



Mathematisches Kolloquium

A Multi-resolution Cluster Analysis of Carcinogenic Air Pollution in the USA

Sprecher: Prof. Dr. Raid Amin

Eingeladen von Prof. Hans-Joachim Zwiesler

12.05.2017 | 15:30 Uhr | R220 HeHo 18

The U.S. Environmental Protection Agency (EPA) provides air quality data through its Risk Screening Environmental Indicators (RSEI). We focused on the carcinogenic components in a cluster analysis that was sequentially done in stages. As a first step, we used data at the county level for the USA's 48 contiguous states in a cluster analysis with the software SaTScan. The Normal model was used, after transforming the raw air quality scores into normal quantiles in order to reduce the impact of any existing outliers. If a very large cluster was identified, it was re-analyzed by itself to identify any existing hot spots. Then, in the next stage, we analyzed the county based hot spots with data at the census tract level. If we find any significant clusters at this stage, we throw away the census tract data for each cluster, and we replace them with data at the block group resolution. In the last stage, we replace block group data with block level data. This is a very efficient methodology for modeling large amounts of environmental data.