



Business

Industry

ISBIS: International Society for Business and Industrial Statistics

An Association of the International Statistical Institute



ISBIS CONFERENCE 2022 on “Statistics and Data Science in Business and Industry”



PROGRAMME AND ABSTRACTS



UNIVERSITÀ DEGLI STUDI DI NAPOLI
FEDERICO II

University of Naples Federico II
20-21 June 2022, Naples (Italy)



ISBIS CONFERENCE 2022 PROGRAMME AND ABSTRACTS

Programme and abstracts of the ISBIS-22 Conference. 20-21 June 2022

ISBIS (<http://isbis2022.dii.unina.it/>)

All rights reserved

Edited by: Antonio Balzanella, Matilde Bini, Christian Capezza, Fabio Centofanti, Michele Gallo, Daniel Jeske, Antonio Lepore, Lucio Palazzo, Biagio Palumbo, Rosaria Romano, Roberto Rondinelli, Violetta Simonacci, Rosaria Simone, Maria Sole Pellegrino, Gianluca Sposito, Amalia Vanacore.
ISBN: 979-12-210-1389-4

Monday June 20, 2022

8:30-9:00: Registration

9:00-9:45: Opening and Welcome

Room: Plenary

9:45-11:00: Bayesian methods for Business and Industrial Statistics (Invited Session)

Session chair and organizer: Refik Soyer

Room: Plenary

Dynamic Stochastic MIDAS Copula Model	14
Hoang Nguyen, Audrone Virbickaite	
Deep Learning Partial Least Squares	15
Nicholas Polson, Vadim Sokolov	
Bayesian Modeling of Multivariate Time Series of Counts: Recent Advances	16
Refik Soyer	

9:45-11:00: Machine Learning and Statistics in Science Technology (Invited Session)

Room: G3

Irradiance Modelling for Vehicle-integrated Photovoltaics Using Gradient Boosting	17
Evgenii Sovetkin, Neel Patel, Andreas Gerber, Bart Pieters	

9:45-11:00: Methods and Models in Health Studies

Session chair: Antonio Lepore

Room: G4

A First Investigation of a Single Shock Impact on Italian Mortality Rates through Short-term Mortality Fluctuations Data	18
Maria Francesca Carfora, Albina Orlando	
Predictive Monitoring Using Machine Learning Algorithms and a Real-life Example on Schizophrenia	19
Leo C.E. Huberts	
Data Science: Testing Hypotheses in Large Complex Data in Cancer Research	20
Sunil Mathur	

11:00-11:45: Keynote Address - Mia Hubert

Room: Plenary

MacroPCA: An All-in-one PCA Method Allowing for Missing Values as well as Cellwise and Rowwise Outliers	21
---	----

11:45-12:15: Coffee Break

12:15-13:30: Statistical Applications in Industry, Finance, and Biostatistics (Invited Session)

Session chair and organizer: Daniel Jeske

Room: Plenary

Functional Clustering Techniques Applied to Resistance Spot Welding Curves in the Automotive Industry 4.0	22
Christian Capezza, Fabio Centofanti, <u>Antonio Lepore</u> , Biagio Palumbo	
Corporate Financing in the Deleveraging Era	23
<u>Alessandro Zeli</u> , Matilde Bini	
Design and Inference in a RCT when Treatment Observations Follow a Two-Component Mixture Model	24
Bradley Lubich, <u>Daniel Jeske</u> , Weixin Yao	

12:15-13:30: Economic and Financial Data

Session chair: Rosaria Romano

Room: G3

A Formal Framework for Statistical Modelling of Big Data Business Value Assessment	25
<u>Mario Angelelli</u> , Serena Arima, Massimiliano Gervasi	
Measuring the Recovery Performance of a Portfolio of NPLs	26
Roberto Rocci, <u>Alessandra Carleo</u> , Maria Sole Staffa	
COVID-19 Effects on the Canadian Term Structure of Interest Rates	27
Federico Severino, <u>Marzia A. Cremona</u> , Éric Dadié	

12:15-13:30: Data Science for Smart Statistics on Businesses (Invited Session)

Session chair and organizer: Mauro Bruno

Room: G4

Cosmopolitics: Use of Google Mobility Data for International Trade Analysis	28
<u>Federico Brogi</u> , Mauro Bruno, Maria Serena Causo , Barbara Guardabascio	
Online Job Advertisement: the WIN experience	29
<u>Francesco Amato</u> , Elena Catanese, Francesca Inglese, Annalisa Lucarelli, Giuseppina Ruocco	
Web Intelligence for Smart Specialization Detection in Enterprises	30
<u>Donato Summa</u> , Fabrizio De Fausti, Sergio Salamone, Mauro Bruno	

