Outlier detection in environmental monitoring networks

Kathy Ensor

Department of Statistics
Rice University

Abstract

Data from environmental monitoring networks often are collected for regulatory purposes but are used for many other purposes such as assessment of the validity of physical models. One example of such a database readily available to the public is the AIRS (Aerometric Information Retrieval System) database maintained by the Environmental Protection Agency. The data within this system undergo rigorous QAQC tests before inclusion but the data may still be inappropriate for these additional investigations. For example, highly influential values within the database can skew the assessment of physical models for the system. However, are influential values in-line with the system under study? We examine spatial-temporal methodologies for identifying unusual observations or sites in a monitoring network based on the observational data. Examples of our methodologies will be applied to segments of the AIRS database for the greater Houston metropolitan area as well as other environmental applications.