Contributed abstracts of statGIS2007 Conference
24. September – 26. September 2007

Department of Statistics
University of Klagenfurt

Extracting Patterns of Individual Movement Behaviour from a Massive Collection of Tracked Positions

Gennady Andrienko and Natalia Andrienko

Fraunhofer Institute IAIS, Germany
gennady.andrienko@iais.fraunhofer.de

Summary. A EU-funded project GeoPKDD develops methods and tools for analysis of massive collections of movement data, which describe changes of spatial positions of discrete entities. Within this project, we design and develop methods of visual analytics, which combine interactive visual displays with database operations and computational methods of analysis. In this article, we demonstrate by example how visual analytics methods can help in acquiring knowledge about the movement behaviour of an individual from a very large set of movement data.

Effects of Aggregation on Spatially Correlated Time Series

Guiseppe Arbia[1], Marco Bee[2] and Guiseppe Espa[3]

[1] Department of the Bussiness, Statistical Technological and Environmental Sciences, University "G. d'Annunzio", Italy
arbia@unich.it
[2] Department of Economics, University of Trento, Italy
marco.bee@economia.unitn.it
[3] Department of Economics, University of Trento, Italy
guiseppe.espa@economia.unitn.it

Summary. The problem of choosing the best forecasting strategy when dealing with disaggregated time series has a long tradition in econometrics. Giacomini and Granger (2004) faced this problem in the specific case of forecasting a national aggregate when disaggregated regional series are available and the individual regional series display spatial correlation. Arbia et al. (2006) extended their results in various directions by including the consideration of larger datasets and the treatment of edge effects and of negative spatial correlation. In this paper we perform a more detailed comparison of the relative efficiency of different forecasting methods of space-time series when variables are spatially and temporally correlated. In particular, we study the case where data are available on a fine grid (e.g. counties within regions) and we have the problem of producing a forecast on a coarser grid (e.g. regions within a country). The motivation comes from the need, at a EU level, to produce forecasts for each member state (the NUTS1 level) but the available time series of data are at a regional level (the NUTS2 level). So our approach is general in the sense that we come up with a multivariate forecast. The outcomes obtained provide operational suggestions on how to choose between alternative forecasting methods in empirical circumstances. More precisely, the simulation experiments show quite clearly that the best strategy both in terms of minimum Mean Squared Error and in terms of smallest Moran index is, in all cases

examined, the Space-Time AutoRegressive model, namely the only one, among the methodologies considered in this work, which explicitly takes into account the spatial dependence of the data.

References

[1] Arbia, G., Bee, M. and Espa, G. (2006), "Aggregation of regional economic time series with different correlation structures", paper presented at the International Workshop on Spatial Econometrics and Statistics, Rome, May 2006.

[2] Giacomini R. and Granger C.W.J. (2004), "Aggregation of space-time processes", Journal of Econometrics, 118, 7-26.

Geostatistical approach to data harmonization

Olivier Baume[1], Jon Olav Skoien[2], Gerard B.M. Heuvelink[3] and Edzer J. Pebesma[4]

[1] Wageningen University, The Netherlands
olivier.baume@wur.nl
[2] Department of Physical Geography, Utrecht University, The Netherlands
j.skoien@geo.uu.nl
[3] Wageningen University, The Netherlands
gerard.heuvelink@wur.nl
[4] Department of Physical Geography, Utrecht University, The Netherlands
e.pebesma@geog.uu.nl

Summary. Globalization in environmental risk assessment and mapping requires valid data at a higher level (e.g. from country level to continent level). In the development of national or even regional measurement networks such as radioactivity exposure assessment networks, different protocols are used. These are bent to specific devices, site specifications, laboratory analysis techniques, accuracy standards and sampling designs. There may even be differences in the exact definition of the target variable between organisations. Hence in the case of dealing with data from different sources, mapping the data across networks can show a systematic difference or bias. It is a well known fact that different protocols may not allow the joint use of the data in statistical models and may demand peculiar effort. Standardization is the prevailing outcome of inter-comparison studies of field protocols in varied environmental risks assessment issues. But a faster and efficient alternative to standardization is harmonization of data. This latter is a "bottom up approach" which should be ending in a state of comparability between data from different sources.

