PROGRAMME AND ABSTRACTS

4th CSDA International Conference on

Computational and Financial Econometrics (CFE 10)

http://www.cfe-csda.org/cfe10

and

3rd International Conference of the
ERCIM (European Research Consortium for Informatics and Mathematics) Working Group on

Computing & Statistics (ERCIM 10)

http://www.cfe-csda.org/ercim10

Senate House, University of London, UK
10-12 December 2010

ERCIM
European Research Consortium
for Informatics and Mathematics

Queen Mary, University of London
http://www.qmul.ac.uk

Birkbeck, University of London
http://www.bbk.ac.uk

London School of Economics
http://www.lse.ac.uk

CFE 10 & ERCIM 10 International Organizing Committee:
A.-M. Fuertes, J. Niland, E. Ruiz, B. Rustem, E.J. Kontoghiorghes and G. Loizou.

CFE 10 Co-chairs:
George Kapetanios, Oliver Linton, Michael McAleer and Esther Ruiz.

CFE 10 Programme Committee:

A. Amendola, F. Audrino, P. Boswijk, M. Billio, M. Binder, C. Chen, D. van Dijk, C. Francq, A.-M. Fuertes, A.B. Galvao, R. Gerlach, D. Guegan, S. Hodges, S.J. Koopman, L. Khalaf, J.F. Kiviet, T. Lux, S. Mittnik, M. Ooms, J. Nankervis, Y. Omori, M. Paolella, D.S.G. Pollock, T. Proietti, M. Reale, M. Riani, B. Rustem, W. Semmler, M. Steel, M. So, E. Tzavalis, D. Veredas, M. Wagner, P. Winker, P.L.H. Yu, P. Zadrozny, P. Zaffaroni, J.-M. Zakoian and A. Zeileis.

CSDA Annals of CFE Managing Editors:
David A. Belsley, Erricos J. Kontoghiorghes and Herman K. Van Dijk.

CSDA Annals of CFE Guest Editors:
Luc Bauwens, Siem Jan Koopman, Oliver Linton and Michael McAleer.

ERCIM 10 Co-chairs:
Ana Colubi, Ana-Maria Fuertes, Irini Moustaki, D. Stephen G. Pollock, Stefan Van Aelst and Peter Winker.

ERCIM 10 Programme Committee:

M. Alfo, W. Bergsma, D. Boehning, E. Ceyhan, A. Christmann, C. Croux, M. Debruyne, M. Deistler, I. Demetriou, T. Denoeux, V. Esposito Vinzi, P. Filzmoser, K. Fokianos, P. Foschi, R. Fried, C. Gatu, M.A. Gil, S.G. Gilmour, J. Godolphin, B. Grun, M. Hubert, D. Kuhn, U. Ligges, A. Luati, D. Maringer, S. Meintanis, A. C. Monti, D. Morales, M.C. Pardo, S. Paterlini, B.U. Park, C. Perna, S. Petrone, I. Poli, I. Prunster, G. Storti, M. Timmerman, K. Triantafyllopoulos, L. Ugarte and A. Zeileis.

Local Organizing Committee:
Diane Loizou, Elizabeth Price, Evgenia Tsinti and Sharon Woolf.

E723: **Data fitting by divided differences and splines**

Presenter: **Evangelos Vassiliou**, University of Athens, Greece

A smooth univariate function is measured at a finite number of points and the measurements contain random errors. We seek an approximation that should be closer than the measurements to the true function values. The approximation task can be done either by a parametric or by a non-parametric approach. The first approach assumes that the underlying function is described by a functional form that depends on a set of parameters, for example by using splines functions. There are several notions of nonparametric approach, and we take the view that the errors should be corrected by making least the sum of squares of the errors by imposing a limit, q say, on the number of sign changes of the divided differences of a prescribed order r. Depending on the values of r and q, interesting properties may be revealed, like monotonicity, piecewise monotonicity, convexity, convexity/concavity, etc. This paper applies initially a divided difference method in order to smooth the data and then applies a spline regression on the smoothed data by following the shape of the smoothed data. Some examples with real and simulated data are presented in order to illustrate this method. They all suggest that the method provides usually quite suitable fits, because the user is directed by the divided difference step in locating the knots of the spline.

