Book of Short Papers SIS 2018

Editors: Antonino Abbruzzo - Eugenio Brentari
Marcello Chiodi - Davide Piacentino
Contents

1 Preface ................................................................. 17

2 Plenary Sessions ..................................................... 19
  2.1 A new paradigm for rating data models. Domenico Piccolo 19
  2.2 Statistical challenges and opportunities in modelling coupled behaviour-
disease dynamics of vaccine refusal. Plenary/Chris T. Bauch 32

3 Specialized Sessions .................................................. 45
  3.1 3.1 - Bayesian Nonparametric Learning 45
   3.1.1 Bayesian nonparametric covariate driven clustering. Raffaele Argiento, Iliaria
          Bianchini, Alessandra Guglielmi and Ettore Lanzarone ......................... 46
   3.1.2 A Comparative overview of Bayesian nonparametric estimation of the size of a
          population. Luca Tardella and Danilo Alunni Fegatelli .......................... 56
   3.1.3 Logit stick-breaking priors for partially exchangeable count data. Tommaso Rigon
          64
   3.2 BDsports - Statistics in Sports 72
   3.2.1 A paired comparison model for the analysis of on-field variables in football matches.
          Gunther Schauberger and Andreas Groll .................................................. 72
   3.2.2 Are the shots predictive for the football results?. Leonardo Egidi, Francesco Pauli,
          Nicola Torelli ................................................................. 81
   3.2.3 Zero-inflated ordinal data models with application to sport (in)activity. Maria
          Iannario and Rosaria Simone ......................................................... 89
  3.3 Being young and becoming adult in the Third Millennium: definition issues
      and processes analysis 97
   3.3.1 Do Social Media Data predict changes in young adults’ employment status?
      Evidence from Italy. Andrea Bonanomi and Emiliano Sironi ................. 97
3.3.2 Parenthood: an advanced step in the transition to adulthood. Cinzia Castagnaro, Antonella Guarneri and Eleonora Meli

3.4 Economic Statistics and Big Data
3.4.1 Improvements in Italian CPI/HICP deriving from the use of scanner data. Alessandro Brunetti, Stefania Fatello, Federico Polidoro, Antonella Simone
3.4.2 Big data and spatial price comparisons of consumer prices. Tiziana Laureti and Federico Polidoro

3.5 Financial Time Series Analysis
3.5.1 Dynamic component models for forecasting trading volumes. Antonio Naimoli and Giuseppe Storti
3.5.2 Conditional Quantile-Located VaR. Giovanni Bonaccolto, Massimiliano Caporin and Sandra Paterlini

3.6 Forensic Statistics
3.6.1 Cause of effects: an important evaluation in Forensic Science. Fabio Corradi and Monica Musio
3.6.2 Evaluation and reporting of scientific evidence: the impact of partial probability assignments. Silvia Bozza, Alex Biedermann, Franco Taroni

3.7 Missing Data Handling in Complex Models
3.7.1 Dependence and sensitivity in regression models for longitudinal responses subject to dropout. Marco Alfo and Maria Francesca Marino
3.7.2 Multilevel analysis of student ratings with missing level-two covariates: a comparison of imputation techniques. Maria Francesca Marino and Carla Rampichini
3.7.3 Multilevel Multiple Imputation in presence of interactions, non-linearities and random slopes. Matteo Quartagno and James R. Carpenter

3.8 Monitoring Education Systems. Insights from Large Scale Assessment Surveys

3.9 New Perspectives in Time Series Analysis
3.9.1 Generalized periodic autoregressive models for trend and seasonality varying time series. Francesco Battaglia and Domenico Cucina and Manuel Rizzo

3.10 Recent Advances in Model-based Clustering
3.10.1 Flexible clustering methods for high-dimensional data sets. Cristina Tortora and Paul D. McNicholas
3.10.2 A Comparison of Model-Based and Fuzzy Clustering Methods. Marco Alfo’, Maria Brigida Ferraro, Paolo Giordani, Luca Scrucca, and Alessio Serafini
3.10.3 Covariate measurement error in generalized linear models for longitudinal data: a latent Markov approach. Roberto Di Mari, Antonio Punzo, and Antonello Maruotti

3.11 Statistical Modelling
3.11.1 A regularized estimation approach for the three-parameter logistic model. Michela Battauz and Ruggero Bellio
3.11.2 Statistical modelling and GAMLSS. Mikis D. Stasinopoulos and Robert A. Rigby and Fernanda De Bastiani

3.12 Young Contributions to Statistical Learning
3.12.1 Introducing spatio-temporal dependence in clustering: from a parametric to a nonparametric approach. Clara Grazian, Gianluca Mastrandanto and Enrico Bibbona
3.12.2 Bayesian inference for hidden Markov models via duality and approximate filtering distributions. Guillaume Kon Kam King, Omiros Papaspiliopoulos and Matteo Ruggiero 248


4 Sollicited Sessions 263

4.1 Advances in Discrete Latent Variable Modelling 264
4.1.1 A joint model for longitudinal and survival data based on a continuous-time latent Markov model. Alessio Farcomeni and Francesco Bartolucci 264
4.1.2 Modelling the latent class structure of multiple Likert items: a paired comparison approach. Brian Francis 273
4.1.3 Dealing with reciprocity in dynamic stochastic block models. Francesco Bartolucci, Maria Francesca Marino, Silvia Pandolfi 281
4.1.4 Causality patterns of a marketing campaign conducted over time: evidence from the latent Markov model. Fulvia Pennoni, Leo Paas and Francesco Bartolucci 289

4.2 Complex Spatio-temporal Processes and Functional Data 297
4.2.1 Clustering of spatio-temporal data based on marked variograms. Antonio Balzanella and Rosanna Verde 297
4.2.2 Space-time earthquake clustering: nearest-neighbor and stochastic declustering methods in comparison. Elisa Varini, Antonella Peresan, Renata Rotondi, and Stefania Gentili 304
4.2.3 Advanced spatio-temporal point processes for the Sicily seismicity analysis. Marianna Siino and Giada Adelfio 312
4.2.4 Spatial analysis of the Italian seismic network and seismicity. Antonino D’Alessandro, Marianna Siino, Luca Greco and Giada Adelfio 320

4.3 Dimensional Reduction Techniques for Big Data Analysis 328
4.3.1 Clustering Data Streams via Functional Data Analysis: a Comparison between Hierarchical Clustering and K-means Approaches. Fabrizio Maturo, Francesca Fortuna, and Tonio Di Battista 328
4.3.2 Co-clustering algorithms for histogram data. Francisco de A.T. de Carvalho and Antonio Balzanella and Antonio Irpino and Rosanna Verde 338
4.3.3 A network approach to dimensionality reduction in Text Mining. Michelangelo Misuraca, Germana Scepi and Maria Spano 344
4.3.4 Self Organizing Maps for distributional data. Rosanna Verde and Antonio Irpino 352