12:15-13:30: Nonparametric Approach

Session chair: Roberta Varriale

Room: G1

Bootstrapping Bivariate INAR Models with Applications	31
Lucio Palazzo, <u>Riccardo Ievoli</u>	
Clustering Consumer Based on Text Sentiment Analysis and Fuzzy Rating	32
Rosa Arboretti, Elena Barzizza, <u>Nicolò Biasetton</u> , Riccardo Ceccato, Marta Disegna, Luca Pegoraro, Luigi Salmaso	
Multivariate Nonparametric Test for Proportions with Application to Circular Economy	33
Stefano Bonnini, <u>Michela Borghesi</u>	

13:30-14:45: Lunch

14:45-16:25: Digitization of Processes, Metrology and Maintenance in Advanced Manufacturing: Focused Madein4 Project Achievements (Invited Session)

Session chair and organizer: Antonino La Magna

Room: Plenary

A Fusion of Electronic Design, Process and Metrology in Semiconductor Manufacturing for Improved Process Optimization and Control	34
<u>Anastasiia Doynychko</u> , Andres Torres, Ivan Kissiov, Melody Tao, Sanghyun Choi	
Predictive Maintenance of Backlash Level for Industrial Robotic Manipulators	35
<u>Alfio Minissale</u> , Eliana Giovannitti	
Transfer Learning Applications in the Context of Virtual Metrology in Semiconductor Manufacturing	36
<u>Rebecca Clain</u> , Valeria Borodin, Agnès Roussy	
Predictive Virtual Processes, Metrology and Maintenance in Power Device Manufacturing	37
<u>Daniele Pagano</u> , Antonino La Magna, Umberto Amato, Patrizia Vasquez, Anastasiia Doynychko, Andrea Calimera, Alex Rosenbaum, Marius Enachescu, Ioannis Deretzis	

14:45-16:25: AI and Advanced Statistics for Digital Transformation: from Product to Service (Invited Session)

Session chair and organizer: Luca Malinverno

Room: G3

Purchasing Price Assessment of Leverage Items: A Data Envelopment Analysis Approach	38
<u>Franco Visani</u> , Filippo Boccali	
Statistical Process Control in Aluminium Coil Production: an Approach Based on Classification and Regression Trees	39
Giuliano Bonollo	
Using Cloud Computing Platforms to Deliver AI Solutions in Industry	40
Luca Malinverno	
Innovation, Connected Products e Digital Servitization: What the Future Holds for Us?	41
Veronica Jagher, <u>Roberto Filipelli</u>	

14:45-16:00: Special Surveys on Enterprises During the Covid-19 Health Emergency: Results and Methods (Invited Session)

Session chair and organizer: Roberta Varriale

Room: G4

The Italian Quick Survey on the Effects of the Covid-19 Health Emergency on Enterprises: Sampling Design and Data Editing	42
Maria Cristina Casciano, Salvatore Filiberti, Mariagrazia Rinaldi, <u>Roberta Varriale</u>	
Special Survey: Situation and Perspectives of Italian Enterprises After the Covid-19 Emergency	43
Alessandro Faramondi	
Enterprises in the Last Two Years of the Health Emergency: the Reactions of the Crisis	44
Valentina Cava, <u>Daniela De Francesco</u> , Emanuela Trinca	

14:45-16:00: Multiblock Data Analysis in Food Science (Invited Session)

Session chair and organizer: Rosaria Romano

Room: G1

An Overview of Recently Proposed Multi-block Techniques for the Analysis of Food Data	45
---	----

Federico Marini

Comparison of Different Ways of Handling L-shaped Data for Integrating Sensory and Consumer Information	46
<u>Daniele Asioli</u> , Quoc Cuong Nguyen, Paula Varela, Tormod Naes	
Multi-block Quantile Regression Applied to Preference Mapping	47
Cristina Davino, Tormod Naes, <u>Rosaria Romano</u> , Domenico Vistocco	

16:25-16:55: **Coffee Break**

16:55-18:10: **Advances in Statistical Modeling and Applications (Invited Session)**

Session chair and organizer: Luca Frigau

Room: Plenary

A Boosted-oriented Fuzzy Clustering of Time Series	48
<u>Carmela Iorio</u> , Antonio D'Ambrosio	
The Algorithm of NeSSC for Semisupervised Clustering	49
<u>Luca Frigau</u> , Giulia Contu, Marco Ortu	
Textual Semi-Supervised Clustering: an Application on Booking.com Data	50
<u>Giulia Contu</u> , Maurizio Romano, Gianpaolo Zammarchi	

