

UTSA College of Business
Department of Management Science and Statistics
Research Seminar Series

January 25

“Clustering of Multivariate Geostatistical Data”

Francky Fouedjio

Department of Geological Sciences

Stanford University

Abstract

Multivariate data indexed by geographical coordinates have become omnipresent in the many scientific and engineering fields like remote sensing, environmental monitoring, mineral exploration, mining operations. This type of data poses substantial analysis challenges. One of them is the grouping of data locations into spatially contiguous clusters so that data locations belonging to the same cluster have a certain degree of homogeneity while data locations in the different clusters must be as different as possible. However, groups of data locations created through classical clustering techniques turn out to show poor spatial contiguity; a feature obviously inconvenient for many geoscience applications. In this talk, we will present a clustering method that overcomes this problem by accounting the spatial dependence structure of data; thus, reinforcing the spatial contiguity of resulting clusters. The proposed clustering method relies on a (dis) similarity matrix built from a non-parametric kernel estimator of the multivariate spatial dependence structure of data. It integrates existing methods to find the optimal cluster number and to evaluate the contribution of variables to the clustering. The capability of the proposed clustering method to provide spatially contiguous and meaningful clusters is illustrated using synthetic and real datasets.