4.4 Environmental Processes, Human Activities and their Interactions 353
4.4.1 Estimation of coral growth parameters via Bayesian hierarchical non-linear models. Crescenza Calciulli, Barbara Catarelli and Daniela Cocchi 353
4.4.2 A Hierarchical Bayesian Spatio-Temporal Model to Estimate the Short-term Effects of Air Pollution on Human Health. Fontanella Lara, Ippoliti Luigi and Valentini Pasquale 361
4.4.3 A multilevel hidden Markov model for space-time cylindrical data. Francesco Lagona and Monia Ranalli 367
4.4.4 Estimation of entropy measures for categorical variables with spatial correlation. Linda Alfieri, Giulia Roli 373

4.5 Innovations in Census and in Social Surveys 381
4.5.1 A micro-based approach to ensure consistency among administrative sources and to improve population statistics. Gianni Corsetti, Sabrina Prati, Valeria Tomeo, Enrico Tucci 381
4.5.2 Demographic changes, research questions and data needs: issues about migrations. Salvatore Strozza and Giuseppe Gabrielli 392
Towards more timely census statistics: the new Italian multiannual dissemination programme. Simona Mastroluca and Mariangela Verrascina

Living Conditions and Consumption Expenditure in Time of Crises

Household consumption expenditure and material deprivation in Italy during last economic crises. Ilaria Arigoni and Isabella Siciliani

Network Data Analysis and Mining

Support provided by elderly Italian people: a multilevel analysis. Elvira Pelle, Giulia Rivellini and Susanna Zaccarini

Data mining and analysis of comorbidity networks from practitioner prescriptions. Giancarlo Ragazini, Giuseppe Giordano, Sergio Pagano, Mario De Santis, Pierpaolo Cavallo

Overlapping mixture models for network data (manet) with covariates adjustment. Saverio Ranciati and Giuliano Galimberti and Ernst C. Wit and Veronica Vinciotto

New Challenges in the Measurement of Economic Insecurity, Inequality and Poverty

Social protection in mitigating economic insecurity. Alessandra Coli

Changes in poverty concentration in U.S. urban areas. Francesco Andreoli and Mauro Mussini

Evaluating sustainability through an input-state-output framework: the case of the Italian provinces. Achille Lemmi, Laura Neri, Federico M. Pulselli

New Methods and Models for Ordinal Data

Weighted and unweighted distances based decision tree for ranking data. Antonella Plaia, Simona Buscemi, Mariangela Sciandra

A dissimilarity-based splitting criterion for CUBREMOT. Carmela Cappelli, Rosaria Simone and Francesca Di Iorio

Constrained Extended Plackett-Luce model for the analysis of preference rankings. Cristina Mollica and Luca Tardella

A prototype for the analysis of time use in Italy. Stefania Capecchi and Manuela Michelin

New Perspectives in Supervised and Unsupervised Classification

Robust Updating Classification Rule with applications in Food Authenticity Studies. Andrea Cappozzo, Francesca Greselin and Thomas Brendan Murphy

A robust clustering procedure with unknown number of clusters. Francesco Dotto and Alessio Farcomeni

 Issues in joint dimension reduction and clustering methods. Michel van de Velden, Alfonso Iodice D’Enza and Angelos Markos

New Sources, Data Integration and Measurement Challenges for Estimates on Labour Market Dynamics

The development of the Italian Labour register: principles, issues and perspectives. C. Baldi, C. Ceccarelli, S. Gigante, S. Pacini

Digging into labour market dynamics: toward a reconciliation of stock and flows short term indicators. F. Rapiti, C. Baldi, D. Ichim, F. Pintaldi, M. E. Pontecorvo, R. Rizzi

How effective are the regional policies in Europe? The role of European Funds. Gennaro Punzo, Mariateresa Ciommi, and Gaetano Musella

Issues in joint dimension reduction and clustering methods. Lucio Masserini and Matilde Bini
4.12 Quantile and Generalized Quantile Methods

4.12.1 Multiple quantile regression for risk assessment. Lea Petrella and Valentina Raponi 547

4.12.2 Parametric Modeling of Quantile Regression Coefficient Functions. Paolo Frumento and Matteo Bottai ........................................... 550

4.12.3 Modelling the effect of Traffic and Meteorology on Air Pollution with Finite Mixtures of M-quantile Regression Models. Simone Del Sarto, Maria Francesca Marino, Maria Giovanna Ranalli and Nicola Salvati ........................................ 552

4.12.4 Three-level M-quantile model for small area poverty mapping. Stefano Marchetti and Nicola Salvati ........................................... 560

4.13 Recent Advances on Extreme Value Theory

4.13.1 Extremes of high-order IGARCH processes. Fabrizio Laurini ..................... 560

4.14 Spatial Economic Data Analysis

4.14.1 Spatial heterogeneity in principal component analysis: a study of deprivation index on Italian provinces. Paolo Postiglione, M. Simona Andreano, Roberto Benedetti, Alfredo Cartone ........................................... 569

4.15 Spatial Functional Data Analysis

4.15.1 Object oriented spatial statistics for georeferenced tensor data. Alessandra Menafoglio and Davide Pigoli and Piercesare Secchi ..................... 578

4.15.2 A Spatio-Temporal Mixture Model for Urban Crimes. Ferretti Angela, Ippoliti Luigi and Valentini Pasquale ........................................... 585

4.16 Statistical Methods for Service Quality

4.16.1 Cumulative chi-squared statistics for the service quality improvement: new properties and tools for the evaluation. Antonello D’Ambra, Antonio Lucadamo, Pietro Amenta, Luigi D’Ambra ........................................... 591

4.16.2 A robust multinomial logit model for evaluating judges’ performances. Ida Camminiatiello and Antonio Lucadamo ........................................... 600

4.16.3 Complex Contingency Tables and Partitioning of Three-way Association Indices for Assessing Justice CourtWorkload. Rosaria Lombardo, Yoshio Takane and Eric J Beh 607

4.16.4 Finding the best paths in university curricula of graduates to improve academic guidance services. Silvia Bacci and Bruno Bertaccini ........................................... 615

4.17 Statistical Modelling for Business Intelligence Problems

4.17.1 A nonlinear state-space model for the forecasting of field failures. Antonio Pievatolo 623

4.17.2 Does Airbnb affect the real estate market? A spatial dependence analysis. Mariangela Guidolin and Mauro Bernardi ........................................... 632

4.17.3 Bayesian Quantile Trees for Sales Management. Mauro Bernardi and Paola Stolfi 640

4.17.4 Discrimination in machine learning algorithms. Roberta Pappadà and Francesco Pauli ........................................... 648

4.18 Statistical models for sports data

4.18.1 Exploring the Kaggle European Soccer database with Bayesian Networks: the case of the Italian League Serie A. Maurizio Carpita and Silvia Golta ........................................... 656

4.18.2 A data-mining approach to the Parkour discipline. Paola Pasca, Enrico Ciavolino and Ryan L. Boyd ........................................... 665

4.18.3 Players Movements and Team Shooting Performance: a Data Mining approach for Basketball. Rodolfo Metulini ........................................... 673
4.19 Supporting Regional Policies through Small Area Statistical Methods 681

4.19.1 Survey-weighted Unit-Level Small Area Estimation. Jan Pablo Burgard and Patricia Dörr ............................................ 681

4.20 The Second Generation at School 681

4.20.1 Resilient students with migratory background. Anna Di Bartolomeo and Giuseppe Gabrielli ........................................ 681

