

L11: Area-wide predictions

“It's my fault. I programmed you to want too much.”

Ed Dillinger, Tron

Things we cover in this session

- Predicting variable values using area wide raster data sets

Things you need for this session

- [W11-1 Satellite-supported predictions](#)

Things to take home from this session

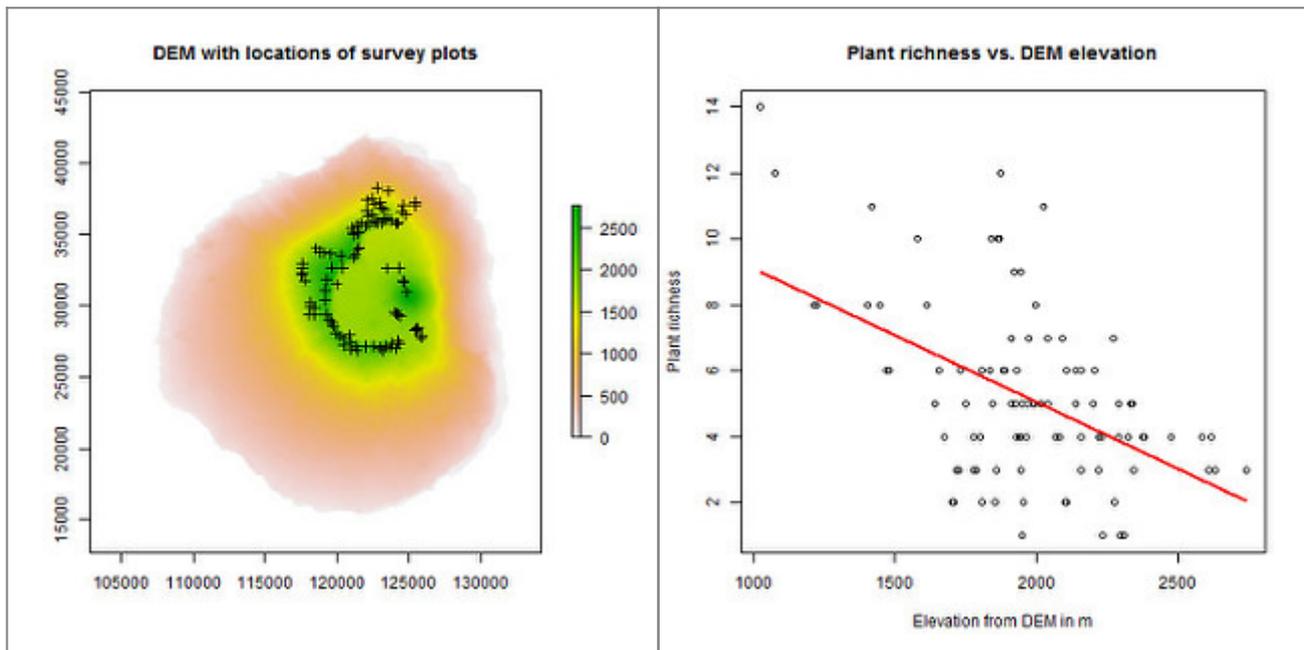
At the end of this session you should be able to

- visually check if the residuals of a linear model are normally distributed
- transform input variables of regression models if necessary
- predict values based on linear models

Area-wide prediction

As already introduced in [L05](#) and [L08](#), the predicting of data values based on a (linear or non-linear) regression model is quite straight forward.

So far we have always used some environmental information as independent variable for the prediction which has also been recorded at the field survey plots (e.g. vegetation coverage). Since we are now able to read raster data sets, we could also try to use the information of these rasters as independent variable.



For example, if we extract the height values from a DEM at the locations of our field survey plots and given that there is actually a relationship between height above sea level and plant diversity, we could use the area-wide digital elevation raster data set to predict the plant diversity for each grid of the raster. Of course we could also try to utilize satellite observations which are also area-wide raster data sets.

Basically, the only (technical) difference between the non-area wide and area-wide prediction and also the validation of such predictions is the type of independent variable we use. For the non-area wide prediction we have used environmental information recorded on the field survey plots as independent variable(s). For area-wide predictions, we try to use information from raster data sets (e.g. DEM, satellite indices) as independent variable. The rest of our analysis workflow stays the same.

Please note that this kind of prediction requires careful field survey planning and data selection and also has some restrictions which must be considered. For today, we keep it as simple as described and will come back to that topic in the school of 2015.

Time for practice

[W11-1 Satellite-supported predictions](#)

Note on data used for illustrating analysis The analysis used for illustration on this site are based on data from a field survey of areas in the Fogo natural park in 2007 by K. Mauer. For more information, please refer to [this report](#).

From:
<http://bisfogo.environmentalinformatics-marburg.de/> - **BIS-Fogo**

Permanent link:
<http://bisfogo.environmentalinformatics-marburg.de/doku.php?id=en:learning:schools:s01:lecture-notes:ba-ln-11&rev=1443424646>

Last update: **2015/09/28 08:17**

