Geophysical Research Abstracts,
Vol. 10, EGU2008-A-06922, 2008
SRef-ID: 1607-7962/gra/EGU2008-A-06922
EGU General Assembly 2008

Predictive Modeling of Geospatial Data with Machine Learning Algorithms

M. Kanevski (1), A. Pozdnoukhov (1), V. Timonin (1), L. Foresti (1)

(1) Institute of Geomatics and Analysis of Risk, University of Lausanne, Switzerland, (Mikhail.Kanevski@unil.ch / Fax: +41 21 6923535 / Phone: +41 21 6923531)

Recent technological advances in measurement systems result in the growing amount of multi-dimensional data coming from environmental monitoring networks. Automatic monitoring stations produce large databases of climatic variables such as air and surface temperatures, precipitation. The spatial domain of the underlying processes, usually considered as 2D or 3D space, is now extended with terrain features available from digital elevation models. Geographical Information Systems can further provide useful sources of information. The multi-band remote sensing images bring even more input dimensions to the analysed information. The relations between the input information and their impact on water quality degradation and public health are complex, non-linear, corrupted with noise and outliers in the real data. The useful methods within this domain are the ones concerned with data-driven modelling. They are known as Machine Learning algorithms and are designed to extract dependencies from empirical data.

This paper first presents the basics of Machine Learning (ML) methods. Mainly developed for high-dimensional data, ML methods aim at being independent of the dimensionality of the input space. They are furthermore designed to deal with nonlinear problems in a robust and non-parametric way. Then, some applications devoted to the modeling of hydrological and water quality related dependencies are presented. The overview of the real case studies includes the exploratory analysis and clustering of the multivariate data on the chemical contamination of the Leman Lake (Switzerland). Other applications include the approach to topo-climatic mapping (spatial maps of temperatures and precipitation) in the mountainous regions of Swiss Alps and Aral

Sea region using the data from automatic monitoring networks.