

parspatstat: An R Package for Large-Scale Spatial Analysis with Parallel Computing

Jonathan S.W. Lee

The University of Western Ontario, London, Ontario, Canada jlee253@uwo.ca

Reg J. Kulperger

The University of Western Ontario, London, Ontario, Canada rjk@stats.uwo.ca

Hao Yu*

The University of Western Ontario, London, Ontario, Canada hyu@stats.uwo.ca

The `parspatstat` package is an extension of the `spatstat` package for spatial analysis. It implements some of the more common functions of `spatstat` in a parallel environment using the `Rmpi` implementation of the message passing interface (MPI) framework. Spatial descriptive statistics such as Ripley's K-function have wide applicability in spatial analysis but current implementations do not scale well for large data sets. Parallel computing (high performance computing) is one solution that can provide almost linear scalability to these applications. The usages of these functions are kept as similar as possible to the current `spatstat` function to aid in updating existing algorithms. Implementation, optimizations, and complications that arise from parallelizing existing algorithms are discussed.

Key Words: Spatial statistics, K-function, Parallel computing, `Rmpi`.