ES27 Room MAL 355 ANALYSIS OF SPATIAL DATA: ESTIMATION, MODELLING, AND INFERENCE Chair: Elvan Ceyhan

E152: **Estimates for geographical domains through geoadditive models in presence of missing information**

Presenter: **Chiara Bocci**, University of Florence, Italy

Co-authors: Emilia Rocco

The implementation of geoadditive models needs the statistical units to be referenced at point locations. If the aim of the study is to analyze the spatial pattern or to produce a spatial interpolation of a studied phenomenon, spatial information are required only for the sampled units. If, however, the geoadditive model is used to produce estimates of a parameter of interest for some geographical domains, the spatial location is required for all the population units. This information is not always easily available. Typically, we know the coordinates for sampled units, but for the non-sampled units we know just the areas - like blocks, municipalities, etc. - to which they belong. In such situation, the classical approach is to locate all the units by the coordinates of their corresponding area centroid. This is obviously an approximation and its effect on the estimates can be strong, depending on the level of nonlinearity in the spatial pattern and on the area dimension. We decided to investigate a different approach: instead of using the same coordinates for all the units, we impose a distribution for the locations inside each area. To analyze the performance of this approach, various MCMC experiments are implemented.

E393: **Linear approximations of individual-level models for infectious disease**

Presenter: **Grace Pui Sze Kwong**, University of Guelph, Canada

Co-authors: Rob Deardon

Individual-level models (ILMs) for infectious diseases, fitted in a Bayesian MCMC framework, are an intuitive and flexible class of models that can take into account population heterogeneity via various individual-level covariates. ILMs containing a geometric distance kernel to account for geographic heterogeneity provide a natural way to model the spatial spread of many diseases. However, in even only moderately large populations, the likelihood calculations required can be prohibitively time consuming. It is possible to speed up the computation via a technique which makes use a linearized distance kernel. Here we examine some methods of carrying out this linearization and compare the performances of these methods.

E182: **Adjusted calibration estimators for sparse spatial data**

Presenter: **Ivan Sciascia**, Universita di Torino, Italy

The technique of calibrated estimation is coupled with an algorithm of vehicles routing, building through the latter, a distance function that adjusts the estimator. The calibration and a vehicle routing algorithm are considered in this paper with the constraint of the limited availability of auxiliary variables and with the use of remotely sensed distances. Remote sensing is an interesting topic due to the development of modern GIS systems which can be useful for sample surveys with geospatial data. In a context of spatial data analysis with a constraint of low-density we propose to adjust the calibrated estimator by a distance function. The Traveling Salesman Problem (TSP) is a basic routing problem stated for visit to P cities with the shortest closed tour. Different algorithms have been developed to resolve heuristically the TSP. This article develops a distance function calculated on the shortest tour distance coming from an heuristic algorithm. The performances of the proposed estimator are evaluated using a simulation study.

ES46 Room MAL B29 COMPONENT ANALYSIS Chair: Marieke Timmerman

E284: **The generic subspace clustering model**

Presenter: **Marieke Timmerman**, University of Groningen, Netherlands

Co-authors: Eva Ceulemans, Kim De Roover

The clustering of high-dimensional data can be troublesome. In the case of high-dimensional data, a subspace clustering model can be used to achieve a proper recovery of the clusters and to obtain an insight into the structure of the variables relevant to the clustering. In such a model the objects are assigned to mutually exclusive classes in low dimensional spaces. In this paper, we present the Generic Subspace Clustering Model. As will be shown, this model encompasses a range of existing (subspace) clustering techniques as special cases. The specific properties of the model variants will be discussed. An algorithm for fitting the Generic Subspace Clustering Model is presented and its performance is evaluated by means of a simulation study. The value of the model for empirical research is illustrated with data from psychiatric diagnosis research.

E340: **Clusterwise SCA-P**

Presenter: **Kim De Roover**, K.U. Leuven, Belgium

Co-authors: Eva Ceulemans, Marieke Timmerman, Patrick Onghena

Numerous research questions in educational sciences and psychology concern the structure of a set of variables. For instance, one can study the structure of a number of variables, measured on multiple occasions for different subjects. In that case, one may wonder whether the same structure is underlying the data of all subjects. Obviously, the crucial question is how such data have to be analyzed to find out whether and in what way the structure of the variables differs across the persons. A number of principal component analysis techniques exist to study such structural differences, for instance, simultaneous component analysis (SCA). However, these techniques suffer from important limitations. Therefore, we propose a novel modeling strategy, called Clusterwise SCA-P, which solves these limitations and which encompasses several existing techniques as special cases. Clusterwise SCA-P partitions the subjects into a number of clusters and specifies a simultaneous component model per cluster. Also, it allows for between subject differences in the variances and the correlations of the cluster specific components, to offer additional insight into interindividual differences. The value of the model for empirical research is illustrated.