RÉPUBLIQUE FRANÇAISE
Liberté Égalité Fraternité

INRAE

DINAMIC

EXPLORATORY PROJECT

2021-2023

Analysing biological networks of mixed-type data with copula models

Coordination
Andrea Rau
UMR GABI
andrea.rau@inrae.fr

Key words
Mixed type data
Network plasticity
Copulas
Differential network analysis

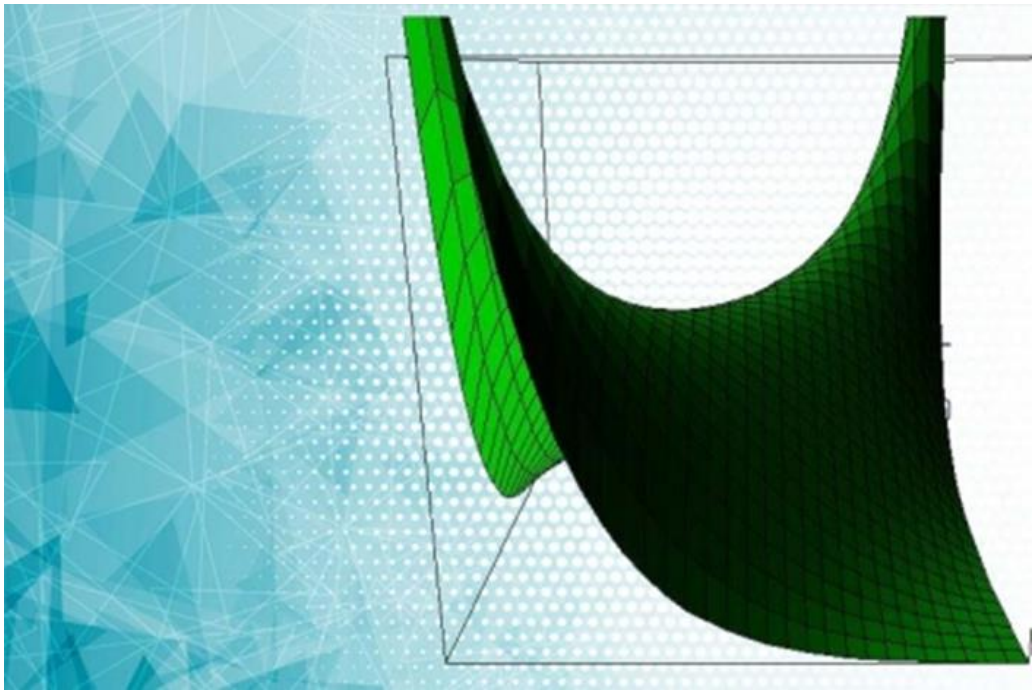
INRAE units involved
GABI
Transfrontalière BIoEcoAgro
GQE-Le Moulon
MaIAGE
BREED
NutriNeurO

Partnerships
Athens University of Economics and Business - Greece

Context and challenges

Integrative biology is based on the study of complex biological networks. Understanding the plasticity of biological interaction networks due to phenotypic, environmental or interventional variability is an important challenge in fields as diverse as genomics or human nutrition. Such studies often include comparisons between contrasting groups, including variables of various natures (continuous, counts, binary, etc.). These so-called "mixed-type" data can be difficult to analyse in a unified way. While multivariate probabilistic models provide a robust framework for inferring interrelationships among continuous variables, an analogous model for mixed-type data has yet to be defined.

A particularly promising but as-yet unexplored approach for this purpose is the use of parametric copula models, which can be used to couple variables of disparate natures. The development of such a model in a computationally efficient graphical form thus represents an open methodological challenge for the inference of generic networks from mixed-type data.



© Charthur - Wikipedia & Starline - freepik

Metaprogramme
DIGIT-BIO

digitbio@inrae.fr
www.inrae.fr/digitbio/

Goals

The DINAMIC project aims to develop and implement an innovative and widely applicable multivariate framework based on copulas and random pairwise likelihood (Mazo et al., 2021) for the differential analysis of mixed-type networks.

These methodological developments will be based on a succession of three applications covering several research themes at INRAE:

- cognitive health networks in seniors following the introduction of nutritional supplements;
- phenotypic networks in response to thermal stress in maize lines structured according to their genetic proximity;
- multi-omic networks in sperm from groups of bulls with contrasting fertility.

Each application will motivate a distinct facet of our approach, highlighting the added value of our interdisciplinary collaboration. To combine theoretically sound and computationally efficient statistical developments with relevant modelling assumptions aligned with the underlying biology, the DINAMIC project relies on a continuous cycle of interactions between methodologists and domain-specific experts.

Our multivariate mixed-type network model will represent a new approach to digital biology, with the potential to generate new insights into network plasticity in a wide variety of scientific disciplines.

Research units involved and partnerships

INRAE scientific division	INRAE research units	Expertises
Plant biology and breeding	GABI	Biostatistics
	Transfrontalière BIoEcoAgro	Quantitative genetics, plant genomics
	GQE-Le Moulon	Omic analysis
Mathematics, computer and data sciences, digital technologies	MaIAGE	Statistics, mathematics
Animal physiology and livestock systems	BREED	Animal genomics
Human nutrition and food safety	NutriNeurO	Human nutrition
External partners		**Expertises**
Athens University of Economics and Business (Greece)		Statistics and methodology

Reference

- **Mazo, G., Karlis, D., and Rau, A. (2021)** A randomized pairwise likelihood method for complex statistical inferences. Under review. ⟨hal-03126621⟩