

Sujet de stage de MASTER 2 en statistiques année 2022

Isabelle Albert, INRAE-AgroParisTech-Université Paris-Saclay & Sébastien Marze, INRAE-

Graphe acyclique dirigé et inférence bayésienne pour identifier les sources alimentaires durables d'oméga-3 et de vitamine D assurant une alimentation humaine saine et riche en ces nutriments

Les acides gras polyinsaturés oméga-3 et la vitamine D font partie des nutriments essentiels à la santé humaine. Cependant, de nombreuses populations sont en état de carence et cette insuffisance n'est pas simple à résoudre car elle dépend de nombreux facteurs. Notamment, les sources de ces nutriments actuellement exploitées sont peu abondantes et peu durables, avec des impacts environnementaux qui restent à quantifier ; leur transformation en aliments et leur conservation posent également des difficultés car ces nutriments se dégradent relativement facilement ; les aliments les plus riches, en particulier poissons gras et fruits de mer, peuvent apporter aussi des contaminants (entre autres les métaux lourds) ; en aval, il est reconnu que l'apport de ces nutriments participe au maintien en bonne santé, cependant les rôles de leurs interactions avec les autres nutriments du régime alimentaire ne sont pas tous compris.

À partir des connaissances d'un groupe d'experts, constitué dans un consortium interdisciplinaire qui réalise un état des lieux des connaissances scientifiques sur les relations entre les systèmes alimentaires et la santé pour de nombreuses sources d'oméga-3 et de vitamine D, et des données quantitatives recueillies par ces experts, nous voulons produire une première modélisation statistique de l'amont (les sources alimentaires) à l'aval (la santé humaine) pour quantifier les différents critères d'intérêt (l'apport, la durabilité, l'acceptabilité et l'impact santé, entre autres) et leurs incertitudes associées pour les différentes sources des deux nutriments.

Le stagiaire sera confronté à des dires d'experts, des données quantitatives et qualitatives hétérogènes, multi-échelles et « sparses ». La construction du modèle statistique avec ses nombreuses variables latentes se fera à l'aide d'un graphe d'indépendances conditionnelles (graphe acyclique dirigé) et l'inférence du réseau ainsi constitué se fera de façon bayésienne de façon computationnelle grâce à des algorithmes de Monte Carlo par Chaîne de Markov de type échantillonnage de Gibbs ou hamiltonien.

Contact et encadrement

Contact : envoyer CV et lettre de motivation aux deux encadrants Isabelle Albert (INRAE - MIA Paris) isabelle.albert@inrae.fr et Sébastien Marze (INRAE -) sebastien.marze@inrae.fr

Collaboration : Cedric Baudrit, Pierre Bisquert, Patrice Buche, chercheurs INRAE en Informatique

Lieu de stage : Paris 5^{ème} puis Paris-Saclay (Palaiseau)

Durée : 6 mois (démarrage février/avril 2022)

Rémunération : Gratification standard des stages de M2 de la fonction publique (environ 560 euros net/mois)

Connaissances et compétences souhaitées :

1. Connaissances approfondies en statistiques, si possible en statistique bayésienne également
2. Aptitudes informatiques : programmation R, Latex et Markdown
3. Intérêt pour les applications en sciences de l'environnement, de l'alimentation et de la santé
4. Lecture de l'anglais scientifique et technique
5. Capacité rédactionnelle
6. Aptitude pour le travail d'équipe, la communication et l'autonomie

Références

[1] Gaspar Massiot, Isabelle Albert. Évaluation des risques liés aux pathogènes émis par l'irrigation de parcelles agricoles avec de l'eau usée traitée en station d'épuration à l'aide d'un réseau bayésien ; *Journées de statistiques de la SFDS*, Jun 2021, Nice, France.

[2] Clémence Rigaux, Sophie Ancelet, Frederic Carlin, Christophe Nguyen The, Isabelle Albert. Inferring an Augmented Bayesian Network to Confront a Complex Quantitative Microbial Risk Assessment Model with Durability Studies: Application to *Bacillus Cereus* on a Courgette Puree Production Chain ; *Risk Analysis*, Wiley, 2013, 33 (5), pp.877-892.