- **Title**

Introduction to the Bayesian Maximum Entropy approach for space/time geostatistical exposure assessment

- **Name and affiliation for instructor**

Marc Serre
University of North Carolina at Chapel Hill

- **Contact information for lead instructor**

Marc L. Serre, Ph.D.
Assistant Professor, Department of Environmental Sciences and Engineering
Room 115 Rosenau Hall, School of Public Health
University of North Carolina at Chapel Hill
Chapel Hill, NC 27599-7431
USA
Phone: +1 (919) 966-7014
URL: http://www.unc.edu/~mserre/ESEhome.html
Email: marc_serre@unc.edu

- **Brief biography**

Dr. Marc L. Serre has been serving as an assistant professor in the Department of Environmental Sciences and Engineering at the University of North Carolina since 2002. He holds a Ph.D. in Environmental Sciences & Engineering (1999), and M.S. degrees in Civil and Environmental Engineering (1992) and Physics Engineering (1989).
Dr. Serre has 16 years of experience in environmental engineering, space/time geostatistics, exposure mapping, health risk assessment, and spatial epidemiology. He has worked as an environmental engineer in the consulting industry, developed Geographic Information System (GIS) software for environmental applications in the Computer Aided Design (CAD) industry, and, over the past decade, has been part of an active group developing the Bayesian Maximum Entropy (BME) method of modern spatiotemporal Geostatistics, applying this method for the spatiotemporal modeling of environmental and health processes.
Dr. Serre is currently involved in the application of the BME method to model the distribution of environmental pollutants across space and time in the groundwater, surface water and atmosphere. He is also using BME for the space/time modeling of infectious diseases, as well as the health impact of environmental contamination. This work has been presented in two books he has co-authored on temporal GIS and on modeling infectious diseases, as well as several journal papers and talks at international conferences or as invited speaker and during numerous workshops on the analysis of space/time data.

- **Description**

Space/time exposure assessment and the integration of data from multiple sources have received increasing interests in recent international conferences (e.g. ISEA06, ISEE06, IAMG06 and many others). Indeed, as exposure assessment experts we are increasingly asked to characterize exposure across space and time in order to better assess the exposure of susceptible populations, to provide more accurate exposure information for environmental epidemiology studies, and to improve health risk assessment under conditions of uncertainty.

Traditionally much of the information available for space/time exposure assessment has consisted of sparse monitoring data, and a general lack of adequate mechanistic models (hydrologic, atmospheric, toxicokinetic, etc.). As a result, the field of exposure assessment has recently seen an increased interest in using a Geostatistical (i.e. spatial statistics based) approach to obtain an efficient and accurate estimator for exposure. However; there is a critical need in exposure assessment for space/time Geostatistics, rather than just spatial Geostatistics.

More recently many scientists have recognized the usefulness of integrating secondary data from multiple sources. Secondary data may include satellite information, land use models, environmental sensors, confidential data, data obtained at different observation scales, empirical secondary data, GPS exposure data, physical models, etc. Such secondary data is very valuable, however it may have high data uncertainty, i.e. it should be treated as soft data. Hence the second critical need in exposure assessment is to integrate soft data from multiple sources.

The Bayesian Maximum Entropy (BME) approach appears to be a potential candidate for achieving this task: it is especially designed for managing simultaneously space/time data of various nature and quality ("hard" and "soft" data, continuous or categorical). It relies on a two-steps procedure that first involves a Maximum Entropy step (the ME part of BME) to objectively obtain a prior distribution in accordance with the general knowledge at hand, and an operational Bayesian conditionalization step that updates this prior probability distribution function (pdf) with respect to the specific data collected on the study site. BME provides a flexible framework that accounts for the wide variety of knowledge bases available, and leads in general to the best non linear space/time estimator. Traditional kriging methods are naturally obtained as a limiting case for linear estimation.

The BME approach thus appears as a kind of new unifying theory, opening new perspectives for space/time exposure assessment. Traditional simple, ordinary and universal kriging methods are derived as limiting linear cases, while the more general BME framework provides a powerful non-linear estimator for space/time Geostatistics using both hard data and a variety of soft data.

The course is intended as a large audience introduction to the concepts driving the BME approach for space/time exposure assessment. The basic concepts are illustrated through a series of conceptual examples and real case studies. The course combines lecture sessions and interactive practical sessions. Comparisons with traditional geostatistical methods are encouraged and open discussions are expected. Each participant receives at set of lecture notes. While BME is developed for the space-time domain, the examples presented in this course are both for the purely spatial case, as well as the space/time case.

The objectives for the theoretical parts of the course include:

- A quick review of the fundamental concepts of geostatistics (random variable, spatial correlation, spatial estimation and uncertainty assessment),
- An introduction to the fundamental concepts of the BME approach (information, entropy, hard and soft data, operational Bayesian conditionalization, …)
- A detailed explanation of the various BME solutions for continuous variables, and the kriging solutions obtained as limiting cases of BME.

The objective for the practical part of the course is to cover real case studies including:

- Spatial exposure mapping of Arsenic in the *groundwater*.
- Space/time exposure assessment of *surface water* quality (e.g. Dissolved Oxygen, PCE and Mercury) in a river network
- Space/time estimation of contaminants in the *air* (e.g. Particulate Matter, Ozone and Ammonia).

The exposure assessment case studies that have been conducted using the *BMElib* library of comprehensive computer programs (written in Matlab®). The case studies provide examples at forefront of modern space/time analysis dealing with subsurface, soil, surface water, and air toxic contaminants. The participants are shown the benefits of using this integrated toolbox for exploratory analysis of the data, modeling of spatial and space/time variability, spatial and space/time analysis and estimation, as well as graphical presentation of maps.

- **Target audience**

This course is intended as a large audience introduction to researchers and professionals involved in the analysis of spatial and space/time data sets for exposure assessment. In order to fully benefit from the whole course, participants should be knowledgeable about linear regression theory, and the concept of spatial auto-correlation.

- **Course level**

Introductory to intermediate.

- **Prerequisites or expected proficiency**

Some knowledge of linear regression theory, or some knowledge of kriging.

- **Number of students**

5 to 30

- **Course length**

One full day.