16:55-18:10: **Enterprises and Sectors**

Session chair: Christian Capezza

Room: G3

Websites' Data: a New Asset for Enhancing Credit Risk Modeling	51
Lisa Crosato, Josep Domènech, <u>Caterina Liberati</u>	
On the Definition of Italian District Inter-firms Network	52
<u>Ilaria Primerano</u> , Marialuisa Restaino, Giuseppe Giordano	
Multinationals in the Territories	53
<u>Giuseppe Cinquegrana</u> , Cristina Donati, Giovanni Fosco, Serena Migliardo, Margherita Lanini, Matteo Potenzieri	

16:55-18:10: **Machine Learning, AI and Questionnaire Design**

Session chair: Amalia Vanacore

Room: G4

Comparative Evaluation of Classifier Accuracy	54
<u>Amalia Vanacore</u> , Maria Sole Pellegrino, Armando Ciardiello	
Oracle-LSTM: a Neural Network Approach to Mixed Frequency Time Series Prediction	55
<u>Alessandro Bitetto</u> , Paola Cerchiello	
Questionnaire Design as a Tool to Enhance Data Quality in the Italian Census of Agriculture	56
Sabrina Barcherini, <u>Barbara Lorè</u>	

18:10-19:10: **ISBIS Assembly**

Room: Plenary

20:30-23:00: **Social Dinner**

Tuesday June 21, 2022

8:15-8:45: Registration

8:45-10:00: ASMBI Session (Invited Session)

Session chair and organizer: Fabrizio Ruggeri

Discussants: Refik Soyer, Fabrizio Ruggeri

Room: Plenary

- A Bounded Transformation of the Gamma Degradation Process 57
Gianpaolo Pulcini, Massimiliano Giorgio

8:45-10:00: Data Science in Business and Industry (Invited Session)

Session chair and organizer: Paulo Canas Rodrigues

Room: G3

- Critical Discussion of Parametric and Machine Learning Approaches for the Classification of Sentinel Satellite Data 58
Elisabetta Carfagna
- Unsupervised Multicriteria Prescription Anomaly Detection 59
Babak Zafari
- Time Series Forecasting Using Ensemble and Hybrid Methodologies 60
Paulo Canas Rodrigues

8:45-10:00: Pricing

Session chair: Matilde Bini

Room: G4

- A Statistical Model to Estimate the Airbnb Price 61
Giulia Contu, Luca Frigau, Gianpaolo Zammarchi
- The Quest for Business Value Drivers: Applying Machine Learning to Performance Management 62
Franco Visani, Anna Raffoni, Emanuele Costa

10:00-10:30: Keynote Address - Corrado Crocetta

Room: Plenary

- Data as public goods: new challenges 63
Corrado Crocetta, Laura Antonucci

10:30-11:00: Coffee Break

11:00-12:40: ENBIS Session: Statistical Models for Complex Data in Industry (Invited Session)

Session organizer: Biagio Palumbo, Antonio Lepore

Session chair: Antonio Lepore

Room: Plenary

- Boosting Diversity in Regression Ensembles 64
Jean-Michel Poggi
- Physics-based Residual Kriging: Combining Physics-based and Data-driven Approaches for Functional Oil Production Forecasting 65

Alessandra Menafoglio, Riccardo Peli, Marianna Cervino, Laura Dovera, Piercesare Secchi

Entropy: a Useful Unified Formulation	66
Narayanaswamy Balakrishnan, Francesco Buono, <u>Maria Longobardi</u>	
Probabilistic Forecasting of Electricity Demand via Additive Stacking	67
<u>Christian Capezza</u> , Biagio Palumbo, Yannig Goude, Simon N. Wood, Matteo Fasiolo	

11:00-12:40: **Young SIS Session: Advances in Statistical Modelling and Process Monitoring (Invited Session)**

Session chair and organizer: Fabio Centofanti

Room: G3

Exploiting Locality in High-dimensional Factorial Hidden Markov Models	68
Lorenzo Rimella	
Uncertainty Assessment of Multiple Redundant Sources in the Development of Sensor Fusion Methods	69
<u>Eugeniu Strelet</u> , Zhenyu Wang, You Peng, Ivan Castillo, Ricardo Rendall , Birgit Braun, Chiang Leo, Marco Reis	
Robust Multivariate Control Chart Based on Shrinkage for Individual Observations . . .	70
<u>Elisa Cabana</u> , Rosa E. Lillo	
Statistical Modeling and Monitoring of Geometrical Deviations in Complex Shapes With Application to Additive Manufacturing	71
<u>Riccardo Scimone</u> , Tommaso Taormina, Bianca Maria Colosimo, Marco Grasso, Alessandra Menafoglio, Piercesare Secchi	