4.20.2 Residential Proximity to Attended Schools among Immigrant-Origin Youths in Bologna. Federica Santangelo, Debora Mantovani and Giancarlo Gasperoni ...... 690

4.20.3 From school to ... future: strategies, paths and perspectives of immigrant immediate descendants in Naples. Giustina Orientale Caputo and Giuseppe Garigli 698

4.21 Tourism Destinations, Household, Firms 706

4.21.1 The Pricing Behaviour of Firms in the On-line Accommodation Market: Evidence from a Metropolitan City. Andrea Guizzardi and Flavio Maria Emanuele Pons 706

4.21.2 The Migration-Led-Tourism Hypothesis for Italy: A Survey. Carla Massidda, Romano Piras and Ivan Ezzo .................................................. 716

4.21.3 Tourism Statistics: development and potential uses. Fabrizio Antolini .............. 724

4.21.4 Tourism attractiveness in Italy. Some empirical evidence comparing origin-destination domestic tourism flows. Francesca Giambona, Emanuela Dreassi, and Alessandro Magrini .................................................. 732

4.22 What’s Happening in Africa 740

4.22.1 Environmental shocks and internal migration in Tanzania. Maria Francesca Marino, Alessandra Petrucci, and Elena Pirani ........................................ 740

4.22.2 Determinants and geographical disparities of BMI in African Countries: a measurement error small area approach. Serena Arima and Silvia Polettini ...... 748

5 Contributed Sessions ........................................................................... 757

5.1 Advanced Algorithms and Computation 758

5.1.1 Brexit in Italy. Francesca Greco, Livia Celardo, Leonardo Salvatore Alaimo . 758

5.1.2 Distance based Depth-Depth classifier for directional data. Giuseppe Pandolfo and Giovanni C. Porzio ................................................. 765

5.1.3 Approximate Bayesian Computation for Forecasting in Hydrological models. Jonathan Romero-Cuéllar, Antonino Abbuzzi, Giada Adeilfo and Félix Frances . 769

5.1.4 Customer Churn prediction based on eXtreme Gradient Boosting classifier. Mohammed Hassan Elbedawi Omar and Matteo Borrotti ............. 775

5.1.5 HPC-accelerated Approximate Bayesian Computation for Biological Science. Rita-brata Dutta .................................................. 781

5.1.6 PC Algorithm for Gaussian Copula Data. Vincenzita Vitale and Paola Vicard 789

5.2 Advances in Clustering Techniques .................................................. 795

5.2.1 On the choice of an appropriate bandwidth for modal clustering. Alessandro Casa, José E. Chacón and Giovanna Menardi ....................... 795

5.2.2 Unsupervised clustering of Italian schools via non-parametric multilevel models. Chiara Masci, Francesca leva and Anna Maria Paganoni .................. 802

5.2.3 Chiara Masci, Francesca leva and Anna Maria Paganoni. Laura Bocci and Donatella Vicari .................................................. 808

5.2.4 Robust Reduced k-Means and Factorial k-Means by trimming. Luca Greco and Antonio Lucadamo and Pietro Amenta ....................... 813

5.2.5 Dirichlet processes, posterior similarity and graph clustering. Stefano Tonellato 819

5.2.6 Bootstrap ClustGeo with spatial constraints. Veronica Distefano, Valentina Mameii, Fabio Della Marra ................................................. 825
5.3 **Advances in Statistical Models** 831

5.3.1 Regression modeling via latent predictors. Francesca Martella and Donatella Vicari ................................................................. 831

5.3.2 Analysis of dropout in engineering BSc using logistic mixed-effect models. Luca Fontana and Anna Maria Paganoni .......................... 838

5.3.3 dglARS method for relative risk regression models. Luigi Augugliaro and Angelo M. Mineo .......................................................... 844

5.3.4 A Latent Class Conjoint Analysis for analysing graduates profiles. Paolo Mariani, Andrea Marletta, Lucio Masserini and Mariangela Zenga .... 844

5.3.5 A longitudinal analysis of the degree of accomplishment of anti-corruption measures by Italian municipalities: a latent Markov approach. Simone Del Sarto, Michela Gnaldi, Francesco Bartolucci ........................................ 856

5.3.6 Modelling the effect of covariates for unbiased estimates in ecological inference methods. Venera Tomaselli, Antonio Forcina and Michela Gnaldi ........................................... 862

5.4 **Advances in Time Series** 868

5.4.1 Filtering outliers in time series of electricity prices. Ilaria Lucrezia Amerise ................................................................. 868

5.4.2 Time-varying long-memory processes. Luisa Bisaglia and Matteo Grigoletto ................................................................. 875

5.4.3 Statistical Analysis of Markov Switching DSGE Models. Maddalena Cavicchioli ................................................................. 881

5.4.4 Forecasting energy price volatilities and comovements with fractionally integrated MGARCH models. Malvina Marchese and Francesca Di Iorio ................................................................. 886

5.4.5 Improved bootstrap simultaneous prediction limits. Paolo Vidoni ................................................................. 892

5.5 **Data Management** 898

5.5.1 Using web scraping techniques to derive co-authorship data: insights from a case study. Domenico De Stefano, Vittorio Fuccella, Maria Prosperina Vitale, Susanna Zaccarin ................................................................. 898

5.5.2 Dealing with Data Evolution and Data Integration: An approach using Rarefaction. Luca Dei Core, Eugenio Montini, Clelia Di Serio, Andrea Calabria ................................................................. 905

5.5.3 Monitoring event attendance using a combination of traditional and advanced surveying tools. Mauro Ferrante, Amit Birenboim, Anna Maria Milito, Stefano De Canton ................................................................. 911

5.5.4 Indefinite Topological Kernels. Tullia Padellini and Pierpaolo Brutti ................................................................. 917

5.5.5 Data Integration in Social Sciences: the earnings intergenerational mobility problem. Veronica Ballerini, Francesco Bloise, Dario Briscolin and Michele Raitano ................................................................. 923

5.5.6 An innovative approach for the GDPR compliance in Big Data era. M. Giacalone, C. Cusatelli, F. Fanari, V. Santarcangelo, D.C. Sinitó ................................................................. 929

5.6 **Developments in Graphical Models** 935

5.6.1 An extension of the glasso estimator to multivariate censored data. Antonino Abbruzzo and Luigi Augugliaro and Angelo M. Mineo ................................................................. 935

5.6.2 Bayesian Estimation of Graphical Log-Linear Marginal Models. Claudia Tarantola, Ioannis Ntzoufras and Monia Lupparelli ................................................................. 942

5.6.3 Statistical matching by Bayesian Networks. Daniela Marella and Paola Vicard and Vincenzina Vitale ................................................................. 948

5.6.4 Sparse Nonparametric Dynamic Graphical Models. Fabrizio Poggioni, Mauro Bernardi, Leo Petrella ................................................................. 954

5.6.5 Non-communicable diseases, socio-economic status, lifestyle and well-being in Italy: An additive Bayesian network model. Laura Maniscalco and Domenica Matranga ................................................................. 960

