Partitioning BD into replacement and richness difference

Pierre Legendre
Département de sciences biologiques
Université de Montréal

© Pierre Legendre 2018
Outline of the presentation

1. Whittaker’s alpha, beta and gamma diversities
2. Compute total beta diversity ($BD_{Total}$)
3. Partitioning $D$ into $Repl$ and $RichDiff$ components
4. Case study: Doubs River fish data
5. Conclusion
6. References
1. Alpha, beta and gamma diversities

- **Alpha diversity** is local diversity – or species diversity at a site. Estimated by species richness or by one of the alpha diversity indices (richness, Shannon, Simpson).

- **Beta diversity** is spatial differentiation – or the variation in species composition among sites within a region of interest.

- **Gamma diversity** is regional diversity – or species diversity in a region of interest. Estimated by pooling observations from a large number of sites in the area and computing an alpha diversity index.

Species

\[ 1 \ 2 \ 3 \ \ldots \ \ p \]

Sums

\[ n \]

\[ \alpha_i = N_0, \ H_1, \ N_2, \ldots \]

\[ \gamma = N_0, \ H_1, \ N_2, \ldots \]

Y = community composition data

\[ \beta = \text{variation in species composition among sites} \]

Legendre & Legendre *Numerical ecology* (2012, Fig. 6.3).
2. Compute total beta diversity ($\text{BD}_{\text{Total}}$)

Transformed species composition

$Y_{\text{tr.nxp}}$ → (1) $S_{\text{nxp}}$

Species composition

$Y_{\text{nxp}}$ → (7) $D_{\text{nxn}}$

Squared differences from column means

$S_{\text{nxp}}$

Species contributions to beta diversity (SCBD)

1 ... p → (4)

1 ... n

Local contributions to beta diversity (LCBD)

Total dispersion

$SS_{\text{Total}}$ → (3) $\text{BD}_{\text{Total}}$

Total beta diversity

Replacement and Richness difference

5
3. Partitioning $D$ into $Repl$ and $RichDiff$

$\Rightarrow$ Partition the dissimilarities $D$ into Replacement and Richness difference components.

Two different methods have been proposed:

(1) by Baselga (in Spain) and

(2) by Podani (in Hungary) and coauthors.

Each group proposed ways of partitioning the Jaccard and Sørensen dissimilarity indices for presence-absence data, as well as their quantitative forms, the Ružička and percentage difference indices.

$\Rightarrow$ I will focus on the Podani approach. I will describe how these new indices can be interpreted through graphical analysis.
The dissimilarity $D$ between two sites contains two components:

- **replacement** (turnover): environmental gradients, competition …

- and **richness difference** due to difference in diversity of niches causing species thinning, or other processes (e.g. physical barriers).
Podani-family indices, presence-absence data

**Components of \( D_J \) and \( D_S \)**

<table>
<thead>
<tr>
<th>Site 1</th>
<th>Site 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Site 1 components" /></td>
<td><img src="image" alt="Site 2 components" /></td>
</tr>
</tbody>
</table>

**Dissimilarity**

<table>
<thead>
<tr>
<th>Replacement</th>
<th>Richness difference</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Presence-absence data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Numerators (num.)</strong></td>
</tr>
<tr>
<td>Replacement num.</td>
</tr>
<tr>
<td>Richness or abund. difference num.</td>
</tr>
<tr>
<td>Dissimilarity num.</td>
</tr>
</tbody>
</table>

Replacement and Richness difference
<table>
<thead>
<tr>
<th>Components of $D_J$ and $D_S$</th>
<th>a</th>
<th>b</th>
<th>c</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site 1</td>
<td>□ □ □ □ □</td>
<td>□ □ □ □ □</td>
<td>□ □ □ □ □</td>
</tr>
<tr>
<td>Site 2</td>
<td>□ □ □ □ □</td>
<td>□ □ □ □ □</td>
<td>□ □ □ □ □</td>
</tr>
</tbody>
</table>

**Dissimilarity**

- $b+c$
- $2\times\min(b,c)$

**Replacement**

- $|b-c|$

**Richness difference**

- $|b-c|$

<table>
<thead>
<tr>
<th>Presence-absence data</th>
<th><strong>Numerators (num.)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Replacement num.</td>
<td>$2\min(b,c) = (b+c) -</td>
</tr>
<tr>
<td>Richness or abund. difference num.</td>
<td>$</td>
</tr>
<tr>
<td>Dissimilarity num.</td>
<td>$(b+c) = 2\min(b,c) +</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Jaccard group ($J$)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Denominator</td>
</tr>
<tr>
<td>$D$</td>
</tr>
<tr>
<td>Repl</td>
</tr>
<tr>
<td>RichDiff</td>
</tr>
</tbody>
</table>
### Components of $D_J$ and $D_S$