We discuss in this paper how to deal with harmonization of data across networks, and how to handle multiple information on measurements, estimation of constant biases and prediction in space. We define a geostatistical model wich includes both constant bias terms (in the form of trend factors) and a random error component (as a separate covariance component). The model ends up in a classical multiple regression problem for which some accuracy diagnostic must be held because of the multiplecollinearity issue. Two application examples are presented. First a synthetic database is created so that the method can be challenged with a known covariance structure. The quality of the estimation is checked through both the result statistics and the Variation Inflation Factor diagnostic. Second a real data example is taken from the radioactivity exposure issue, a european-wide exchange data base platform (EURDEP - http://eurdep.jrc.it/) which provides data from most of European countries.

Spatial Vector Error Correction

Michael Beenstock

Department of Economics, Hebrew University of Jerusalem, Israel
msbin@mscc.huji.ac.il

Summary. We "spatialize" residual-based panel cointegration tests for nonstationary spatial panel data in terms of a spatial vector error correction model (SVECM). Local panel cointegration arises when the data are cointegrated within spatial units but not between them. Spatial panel cointegration arises when the data are cointegrated through spatial lags between spatial units but not within them. Global panel cointegration arises when the data are cointegrated both within and between spatial units. Spatial error correction arises when the SVECM contains error correction between spatial units in addition to within units. We use nonstationary spatial panel data on the housing market in Israel to illustrate the methodology. We show that although regional house prices in Israel are globally cointegrated, there is no evidence of spatial error correction.

Keywords: panel cointegration, spatial panel data, regional house prices

Analysing Regional Firm Startup Activity Using Geographically Weighted Regression: The case of Austria

Robert Breitenecker

Department of Innovation Management and Entrepreneurship, University of Klagenfurt, Austria
robert.breitenecker@uni-klu.ac.at

Summary. One subject in entrepreneurship research is to find factors which explain regional variations in firm startup activity. Reviewing the entrepreneurship literature points out that in most studies classical linear models are used to identify influencing factors on regional founding rates and that the spatial component of data is mostly disregarded. Estimating a global model assumes that the influence of the independent variables are constant over the whole study region. I will test in this paper if the global linear model is appropriate to estimate the founding activity. Therefor global and local indicators for spatial autocorrelation (Moran I, [1]) are calculated first, to show the spatial dependence in regional firm startup rates and additional regional data on the level of 93 counties of Austria. Further a geographically weighted regression (GWR) model is estimated to test whether a constant relationship between the founding rate and dependent variables exists over the whole study region [2]. The test of spatial dependence results in significant positive autocorrelation for all regional variables. The local parameter values of the GWR model indicate significant regional variations which indicates that a global linear model is not appropriate to explain the startup activity in Austria.

Keywords: spatial models, geographically weighted regression, founding activity, AMS: 91B72

References

[1] Anselin, L. (1995), "Local Indicators of Spatial Association-LISA", Geographical Analysis, 27(2), 93-115.

[2] Fotheringham, S. A., Brunsdon, C., und Charlton, M. (2002), "Geographically Weighted Regression: The Analysis of Spatially Varying Relationships", Wiley, Chichester.

Automatic Detection of Spatial Anisotropy in Environmental Data Sets

Ersi Chorti[1] and Dionissios T. Hristopulos[2]

[1] Department of Mineral Resources Engineering, Technical University of Crete, Greece
achorti@mred.tuc.gr

[2] Department of Mineral Resources Engineering, Technical University of Crete, Greece
dionisi@mred.tuc.gr

Summary. The interpolation and filtering of environmental data is often based on the assumption that the respective random field is isotropic. However, this assumption fails in most natural data sets. The determination of anisotropy parameters has so far lain upon either empirical methods such as the visual inspection of geological maps or computationally expensive methods such as the maximum likelihood and directional semivariogram analysis. In this paper we propose a novel method for the automatic detection of geometric anisotropy in spatial scalar fields. The method computational complexity is linear to the sample size and it does not assume a specific model for the field correlation function. It applies to differentiable random fields with normal or lognormal probability density functions. It uses sample based estimates of the field spatial derivatives that we relate through closed form expressions to the geometric anisotropy parameters. We investigate the behavior of the method on simulated data according to commonly used spatial correlation models, both on regular and irregular lattices. The automatic anisotropy detection provides an important pre-processing stage of the data. Knowledge of the anisotropy parameters, followed by the "isotropic transformations" restores isotropy thus allowing kriging and other processing methods to be applied.

