Spatial modelling with the *R*–GRASS Interface

Rainer M Krug[1,⋆]

1. Centre of Excellence for Invasion Biology, Stellenbosch University, South Africa
⋆Contact author: Rainer@krugs.de

Keywords: Ecological modelling, spatial, monte carlo simulation, GRASS, GIS

R is a widely used tool for analysis and simulation of non-spatial data, even considered by many as the "lingua franca" for statistical modelling. But for working with spatial data, the toolbox provided by *R* does not consist of all the functionality to do complex analysis in *R* directly, unless one is prepared to spend a substantial amount of time with programming routines. This problem extends to the storage of spatial data, where, especially when dealing with large areas, the data sets can become extremely large and one is quickly confronted with memory problems. As a solution to these problems, interfaces with different GIS programs are provided for in R. These interfaces (for linking to GIS **spgrass6**, **rsaga**, importing and exporting **rgdal**, **sqlitemap** and others, but also **spgrass6**, and **rsaga**) provide functions to import the data into R and export it again into a format readable by the GIS, but also to utilise the functions of the GIS applications from within *R*, i. e. to use *R* as a scripting language for those GIS applications. This scripting becomes an important tool in the analysis and simulation of spatial data, as large maps can be manipulated without reading them into *R*, instead using functions of GIS programs which were written to accommodate large data sets (there is a package in *R* (**raster**) which uses the same approach, i. e. only loading smaller sections of the whole map into memory and processing those, but it obviously does not have all the functions included in dedicated GIS applications).

As elegant the solution is of using *R* as a scripting language to combine the power of *R* with the power of GIS applications like GRASS and SAGA, it has still problems. These, however, show only when these interfaces are used intensively.

To illustrate my points and to highlight a few areas which the *R*–GRASS interface could be improved, I will present a study conducted to investigate the impact of different budgets and management scenarios on the effectiveness of alien plant clearing. The study was conducted using a spatio-temporal explicit simulation model, covering areas of up to 215.000 hectares with a cell size of one hectare. The simulation used several vector and raster input layers, and generated hundreds of raster output layers. The interface between *R* and GRASS (**spgrass6**) was used intensively, as certain processes were not available in *R* or the implementation in GRASS was considerably faster due to implementation in *C*.

Main aspects identified during the project as "could be improved" range from direct reading of GRASS data without the need of additional software and simple implementation of processing with a MASK to the general question of backends for spatial data to parallisation especially of spatial routines.