<table>
<thead>
<tr>
<th>Site 1</th>
<th>Site 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Site 1 Components" /></td>
<td><img src="image2.png" alt="Site 2 Components" /></td>
</tr>
</tbody>
</table>

### Dissimilarity

- $b+c$
- $2 \times \min(b,c)$

### Replacement

- $|b-c|$ 

### Richness difference

### Presence-absence data

<table>
<thead>
<tr>
<th>Numerators (num.)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Replacement num.</td>
<td>$2 \times \min(b,c) = (b+c) -</td>
</tr>
<tr>
<td>Richness or abund. difference num.</td>
<td>$</td>
</tr>
<tr>
<td>Dissimilarity num.</td>
<td>$(b+c) = 2 \times \min(b,c) +</td>
</tr>
</tbody>
</table>

### Sørensen group ($S$)

| Denominator | $(2a+b+c)$ |

### $D$

Sørensen dissimilarity:

$$D_S = \frac{(b+c)}{(2a+b+c)}$$

### $Repl$

$$Repl_S = \frac{2 \times \min(b,c)}{(2a+b+c)}$$

### RichDiff

$$RichDiff_S = \frac{|b-c|}{(2a+b+c)}$$
Podani-family indice, species abundance data

\[ A = \text{abun. common to sites 1 and 2} = A_1 + A_2 + 0 + A_4 \]
\[ B = \text{abundances unique to site 1} = 0 + B_2 + B_3 + 0 \]
\[ C = \text{abundances unique to site 2} = C_1 + 0 + 0 + 0 \]

<table>
<thead>
<tr>
<th></th>
<th>Presence-absence data</th>
<th>Species abundance data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Numerators (num.)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Replacement num.</td>
<td>2 ( \min(b,c) )</td>
<td>2 ( \min(B,C) )</td>
</tr>
<tr>
<td>Richness or abund. difference num.</td>
<td>(</td>
<td>b−c</td>
</tr>
<tr>
<td>Dissimilarity num.</td>
<td>( (b+c) = 2 \min(b,c) +</td>
<td>b−c</td>
</tr>
</tbody>
</table>
### Numerators (num.)

<table>
<thead>
<tr>
<th>Presence-absence data</th>
<th>Species abundance data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replacement num.</td>
<td>2 ( \min(b,c) )</td>
</tr>
<tr>
<td>Richness or abund. difference num.</td>
<td>(</td>
</tr>
<tr>
<td>Dissimilarity num.</td>
<td>( b+c )</td>
</tr>
</tbody>
</table>

### Jaccard group (\( J \))

<table>
<thead>
<tr>
<th>Denominator</th>
<th>( (a+b+c) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( D )</td>
<td>Jaccard dissimilarity:</td>
</tr>
<tr>
<td></td>
<td>( D_J = (b+c) / (a+b+c) )</td>
</tr>
<tr>
<td>( Repl )</td>
<td>( Repl_J = 2 \min(b,c) / (a+b+c) )</td>
</tr>
<tr>
<td>( RichDiff ) or ( AbDiff )</td>
<td>( RichDiff_J =</td>
</tr>
</tbody>
</table>

**Ružička dissimilarity:**

<table>
<thead>
<tr>
<th>Denominator</th>
<th>( (A+B+C) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( D )</td>
<td>Ružička dissimilarity:</td>
</tr>
<tr>
<td></td>
<td>( D_R = (B+C) / (A+B+C) )</td>
</tr>
<tr>
<td>( Repl )</td>
<td>( Repl_R = 2 \min(B,C) / (A+B+C) )</td>
</tr>
<tr>
<td>( RichDiff ) or ( AbDiff )</td>
<td>( AbDiff_R =</td>
</tr>
</tbody>
</table>
### Numerators (num.)

