



DINAMIC

PROJET
EXPLORATOIRE

2021-2023

Coordination

Andrea Rau

GABI

andrea.rau@inrae.fr

Mots clés

Données de type mixte

Plasticité des réseaux

Copules

Analyse différentielle de réseaux

Unités INRAE impliquées

[GABI](#)

[Transfrontalière BloEcoAgro](#)

[GQE-Le Moulon](#)

[MaIAGE](#)

[BREED](#)

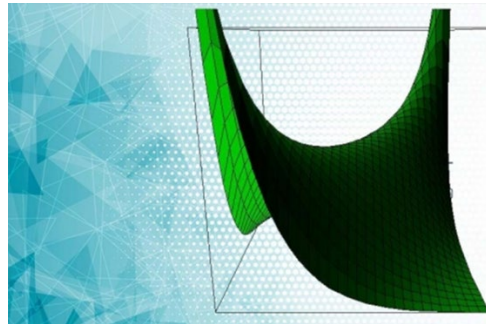
[NutriNeuro](#)

[AGAP](#)

Partenaires

[Athens University of Economics and Business](#) - Grèce

Étude de la plasticité des réseaux biologiques hétérogènes par les modèles à copules



© Chartur - Wikipedia & Starline

Contexte, enjeux et objectifs

L'étude d'un système biologique requiert la collection d'un grand nombre de données, et de tout type (mesure, comptage, absence/présence...). Cela donne donc naissance à des jeux de données hétérogènes, dits de « type mixte », dont l'analyse doit permettre l'inférence de réseaux biologiques, c'est-à-dire la reconstruction des connexions entre les unités d'un organisme (molécules, cellules, organes) en se basant sur les données expérimentales. Problème : les outils mathématiques classiquement utilisés par les biologistes et analystes ont du mal à faire des liens entre des données de nature différentes. Pourtant, l'inférence de réseaux biologiques est un enjeu dans divers domaines car elle permet non seulement de mieux comprendre l'organisme en question, mais aussi d'étudier sa plasticité, c'est-à-dire sa capacité à réagir face à des variations phénotypiques, environnementales et interventionnelles. Il y a donc un réel besoin d'outils efficaces pour l'inférence de réseaux.

L'utilisation de modèles à « copules » est particulièrement prometteuse dans ce contexte. La copule est un outil mathématique capable d'établir des relations entre des données de natures différentes. Le défi est donc le suivant : développer un modèle à copules sous forme graphique et efficace en calcul.

Le projet DINAMIC a regroupé plusieurs unités INRAE afin de développer et d'implémenter un modèle multivarié basé sur des copules pour l'appliquer à l'analyse différentielle de réseaux de type mixte. Ce modèle innovant avait la particularité de pouvoir s'appliquer à divers domaines.

Résultats

Preuve de concept : utilisation d'outils d'analyse différentielle de réseaux pour le traitement des données

Des outils d'inférence de réseaux déjà existants dans la littérature ont été utilisés pour explorer les caractéristiques de trois jeux de données collectées dans des projets INRAE précédents. Ces projets concernaient la santé humaine, la génomique du maïs et la génomique du bétail. Les analyses menées dans le cadre de DINAMIC ont permis de tirer des conclusions concrètes pour les trois domaines et donc de démontrer l'utilité de telles méthodes d'analyses. L'étude des données génomiques du bétail a notamment été utilisée pour la création d'une puce d'épigénotypage



(projet RUMIGEN). Les résultats obtenus pour les trois projets seront utilisés pour de futures comparaisons avec ceux que donneront les modèles à copules.

Des données supplémentaires ont également été collectées pour le projet de génomique du maïs ainsi que pour le projet de génomique du bétail afin d'approfondir les analyses, et donc la connaissance de ces domaines, et de préparer le terrain pour de futures validations biologiques.

