

Generalized spatial conditional overdispersion models: Semiparametric extensions proposals

*Mabel Morales-Otero*¹, *María Durbán*², *Vicente Núñez-Antón*³

¹mmoralesote@unav.es, Institute of Data Science and Artificial Intelligence, University of Navarra

²mdurban@est-econ.uc3m.es, Department of Statistics, University Carlos III of Madrid

³vicente.nunezanton@ehu.eus, University of the Basque Country UPV/EHU

Generalized spatial conditional overdispersion models represent an excellent choice to fit spatial count data, since they are able to account for overdispersion, capture the possible existing spatial correlation, and they are also flexible enough as to allow for the dispersion to vary according to covariates and/or spatial terms. However, when including covariates in the regression structures for these models, we are assuming that the possible existing relationship between each covariate and the predictor is linear, which may not be necessarily the case, since they could be given by another, maybe non linear pattern. In this sense, smoothing methods should be considered in generalized linear models (GLM), so that the linearity hypothesis can be relaxed. Therefore, in this work we propose a semiparametric extension of the generalized spatial conditional overdispersion models that will allow us to capture such non linear relationships. In particular, for the smoothing of such variables, we have specified P-splines in their mixed model representation. We illustrate their usefulness by fitting them to the study of infant mortality rates and mother's postnatal period screening test in Colombia, where we investigate the possible existence of such non linear relationships. In these applications, we have found evidence of a non linear relationship between the mortality rates and the variable representing the amount of resources provided by the government for academic achievement or education.

Keywords: Spatial models, Semiparametric models, P-splines.