<table>
<thead>
<tr>
<th>Presence-absence data</th>
<th>Species abundance data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replacement num.</td>
<td>2 min(b,c)</td>
</tr>
<tr>
<td>Richness or abund. difference num.</td>
<td>(</td>
</tr>
<tr>
<td>Dissimilarity num.</td>
<td>((b+c))</td>
</tr>
</tbody>
</table>

### Sørensen group (S)

<table>
<thead>
<tr>
<th>Denominator</th>
<th>((2a+b+c))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2A+B+C)</td>
<td></td>
</tr>
</tbody>
</table>

### \(D\)

- **Sørensen dissimilarity:**  
  \[D_S = \frac{(b+c)}{(2a+b+c)}\]
- **Percentage difference dissimilarity:**  
  \[D_{\%\text{diff}} = \frac{(B+C)}{(2A+B+C)}\]

### \(Repl\)

- **Repl**  
  \[Repl_S = 2 \text{ min}(b,c) / (2a+b+c)\]
- **Repl\%\text{diff}**  
  \[Repl_{\%\text{diff}} = 2 \text{ min}(B,C) / (2A+B+C)\]

### \(RichDiff\) or \(AbDiff\)

- **RichDiff** or **AbDiff**  
  \[RichDiff_S = \frac{|b-c|}{(2a+b+c)}\]
  \[AbDiff_{\%\text{diff}} = \frac{|B-C|}{(2A+B+C)}\]
4. Case study: Doubs River fish data

Fish observed at 29 sites along the Doubs river, a tributary of the Saône River running near the France-Switzerland border in the Jura Mountains in eastern France.

• Data from Verneaux (1973), available on the Web page of the book *Numerical ecology with R* (Borcard et al. 2011):

  http://adn.biol.umontreal.ca/~numericalescology/numecolR/

Les sources du Doubs à Mouthe

Entre Laissey et Deluz, peu avant Besançon

Besançon
Decomposing dissimilarity matrices

\[ \text{Repl} + \text{RichDiff} = D \]
A development –

We can compute \( BD_{\text{total}} \) for the \textbf{Repl} and \textbf{RichDiff} matrices, just as we can do it for the \( D \) matrix:

\[
BD_{\text{Total}} = \frac{1}{n(n-1)} \sum_{h=1}^{n} \sum_{i=h+1}^{n} D_{hi}
\]

With this equation, we obtain \( \text{Repl}_{\text{total}} \) and \( \text{RichDiff}_{\text{total}} \).
• In the Podani family of indices used in the examples (below), the partitioning is additive: $BD_{\text{total}}(\text{Repl}) + BD_{\text{total}}(\text{RichDiff}) = BD_{\text{total}}(D)$.

• In the Baselga family (not developed further in this talk), RichDiff is replaced by Nestedness (Nes). The partitioning is not additive: $BD_{\text{total}}(\text{Repl}) + BD_{\text{total}}(\text{Nes}) \neq BD_{\text{total}}(D)$. 
Decomposition of the fish beta diversity for presence-absence data, Podani family

<table>
<thead>
<tr>
<th></th>
<th>Jaccard</th>
<th>Sørensen</th>
</tr>
</thead>
<tbody>
<tr>
<td>BD$_{\text{Total}}$</td>
<td>0.326</td>
<td>0.267</td>
</tr>
<tr>
<td>BD$_{\text{max}} = 0.5$</td>
<td>65% of BD$_{\text{max}}$</td>
<td>53% of BD$_{\text{max}}$</td>
</tr>
<tr>
<td>Repl$<em>{\text{Total}}$ / BD$</em>{\text{Total}}$</td>
<td>0.28</td>
<td>0.28</td>
</tr>
<tr>
<td>RichDiff$<em>{\text{Total}}$ / BD$</em>{\text{Total}}$</td>
<td>0.72</td>
<td>0.72</td>
</tr>
</tbody>
</table>

=> Beta diversity in the Doubs River is dominated by richness differences among sites
(a) Jaccard dissimilarity ($D$, ●), replacement (■) and richness difference (▲) indices for presence-absence data: sites 1–29 are compared to site 30. (b) Species richness. Grey: polluted #23 to 25.
Left: Ordination of the sites, Sørensen dissimilarity ($D_S$).
Right: Local contributions to beta ($LCBD_S$) on map of Doubs River.
=> Sites 1-3: pristine, few species. Sites 23-25: urban pollution.
Left: Ordination of the sites, $Repl_S$ (28% of beta diversity).

Right: Replacement LCBD ($Repl_{LCBD}$) on map of Doubs River.

$\Rightarrow$ Higher replacement LCBD at sites 11-15 and 23-25.
Left: Ordination of the sites, $RichDiff_S$ (72% of beta diversity).

Right: Richness difference LCBD ($RichDiff_{LCBD}$) on Doubs River map.

