Title: Using open source data analysis environments for prototyping spatial modeling implementations in R

Authors: Yongwan Chun, Roger Bivand and Michael Tiefelsdorf

Abstract:

When exploring some of the models proposed for analyzing spatial lattice data, it is arguably useful to prototype implementations. Using open source language environments, such as the R language, it is possible to try a variety of approaches to the different methods. This permits cross-checking to see whether the choice of implementation technique impacts the outcome in unexpected ways, as well as providing a rich selection of tools for examining output for artefacts.

We demonstrate the potential value of this approach in relation to several cases. The first is the broadening of functions in the R spdep package for simultaneous autoregressive models to accommodate weights over and above the spatial neighbor weights used at present. This will be extended to cover the introduction of conditional autoregressive and moving average models in the same setting, also with weights. Since the approach adopted is for prototype implementation, scaling up the functions produced to large data sets will come as a second stage, which can be compared with the full spatial weights matrix approach used here. A third area in which full spatial weights matrices are used is in spatial filtering, and an implementation for this case will also be discussed.

We will argue that this approach is well-supported when carried out in open source data analysis environments such as R, because, when necessary, the underlying source code of the built-in functions being used is available for reading. If need be, it can even be modified, although in practice such modifications are usually limited to accompanying contributed package code, sometimes in the same package, sometimes in other packages linked by dependency relationships. Here the term open source expresses not just the fact that the source code is accessible, but also the style of collaborative work that this accessibility makes possible and advances. Truly, the driving force behind the development of contributed packages like spdep is the stream of questions, bug-reports, and code ideas presented by users.