Avancées méthodologiques, informatiques et théorique sur les modèles à copules

Les scientifiques portant le projet ont développé une méthode permettant de réaliser de l'inférence sur des modèles multivariés complexes. Cette méthode générique est applicable à plusieurs modèles, dont les modèles à copules. La méthode a notamment été appliquée sur des données de comptage multivariées issues d'une étude sur le transcriptome d'un parasite de l'abeille et la variation de ce transcriptome au cours des cycles de vie dudit parasite. L'approche utilisant une copule gaussienne a également été intégrée dans les packages R *rpl* et *heterocop*. Enfin, deux projets de master se sont penchés en particulier sur l'utilisation de modèles à copules gaussiennes pour le traitement de données de type mixte (application aux données prélevées sur le bétail), fournissant une preuve de concept et conduisant à la préparation de deux papiers.

Perspectives

Une communauté enthousiaste et souhaitant poursuivre les travaux

Le projet DINAMIC a regroupé sept unités INRAE réparties dans cinq départements différents et a posé les bases pour une nouvelle collaboration avec le Cirad. De solides avancées méthodologiques ont été faites mais un certain nombre de problèmes restent à résoudre pour utiliser les modèles à copules dans le traitement de jeux de données de type mixte. Ces problèmes continueront d'être explorés via plusieurs actions permises par les résultats déjà obtenus, mais aussi par l'émulation générée par le projet qui a créé une véritable communauté pluridisciplinaire enthousiaste face à l'utilisation des modèles à copules pour l'inférence de réseaux biologiques. Ces actions sont d'ordre divers :

- Le financement de la thèse d'Ekaterina Tomilina (2022-2025, financement DIGIT-BIO EDHD) sur le développement d'outils théoriques et méthodologiques pour l'utilisation de modèles à copules pour le traitement de données multi-omiques.
- Un complément de financement pour la thèse d'Amrita Raja-Rai-Shankar (2023-2026) qui exploite notamment les données de méthylation collectées par DINAMIC.
- Le projet JCJC ExonuTriAge (2024-2027, financement ANR) porté par Jean-Christophe Delpech sur l'analyse des vésicules extracellulaires¹ extraites des échantillons sanguins du projet de santé humaine, projet qui inclut l'utilisation des méthodes d'analyse développées par DINAMIC.
- Le projet ADAAPT (2024-2028, PEPR Agroécologie et Numérique) piloté par Laurence Drouilhet qui s'appuie surtout sur la puce d'épigénotypage et qui pourra bénéficier des outils et méthodes développées dans DINAMIC.

¹ Sortes de sacs naturellement produit par les cellules, contenant et transportant divers éléments contenus dans la cellules (protéine, ARN, métabolite...).

Publications

- Gildas Mazo, Dimitris Karlis, Andrea Rau. A randomized pairwise likelihood method for complex statistical inferences, *Journal of the American Statistical Association*, 2023, 119 (547), p 2317-2327 [10.1080/01621459.2023.2257367](https://doi.org/10.1080/01621459.2023.2257367) [hal-03126620](https://hal.inrae.fr/hal-03126620)
- Delpech, J.-C., Yeh, H., Kalavai, S. V., You, Y., Ruan, Z., Touch, N., Hersh, S., Monguillon, P., Johnson, W. E., Rau, A., Madore, C., Ikezu, T. and Ikezu, S. Sex specific correction of maternal inflammation-induced behavioral abnormalities by the inhibition of colony-stimulating factor 1 receptor, *Brain Behavior and Immunity*, 2025, 131:106163, <https://doi.org/10.1016/j.bbi.2025.106163>.
- Tomilina, E., Jaffrézic, F., Mazo, G. Gaussian copula correlation network analysis of mixed-type data using a semi-parametric pairwise likelihood with a multi-omics application, *Computational Statistics & Data Analysis*, 2026, 223:108414, <https://doi.org/10.1016/j.csda.2026.108414>, <https://hal.inrae.fr/hal-04847648v4/>



Voir l'ensemble des publications et communications issues du projet sur <https://digitbio.hub.inrae.fr/>

Métaprogramme
DIGIT-BIO



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