Assessing Predictive Performance in Spatial Count Data: A Comparison of Advanced Modeling Approaches

曾聖澧[1], 張家輔[2], 施嘉翰[2]

[1] 國立中興大學統計學研究所 [2]國立中山大學應用數學系

Abstract

Spatial count data, common in fields like ecology and epidemiology, often exhibit heterogeneity and spatial dependence, challenging traditional assumptions.

This study conducts a comparative analysis of spatial count data models, focusing on predictive performance under varying data characteristics and parameter settings. Seven modeling approaches are evaluated: Gaussian Copula Negative Binomial (GC-NB) models using GHK and GQT methods, spatial Poisson and negative binomial random fields, Gaussian misspecification models, and Gaussian random fields with different likelihood methods. Through simulations and real datasets, predictive accuracy is assessed using RMSE metrics, considering factors such as overdispersion, spatial correlation, and range parameters.

Results show GC-NB models consistently outperform others in accuracy and robustness, even when assumptions are violated. Other models struggle with high overdispersion or sensitivity to data structure. Real data analysis supports the effectiveness of GC-NB models in capturing spatial clustering and high-density areas, offering valuable guidance for spatial count data modeling.

Keyword:

Spatial Count Data, Predictive Performance, Gaussian Copula, Overdispersion