Topo-Climatic Data: Analysis, Modelling and Geovisualisation

Loris Foresti[1], Alexei Pozdnoukhov[2] and Mikhail Kanevski[3]

[1] Institute of Geomatics and Analysis of Risk, University of Lausanne, Switzerland
loris.foresti@unil.ch
[2] Institute of Geomatics and Analysis of Risk, University of Lausanne, Switzerland
mikhail.kanevski@unil.ch
[3] Institute of Geomatics and Analysis of Risk, University of Lausanne, Switzerland
alexei.pozdnoukhov@unil.ch

Summary. Geomatics and Geographical Information Sciences are widely used in many fields like urban planning, infrastructure management, telecommunications, remote sensing, epidemiology and environment (forestry, natural resources monitoring, pollution, natural hazards). One of the most recent fields in which geomatics have grown up is meteorology and climatology. With growing amount of information coming from meteorological monitoring networks, the use of data-driven modeling techniques becomes more important. The visualization of point measurements in real time provides useful information which can be done accessible through the web. However, the methods for producing and visualization spatially interpolated maps of climatic variables such as temperature, precipitation, wind speeds still need to be developed. This study aims to exploit machine learning algorithms (Multi-Layer Perceptron and Support Vector Machines) to model meteorological variables. The presented modeling scheme uses geographical information systems and especially the digital elevation models (DEM) with the terrain geo-features extracted from them. These features are essential to model climatic variables when phenomena are complex and non-linear. Feature selection techniques were applied to choose the characteristics derived from DEM which are (possibly, non-linearly) correlated

with the variable one wants to model. Multi-Layer Perceptron and Support Vector Machines are very powerful universal tools to model spatio-temporal data (meteorological fields) which depend non-linearly on several complementary variables (geo-features). In order to demonstrate the importance of these methods for topo-climatic modeling in Switzerland, the following case studies were chosen: monthly mean air temperature, hourly mean air temperature (föhn situation and temperature inversion), monthly sum of precipitation, daily sum of precipitation in a storm, monthly mean wind speed, daily mean wind speed, gusts and direction. The study of these cases has shown that the temporal scale is also an important factor to be taken into account. In fact, depending on the temporal scale, the same meteorological parameters will be influenced by different topographical features. Machine learning methods perform well in producing mean monthly and mean annual climatic maps which can be then incorporated into a GIS system. However, the task of real-time spatial modeling of instant measurements is a more challenging problem. Ideally, the fully automatic procedure is desired in order to produce spatial topo-climatic maps in a real time. The use of the developed modeling methodology based on machine learning algorithms is discussed in this context. An important part of the presentation deals with visualization of raw data, geo-features, and the results using contemporary GIS- and web-based technologies.

Mixing Additive Voronoi Mosaics and Kriging

Albrecht Gebhardt

Department of Statistics, University of Klagenfurt, Austria
albrecht.gebhardt@uni-klu.ac.at

Summary. Prediction methods for correlated spatial data like kriging are based on the knowledge of both a global trend function and an overall covariance structure. As this knowledge may not be applicable globally to a real data set, we try to combine a tessellation approach with the kriging predictor. The tessellation method is based on a generalized Voronoi mosaic, which splits the area of interest into non-overlapping tiles. Points belonging to a specific tile carry some nearest neighbour property with respect to the cell center. Model parameters for trend and covariance are treated as constant per tile of the tessellation. Finally kriging prediction will be performed per tile using these model parameters. The geometry of the tiles can be used in trend and covariance modelling e.g. if the data is assumed to be generated by some contamination process starting at the cell centers.

References

[1] Hyoung-Moon Kim, Mallick, B.K., and Holmes C.C. (2005), "Analyzing nonstationary spatial data using piecewise Gaussian processes.", Journal of the American Statistical Association, 100(470):p.653–668.

