

## Supplementary file 1

### Monte Carlo analysis of power (simulation method description)

Start by defining alternative sampling designs by varying the number of stands and number of samples within stand. Each design should be equally expensive to conduct on the field.

Assume that the between-stand variance  $\text{var}(b_i)=\sigma_b^2$ , within-stand-variance  $\text{var}(\varepsilon_{ij})=\sigma^2$  and regression coefficients  $\beta_1, \beta_2, \beta_3, \beta_4$  of model

$$y_{ij}=\beta_1+\beta_2T_{2ij}+\beta_3T_{3ij}+\beta_4T_{4ij}+b_i+\varepsilon_{ij}$$

are known. Here  $\beta_1$  is the mean of the first type of disturbance treatment,  $\beta_2$  - of the second, and so on. Element  $T_{2ij}$  is a binary indicator taking value 1 if the treatment is of the second type and 0 otherwise,  $T_{3ij}$  and  $T_{4ij}$  likewise binary indicators for third and fourth type treatments. The question is: what are the optimal number of stands ( $k$ ) and plots within stand ( $n/k$ ) for a given pre-specified total number of plots. The Monte Carlo simulation with  $N$  iterations is done as follows.

Repeat stages A and B for each alternative design:

Stage A. Repeat steps A1-A4  $N$  times for the selected design:

A1. Simulate  $k$  independent random effects  $b_i$  from distribution  $N(0, \sigma_b^2)$

A2. Simulate  $n$  residuals  $\varepsilon_{ij}$  from distribution  $N(0, \sigma^2)$

A3. Construct a simulated dataset from model for the effects of different types (in our case four) of experimental disturbance treatments using the simulated random effects and residuals. The same random effect is used for all observations of the same stand.

A4. Fit a linear mixed effect model of form as outlined above to the simulated data (the model

can be of different form than the one presented here based on the study setup in question).

Test the null hypothesis (in our case that means of soil and vegetation parameters are not different between the experimental treatment types) against the alternative (the means are different).

B. After the  $N$  iterations, calculate the proportion of cases where the null hypothesis was rejected. It gives you a Monte Carlo estimate of the power of the test under the applied design.

C. Select the design with the highest estimated power.