5.6.6 Using Almost-Dynamic Bayesian Networks to Represent Uncertainty in Complex Epidemiological Models: a Proposal. Sabina Marchetti ................................................................. 966
<table>
<thead>
<tr>
<th>5.7</th>
<th>Educational World</th>
<th>972</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.7.1</td>
<td>How to improve the Quality Assurance System of the Universities: a study based on compositional analysis. Bertaccini B., Gallo M., Simonacci V., and Menini T.</td>
<td>972</td>
</tr>
<tr>
<td>5.7.3</td>
<td>Testing for the Presence of Scale Drift: An Example. Michela Battauz</td>
<td>983</td>
</tr>
<tr>
<td>5.7.4</td>
<td>The evaluation of Formative Tutoring at the University of Padova. Renata Clerici, Lorenza Da Re, Anna Giraldo, Silvia Meggiolaro</td>
<td>988</td>
</tr>
<tr>
<td>5.7.5</td>
<td>Benefits of the Erasmus mobility experience: a discrete latent variable analysis. Silvia Bacci, Valeria Caviezel and Anna Maria Falzoni</td>
<td>993</td>
</tr>
<tr>
<td>5.7.6</td>
<td>University choice and the attractiveness of the study area. Insights from an analysis based on generalized mixed-effect models. Silvia Columbu, Mariano Porcu and Isabella Sulis</td>
<td>999</td>
</tr>
<tr>
<td>5.8</td>
<td>Environment</td>
<td>1005</td>
</tr>
<tr>
<td>5.8.1</td>
<td>The climate funds for energy sustainability: a counterfactual analysis. Alfonso Carfora and Giuseppe Scandurra</td>
<td>1005</td>
</tr>
<tr>
<td>5.8.2</td>
<td>Exploratory GIS Analysis via Spatially Weighted Regression Trees. Carmela Iorio, Giuseppe Pandolfo, Michele Staino, and Roberta Siciliano</td>
<td>1012</td>
</tr>
<tr>
<td>5.8.3</td>
<td>A functional regression control chart for profile monitoring. Fabio Centofanti, Antonio Lepore, Alessandra Menafoglio, Biagio Palumbo and Simone Vantini</td>
<td>1018</td>
</tr>
<tr>
<td>5.8.4</td>
<td>Understanding pro-environmental travel behaviours in Western Europe. Gennaro Punzo, Rosalia Castellano, and Demetrio Panarello</td>
<td>1023</td>
</tr>
<tr>
<td>5.9</td>
<td>Family &amp; Economic issues</td>
<td>1029</td>
</tr>
<tr>
<td>5.9.1</td>
<td>Measuring Economic Uncertainty: Longitudinal Evidence Using a Latent Transition Model. Francesca Giambona, Laura Grassini and Daniele Vignoli</td>
<td>1029</td>
</tr>
<tr>
<td>5.9.2</td>
<td>Intentions to leave Italy or to stay among foreigners: some determinants of migration projects. Ginevra Di Giorgio, Francesca Dota, Paola Muccitelli and Daniele Spizzichino</td>
<td>1036</td>
</tr>
<tr>
<td>5.9.3</td>
<td>Wages differentials in association with individuals, enterprises and territorial characteristics. S. De Santis, C. Freguja, A. Masi, N. Pannuzi, F. G. Truglia</td>
<td>1042</td>
</tr>
<tr>
<td>5.9.4</td>
<td>The Transition to Motherhood among British Young Women: Does housing tenure play a role?. Valentina Tocchioni, Ann Berrington, Daniele Vignoli and Agnese Vitali</td>
<td>1048</td>
</tr>
<tr>
<td>5.10</td>
<td>Finance &amp; Insurance</td>
<td>1054</td>
</tr>
<tr>
<td>5.10.1</td>
<td>Robust statistical methods for credit risk. A. Corbellini, A. Ghiretti, G. Morelli and A. Talignani</td>
<td>1054</td>
</tr>
<tr>
<td>5.10.2</td>
<td>Depth-based portfolio selection. Giuseppe Pandolfo, Carmela Iorio and Antonio D’Ambrosio</td>
<td>1061</td>
</tr>
<tr>
<td>5.10.3</td>
<td>Estimating large-scale multivariate local level models with application to stochastic volatility. Matteo Pelagatti and Giacomo Sbrana</td>
<td>1067</td>
</tr>
<tr>
<td>5.11</td>
<td>Health and Clinical Data</td>
<td>1073</td>
</tr>
<tr>
<td>5.11.1</td>
<td>Is retirement bad for health? A matching approach. Elena Pirani, Marina Ballerini, Alessandra Mattel, Gustavo De Santis</td>
<td>1073</td>
</tr>
<tr>
<td>5.11.2</td>
<td>The emergency department utilisation among the immigrant population resident in Rome from 2005 to 2015. Eleonora Trappolini, Laura Cacciani, Claudia Marina, Cristina Giudici, Nera Agabiti, Marina Davoli</td>
<td>1080</td>
</tr>
<tr>
<td>5.11.3</td>
<td>Multi-State model with nonparametric discrete frailty. Francesca Gasperoni, Francesca Ieva, Anna Maria Paganoni, Chris Jackson and Linda Sharples</td>
<td>1087</td>
</tr>
<tr>
<td>5.11.4</td>
<td>A Functional Urn Model for CARA Designs. Giacomo Aletti, Andrea Ghiglietti, and William F. Rosenberger</td>
<td>1093</td>
</tr>
</tbody>
</table>
5.11.5 Assessment of the INLA approach on gerarchic bayesian models for the spatial disease distribution: a real data application. Paolo Girardi, Emanuela Bovo, Carmen Stocco, Susanna Baracco, Alberto Rosano, Daniele Monetti, Silvia Rizzato, Sara Zamberlan, Enrico Chinellato, Ugo Fedeli, Massimo Rugge

5.12 Medicine

5.12.1 Hidden Markov Models for disease progression. Andrea Martino, Andrea Ghiglietti, Giuseppina Guatteri, Anna Maria Paganoni

5.12.2 A simulation study on the use of response-adaptive randomized designs. Anna Maria Paganoni, Andrea Ghiglietti, Maria Giovanna Scarale, Rosalba Miceli, Francesca Ieva, Luigi Mariani, Cecilia Gavazzi and Valeria Edefonti

5.12.3 The relationship between health care expenditures and time to death: focus on myocardial infarction patients. Luca Grasetti and Laura Rizzi

5.12.4 A multivariate extension of the joint models. Marcella Mazzoleni and Mariangela Zenga

5.12.5 Multipurpose optimal designs for hypothesis testing in normal response trials. Marco Novelli and Maroussa Zagoraiou

5.12.6 Additive Bayesian networks for an epidemiological analysis of swine diseases. Marta Pittavino and Reinhard Furrer

5.13 Population Dynamics

5.13.1 Employment Uncertainty and Fertility: a Meta-Analysis of European Research Findings. Giammarco Alderotti, Daniele Vignoli and Michela Baccini

5.13.2 What Shapes Population Age Structures in the Long Run. Gustavo De Santis and Giambattista Salinari

5.13.3 The impact of economic development on fertility: a complexity approach in a cross-country analysis. Niccolò Innocenti, Daniele Vignoli and Luciana Lazzeretti