The Modified SIMPAT Image Processing Method for Reproducing Geological Patterns

Mojtaba Ghoraishy

Department of Petroleum Engineering, University of Kansas, USA
seyed@ku.edu

Summary. Reproducing geological patterns from the available set of data is one of the most challenging research in the Earth science. Such reproduction requires multiple-point statistics that correlates three points (plus) at a time. However, inference of multiple-point statistics is usually not practical, or well developed. The random function models have not been developed explicitly to account for multiple-point statistics. The multiple-point geostatistics use the concept of training image. This method was introduced to overcome restrictions and complexities of multiple-point statistics. A training image represents a similar pattern of field study geological heterogeneities. A training image does not need to honor any locally accurate information. It represents prior geological and structural concepts. The training image method is applied to model the geological patterns as image construction problems. The image construction method relies on the concept of similarity of available data and the patterns of a training image. Similarity distance is used as selection criteria to find the most consistent and similar pattern for to the existing data. This research covers a modification to the SIMPAT algorithm (SIMulation with PATtern). This modification employs the NCC similarity distance instead of the Manhattan distance. The Comparison of the Modified SIMPAT method with the original SIMPAT method demonstrates a great robustness and improvement in the generated results. The Modified SIMPAT is used to improve the flow simulation studies in the Petroleum Engineering. Multiple cases are studied and results of both the SIMPAT and the Modified SIMPAT are

compared. The results show that the Modified SIMPAT has the capabilities to improve geological patterns simulation dramatically.

Resampling SRTM 3"-data with kriging

Carlos H. Grohmann[1] and Samar S. Steiner[2]

[1] Institute of Geosciences, University of Sao Paulo, Brazil
carlos.grohmann@gmail.com
[2] Institute of Geosciences, University of Sao Paulo, Brazil
samar.steiner@gmail.com

Summary. SRTM data is distributed at horizontal resolution of 1arcsec (aprox. 30m) for areas within the USA and at 3arcsec (aprox. 90m) resolution for the rest of the world. A resolution of 90m can be considered suitable for small or medium-scale analysis, but it is too coarse for more detailed purposes. One alternative is to interpolate the SRTM data at a finer resolution; it won't increase the level of detail of the original DEM, but it will lead to a surface where there is coherence of angular properties (i.e., slope, aspect) between neighbouring pixels, an important characteristic when dealing with terrain analysis. This work intents to show how the proper adjustment of variogram and kriging parameters, namely the nugget effect and the maximum distance within which values are used in interpolation, can be set to achieve quality results on resampling SRTM data from 3arcsec to 1arcsec. The basic points of using kriging to interpolate terrain data are:

- Work only with the immediate neighbourhood of the predicted point, due the high spatial correlation of the topographic surface and omnidirectional behaviour of variogram in short distances;
- Add a very small random variation to the coordinates of the points prior to interpolation;
- Use a small value of nugget effect, to avoid smoothing that can obliterate terrain features; in our tests, a value of $10m^2$ was sufficient to eliminate noise.

Keywords: SRTM; DEM; kriging; geostatistics; nugget value; variogram; interpolation.

Spatial air quality analysis and assessment in Europe

Jan Horalek[1], Peter de Smet[2], Bruce Denby[3], Frank de Leeuw[4], Pavel Kurfürst[5]

[1] Czech Hydrometeorological Institute, Praha, Czech Republic
horalek@chmi.cz
[2] Netherlands Environmental Assessment Agency, Bilthoven, The Netherlands
[3] Norwegian Institute of Air Research, Kjeller, Norway
denby@ieee.org
[4] Netherlands Environmental Assessment Agency, Bilthoven, The Netherlands
[5] Czech Hydrometeorological Institute, Praha, Czech Republic

Summary. Spatial air quality analysis and assessment in Europe The European Topic Centre for Air and Climate Change (ETC/ACC) is currently carrying out a task "Spatial air quality data" for the European Environmental Agency (EEA) that reviews and further develops interpolation methods for use in European wide air quality mapping. Among the activities of this task, emphasis is on the development of interpolation methodologies for O3 and PM10, as well as on the preparation of high resolution air quality maps. Air quality data selected from the database AirBase for the years 2000-2004 has been interpolated by combining the data with different sources of supplementary data. The project started with mapping of health- and vegetation related ozone indicators (SOMO35, AOT40), and with PM10 indicators (annual average, max 36th daily value), and expanded to mapping of other health- (26th highest maximum daily value) and vegetation related indicators (AOT40 for forests, SO2 annual mean and NOx annual mean). The urban and rural air quality is dealt with separately; the final maps are created by merging the rural and the urban maps. All the mapping work is carried out in GIS.

