

21st International Conference on Computational Statistics

(COMPSTAT 2014)

Hosting the 5th IASC World Conference

Geneva, Switzerland
19-22 August 2014

Editors:

Manfred Gilli
Alicia Nieto-Reyes

Gil Gonzalez-Rodriguez

ISBN: 978-1-63439-585-4

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2014) by the International Statistical Institute
All rights reserved.

Printed by Curran Associates, Inc. (2015)

For permission requests, please contact the International Statistical Institute
at the address below.

International Statistical Institute
P.O. Box 24070
2490 AB The Hague
The Netherlands

Phone: 31-70-3375737
Fax: 31-70-3860025

isi@cbs.nl

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2634
Email: curran@proceedings.com
Web: www.proceedings.com

Contents

Jan Kalina, Zdeněk Valenta and Jurjen Duintjer Tebbens	
Computation of Regularized Linear Discriminant Analysis	1
Paul Fischer and Astrid Hilbert	
Fast Detection of Structural Breaks	9
Anthony C. Atkinson, Marco Riani, Andrea Cerioli and Domenico Perrotta	
Random Start Forward Searches for Detecting Mixtures of Regression Models	17
M. Helena Gonçalves and M. Salomé Cabral	
Incomplete longitudinal binary responses in marginal model	25
Grzegorz Konczak	
On the modification of the non-parametric test for comparing locations of two populations	35
Joan del Castillo, Maria Padilla and Isabel Serra	
Comparison of techniques for extreme values using financial data	45
Paulo C. Rodrigues, Andreia Monteiro and Vanda Lourenço	
New insights into the usefulness of robust singular value decomposition in statistical genetics	53
Borja Lafuente–Rego and Jose Antonio Vilar	
Time series clustering based on quantile autocovariances	61

Frederick Kin Hing Phoa	
A Graphical User Interface Platform of the Stepwise Response Refinement Screener for Screening Experiments	69
Helmut Vorkauf	
Unravel: A Method and a Program to Analyze Contingency Tables, Unveiling Confounders.	81
Juan Eloy Ruiz-Castro	
Preventive maintenance in a complex warm standby system. A transient analysis	89
Pranesh Kumar and Faramarz Kashanchi	
Linear Regression Models Using L_1, L_2 and L_∞-Norms	97
Simon Wilson et al.	
Using Storm for scaleable sequential statistical inference	103
M. Salomé Cabral and M. Helena Gonçalves	
A simulation study to assess statistical approaches for longitudinal count data	111
Matthieu Marbac, Christophe Biernacki and Vincent Vandewalle	
Mixture model of Gaussian copulas to cluster mixed-type data	119
Miguel Casquilho and Elisabete Carolino	
Sampling inspection by (Gaussian) variables via estimation of the lot fraction defective: a computational approach	127
José Antonio Roldan-Nofuentes	
Estimation of the weighted kappa coefficient subject to case-control design	135
Leyla Azarang and Jacobo de Uña-Álvarez	
The jackknife estimate of variance for transition probabilities in the non-Markov illness-death model	143

Ana M. Aguilera and M. Carmen Aguilera-Morillo	
Linear discriminant analysis based on penalized functional PLS	151
D. Ferrari, M. Giuzio and S. Paterlini	
A generalized Description Length approach for Sparse and Robust Index Tracking	157
Paolo Ghisletta, Stephen Aichele, Patrick Rabbitt	
Longitudinal data mining to predict survival in a large sample of adults	167
Fastrich, Paterlini and Winker	
Penalized Least Squares for Optimal Sparse Portfolio Selection	177
Alessandra Amendola and Giuseppe Storti	
Combining information at different frequencies in multivariate volatility prediction	187
Kohei Adachi and Nickolay T. Trendafilov	
Penalty-free sparse PCA	197
Ali Charkhi, Gerda Claeskens and Bruce E. Hansen	
Weight choice by minimizing MSE for general likelihood averaging	205
Jean-Baptiste Durand and Yann Guédon	
Quantifying and localizing state uncertainty in hidden Markov models using conditional entropy profiles	213
S.K. Ng and G.J. McLachlan	
Mixture of regression models with latent variables and sparse coefficient parameters	223
Souleyman Sahnoun and Pierre Comon	
Tensor polyadic decomposition for antenna array processing	233

Caren Hasler and Alina Matei	
Adjustment for nonignorable nonresponse using latent homogeneous response groups	241
Nobuhiro Taneichi, Yuri Sekiya and Jun Toyama	
Bartlett adjustment of deviance statistic for three types of binary response models	249
Yuichi Mori, Masahiro Kuroda, Masaya Iizuka and Michio Sakakihara	
Performance of acceleration of ALS algorithm in nonlinear PCA	257
Niklas Ahlgren and Paul Catani	
Finite-Sample Multivariate Tests for ARCH in Vector Autoregressive Models	265
Miguel Casquilho and Fátima Rosa	
Behaviour of the quality index in acceptance sampling by variables: computation and Monte Carlo simulation	273
Sara Fontanella, Nickolay T. Trendafilov and Kohei Adachi	
Sparse exploratory factor analysis	281
M. Ivette Gomes and Frederico Caeiro	
Efficiency of partially reduced-bias mean-of-order-p versus minimum-variance reduced-bias extreme value index estimation	289
J.R. Wishart	
Data-driven wavelet resolution choice in multichannel box-car deconvolution with long memory	299
Hirohito Sakurai and Masaaki Taguri	
Comparison of block bootstrap testing methods of mean difference for paired longitudinal data	309
M. Arisido	
Functional data modeling to measure exposure to ozone	319

