Advances in plant trend analyses

sMon – Analyzing trends in German biodiversity data

Workshop 3rd – 7th December 2018
iDiv - Leipzig
Data source:

Plant occurrence records (i.e. presence only data) from different sources:

- FlorKart Database (BfN)
- Occurrence records from federal states
- Occurrence records deduced from plant relevées
- approx. 28 Mio occurrence records from 1960 - 2017

Standardized taxonomy
Deduction of trends:

FreScaLo Algorithm (Hill 2012)

Correction for observer bias through neighbourhood occurrences.

Calculation of „probability of being present“ of species $i$ at location $j$ and time $t$
FreScaLo Algorithm (Hill 2012)

Ecological similarity of a region

- edaphic
- climatic
- topographic
FreScaLo Algorithm (Hill 2012)

- Proportion of species \(i\) being present in the neighbourhood of site \(j\)
- Ecological similarity of neighbouring sites
- Distances of neighbouring sites
- Probability of being present in site \(j\)
FreScaLo Algorithm (Hill 2012)

**Probability of being present in site j**

- Number of characteristic (i.e. most frequent) species in the neighbourhoods
- "Wants list" of so-called benchmark-species
- Proportion of number of benchmark specs found in site j
- Measure of sampling effort
FreScaLo Algorithm (Hill 2012)

- Proportion of species i being present in the neighbourhood of site j
- Proportion of number of benchmark specs found in site j

Probability of being present in site j
Probability of species $i$ being present at site $j$

Scales between 0 and 1 (percentage)

Computed for all time steps $t$
Probability of species $i$ being present at site $j$
Probability of species $i$ being present at site $j$

Spatial autocorrelation!

Across time?!??!
FreScaLo Algorithm (Hill 2012)

Calculation of the TFactor:

The estimated frequency of species $j$ at time $t$

Linear Model

$$y = x^*b + a + e$$
FreScaLo Algorithm (Hill 2012)

What about all this precious data?
What about all this precious data? Complicated!

Spatial autocorrelation! Across time!??!!

0 – 1 !??????!!!! Bugger!!!
FreScaLo Algorithm (Hill 2012)

Calculation of the TFactor:

The estimated frequency of species j at time t

Linear Model

\[ y = x^*b + a + e \]
Spatial autocorrelation 

Integrated Nested Laplace Approximation!

Temporal autocorrelation 

INLA is my Friend!!!!

Among them:

Beta distribution (0-1)

Heaps and heaps of beautiful probability density functions!

0 – 1?
Model with timestep and coordinates

Check the residuals

v.s. fitted
Model with timestep and coordinates

Check the residuals

v.s. response

Cols = Timesteps!
Model with timestep and coordinates;
Including spatial random term (i.e. latent process)

Check the residuals

v.s. fitted

Cols= Timesteps!
Model with timestep and coordinates;  
Including spatial random term (i.e. latent process)  

Check the residuals  

v.s. response  

Cols= Timesteps!
Model with timestep and coordinates; Including spatial random term (i.e. latent process)

Check the residuals

v.s. timestep (predictor)

Cols= Timesteps!
Model with timestep and coordinates; Including spatial random term (i.e. latent process)

Check the predictions v.s. response

Cols= Timesteps!
Model with timestep and coordinates; Including spatial random term (i.e. latent process)
For each time step
Check the residuals

v.s. fitted

Cols= Timesteps!
Model with timestep and coordinates;
Including spatial random term (i.e. latent process)
For each time step

Check the residuals

v.s. Timestep (predictor)

Cols= Timesteps!
Model with timestep and coordinates; Including spatial random term (i.e. latent process) For each time step

Check the predictions

v.s. response

Cols= Timesteps!
### Model comparison

<table>
<thead>
<tr>
<th>Model description</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model with timestep and coordinates</td>
<td>-91636.01</td>
</tr>
<tr>
<td>Model with timestep and coordinates; Including spatial random term (i.e. latent process)</td>
<td>-96220.72</td>
</tr>
<tr>
<td>Model with timestep and coordinates; Including spatial random term (i.e. latent process) For each time step</td>
<td>-372525.12</td>
</tr>
</tbody>
</table>
The „latent process“ (i.e. spatial component in the model)

Model with timestep and coordinates;
Including spatial random term (i.e. latent process)
The „latent process“ (i.e. spatial component in the model)

Model with timestep and coordinates;
Including spatial random term (i.e. latent process)
The „latent process“ (i.e. spatial component in the model)

Model with timestep and coordinates; Including spatial random term (i.e. latent process)

For each time step:

Fixed effects:

<table>
<thead>
<tr>
<th>Effect</th>
<th>mean</th>
<th>sd</th>
<th>0.025quant</th>
<th>0.5quant</th>
<th>0.975quant</th>
<th>mode</th>
<th>kld</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>8.2992</td>
<td>0.2317</td>
<td>7.9443</td>
<td>8.3001</td>
<td>8.7532</td>
<td>8.2992</td>
<td>0</td>
</tr>
<tr>
<td>Timestep</td>
<td>-8.4227</td>
<td>0.1072</td>
<td>-8.6333</td>
<td>-8.4227</td>
<td>-8.2123</td>
<td>-8.4227</td>
<td>0</td>
</tr>
</tbody>
</table>
The „latent process“ (i.e. spatial component in the model)

Model with timestep an
For each time step

What drives the change in spatial dependency? (i.e. what is the „latent process“?)