**Assignment: Point Pattern Analysis**

In this assignment, you will work with **R** software on Point Pattern Analysis. Input data concerns natural and simulated, individual tree locations.

The knowledge that is necessary to successfully complete this assignment can be gathered through Lessons:

* Review – Foundations of Statistical Description and Analysis
* Descriptive Spatial Statistics
* Positional Error Analysis
* Explorative (spatial) Data Analysis – EDA & ESDA
* Spatial Autocorrelation
* Spatial Distributions
* Point Pattern Analysis

So, do not proceed unless you have not finished these lessons.

The points in the dataset ***PPA\_data*** refer to single trees located at the alpine tree line in the Passeier valley (South Tyrol, in North Italy). Data are from Gudrun Wallentin’s thesis on individual based tree modelling (IBM) of the alpine tree line (appreciating your help, Gudrun!). Similar to agent based models, IBMs are able to simulate the growth of individual trees and their local interferences. Knowledge of the process variables (growth rate, seed production, mutual shadowing, terrain height) is used to simulate a tree pattern.

The simulated patterns are compared with the real-world tree distribution pattern to validate the results.

The following data sets are available in the “***PPA\_data***” folder:

1. Ortho image: ***orthophoto\_1999.img***
2. Trees as digitized from (1): ***digitised\_trees***
3. Trees as simulated by IBM: ***simulated\_trees***
4. Study area: ***study\_area***

**Instructions for R**

1. Include necessary libraries (spatstat, maptools). Most functionalities used in this assignment are taken from the spatstat package. The manual of the spatstat package is a valuable resource of information, while working on this assignment. The manual is available for download on the Cran R website:   
   <http://cran.r-project.org/web/packages/spatstat/spatstat.pdf>
2. For the digitised and the simulated tree data sets, perform a nearest neighbour point pattern analysis.
   1. Read the shapefiles digitised\_trees.shp and simulated\_trees.shp and display them. Use readShapePoints to load the point dataset:

* dt <- readShapePoints("C:/…/digitised\_trees.shp")
  1. Calculate the observed average nearest neighbour distances of the two datasets with the nndist command:
* dt\_nndist <- nndist(dt@coords[,1], dt@coords[,2])
  1. Calculate the mean of the theoretical (completely randomly) distributed nn distances. Remember, the equation of the random nn distances (CSR pattern) is: .   
     The number of points equals the length of the point list:
* length(dt)

The investigation area (A) can be calculated from the corners of the bounding box:

* dt@bbox
* dt\_area <- (dt@bbox[1,2] - dt@bbox[1,1]) \* …
  1. Now, determine the Average Nearest Neighbour Ratio between the empirical and the theoretical mean nn-distances. What do the resulting values significate?

1. Calculate a histogram of simulated tree heights to get an overview of tree heights (don’t forget the screenshot).

* st\_table <- read.dbf("C:/…/simulated\_trees.dbf")

Based on the histogram, define two meaningful subgroups of tree-heights that together cover the full range of heights. Then, for each subgroup, redo the nearest neighbour analysis and store the results.

* st\_small <- st[st$HEIGHT < xx,]

1. Based on visual inspection, delineate areas that are typical for each of the three point distribution patterns (dispersed, random and clustered).

* st\_rand <- st\_rand[st\_rand@coords[,1] > xxx,]

For the identification of coordinates, the map.axes() command of the *maps* package in the library proves useful.   
Redo the nearest neighbour analysis for all three selected subareas and store the results.

1. For each of the three delineated subareas, perform a multi distance spatial cluster analysis (Ripley’s K), plot it and store the graphical output.   
   In R, the Ripley’s K is calculated with the command ‘Kest’ that needs a point-pattern (ppp) object as input:

* st\_ppp <- as (st, "ppp")
* K <- Kest(st\_ppp)

1. Give a short report on the above steps, including screenshots, graphics of the areas you selected and the graphical outputs of nearest neighbour and Ripley’s pattern analyses. Give a compact interpretation of the achieved results (~500 words).
2. Submit the results of this assignment in a single PDF.

**References**

Wallentin, G., et al. 2008. *Alpine tree line dynamics: an individual based model*. Ecological Modelling, 218(3-4), 235-246.