Several mapping technologies were compared:

- interpolation of measured air quality data only (inversed distance weighting (IDW), different types of kriging),
- the combination of measured air quality data with the results of a dispersion model and
- the combination of measured air quality data with different supplementary data, such as climatological or meteorological parameters, station altitude, population density)

For the purposes of mapping, the relations of concentrations with supplementary data were examined. These relations are utilized in the mapping technology.

For all the years, the rural maps of selected pollutants and indicators have been created by several methodologies. The quality of these maps was compared using cross-validation and RMSE. The best results are given by the combination of measurement data with EMEP dispersion model results, altitude and different meteorological parameters. The other findings are that the kriging gives better results than IDW and that the use of logarithmic transformation gives more precise results for PM10.

Special attention was given to mapping of urban air quality, especially for those urban agglomerations for which no measurements are available. These values are interpolated across the border of the cities. In this way, a map of European urban air quality is obtained. Finally, the rural and the urban maps are merged together with the help of population density map.

The final maps are further used to estimate potential exposure of "stock at risk": the maps of AOT40 are used by the combination with land cover data in order to provide information on the ozone exposure of crops and forests. The final maps of O3 and PM10 are used by the combination with the population density for estimation the "population at risk".

References

[1] de Smet P., Horalek J., Denby B. (2007), "European air quality mapping through interpolation with application to exposure and impact assessment", In: Arno Scharl, Klaus Tochtermann (Eds.), The Geospatial Web. Springer, London

[2] Horalek J., Kurfürst P., Denby B., de Smet P., de Leeuw F., Swart R., van Noije T. (2007), "Spatial mapping of air quality for European scale assessment", ETC/ACC Technical paper 2006/6, http://air-climate.eionet.europa.eu/reports/ETCACC_TechnPaper_2006_6_Spat_AQ

Multiscale Spatio-Temporal Ecoclimatic Ecoregioning Pattern Analysis

Didier Leibovici

Centre for Geospatial Sciences, University of Nottingham, United Kingdom
didier.leibovici@nottingham.ac.uk

Summary. In a recent paper [Leibovici et al., 2007] we proposed a methodology to spatially identify different pattern of ecoclimatic spatio-temporal variations in circum-Saharan Africa. Using a multiway method enabled us to take into account the dynamic of multiple climatic indicators into the clustering problem. Here we focus on the scale issue well known in geoscience [Marceau and Hay, 1999] and particularly about its impact on ecoregion delineation. We propose to extend the above methodology to integrate a multiscale approach. This allow to analyse the impact of changing scale in a spatio-temporal description and also to use a multiscale approach into the process of spatio-temporal clustering of the regions. Two different scale changes can be considered: one implied by change of resolution in raster data and one implied by aggregation of levels in given tree hierarchy of locations.

Keywords: MultiScale, Scaling Effect, Ecoregion, Hierarchy, Ecoclimatic, desertification, classification, Pattern Recognition, Multiway analysis, Tensor decomposition, PTAk.

References

[1] Leibovici D., Quillevere G., and Desconnets J.C. (2007), "A Method to Classify Ecoclimatic Arid and Semi-Arid Zones in Circum-Saharan Africa Using Monthly Dynamics of Multiple Indicators", IEEE Trans-

actions on Geoscience and Remote Sensing, Special Issue on Pattern Recognition (accepted)

[2] Marceau D.J. and Hay G.J. (1999), "Remote sensing contributions to the scale issue", Canadian Journal of Remote Sensing, 25(4):357-366

Relating climate, habitat, immigration, and establishment processes to a bird species. range expansion

Stephanie Joy Melles[1], M.-J. Fortin, D. Badzinski and K. Lindsay

Environmental Sciences Group, Wageningen University and Research Centre, The Netherlands
stephajm@zoo.utoronto.ca

