PROJECT INFORMATION

Project title: Deleterious effects of high nitrogen (N) availability on stand health

Project ID: 28

Contact person: Stavros D. Veresoglou (sveresoglou@zedat.fu-berlin.de)

PROJECT DESCRIPTION

We are a group of ecologists specializing on ecological synthesis. The focus of the entire group of professor Rillig is arbuscular mycorrhizas and fungal pathogens. Personally I work on the impact of nitrogen (N) availability on plant susceptibility to pathogens. Professor Rillig and I have recently published a meta-analysis in Plant Pathology (still in press) on the impact of fertilizer additions to plant susceptibility to fungal pathogens (citation: Veresoglou SD, Barto EK, Menexes G, Rillg MC. 2013 Fertilization affects severity of disease caused by fungal plant pathogens. Plant Pathology).
In the article we were able to demonstrate that N additions increase plant susceptibility to biotrophic and hemibiotrophic pathogens in agreement with the N disease hypothesis (Mitchel et al. 2003 Global Change Biology 9: 438-151).

We would like to build on these results and test the N disease hypothesis for forest ecosystems. Specifically we want to assess the effect of N availability on stand health. Because it is impossible to retrieve direct assessments of N availability we are planning to focus on foliar N:P ratios as a surrogate of N availability (for a review check Güssewell 2004 New Phytologist 164: 243-266). We understand that most of the N:P ratio literature is focused on bog ecosystems and is dependent on plant species. However, we are planning to use the N:P variable (in combination with foliar N status) as a continuous variable specific to each plant species in a modeling procedure.
We hypothesize that all else being equal (we will not explicit model other causes of declined stand health such as pollution or acidification) we will get a hump-shaped N:P ratio vs. stand health curve with minima for low and high availability of N. To account for climatic factors we are planning to use online recourses to retrieve at a crude scale temperature and precipitation data to be included in our model.

Our modelling procedure will be carried out in three steps:

(1) **Preliminary phase**: We are planning to implement a data mining technique called boosted regression trees (Elith et al. 2008 J. Animal Ecology 77: 802-813) to the entire dataset in order to evaluate the relative importance of climatic variables, nutrient stress and high N availability. This step will help us decide on the way nutrient stress should be incorporated in the model. The use of N:P ratio to model nutrient stress is suboptimal as this may not disentangle instances when plants have been limited by both N and P and due to complications in modelling humped curves (this is usually done through high order polynomials that can easily lead to overfitting of the models – N:P ratio will

only be used to model relative availability of N over P which is expected to have a monotonic negative effect on stand health). To avoid issues with collinearity [N] and [P] leaf concentrations cannot be simultaneously included in a model. [P] may be expected to represent a better surrogate for nutrient stress as plants can make luxurious utilization of P to a considerably larger extent than N. Based on the outcome of our boosted regression trees procedure we will be able to answer this question as well as visualize the relationship of each individual variable with stand health (this is important in deciding whether first order polynomials are sufficient to model each variable or higher order terms are also required).

(2) **Modelling phase**: Based on the results of the data mining procedure we will model stand health using logistic regression. A random subset of the data will be used to develop candidate models (*training data*) and the rest of the data will be used for evaluation purposes (*test data*) in agreement with earlier studies (e.g. Anderson et al. 2003 Ecological Modelling 162: 211-232). We anitcipate that there are at least two additional factors that need to be considered in our model. A random effects categorical factor, "*vegetation*, that will allow for variability in the stand health as a consequence of the plant species composition (for example the typical stand health for a pine forest may differ from an oak forest). An additional consideration of the categorical factor *vegetation* to account for the slope of the variable N:P ratio as the plasticity of the N:P bioindicator differs among plant species. The model in this simple form can be easily implemented using frequentistic statistics. In the case that additional complexity is introduced (depending on the form of the data that will be released) we are willing to carry out the modelling using Bayesian techniques in *Winbugs*. Parsimonious models will be selected using information criteria approaches (e.g. Burnham and Anderson 2002 Springer) unless Bayesian inferesence is used – *Winbugs* does not adequately support the DIC criterion for complex models (Kerry and Schaub 2011 Academic Press).

(3) **Evaluation phase**: We will compare relative performance of the optimals models that account for high N availability and those that do not to assess the ecological importance of considering high N availability as an additional variable when modelling stand health.
In the article we are willing to consider coauthorship of up to two contributors to the database provided they will be willing to comment on the manuscript (it is required by many journals all authors to have contributed to the writing of the paper).