

Spatial data

Data

The `bryo_belg.csv` dataset is adapted from the data of this study:

Neyens, T., Diggle, P.J., Faes, C., Beenaerts, N., Artois, T. et Giorgi, E. (2019) Mapping species richness using opportunistic samples: a case study on ground-floor bryophyte species richness in the Belgian province of Limburg. *Scientific Reports* 9, 19122. <https://doi.org/10.1038/s41598-019-55593-x>

This data frame shows the specific richness of ground bryophytes (*richness*) for different sampling points in the Belgian province of Limburg, with their position (x , y) in km, in addition to information on the proportion of forest (*forest*) and wetlands (*wetland*) in a 1 km² cell containing the sampling point.

```
bryo_belg <- read.csv("../donnees/bryo_belg.csv")
head(bryo_belg)
```

##	richness	forest	wetland	x	y
## 1	9	0.2556721	0.5036614	228.9516	220.8869
## 2	6	0.6449114	0.1172068	227.6714	219.8613
## 3	5	0.5039905	0.6327003	228.8252	220.1073
## 4	3	0.5987329	0.2432942	229.2775	218.9035
## 5	2	0.7600775	0.1163538	209.2435	215.2414
## 6	10	0.6865434	0.0000000	210.4142	216.5579

Fitting a geostatistical model

For this exercise, we will use the square root of the specific richness as the response variable. The square root transformation often allows to homogenize the variance of the count data in order to apply a linear regression.

Note: To model the count data directly, for example with a Poisson distribution, and also include spatial dependence, we would need to use a custom likelihood model or a Bayesian model.

- Fit a linear model of the transformed species richness to the proportion of forest and wetlands, without taking into account spatial correlations. What is the effect of the two predictors in this model?
- Calculate the empirical variogram of the model residuals in (a). Does there appear to be a spatial correlation between the points?

Note: The `cutoff` argument to the `variogram` function specifies the maximum distance at which the variogram is calculated. You can manually adjust this value to get a good view of the sill.

- Re-fit the linear model in (a) with the `gls` function in the *nlme* package, trying different types of spatial correlations (exponential, Gaussian, spherical). Compare the models (including the one without spatial correlation) with the AIC.
- What is the effect of the proportion of forests and wetlands according to the model in (c)? Explain the differences between the conclusions of this model and the model in (a).