Summary. For the past 40 years, data on the distribution of bird species have been accumulating across North America demonstrating that species' ranges are dynamic, shifting in response to a large number of interrelated ecological and anthropogenic processes. Climate warming is one of the most important drivers of species' range shifts, but the effects of other confounded ecological processes are often ignored. To determine the relative effects of climate, habitat fragmentation, and biotic processes on observed expansion rates, we examine hooded warbler (Wilsonia citrina) range changes. We focus predominantly on the periphery of the species' northern range in Ontario, Canada, but we also examine data from the entire species' range. Ten hypotheses were generated to describe the relationship between the rate of range expansion and factors known to influence range change, and these hypotheses were compared using model-selection techniques. Zero-inflated regression (ZIP) provided a better functional relationship than Poisson regression as evidenced by lower AICc scores and smaller predictive errors (RMSE) with withheld validation data. Climate warming was negatively related to the probability of 'zero' range expansion indicating that climate constrains the species' distribution. Establishment probability, based on the number of occupied neighbouring grid squares, and immigration pressure from populations to the south were also important predictors of an increased rate of range expansion. Both immigration pressure and population establishment can mask the effects of habitat availability and fragmentation. Range expansion due to climate change may be

slower in fragmented systems, but the rate of expansion will be influenced largely by immigration from nearby populations and the probability of species' establishment in an area; hence these are important variables to consider in a conservation context.

Long-term regional forecasting with spatial equation systems

Wolfgang Polasek[1], Richard Sellner[2] and Wolfgang Schwarzbauer[3]

[1] Institute for Advanced Studies, Vienna, Austria
polasek@ihs.ac.at
[2] Institute for Advanced Studies, Vienna, Austria
[3] Institute for Advanced Studies, Vienna, Austria

Summary. Long-term predictions with a system of dynamic panel models can have tricky properties since the time dimension in regional (cross) sectional models is usually short. This paper describes the possible approaches to make long-term-ahead forecast based on a dynamic panel set, where the dependent variable is a cross-sectional vector of growth rates. Since the variance of the forecasts will depend on number of updating steps, we compare the forecasts behavior of a aggregated and a disaggregated updating procedure. The cross section of the panel data can be modeled by a spatial AR (SAR) or Durbin model, including heteroscedasticity. Since the forecasts are non-linear functions of the model parameters we show what MCMC based approach will produce the best results. We demonstrate the approach by a example where we have to predict 20 years ahead of regional growth in 99 Austrian regions in a space-time dependent system of equations.

Keywords: SAR models, Bayesian dynamic panel models, forecasting stable equation systems, regional econometrics.

Interactive Monitoring Network Optimization using Support Vector Machines

Alexei Pozdnoukhov[1] and Mikhail Kanevski[2]

[1] Institute of Geomatics and Analysis of Risk, University of Lausanne, Switzerland
mikhail.kanevski@unil.ch
[2] Institute of Geomatics and Analysis of Risk, University of Lausanne, Switzerland
alexei.pozdnoukhov@unil.ch

Summary. Monitoring Network Optimization is a wide interdisciplinary field aimed at resources optimization and cost minimization in environmental monitoring. The problem is usually considered in spatial domain, with the aim to increase the accuracy of predictive models by taking a small number of additional measurements. Two main approaches (design-based and model-based) are usually considered, which rely on either geometrical/topological characteristics of the network or on the uncertainty-based outputs of the model. This setting can be extended to the spatio-temporal domain if the modification of the network in time is required following the evolution of the spatial process or changing conditions. Practically, monitoring network optimization faces many real-life constraints and the optimization scheme includes intensive interaction with an expert. In operational use, the decision-support system for monitoring network optimization must provide tools for fast and convenient interaction and easily interpretable results.

This paper develops a novel method for monitoring network optimisation for multi-class spatial classification problems. Spatial classification covers a wide range of tasks in environmental modelling. The typical problems are spatial mapping of categorical variables (soil types, hydro-geological units,

land cover types, etc.) and decision-support mapping aimed at identifying the regions where some quantity (pollutant concentration) exceeds a predefined level. Traditional approach to network optimization by means of the analysis of the kriging variances faces a number of difficulties in this domain. The most frequent one is the insufficient number of data samples of particular class to model the variogram.

