Modelling spatiotemporal change using multidimensional arrays Meng

Meng Lu, Marius Appel, and Edzer Pebesma
University at Muenster, Geoinformatics, Germany (menglu@buffalo.edu)

The large variety of remote sensors, model simulations, and *in-situ* records provide great opportunities to model environmental change. The massive amount of high-dimensional data calls for methods to integrate data from various sources and to analyse spatiotemporal and thematic information jointly. An array is a collection of elements ordered and indexed in arbitrary dimensions, which naturally represent spatiotemporal phenomena that are identified by their geographic locations and recording time. In addition, array regridding (e.g., resampling, down-/up-scaling), dimension reduction, and spatiotemporal statistical algorithms are readily applicable to arrays. However, the role of arrays in big geoscientific data analysis has not been systematically studied: How can arrays discretise continuous spatiotemporal phenomena? How can arrays facilitate the extraction of multidimensional information? How can arrays provide a clean, scalable and reproducible change modelling process that is communicable between mathematicians, computer scientist, Earth system scientist and stakeholders?

This study emphasises on detecting spatiotemporal change using satellite image time series. Current change detection methods using satellite image time series commonly analyse data in separate steps: 1) forming a vegetation index, 2) conducting time series analysis on each pixel, and 3) post-processing and mapping time series analysis results, which does not consider spatiotemporal correlations and ignores much of the spectral information. Multidimensional information can be better extracted by jointly considering spatial, spectral, and temporal information. To approach this goal, we use principal component analysis to extract multispectral information and spatial autoregressive models to account for spatial correlation in residual based time series structural change modelling. We also discuss the potential of multivariate non-parametric time series structural change methods, hierarchical modelling, and extreme event detection methods to model spatiotemporal change. We show how array operations can facilitate expressing these methods, and how the open-source array data management and analytics software SciDB and R can be used to scale the process and make it easily reproducible.