Raquel Caballero-Águila, Aurora Hermoso-Carazo and Josefa Linares-Pérez	
Estimation based on covariances from multiple one-step randomly delayed measurements with noise correlation	327
C. Mante	
Density and Distribution Function estimation through iterates of fractional Bernstein Operators	335
Craig Anderson, Duncan Lee, Nema Dean	
Bayesian cluster detection via adjacency modelling	343
Jan Amos Visek	
Robust test of restricted model	351
Isabelle Charlier, Davy Paindaveine and Jérôme Saracco	
Conditional quantile estimation using optimal quantization: a numerical study	361
Robert M. Kunst	
A combined nonparametric test for seasonal unit roots	369
Hannah Frick, Carolin Strobl and Achim Zeileis	
To Split or to Mix? Tree vs. Mixture Models for Detecting Subgroups	379
Massimo Cannas and Bruno Arpino	
Propensity score matching with clustered data: an application to birth register data	387
Fernanda Figueiredo, M. Ivette Gomes and Adelaide Figueiredo	
Monitoring the shape parameter of a Weibull distribution	395
Mingfei Qiu, Vic Patrangenaru and Leif Ellingson	
How far is the Corpus Callosum of an Average Individual from Albert Einstein's?	403

Kosuke Okusa and Toshinari Kamakura	
Statistical Registration of Frontal View Gait Silhouette with Application to Gait Analysis	411
Pavel Mozgunov	
Application of Kalman Filter with alpha-stable distribution	419
Bernard Fichet	
Combining sub(up)-approximations of different type to improve a solution	427
Elvira Pelle <i>et al.</i>	
Log-linear multidimensional Rasch model for capture-recapture	435
Adelaide Figueiredo and Fernanda Figueiredo	
Monitoring the process variability using STATIS	443
Stefano M. Iacus and Lorenzo Mercuri	
Estimation of Lévy CARMA models in the yuima package: application on the financial time series	451
Christian Derquenne	
Modelling multivariate time series by structural equations modelling and segmentation approach	459
Martin Schindler, Jan Picek and Jan Kyselý	
Study on the choice of regression quantile threshold in a POT model	467
Ryo Takahashi	
Reduced K-means with sparse loadings	475
Antonio Irpino, Antonio Balzanella and Rosanna Verde	
Spatial dependence monitoring over distributed data streams	483

F. Marta L. Di Lascio, Simone Giannerini and Alessandra Reale	
Imputation of complex dependent data by conditional copulas: analytic versus semiparametric approach	491
Antonio Abbruzzo, Luigi Augugliaro, Angelo M. Mineo and Ernst C. Wit	
Cyclic coordinate for penalized Gaussian graphical models with symmetry restrictions	499
Sujung Kim, Kuniyoshi Hayashi and Koji Kurihara	
The optimal number of lags in variogram estimation in spatial data analysis	507
F. Giordano, S.N. Lahiri and M.L. Parrella	
GRID for variable selection in high dimensional regression	515
Sakyajit Bhattacharya and Vaibhav Rajan	
Unsupervised Learning using Gaussian Mixture Copula Model	523
Francesco Bartolucci, Giorgio E. Montanari and Silvia Pandolfi	
A comparison of some estimation methods for latent Markov models with covariates	531
Fernández-Pascual, R. Espejo, R and Ruiz-Medina, M.D.	
Estimation of spatially correlated ocean temperature curves including depth dependent covariates	539
Frederico Caeiro and M. Ivette Gomes	
On the bootstrap methodology for the estimation of the tail sample fraction	545
Marco Di Marzio <i>et al.</i>	
Local likelihood estimation for multivariate directional data	553
Pierre Fernique, Jean-Baptiste Durand and Yann Guédon	
Estimation of Discrete Partially Directed Acyclic Graphical Models in Multitype Branching Processes	561

Manuela Neves and Clara Cordeiro	
Statistical modelling in time series extremes: an overview and new steps	569
Luca Frigau, Claudio Conversano and Francesco Mola	
A bivariate cost-sensitive classifier performance index	577
Ronald Hochreiter and Christoph Waldhauser	
Effects of Sampling Methods on Prediction Quality. The Case of Classifying Land Cover Using Decision Trees.	585
Stasi <i>et al.</i>	
β models for random hypergraphs with a given degree sequence	593
A.Wawrzynczak, P.Kopka, M.Jaroszynski and M.Borysiewicz	
Efficiency of Sequential Monte Carlo and Genetic algorithm in Bayesian estimation of the atmospheric contamination source	601
Muhammad-Anas Knefati and Farid Beninel	
Transfer of semiparametric single index model in binary classification	609
Pierre Michel and Badih Ghattas	
Clustering ordinal data using binary decision trees	617
Katrin Illner <i>et al.</i>	
Bayesian blind source separation applied to the lymphocyte pathway	625
Manuela Cattelan	
Maximum simulated likelihood estimation of Thurstonian models	633
Manuela Souto de Miranda, Conceição Amado and Margarida Silva	
Robust profiling of Site Index	641
Charalampos Chaniavidis, Ludger Evers and Tereza Neocleous	
Bayesian density regression for count data	649

Contents	xvii
Zdeněk Fabián	
Score Function of Distribution and Heavy-tails	657
Ayca Yetere Kursun, Cem Iyigun and Inci Batmaz	
Consensus Clustering of Time Series Data	665
Yusuke Matsui and Masahiro Mizuta	
SDA for mixed-type data and its application to analysis of environmental radio activity level data	673
Pasquale Dolce, Vincenzo Esposito Vinzi and Carlo Lauro	
Predictive Component-based Multi-block Path Modeling	681
Sadika Rjiba, Mireille Gettler Summa and Saloua Benammou	
Joint analysis of closed and open-ended questions in a survey about the Tunisian revolution	685