5.13.4 A Probabilistic Cohort-Component Model for Population Forecasting - The Case of Germany. Patrizio Vanella and Philipp Deschermeier


5.14 Recent Developments in Bayesian Inference

5.14.1 Posterior distributions with non explicit objective priors. Erlis Ruli, Nicola Sartori and Laura Ventura

5.14.2 A predictive measure of the additional loss of a non-optimal action under multiple priors. Fulvio De Santis and Stefania Gubbiotti

5.14.3 Bayesian estimation of number and position of knots in regression splines. Giola Di Credico, Francesco Pauli and Nicola Torelli

5.14.4 The importance of historical linkages in shaping population density across space. Ilenia Epifani and Rosella Nicolini

5.15 Recent Developments in Sampling

5.15.1 Species richness estimation exploiting purposive lists: A proposal. A. Chiarucci, R.M. Di Biase, L. Fattorini, M. Marcheselli and C. Pisani

5.15.2 Design-based exploitation of big data by a doubly calibrated estimator. Maria Michela Dickson, Giuseppe Espa and Lorenzo Fattorini

5.15.3 Design-based mapping in environmental surveys. L. Fattorini, M. Marcheselli and C. Pisani

5.15.4 Testing for independence in analytic inference. Pier Luigi Conti and Alberto Di Iorio

5.15.5 On the aberrations of two-level Orthogonal Arrays with removed runs. Roberto Fontana and Fabio Rapallo
## 5.16 Recent Developments in Statistical Modelling

### 5.16.1 Quantile Regression Coefficients Modeling: a Penalized Approach  
**Gianluca Sottile, Paolo Frumento and Matteo Bottai** .................................................. 1200

### 5.16.2 Simultaneous calibrated prediction intervals for time series  
**Giovanni Fonseca, Federica Giummolé and Paolo Vidoni** .............................................. 1207

### 5.16.3 Reversibility and (non)linearity in time series  
**Luísa Bissaglia and Margherita Gerolimetto** ................................................................. 1213

### 5.16.4 Heterogeneous effects of subsidies on farms’ performance: a spatial quantile regression analysis  
**Marusca De Castris and Daniele Di Gennaro** ................................................................. 1219

### 5.16.5 On the estimation of high-dimensional regression models with binary covariates  
**Valentina Mameli, Debora Slanzi and Irene Poli** ............................................................ 1226

## 5.17 Social Indicators

### 5.17.1 Can a neighbour region influence poverty? A fuzzy and longitudinal approach  
**Gianni Betti, Federico Crescenzi and Francesca Gagliardi** ............................................... 1232

### 5.17.2 Weight-based discrimination in the Italian Labor Market: how do ethnicity and gender interact?  
**Giovanni Bussetta, Maria Gabriella Campolo, and Demetrio Panarello** ............................ 1239

### 5.17.3 The Total Factor Productivity Index as a Ratio of Price Indexes  
**Lisa Crosato and Biancamaria Zavanella** ............................................................................. 1245

### 5.17.4 Monetary poverty indicators at local level: evaluating the impact of different poverty thresholds  
**Luigi Biggeri, Caterina Giusti and Stefano Marchetti** ....................................................... 1251

### 5.17.5 A gender inequality assessment by means of the Gini index decomposition  
**Michele Costa** ....................................................................................................................... 1257

## 5.18 Socio-Economic Statistics

### 5.18.1 The NEETs during the economic crisis in Italy, Young NEETs in Italy, Spain and Greece during the economic crisis  
**Giovanni De Luca, Paolo Mazzocchi, Claudio Quintano, Antonella Rocca** .......................... 1263

### 5.18.2 Camel or dromedary? A study of the equilibrium distribution of income in the EU countries  
**Crosato L., Ferretti C., Ganugi P.** ......................................................................................... 1270

### 5.18.3 Small Area Estimation of Inequality Measures  
**Maria Rosaria Ferrante and Silvia Pacei** .............................................................................. 1276

### 5.18.4 Testing the Learning-by-Exporting at Micro-Level in light of influence of “Statistical Issues” and Macroeconomic Factors  
**Maria Rosaria Ferrante and Marzia Freo** ............................................................................ 1281

### 5.18.5 The mobility and the job success of the Sicilian graduates  
**Ornella Giambalvo and Antonella Plaia and Sara Binassi** .................................................. 1287

## 5.19 Statistical Analysis of Energy Markets

### 5.19.1 Forecasting Value-at-Risk for Model Risk Analysis in Energy Markets  
**Angelica Gianfreda and Giacomo Scandolo** ......................................................................... 1293

### 5.19.2 Prediction interval of electricity prices by robust nonlinear models  
**Lisa Crosato, Luigi Grossi and Fany Nan** ............................................................................. 1300

### 5.19.3 Bias Reduction in a Matching Estimation of Treatment Effect  
**Maria Gabriella Campolo, Antonino Di Pino and Edoardo Otranto** .................................... 1306

## 5.20 Statistical Inference and Testing Procedures

### 5.20.1 Comparison of exact and approximate simultaneous confidence regions in nonlinear regression models  
**Claudia Furlan and Cinzia Mortarino** .................................................................................. 1311

### 5.20.2 Tail analysis of a distribution by means of an inequality curve  
**E. Taufer, F. Santi, G. Espe and M. M. Dickson** ................................................................. 1318

### 5.20.3 Nonparametric penalized likelihood for density estimation  
**Federico Ferraccioli, Laura M. Sangalli and Livio Finos** .................................................... 1324

### 5.20.4 Rethinking the Kolmogorov-Smirnov Test of Goodness of Fit in a Compositional Way  
**G.S. Monti, G. Mateu-Figueras, M. I. Ortego, V. Pawlowsky-Glahn and J. J. Egozcue** .... 1330
5.20.5 Stochastic Dominance for Generalized Parametric Families. Tommaso Lando and Lucio Bertoli-Barsotti .................................................. 1336

5.21 Statistical Models for Ordinal Data ......................................................... 1341
5.21.1 A comparative study of benchmarking procedures for interrater and intrarater agreement studies. Amalia Vanacore and Maria Sole Pellegrino ............... 1341
5.21.2 Measuring the multiple facets of tolerance using survey data. Caterina Liberati and Riccarda Longaretti and Alessandra Michelangeli ......................... 1348
5.21.3 Modified profile likelihood in models for clustered data with missing values. Claudia Di Caterina and Nicola Sartori ......................................................... 1352
5.21.4 Worthiness Based Social Scaling. Giulio D’Epifanio ................................. 1358
5.21.5 Direct Individual Differences Scaling for Evaluation of Research Quality. Gallo M., Trendafilov N., and Simonacci V. ......................................................... 1363
5.21.6 A test for variable importance. Rosaria Simone ........................................ 1367

