. species not shared $\overline{\delta}$ by cells i and j in neighbourhood $\mathbf{N}$ can be derived using Jaccard distance (β_{d_J}) as follows

$$\overline{\delta_{ij}}(\mathbf{N})_{\beta_{d_J}} = \gamma(\mathbf{N}) \times \overline{\beta_{d_J ij}}(\mathbf{N})$$

The difference between γ and average α in $\mathbf{N}$ is simply

$$\overline{\delta_{ij}}(\mathbf{N})_{\alpha} = \gamma(\mathbf{N}) - \overline{\alpha_i}(\mathbf{N})$$

Additional definitions are in Table 2.

Suppose the neighbourhood $\mathbf{N}$ has four constituent cells each with richness 10. If you were to walk from cell no. 1 to 2, then 2 to 3 and then 3 to 4, three of the species you encounter as you cross into a new cell will differ from those in the previous cell. Thus, $\gamma(\mathbf{N}) = 19$.

The pairwise δ-values are 6, 6, 6, 12, 12, 18. The mean of these ($\overline{\delta_{ij}} = 10$, i.e. the no. species we gain or lose on average when moving between cells in $\mathbf{N}$) is the true value we wish to encapsulate.

From the $\mathbf{N}$ we have set up, we can calculate $\overline{\delta_{ij}}(\mathbf{N})_{\beta_{d_J}}$ and $\overline{\delta_{ij}}(\mathbf{N})_{\alpha}$.

The pairwise β_{d_J}-values are 0.6, 0.6, 0.6, 1.2, 1.2, 1.8, meaning that $\overline{\delta_{ij}}(\mathbf{N})_{\beta_{d_J}} = 19 \times 1 = 19$.

We know the α-values for $\mathbf{N}$, so $\overline{\delta_{ij}}(\mathbf{N})_{\alpha} = 19 - 10 = 9$.

Although it can deviate from $\overline{\delta_{ij}}$, $\overline{\delta_{ij}}(N)_{\beta_{d_J}}$ seems a much better definition of the notion of δ as it necessarily results in $\overline{\beta_{ij}}(\mathbf{N}) = \frac{\overline{\delta_{ij}}(\mathbf{N})}{\gamma(\mathbf{N})}$ (i.e. proportional turnover is the absolute no. species gained/lost relative to the neighbourhood richness), which is desirable.

Exploring $\overline{\delta_{ij}}(N)$'s components & definition

$$\mathbf{N} = \{\mathbf{N}_i\} = \{\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}\}$$

Where

$$\begin{aligned} \mathbf{N}_i &\subset \ \mathbf{N} \\ \mathbf{N}_i &= \{l, m, n, \dots\} \\ s.t.\ |\mathbf{N}_i| &= \alpha_i \end{aligned}$$

$$\gamma(\mathbf{N}) = \left| \bigcup_{i=1}^{n} \mathbf{N}_i \right|$$

$$\begin{aligned} \delta_{ij} &= |A + B - 2J| \\ &= ||\mathbf{N}_i \cup \mathbf{N}_j| - 2|\mathbf{N}_i \cap \mathbf{N}_j|| \end{aligned}$$

$$\begin{aligned} \mathbf{D} &= \delta(\mathbf{N}) \\ &= \{\delta_{12}, \delta_{13}, \dots, \delta_{ij}, \dots, \delta_{n,n-1}\} \\ &\quad where\ 2 \le n \le 4\ and\ j \neq i \\ s.t.\ p = |\mathbf{D}| &= \binom{n}{2} \end{aligned}$$

Thus, $\overline{\delta_{ij}}(\mathbf{N})$ being the average $\mathbf{D}_k$ for a neighbourhood $\mathbf{N}$, we can define $\overline{\delta_{ij}}(\mathbf{N})$ in two ways.

$$\begin{aligned} \overline{\delta_{ij}}(\mathbf{N}) &= \frac{1}{p} \sum_{i=1}^{n} \sum_{j=1}^{n, j \neq i} \delta_{ij} \\ &= \frac{1}{p} \sum_{k=1}^{p} \mathbf{D}_k \end{aligned}$$

ASIDE: ***WHAT ABOUT THIS??***

$$\overline{\delta_{ij}}(\mathbf{N})_{??} = \overline{\alpha_i}(\mathbf{N}) \times \overline{\beta_{d_J\,ij}}(\mathbf{N}) = 10 \times 1 = 10$$

Is $\overline{\delta_{ij}}(\mathbf{N})$ derivable from $\overline{\beta_{d_J ij}}(\mathbf{N})$ (truly)?

We have two definitions of $\overline{\delta_{ij}}$.

First, deriving it from first principles and raw calculations of $\mathbf{D}$ from $\mathbf{N}$:

$$\overline{\delta_{ij}}(\mathbf{N})_{\mathbf{D}} = \frac{1}{p}\sum_{k=1}^{p}\mathbf{D}_k$$

Second, deriving "after-the-fact" from averaged Jaccard distances

$$\overline{\delta_{ij}}(\mathbf{N})_J = \gamma(\mathbf{N}) \times \overline{\beta_{d_J ij}}(\mathbf{N})$$

We wish to assess whether $\overline{\delta_{ij}}(\mathbf{N})_J$ actually equals (or at least estimates) $\overline{\delta_{ij}}(\mathbf{N})_{\mathbf{D}}$.