Statistica Sinica **25** (2015), 1-3

Editorial

Increasing demand for new statistical models and techniques in the analysis of spatial or spatio-temporal data has provided us with challenges and opportunities. New research efforts have been devoted to the development of scalable models and algorithms, as well as new inferential procedures that account for the topology of unique data structures. This special issue on the analysis of spatial or spatio-temporal data represents a wide spectrum of the current research in this exciting area of statistics.

The methodology development in spatial statistics is always data-driven. Several papers in this issue focus on data-oriented methodologies. Motivated by the massive structure in a polar-orbiting satellite data, Horrell and Stein developed a novel likelihood method which incorporates the composite likelihoods with an approximate likelihood of interpolated points within each block. The new method is statistically efficient and computationally scalable. Motivated by data on wave heights and outgoing wave directions over a region in the Adriatic Sea during the time of a storm, Wang, Gelfand, and Jona-Lasinio developed a fully model-based approach to capture joint structured spatial and temporal dependence between a linear variable and an angular variable. In response to the need for estimation and prediction of the Atlantic cod in the Gulf of Maine, Wang et al. developed a computationally efficient Bayesian hierarchical model for zero-inflated count data with dynamic spatial random effects. Castellanos et al. proposed a Multivariate Gaussian Process Factor (MGPF) model to estimate low dimensional spatio-temporal patterns of finger motion in repeated reach-to-grasp movements. The MGPF model decomposes and reduces the dimensionality of variation of the multivariate functional data. Jiang et al. introduced a new approach to spatio-temporal data analysis that is appropriate for monitoring or managing a physical system. Based on a modeling framework for complex predictor-response and spatio-temporal relationships, the proposed method includes an automated criterion for choosing between parametric and nonparametric spatial covariance models to accommodate a wide range of practical situations.

Projection of future sea levels under different climate change scenarios is an important task. Bolin et al. constructed a time series regression model to predict global sea levels from global temperature. Li et al. proposed nonparametric Bayesian models for areal data that can formally identify boundaries between disparate neighbors. In their analysis of Pneumonia and Influenza hospitalization maps from the SEER-Medicare program in Minnesota, they were able to detect and report highly disparate neighboring counties.

Based on the GMRF approximation of Gaussian random fields (GRF) via stochastic partial differential equations (SPDE), Fuglstad et al. explored a new approach to construct a class of non-stationary GRFs by allowing spatially varying anisotropy in the SPDE. Demel and Du proposed two new classes of spatio-temporal models for discrete time lags, where discrete temporal margins follow autoregressive and moving average models. They derived necessary and sufficient conditions for these classes of covariance functions to be valid.

Large data computation is often an inherent part of spatial or spatio-temporal data analysis. Xu et al. proposed a computationally efficient Bayesian hierarchical spatio-temporal model for both regularly and irregularly spaced data by introducing an auxiliary lattice on the spatial region of interest so the spatial dependence can be approximated by a Gaussian Markov random field. Zhang et al. tackled the spatio-temporal covariance model for large datasets from a different perspective by extending their early proposed reduced rank model to space-time data and proposing a reversible jump Markov chain Monte Carlo algorithm for knots selection in the new model. Ahn et al. developed a new sparse reduced rank modeling framework for carrying out functional connectivity analysis across multiple groups of subjects in the frequency domain. Their proposed framework extracts both frequency and spatial factors across subjects and imposes sparse constraints for the identification of important frequencies with high power spectra. Yang et al. developed multi-dimensional spatial functional models that employ low-rank basis function expansions to facilitate model implementation. Chang et al. introduced a composite likelihood-based approach to perform computer model calibration with high-dimensional spatial data and proposed a computationally efficient approach for Bayesian computer model calibration. They further developed a new methodology based on the asymptotic theory for adjusting the composite likelihood posterior distribution to ensure that it accurately represents posterior uncertainties. Guinness and Fuentes proposed a computationally efficient nonstationary Gaussian likelihood approximation for the class of evolutionary spectral models for lattice data on a regular lattice. The covariance matrix implied by the likelihood approximation was examined and its asymptotic rate of approximation to the exact covariance matrix was derived.

Classical problems of parameter estimation and hypothesis testing arise from modeling of spatial data but require additional investigations to account for the correlation in the data. Tzeng and Huang introduced a regularization approach for multivariate spatial covariance estimation based on a spatial random effect model. This new approach allows for both non-nonstationarity and asymmetry in the spatial cross covariances, and effectively controls the estimation variability.

Lai and Lim considered parameter estimation of Markov random field models for spatial and temporal data. They developed composite likelihood estimators that are analytically and computationally tractable as well as asymptotically efficient under some mild correlation decay assumptions. Thurman et al. studied model selection and parameter estimation for clustered spatial point processes using penalized quasi-likelihood with an adaptive lasso penalty. Yang and Zhu studied estimation and prediction for a heteroskedastic spatial process. A difference-based kernel estimator of the variance function and a modified likelihood estimator of the measurement variance are proposed for the estimation and a new kriging predictor is given. Bai et al. proposed polynomial spline approximation to estimate both parameters and nonparametric functions in varying-coefficient partially linear models with spatially and temporally correlated errors, and further demonstrated the asymptotic properties of proposed estimators. Yang and He developed Bayesian spatial quantile regression using empirical likelihood as a working likelihood. The proposed approach naturally incorporates spatial priors to smooth the conditional quantile functions across locations and across quantiles.

There are two articles in this special issue on scan statistics. Chang and Rosychuk developed a spatial scan test based on a special compound Poisson model that identifies clusters of events in a study region. The proposed model is advantageous in computation over the general compound Poisson model. Their proposed method was applied to analyze administrative health data on asthma for cluster detection on emergency department visits. de Lima et al. proposed a spatial scan statistic with the Zero Inated Double Poisson (ZIDP) model in order to reduce type I error and to accommodate simultaneously the excess of zeroes and overdispersion. The null and alternative model parameters were estimated by the Expectation-Maximization algorithm and the p-value was obtained through the Fast Double Bootstrap.

The spatial and spatio-temporal data analysis is a growing field. We hope that this special issue will help bring together researchers of diverse background to better understand many of the problems and challenges in this field. We also hope that this issue will make Statistica Sinica a friendly home to many more advancements in spatial data analysis in the years to come.

Guest Editors
Ming-Hui Chen[1], Xuming He[2] and Bo Li[3]

1. University of Connecticut
2. University of Michigan
3. University of Illinois, Urbana-Champaign