5.22 Statistical Models New Proposals .......................................................... 1373
5.22.1 Decomposing Large Networks: An Approach Based on the MCA based Community Detection. Carlo Drago ................................................................. 1373
5.22.2 On Bayesian high-dimensional regression with binary predictors: a simulation study. Debra Slanzi, Valentina Mameli and Irene Poli ............................................ 1380
5.22.3 On the estimation of epidemiological parameters from serological survey data using Bayesian mixture modelling. Emanuele Del Fava, Piero Manfredi, and Ziv Shkedy ................................................................. 1386
5.22.4 An evaluation of KL-optimum designs to discriminate between rival copula models. Laura Deldossi, Silvia Angela Osmetti, Chiara Tommassi ......................... 1392
5.22.5 Variational Approximations for Frequentist and Bayesian Inference. Luca Maestrini and Matt P Wand ................................................................. 1398
5.22.6 Node-specific effects in latent space modelling of multidimensional networks. Silvia D’Angelo and Marco Alfó and Thomas Brendan Murphy......................... 1404

5.23 Statistics for Consumer Research .......................................................... 1410
5.23.1 A panel data analysis of Italian hotels. Antonio Giusti, Laura Grassini, Alessandro Viviani ................................................................. 1410
5.23.2 A Bayesian Mixed Multinomial Logit Model for Partially Microsimulated Data on Labor Supply. Cinzia Carota and Consuelo R. Nava ......................... 1417
5.23.3 Comparison between Experience-based Food Insecurity scales. Federica Onori, Sara Viviani and Pierpaolo Bruttì ................................................................. 1423
5.23.4 Sovereign co-risk measures in the Euro Area. Giuseppe Arbia, Riccardo Bramonte, Silvia Facchinetti, Diego Zappa ................................................................. 1429
5.23.5 Simultaneous unsupervised and supervised classification modelling for clustering, model selection and dimensionality reduction. Mario Fordellone and Maurizio Vichi ................................................................. 1435
5.23.6 Consumers’ preference for coffee consumption: a choice experiment including organoleptic characteristics and chemical analysis Rossella Berni, Nedka D. Nikiforova and Patrizia Pinelli ................................................................. 1442

5.24 Statistics for Earthquakes ................................................................. 1449
5.24.1 How robust is the skill score of probabilistic earthquake forecasts? Alessia Caponera and Maximilian J. Werner ................................................................. 1449
5.24.2 Functional linear models for the analysis of similarity of waveforms. Francesca Di Salvo, Renata Rotondi and Giovanni Lanzano ................................................................. 1456
5.24.3 Detection of damage in civil engineering structure by PCA on environmental vibration data. G. Agró, V. Carlisi, R. Mantione ................................................................. 1462
### 5.25 Statistics for Financial Risks

| 5.25.1 | Conditional Value-at-Risk: a comparison between quantile regression and copula functions. Giovanni De Luca and Giorgia Rivieccio | 1468 |
| 5.25.2 | Systemic events and diffusion of jumps. Giovanni Bonaccolto, Nancy Zambon and Massimiliano Caporin | 1474 |
| 5.25.3 | Traffic Lights for Systemic Risk Detection. Massimiliano Caporin, Laura Garcia-Jorcano, Juan-Angel Jiménez-Martín | 1480 |
| 5.25.4 | Bayesian Quantile Regression Trees. Mauro Bernardi and Paola Stolfi | 1487 |
| 5.25.5 | Model Selection in Weighted Stochastic Block models. Roberto Casarin, Michele Costola, Erdem Yenerdag | 1492 |

### 5.26 Tourism & Cultural Participation

| 5.26.1 | The determinants of tourism destination competitiveness in 2006-2016: a partial least square path modelling approach. Alessandro Magrini, Laura Grassini | 1496 |
| 5.26.2 | Participation in tourism of Italian residents in the years of the economic recession. Chiara Bocci, Laura Grassini, Emilia Rocco | 1503 |
| 5.26.3 | Cultural Participation in the digital Age in Europe: a multilevel cross-national analysis. Laura Bocci and Isabella Mingo | 1509 |
| 5.26.4 | Tourist flows and museum admissions in Italy: an integrated analysis. Lorenzo Cavallo, Francesca Petrei, Maria Teresa Santoro | 1516 |
| 5.26.5 | Posterior Predictive Assessment for Item Response Theory Models: A Proposal Based on the Hellinger Distance. Mariagililia Matteucci and Stefania Mignani | 1522 |

### 5.27 Well-being & Quality of Life

| 5.27.1 | Is Structural Equation Modelling Able to Predict Well-being? Daniele Toninelli and Michela Cameletti | 1528 |
| 5.27.2 | The well-being in the Italian urban areas: a local geographic variation analysis. Eugenia Nissi and Annalina Sarra | 1535 |
| 5.27.3 | Comparing Composite Indicators to measure Quality of Life: the Italian "Sole 24 Ore" case. Gianna Agró, Mariantonietta Ruggieri and Erasmo Vassallo | 1541 |
| 5.27.4 | Quality of working life in Italy: findings from Inapp survey. Paolo Emilio Cardone | 1547 |
| 5.27.5 | Well-being indices: what about Italian scenario? Silvia Facchinetti and Elena Siletti | 1554 |
| 5.27.6 | How can we compare rankings that are expected to be similar? An example based on composite well being indicators. Silvia Terzi e Luca Moroni | 1560 |

### 6 Poster Sessions

| 6.0.1 | A distribution curves comparison approach to analyze the university moving students performance. Giovanni Boscalino, Giada Adelfio, Gianluca Sottile | 1567 |
| 6.0.2 | A Partial Ordering Application in Aggregating Dimensions of Subjective Well-being. Paola Conigliaro | 1574 |
| 6.0.3 | A note on objective Bayes analysis for graphical vector autoregressive models. Lucia Paci and Guido Consonni | 1580 |
| 6.0.4 | Bayesian Population Size Estimation with A Single Sample. Pierfrancesco Alaimo Di Loro and Luca Tardella | 1586 |
| 6.0.5 | Classification of the Aneurisk65 dataset using PCA for partially observed functional data. Marco Stefanucci, Laura Sangalli and Pierpaolo Brutt | 1592 |
| 6.0.6 | Deep Learning to the Test: an Application to Traffic Data Streams. Nina Deliu and Pierpaolo Brutt | 1597 |
| 6.0.7 | Estimating the number of unseen species under heavy tails. Marco Battiston, Federico Camerlenghi, Emanuele Dolera and Stefano Favaro | 1603 |
| 6.0.8 | How to measure cybersecurity risk. Silvia Facchinetti, Paolo Giudici and Silvia Angela Osmetti | 1609 |
6.0.9 Implementation of an innovative technique to improve Sauvignon Blanc wine quality. Filippa Bono, Pietro Catanaia and Mariangela Vallone ............ 1613

6.0.10 Investigating the effect of drugs consumption on survival outcome of Heart Failure patients using joint models: a case study based on regional administrative data. Marta Spreafico, Francesca Gasperoni, Francesca leva ................. 1619

6.0.11 Mapping the relation between University access test and student's university performance. Vincenzo Giuseppe Genova, Antonella Plaia .................. 1625

6.0.12 Multivariate analysis of marine litter abundance through Bayesian space-time models. C. Calcuilli, A. Pollice, L. Sion, and P. Maiorano .................. 1631

6.0.13 Power Priors for Bayesian Analysis of Graphical Models of Conditional Independence in Three Way Contingency Tables. Katerina Mantzouni, Claudia Tarantola and Ioannis Ntzoufras .................................................. 1635

6.0.14 Random Garden: a Supervised Learning Algorithm. Ivan Luciano Danesi, Valeria Danese, Nicola Russo and Enrico Tonini .................. 1641