The developed method is based on the recent Machine Learning technique known as Support Vector Machine. It is a robust non-parametric data-driven classifier, which gained considerable success and promising generalization abilities in many data modelling fields. The network optimization method, based on Support Vector Machine, is problem-oriented in the sense that it directly answers the question of whether the advised spatial location is important for the classification model. The developed method is aimed to effectively increase the accuracy of spatial classification models by taking a small number of additional measurements. The method is based on a solid mathematical background of Statistical Learning Theory and purely data-driven. The queries for perspective measurements, assigned by an expert, are evaluated using Support Vector Machine. The quires are ranked and the potential changes in the model are visualized immediately. The predictive model is built directly from data and little or no user interaction (such as variogram modelling at intermediate step) is required. The method opens promising perspectives for fast and convenient interactive process of monitoring network optimization.

The use of the method and the comparison with traditional approach is presented on a real case study with data on soil types.

Estimation of parameters of interspike intervals observed in a small window

Martin Prokop[1], Zbynek Pawlas[2] and Petr Lansky[3]

[1] Department of Probability and Mathematical Statistics, Charles University, Praha, Czech Republic
maris@post.cz
[2] Department of Probability and Mathematical Statistics, Charles University, Praha, Czech Republic
[3] Institute of Psychology, Academy of Sciences of the Czech Republic, Praha, Czech Republic

Summary. We simulate transfer of information in the nervous systems by a temporal organization of interspike intervals. Various models of the random point process of spikes are used. For the representation we suppose n time windows of equal length with independent realizations of the process. A special problem is investigated where the length of window is very short which leads to a low number of spikes observed in each window. Based on the observations we can reconstruct the distribution of the process. The first process type we consider is the stationary renewal process and we suppose the case when interspike intervals are random variables with Gamma distribution. The second type is a stationary Cox point process with random driving intensity in a parametric form. We simulate the processes in n windows for several combinations of parameters. Then we calculate the maximum likelihood and moment estimators of the mean length of interspike interval and the same procedure is used for the estimation of the coefficient of variation. The simulation is repeated many times and we compute the relative mean square errors of these estimators. While for the estimation of the mean length both methods are comparable, for the coefficient of variation the maximum likelihood method is better.

A County-Level Simultaneous Spatial Econometric Model of Transportation Infrastructure Provision and Economic/Population Growth: The Case of Missouri

Dennis P. Robinson[1], Thomas G. Johnson[2] and Yong-Lyoul Kim[3]

[1] Policy Analysis Center, University of Missouri, USA
robinsonden@missouri.edu
[2] Policy Analysis Center, University of Missouri, USA
[3] Korea Rural Economic Institute, South Korea

Summary. The purpose of this paper is to investigate the effects of highway infrastructure provision decisions on the economic and demographic performance of MissouriaEUR(tm)s counties. Using data on the counties of Missouri for 1990 and 2000, we estimate a simultaneous spatial econometric model of economic and population growth and transportation infrastructure provision decision-making. In addition, we investigate several spatial econometric specification issues. First, most recent studies of small-area economic and population growth have considered a spatial cross-regressive equation structure (i.e., each equation contains spatial lags of other endogenous variables but not the spatial lag of the equationsaEUR(tm) dependent variable). We evaluate the efficacy of this specification relative to a simpler formulation of just a spatial lag in the dependent variables and a more complex specification containing both the dependent variablesaEUR(tm) spatial lag and the spatial lags of the other endogenous variables. Second, we compare the use of a variety of alternative spatial weights matricesaEUR"contiguity-based, row and not-row normalized, characteristic-based (e.g., county-size). Third, we consider the aEURoeboarder issueaEUR created by only using data for MissouriaEUR(tm)s counties to estimate our mode.

Identifying and removing heterogeneities between monitoring networks

Jon Olav Skoien[1], Olivier Baume[2], Edzer J. Pebesma[3] and Gerard B.M. Heuvelink[4]

[1] Department of Physical Geography, Utrecht University, The Netherlands
j.skoien@geo.uu.nl
[2] Wageningen University, The Netherlands
olivier.baume@wur.nl
[3] Department of Physical Geography, Utrecht University, The Netherlands
e.pebesma@geog.uu.nl
[4] Wageningen University, The Netherlands
gerard.heuvelink@wur.nl

Summary. Environmental variables are independent of country borders, and there is an increasing interest in combining national observations into international data bases. One common problem while merging data from different networks is the fact that different countries use different measurement devices, or treat the measurements differently before uploading the values into the central database. As a result, the international database will contain heterogeneities that will complicate the use of the database for international comparisons and mapping. Although easy in theory, it will in many cases be difficult or impossible to collect the correct information to determine these heterogeneities.

