

Sélection de modèles d'équations structurelles par critères d'information : application à de faibles tailles d'échantillon en agro-écologie.

Anaïs Widmer^{1,2}, Mathieu Emily^{1,2}

¹IRMAR, Université de Rennes 1, Rennes, France

²Institut Agro Rennes Angers, Rennes, France

E-mail for correspondence: anais.widmer@agrocampus-ouest.fr

Abstract: La modélisation holistique de systèmes complexes et en particulier d'agrosystèmes doit pouvoir tenir compte du caractère multi-échelle et multi-compartiment du système ainsi que des relations causales qui pilotent son évolution. Les modèles d'équations structurelles offrent un cadre mathématique et théorique particulièrement adapté à ces caractéristiques, notamment en s'appuyant sur la définition d'un graphe orienté acyclique (DAG) et sur l'utilisation d'une structure latente permettant une hiérarchisation du système [Bollen, 1989]. L'obtention des différentes relations caractérisant un DAG se fait par la combinaison de connaissances *a priori*. Toutefois, ces connaissances peuvent être partielles et/ou controversées générant un ensemble de modèles possibles qu'il convient de pouvoir comparer en les confrontant à des observations [Shipley, 2016].

Dans cette étude, nous nous intéressons à la comparaison de modèles structuraux à partir de critères d'information [Lin, et al. , 2017]. Nous nous focaliserons sur la comparaison de modèles à structures latentes dans le cadre de données à faible taille d'échantillon. Ce cadre s'inspire de celui des études des agrosystèmes pour lesquelles un enjeu majeur est de proposer une approche de modélisation à la fois globale et parcimonieuse.

Par une étude par simulation, nous avons tout d'abord étudié la capacité de chaque critère d'information à sélectionner correctement le vrai modèle sous-jacent. Nous avons pu constater que la taille d'échantillon ainsi que la structure latente avait un impact sur le meilleur critère d'information. Nous avons également identifié que la prise en compte de la structure latente dans les critères d'information n'était pas optimale.

Nous avons ensuite appliqué la sélection de modèles à l'étude de la structure causale de la culture du blé en focalisant sur les relations d'interactions entre le sol et la plante. Cette étude a notamment permis d'identifier l'impact direct et indirect de la composition et du travail du sol sur les composantes de rendement et notamment le nombre d'épis.

Key words: Sélection de variables, Agro-écologie, Modèle d'équations structurelles, critère d'information

K. A. Bollen (1989). Structural Equations with Latent Variables. Wiley.

B. Shipley (2016). Cause and correlation in biology: A user's guide to path analysis, structural equations and causal inference. Cambridge university.

L-C Lin, P-H Huang and L-J Weng (2017). Selecting Path Models in SEM - A Comparison of Model Selection Criteria. Structural Equation Modeling: A Multidisciplinary Journal, 24:6, 855-869.