6.0.15 Spatiotemporal Prevision for Emergency Medical System Events in Milan. Andrea Gilardi, Riccardo Borgoni, Andrea Pagliosa, Rodolfo Bonora ............ 1647

6.0.16 Spatial segregation immigrant households in Messina. Angelo Mazza and Massimo Mucciardi ...................................................... 1653

6.0.17 Supervised Learning for Link Prediction in Social Networks. Riccardo Giubilei, Pierpaolo Brutti ..................................................... 1657

6.0.18 Women’s empowerment and child mortality: the case of Bangladesh. Chiara Puglisi, Annalisa Busetta ........................................ 1663
This book includes the papers presented at the "49th Meeting of the Italian Statistic Society". The conference has registered 445 participants, 350 reports divided into 4 plenary sessions, 20 specialised sessions, 25 sessions solicited, 27 sessions spontaneous, 2 poster sessions. The high number of participants, the high quality of the interventions, the productive spirit of the conference, the ability to respect the time table, are the main indices of the full success of this conference. The meeting hosted also, as plenary sessions, the ISTAT annual report 2018, and a round table on statistics and job markets. Methodological plenary sessions concerned with ordinal data, the dynamics of climate change and models in biomedicine. Moreover, two related events were held: Start-up Research (SUR) and Stats Under the Stars (SUS4). The SUS4 event attracted many sponsors of statistical, financial, editorial fields as well as numerous students, not only from Italy but also from abroad (Groningen, Tyumen, Barcelona, and Valencia): 98 students for a total of 25 teams. The SUR was a 2-day meeting where small research groups of young scholars, advised by senior researchers with a well-established experience in different areas of Statistics, was asked to develop innovative methods and models to analyse a common dataset from the Neurosciences.
Spatial analysis of the Italian seismic network
and seismicity

Analisi spaziale della rete sismica italiana e della sismicità

Antonino D’Alessandro, Marianna Siino, Luca Greco and Giada Adelfio

Abstract Seismic networks are powerful tools for understanding active tectonic processes in a monitored region. Their numerous applications, from monitoring seismicity to characterizing seismogenic volumes and generated seismicity, make seismic networks essential tools for assessing seismic hazard in active regions. The ability to locate earthquakes hypocenters requires a seismic network with a sufficient number of optimally distributed, stations. It is important to assess existing network geometry, to identify seismogenic volumes that are not adequately monitored, and to quantify measures that will allow network improvement. In this work we have studied the spatial arrangement of the stations of the Italian National Seismic Network by means of several Point Pattern techniques. The results of the point pattern analysis were compared with the spatial distribution of the historical and instrument seismicity and with the distribution of the well-known seismogenic sources of the Italian peninsula. Some considerations have also been made on some models of seismic hazard of the Italian territory. Our analysis allowed us to identify some critical areas that could require an optimization of the monitoring network.

Abstract Le reti sismiche permettono di misurare e comprendere i processi tectonici attivi in una regione monitorata. Le informazioni acquisite mediante le reti vengono utilizzate per il monitoraggio della sismicità, per la caratterizzazione dei volumi sismogenetici e della sismicità generata, di fatto sono uno strumento essenziale per la valutazione del rischio sismico. Per localizzare gli ipocentri bisogna disporre di una rete sismica con un numero sufficiente di stazioni distribuite in modo ottimale. È importante valutare la geometria della rete esistente, identificare i volumi sismogenetici che non sono adeguatamente monitorati e quantificare le misure che consentiranno il miglioramento della rete. In questo lavoro abbiamo studiato la
disposizione spaziale delle stazioni della Rete Sismica Nazionale Italiana attraverso diverse tecniche per lo studio dei processi puntuali. I risultati dell’analisi spaziali delle stazioni sono stati confrontati con la distribuzione spaziale della sismicità storica e strumentale e con la distribuzione delle ben note sorgenti sismiche della penisola italiana. Alcune considerazioni sono state fatte anche su alcuni modelli di pericolosità sismica del territorio italiano. La nostra analisi ci ha permesso di identificare alcune aree critiche che potrebbero richiedere un’ottimizzazione della rete di monitoraggio.

Key words: earthquakes; point process; seismic network; spatial correlation

1 Introduction

The Italian seismic network was developed immediately after the Irpinia seismic crisis in 1980. After this disastrous event, the National Seismic Network (NSN) was established. Initially, the NSN consisted of only few stations spread over Italy. In the 90s, the ING (Istituto Nazionale di Geofisica), then became INGV (Vulcanologia), upgraded progressively the monitoring network leading, in about thirty years, to the current NSN consisting of about 500 seismic stations. Over the years, the spatial and temporal network development continued, depending on the funds availability for the purchase of the instrumentation and the implementation of monitoring nodes. As it is well known, the quality of the estimate of focal parameters depends on the density and geometry of the monitoring stations. Therefore, the development of a seismic network should be carried out according to precise criteria, designed to guarantee a rational development of the monitoring infrastructure. The most correct criterion is to adapt the seismic network, that is to increase the density of monitoring stations, in the areas with the greatest number of seismogenic structures, historical strong earthquakes and with the greater release of seismic energy observed as instrumental seismicity. Unfortunately, the development of the NSN did not follow a precise criterion, uniformly applied to the whole territory. The result is that the quality of the location and the magnitude of completeness is very inhomogeneous and sometimes does not seem to be rational or proportional to the seismic rate (Schorlemmer et al, 2010; D’Alessandro et al, 2011). This clearly can have repercussions on the quality of the seismic monitoring and studies and it is therefore necessary to identify a simple and effective criterion that could be used for the future NSN optimization. In the following, we propose a statistical approach based on the spatial distribution of the NSN, instrumental and historical seismicity, to evaluate the current degree of the network coverage and plan for future optimization.
2 Data description

We integrate seismological, geological and seismic network data in order to address the goals of the analysis, and in this paragraph they are briefly described. We restrict the study window \((W)\) to the Italian peninsula and Sicilian land boundary because only few stations are placed in offshore areas and Sardinia island is not characterised by a high seismicity. As a matter of fact, at this point of the analysis, we want to relate the spatial distribution of the stations with the seismicity and geological information available on the mainland.

More than 500 stations of the National Seismic Network managed by the National Institute of Geophysics and Volcanology and other networks managed by other bodies are used to locate earthquakes, sending data to the central branch in Rome in real time to monitor the seismicity in Italy. The seismic network is composed by 363 stations (Figure 1a).

The Italian catalogue contains events since 1985 and we considered the earthquake since 2005, when the network was upgraded and earthquake location was sensibly improved. A subset of the catalogue consisting by 2936 events is analysed, selecting the earthquakes in \(W\) with a threshold magnitude equal to 3 and a focal depth less than 50 km (Figure 1b). Furthermore, the historical seismicity is selected starting from the Parametric Catalogue of the Italian Earthquakes, (Rovida et al, 2016) containing 4584 events in the time window 1000-2014; in particular, we select a subset of the events with location, magnitude and depth corresponding to the ones of the catalogue data (Figure 1c).

