CLADAG - VOC 2025

15th Scientific Meeting of the Classification and Data Analysis Group
1st International Scientific Joint Meeting of the Italian and
Dutch/Flemish Classification Societies

BOOK OF ABSTRACTS

Naples, 8–10 September 2025

Edited by

The Local Organizing Committee

9 788899 594244

Zaccaria Editore, Napoli

Scientific Committee

Michele La Rocca (Chair) (University of Salerno, Italy)
Mark de Rooij (Chair) (Leiden University, The Netherlands)

Andreas Alfons (Erasmus University Rotterdam, The Netherlands)
Zsuzsa Bakk (Leiden University, The Netherlands)
Ruggero Bellio (University of Udine, Italy)
Paula Brito (University of Porto, Portugal)
Carlo Cavicchia (Erasmus University Rotterdam, The Netherlands)
Petros Dellaportas (University College London, UK)
Katrijn van Deun (Tilburg University, The Netherlands)
Jeffrey Durieux (Erasmus University Rotterdam, The Netherlands)
Stefania Fensore (University of Chieti-Pescara “G. d’Annunzio”, Italy)
Luca Frigau (University of Cagliari, Italy)
Sylvia Frühwirth-Schnatter (WU – Vienna University of Economics and Business, Austria)
Sabrina Giordano (University of Calabria, Italy)
Claire Gormley (University College Dublin, Ireland)
Leonardo Grilli (University of Florence, Italy)
Jos Hageman (Wageningen University, The Netherlands)
Krzysztof Jajuga (Wroclaw University of Economics and Business, Poland)
Agustin Mayo Iscar (University of Valladolid, Spain)
Mattieu Marbac (ENSAI/CREST Bruz, France)
Volodymyr Melnikov (University of Alabama, USA)
Marta Nai Ruscone (University of Genoa, Italy)
Roberto Rocci (Sapienza University of Rome, Italy)
Kim de Roover (KU Leuven, Belgium)
Laura Sangalli (Polytechnic University of Milan, Italy)
Michel van de Velden (Erasmus University Rotterdam, The Netherlands)
Tim Verdonk (KU Leuven, Belgium)
Adalbert Wilhelm (Constructor University Bremen, Germany)

Local Organizing Committee

Antonio D’Ambrosio (Chair) (University of Naples Federico II, Italy)

Enrico Cafaro (University of Naples Federico II, Italy)
Christian Capezza (University of Naples Federico II, Italy)
Rosanna Cataldo (University of Naples Federico II, Italy)
Luca Coraggio (University of Naples Federico II, Italy)
Anna Crisci (University of Naples Federico II, Italy)
Luca D’Aniello (University of Naples Federico II, Italy)
Rosa Fabbricatore (University of Naples Federico II, Italy)
Agostino Gnasso (University of Naples Federico II, Italy)
Alfonso Iodice D’Enza (University of Naples Federico II, Italy)
Carmela Iorio (University of Naples Federico II, Italy)
Lucio Palazzo (University of Naples Federico II, Italy)
Giuseppe Pandolfo (University of Naples Federico II, Italy)
Alfonso Piscitelli (University of Naples Federico II, Italy)

Massimiliano Politano (University of Naples Federico II, Italy)
Germana Scepi (University of Naples Federico II, Italy)
Rosaria Simone (University of Naples Federico II, Italy)
Maria Spano (University of Naples Federico II, Italy)
Giulia Vannucci (University of Naples Federico II, Italy)

Volunteers

Marco Cardillo (University of Naples Federico II, Italy)
Alessandra D'Alessio (University of Naples Federico II, Italy)
Alessia D'Ambrosio (University of Naples Federico II, Italy)
Giuseppe Gismondi (University of Naples Federico II, Italy)
Fabio Leone (University of Naples Federico II, Italy)
Antonella Meccariello (University of Naples Federico II, Italy)
Rebecca Rivieccio (University of Naples Federico II, Italy)
Daniele Rossi (University of Naples Federico II, Italy)
Luca Ruocco (University of Naples Federico II, Italy)
Marco Russo (University of Naples Federico II, Italy)
Carmine Santone (University of Naples Federico II, Italy)
Maria Tomas (University of Naples Federico II, Italy)
Roberta Accardo (University of Naples Federico II, Italy)

Clustering Metabarcoding Data: a Model-Based Approach

Authors:

Luisa Ferrari[1*†], Maria Franco-Villoria[1], Garritt L. Page[2]
Alex Laini[3]

[1] University of Modena and Reggio Emilia
[2] Brigham Young University
[3] Università di Torino

[*] Corresponding author [†] Presenter

Contact: luisa.ferrari5@unibo.it

Keywords:

Metabarcoding, Bernoulli Mixture Model, Informative Priors

Abstract:

Analyzing species occurrence data is extremely relevant for monitoring biodiversity and designing conservation plans. The traditional method used to collect this type of data consists of capturing a sample using traps, followed by morphological identification by an expert. Obviously, this methodology is very time-consuming, and for this reason, the datasets collected in the field of ecology are usually quite small. Metabarcoding is a relatively novel technique based on genetic identification, which makes it possible to rapidly retrieve a large amount of information about the occurrence of multiple species from a single sample. The deployment of metabarcoding represents a complete paradigm shift in the study of ecological communities. Metabarcoding-generated datasets represent the future of research in this field, and therefore, new statistical tools specifically designed to address their challenges must be developed.In particular, these datasets consist of binary entries representing the presence or absence of a large number of species across multiple sampled sites. In this work, we focus on model-based clustering methods to explore ecological patterns within metabarcoding data. This can be useful to discover groups of species with similar occurrence patterns, as well as to identify sites with similar species composition-that is, biogeographical areas. Specifically, we consider the Bernoulli Mixture Model, which has already been successfully applied in the fields of image and text mining. We propose to extend the basic model to include environmental covariates collected at the different sampling locations. This allows us to investigate the relationship between the identified clusters and habitat characteristics. Additionally, we consider an informative prior on the number of clusters using the asymmetric Dirichlet prior recently proposed in the literature. To evaluate the soundness of our approach, we first conduct a simulation study comparing the performance of the current standard model in the literature-that is, the basic Bernoulli Mixture Model-with our extended version. Finally, we apply our model to a metabarcoding dataset recording the presence or absence of dung beetle species in the Northwestern Alps of Italy, in order to uncover potential clustering patterns and support ecologists in better understanding the ecological structure of the area.