11:00-12:40: **Statistics and Data Science of ISS (Invited Session)**

Session chair and organizer: Rosanna Verde

Room: G4

Scalable Variational Inference for Network Factor Models	72
Emanuele Aliverti	
Clustering Data Recorded by Distributed Acoustic Sensors for Vehicle Traffic Monitoring	73
Antonio Balzanella	
A Multilevel Blockmodeling Approach to Explore Local Based Knowledge Networks . .	74
Domenico De Stefano, <u>Elvira Pelle</u> , Ales Ziberna	
SAFE Machine Learning	75
<u>Alessandro Bitetto</u> , Paolo Giudici	

12:40-13:40: **Lunch**

13:40-14:55: **γ -BIS: Recent Statistical Advances by Young Statisticians in Italy (Invited Session)**

Session chair and organizer: Luca Frigau

Room: Plenary

An Artificial Neural Network Approach to Fault Detection in Multiple Stream Binomial Processes. A Real-case Study Involving the Monitoring of Railway HVAC Systems . . .	76
Giuseppe Giannini, Antonio Lepore, Biagio Palumbo, <u>Gianluca Sposito</u>	
Robust Analysis of Variance for Functional Data	77
<u>Fabio Centofanti</u> , Bianca Maria Colosimo, Marco Grasso, Alessandra Menafoglio, Biagio Palumbo, Simone Vantini	

Monitoring Road Infrastructures from Satellite Images: an Object-oriented Classification Approach for Road Pavement Type	78
<i>Arianna Burzacchi, Matteo Landrò, Simone Vantini</i>	

13:40-14:55: Advances in Statistical Modeling and Computing for Complex Data (Invited Session)

Session organizer: Weixin Yao

Session chair: Hanbaek Lyu

Room: G3

Supervised Dictionary Learning: Algorithms and Applications	79
<i>Joowon Lee</i>	
Introduction to Online Dictionary Learning: Learning Essential Features from Streaming Data	80
<i>Hanbaek Lyu</i>	
Robust Transformations for Multiple Regression via Additivity and Variance Stabilization	81
<i>Marco Riani, Anthony C. Atkinson, Aldo Corbellini, Andrea Cerioli, <u>Fabrizio Laurini</u></i>	

13:40-14:55: Sentiment Analysis and Categorical Data Management

Session chair: Amalia Vanacore

Room: G4

Student Satisfaction Assessments and Natural Language Processing: A Validation Study	82
<i><u>Marco Ortu</u>, Maurizio Romano, Gianpaolo Zammarchi</i>	
Permutation Tests for Model Selection	83
<i>Rosa Arboretti, Elena Barzizza, <u>Nicolò Biasetton</u>, Riccardo Ceccato, Marta Disegna , Luca Pegoraro, Luigi Salmaso</i>	
Predicting Default Risk by GEV Regression Using Massive Data with Rare Events . . .	84
<i>Michele La Rocca, Marcella Niglio, <u>Marialuisa Restaino</u></i>	

15:15-20:00: Visit to Pompeii

A Statistical Model to Estimate the Airbnb Price

Giulia Contu, Luca Frigau, Gianpaolo Zammarchi

University of Cagliari

Airbnb became in time one of the most important online platforms where it is possible to book accommodation for vacation. Different aspects influence the guest when he chooses the accommodation. One of the most important is price. It is a strategic factor influenced by different aspects such as the specific characteristics of their accommodation, the provided services, and the location.

The identification of the determinants of price is so relevant that numerous articles have previously investigated this aspect.

We introduce in the literature proposing the estimation of two different two price indicators for Airbnb accommodation. They can evaluate both the impact that different dimensions can have on Airbnb accommodation and the relative importance of the dimensions concerning neighborhoods. They are built in three different phases. In the first, the probability of accommodations belonging to a specific class of price is estimated using the Proportional Odds Model. In the second phase, three different indicators are estimated to evaluate the ability of the model to predict correctly They are the complementary of Gini's Index normalized, the normalized, and the index. In the third phase, the results of the three indicators are combined to estimate the two indicators q and r .

We focus the analysis on 61 neighborhoods of Rome. The first findings show differences with respect to the impact level of the dimension on the price for each neighborhood of Roma.