$\Rightarrow$ Higher richness difference LCBD at sites 1-3 and 23.
R code: beta diversity partitioning, Doubs River fish data

```r
# Load the Doubs.RData file
# Data frame "spe" contains the fish data (30 sites x 27 species)
spe <- spe[-8,]            # Remove site 8 were no fish were caught
# Compute partitioning with function beta.div.comp() of adepsatial
res1 <- beta.div.comp(spe, coef="S", quant=FALSE) # Sørensen coeff.  
summary(res1)            # Examine the structure of the output file
res1$part

<table>
<thead>
<tr>
<th></th>
<th>Repl</th>
<th>RichDif</th>
<th>Repl/BDtotal</th>
<th>RichDif/BDtotal</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDtotal</td>
<td>0.26738159</td>
<td>0.19180774</td>
<td>0.28264417</td>
<td>0.71735583</td>
</tr>
</tbody>
</table>

Compare the results in this vector to the Doubs River results computed with the Sørensen index, shown 5 slides back.
Function `beta.div.comp()` also computes the D matrix (Jaccard, Sørensen, Ružička, percentage difference) that was requested by the user, as well as the Replacement and Richness difference matrices that can be used for ordination or other representations of the results.

These three matrices can be inputted into the `LCBD.comp()` function of `adespatial` to compute \( SS_{\text{total}} \) and \( BD_{\text{total}} \), as well as the LCBD indices associated with the input matrix.

```
# LCBD indices computed for the Richness difference matrix
res2 <- LCBD.comp(res1$rich, sqrt.D=FALSE)
res2$beta
    SS_{\text{total}}  BD_{\text{total}}
2.978053  0.106359

res2$LCBD
[1] 0.178 0.077 0.052 0.012 0.007 0.008 0.035 0.035 0.024 0.024
[11] 0.024 0.024 0.008 0.007 0.018 0.032 0.032 0.035 0.032 0.032
[21] 0.032 0.077 0.012 0.012 0.029 0.032 0.032 0.044 0.029
```

The values in the matrix are not square-rooted because the RichDiff matrix computed from the Sørensen D is Euclidean. See Legendre (2014, Table S1.4).

These values were used to draw the map of RichDiff LCBD indices a few slides back. *Note:* LCBD #22 in the list corresponds to site #23 because site #8 was removed from the data file.
Test a hypothesis about the response of community data to a factor

The environmental data were used to classify the sites in two groups (Ward clustering): sites 1-22 (upper course) and 23-30 (lower course).

The $D_J$, $\text{Repl}_J$ and $\text{RichDiff}_J$ matrices were tested against this classification using McArdle & Anderson’ s (2001) $F$-test in RDA for non-Euclidean matrices. Function dbrda() in vegan was used.

Results –

$D_J$ was significantly explained by the classification ($p = 0.005$)
$\text{Repl}_J$ was significantly explained by the classification ($p = 0.001$)
$\text{RichDiff}_J$ was not significantly explained by the classification. ($p = 0.943$)

$=>$ This analysis could be carried out with an experimental factor.
Identify the environmental factors responsible for variation in the matrices

The $D_J$, $\text{Repl}_J$ and $\text{RichDiff}_J$ matrices were transformed into rectangular data matrices by principal coordinate analysis (PCoA). Environmental variables that significantly explained the variation in each matrix were selected (forward selection).

Results –
$D_J$ – was explained by $\{\text{slope, hardness, nitrate, } O_2\}$
$\text{Repl}_J$ was explained by $\{O_2\}$
$\text{RichDiff}_J$ was explained by $\{\text{slope, hardness, nitrate}\}$

$\Rightarrow$ The two components, $\text{Repl}_J$ and $\text{RichDiff}_J$, are influenced by different environmental processes, which all influence $D_J$. 


Other methods of graphical analysis have been suggested – for example the triangular plots proposed by Podani & Schmera (2011) and Podani et al. (2013).

The analysis of the Doubs River fish data using triangular plots is shown in the paper.
Replacement and richness difference indices can be interpreted and related to ecosystem processes.

**Innovations in the Legendre (2014) paper** –

- The index values can be summed across all pairs of sites; these sums decompose *total beta diversity* into *total replacement* and *total richness difference* components.

- Local contributions of replacement and richness difference to total beta diversity can be computed and mapped.

- Within a region, differences among sites measured by these indices can be analysed and interpreted using explanatory variables.

- Replacement and richness difference matrices can be analysed by all methods of multivariate data analysis for dissimilarity matrices.
6. References


End of section