Additionally, we consider geological information to understand dependence between stations and the sources of earthquakes. The dataset of the Composite Seismogenic Sources (CSS) (Group et al, 2010) is plotted in Figure 2. A composite source represents a complex fault system with an unspecified number of aligned individual seismogenic sources that cannot be separated spatially. In particular for the analysis, we consider the upper edges of the composite sources that for the sake of simplicity are named faults.

The seismic stations, the catalogue events (instrumental seismicity) and the historical seismicity are treated as three spatial point patterns in the region \(W\). Then, we use R (R Development Core Team, 2005) packages for the statistical analysis of spatial patterns of points in two-dimensional space, spatstat (Baddeley and Turner, 2005) and spatstat.local (Baddeley, 2018).

3 Main results

More formally, a spatial point pattern \(\mathbf{y} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}\) is an unordered set of points in the region \(W \subset \mathbb{R}^d\) where \(n(\mathbf{y}) = n\) is the number of points, \(|W| < \infty\) and \(d = 2\). The first-order intensity is assumed inhomogeneous \((\lambda(\mathbf{u}))\) and for our point patterns is estimated non-parametrically to understand the spatial trend (Figure 1). The usual kernel estimator of the intensity function is (Baddeley et al, 2015)
Fig. 1: Kernel estimate of intensity, with smoothing bandwidth selected by Scott’s rule for: (a) the seismic monitoring stations, (b) earthquakes occurred between 2005 and 2018 and (c) the historical seismicity. In (a) and (b), the selected events have a magnitude greater than or equal to 3 and a focal depth less than 50 km. Black points are the events.

Fig. 2: Faults, upper edges of the composite sources

\[ \hat{\lambda}(u) = 1/e(u) \sum_{j=1}^{n} k(u - u_j, h) \]

where \( e(u) \) is the edge corrections and \( k(\cdot) \) is a Gaussian density with standard deviation (smoothing bandwidth) equal to \( h \).

From Figure 1a, there is a high concentration of stations in the centre of Italy and in the neighbourhood of the Vesuvio and Etna volcanoes. As for the spatial distribution of instrumental seismicity (Figure 1b), the areas with the higher number of events are in centre of Italy (referring to the Aquila sequence in 2009 and Amatrice-
Norcia-Vissio sequence between 2016 and 2018) and in the north centre of Italy for the Emilia sequence in 2012. The historical seismicity (Figure 1c) indicates that most of the seismic activity in Italy is along the alpine and Apennines areas and in Sicily.

In this paper, we want to study the relationship between the spatial distribution of the stations with respect to the two types of seismicity (instrumental and historical) under two aspects, their spatial dependence and the global and local characterization. With respect to the first aspect, we compute the inhomogeneous version of the bivariate K-function, to assess if the pair of point patterns are spatially dependent. As it concerns the second aspect, the global and local correlation coefficients are computed to compare the estimated intensities in Figure 1.

Generally for any pair of types i and j, the multitype K-function $K_{ij}(r)$ is the expected number of points of type j lying within a distance r of a typical point of type, i, dividing by the intensity of points of type j. It takes the form

$$K_{ij}(r) = \frac{1}{\lambda_j} E\left[t(u, r, X^{(j)})|u \in X^{(i)}\right]$$

where $X^{(j)}$ is the sub-process of points of type j, with intensity $\lambda_j$, and $t(\cdot)$ is the the number of points in the point pattern $X^{(j)}$ that lie within a distance r of the location $u$, but not at $u$ itself. If the process of type i points are independent of the process of type j points, then $K_{ij}(r)$ is equal to $\phi r^2$. Deviations between the empirical value of the curve and the theoretical one suggest dependence between the points of types i and j.

For instance, Figure 3a shows the K-cross where the point pattern of stations is type i and the point patterns of earthquake events from 2005 to 2018 is type j. Clearly the two point patterns are dependent and since the observed cross K-function is below the theoretical value, type j events are farther to type i events than it would be expected under complete spatial randomness. Similarly conclusions can be drawn when the K-cross is computed between the point pattern of the stations (type i) and the point pattern of historical seismicity (type j), see Figure 3a. However, for distances up to 0.5 degree, the two distributions (the station and the historical seismic events) seem to be independent.

As expected, the overall correlation between the intensity of stations with respect to the instrumental seismicity and historical seismicity is positive, the Pearson correlation coefficients are 0.423 and 0.643, respectively.

Moreover, considering the estimated intensities as raster data, we want to check if there are variations of correlation in space. Around each cell of the rasters, we define a focal squared area 5x5 and the correlation between the 25 values of each raster in this square is recorded for the central cell, see Figure 4. Comparing the two plots, it seems that there are more areas with negative correlation between the stations and instrumental seismicity (Figure 4a) than the stations and the historical seismicity (Figure 4b). It would suggest the necessity of a development of the current network, not completely concordant with the spatial seismicity evolution.

Finally, the main interest lies in deciding whether the spatial arrangement of the stations occurs more frequently near faults. The geological information in Figure 2
Fig. 3: (a) The black line is the inhomogeneous bivariate K-function for the seismic station point pattern (type i) and the earthquake events from 2005 and 2018 (type j). (b) The black line is the bivariate inhomogeneous K-function for the seismic station point pattern (type i) and the historical seismicity (type j). In both the plots, the red line corresponds to the theoretical value $\pi r^2$.

Fig. 4: (a) Local correlation coefficient for the raster objects, between the stations and the instrumental seismicity. (b) Local correlation coefficient for the raster objects, between the stations and the historical seismicity. In both cases, the focal squared area has a dimension 5x5.
is transformed into a spatial variables defined at all locations \( u \in W \), namely \( D(u) \) distance to the nearest fault. Assuming an inhomogeneous Poisson model with a parametric log-linear form with respect to the covariate \( D \), we estimate the following intensity

\[
\lambda(u) = \exp(\beta_0 + \beta_1 D(u))
\]

where the estimates are \( \hat{\beta}_0 = 2.84 \) and \( \hat{\beta}_1 = -1.69 \).

Moreover, we obtain with local inference (Baddeley, 2017) spatially-varying estimates of the parameters of the previous inhomogeneous Poisson process model,

\[
\lambda(u) = \exp(\beta_0(u) + \beta_1(u)D(u))
\]

This approach has the potential to detect and model gradual spatial variation of the parameters that govern the intensity of the stations and the estimates are in Figure 5.

Generally increasing the distance to the nearest fault the station intensity decreases, however according to the spatially varying slope coefficient this reduction is higher in the centre and north-east of Italy.

![Fig. 5: Spatially varying estimates of intercept (Left) and slope coefficient (Right) from the local likelihood fit of the log-linear model to the seismic station data where the covariate \( D(u) \) is distance to the nearest fault.](image)

4 Conclusions

In this paper we use statistical methods and tools for the description and characterization of the current degree of the network coverage based on the spatial distribution of the NSN and of the instrumental and historical seismicity, in order to suggest directions for planning future optimization.

As observed from Figure 4, a further upgrade of the network allocation is necessary, in order to get homogeneous and positive correlation in all the Italian area.
In particular, for instance, along the Apennines as well as the West Sicily area, the different correlation between the NSN and instrumental and historical seismicity respectively, suggests the necessity of a network strengthening in those areas.

References