We will here present a new test statistic that can be used to indicate whether two neighbouring monitoring networks have a discontinuity on the border between the countries. The proposed method does also quantify this difference, and we suggest how the measurements of the different countries can be jointly adjusted to achieve a data set that is spatially more homogeneous than the original data set. The method is based on a

kriging approach.

The method is applied on a data set from the EURDEP database of radioactive gamma dose rate measurements, collected from national gamma dose rate networks in more than 30 European countries (http://eurdep.jrc.it). As different countries use different measurement devices, and have different methods for treatments of the data, there are strong heterogeneities across the country borders, which we will identify with the proposed method.

Bayesian trans-Gaussian Kriging and the Uncertainty of Variogram Estimates

Gunter Spöck

Department of Statistics, University of Klagenfurt, Austria
gunter.spoeck@uni-klu.ac.at

Summary. The spatial prediction method of kriging is based on the Gaussian assumption for the underlying random field and the assumption that the variogram function is known exactly. In most practical applications these assumptions do not hold; the distribution of the observed phenomenon is not symmetric and the used variogram function is just an estimate. The consequence is that the conventional report of the plug-in kriging variances underestimates the true, unknown uncertainty of the kriging predictions. One way out of this dilemma is to use transformations of the original data and make use of the Bayesian paradigm to specify uncertainties about the unknown transformation and variogram parameters. Unfortunately nobody actually is able to get subjective prior knowledge about these parameters. So one has to ask for an approach to specifying noninformative prior knowledge.

The approach to follow in this presentation is based on a parametric bootstrap by simulating both variogram and Box-Cox transformation estimates. The bootstrapped distribution of these estimates is then used in a Bayesian predictive way as a posteriori distribution specifying uncertainty about corresponding unknown parameters. Thus, the sampling distribution of variogram and transformation parameter estimates is taken into account when calculating posterior predictive distributions at unsampled locations. Concerning the variogram both isotropic and anisotropic ones in this way are considered fully in their uncertainty. As bootstrapping replicas maximum-likelihood and weighted least fitting (variogram) estimates are considered.

A side result of bootstrapping is that the sampling distribution of variogram, anisotropy and transformation parameters can be fully investigated. The theoretical developments are examplified then by both, results from simulated skew, anisotropic data sets and a real data set on radioactivity taken 10 years after the Chernobyl accident in the region of Gomel, Belarus.

Modified method of moments - a fast alternative to Maximum likelihood estimation

Milan Zukovic[1] and Dionisis Hristopulos[2]

[1] Department of Mineral Resources Engineering, Technical University of Crete, Greece
mzukovic@mred.tuc.gr
[2] Department of Mineral Resources Engineering, Technical University of Crete, Greece
dionisi@mred.tuc.gr

Summary. This study focuses on the inference of the correlation model for spatially or temporally dependent data. We present a comparative study of two methods of model parameter inference: the established method of maximum likelihood estimation (MLE) and the recently proposed modified method of moments (MMoM) [1]. For reasons of computational efficiency the comparisons are based on correlated data arranged on regularly spaced 1D samples. We consider four commonly used correlation models: the Gaussian, the exponential, the spherical, and the Whittle-Matern models. We find that as the domain size increases, the MMoM estimates exhibit lower bias and dispersion than the MLE estimates for the Gaussian and the exponential models. Both methods produce comparable results in the case of the spherical model. With either the MMoM or the MLE method, The Whittle-Matern model displays a peculiar behavior, characterized by slow or no convergence of the estimates to the true parameters, relatively large dispersion of the estimates, and large optimization times. In all the cases, the CPU time required by MMoM is much lower than that of MLE. Unlike MLE that is significantly slowed-down for large samples, the CPU time of MMoM is not sensitive to the domain size. In fact, our simulations show that the MMOM computational time can even decrease with increasing

sample size. Hence, the MMoM may prove useful for the analysis of large data sets.

References

[1] Hristopulos, D.T. (2003) "Spartan Gibbs random field models for geostatistical applications", SIAM Journal in Scient. Computation, v. 24, p. 2125-2162