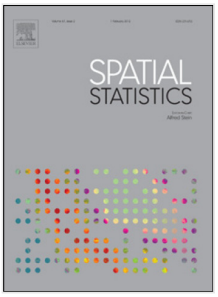
Accepted Manuscript



Penalized composite link models for aggregated spatial count data: A mixed model approach

Diego Ayma, María Durbán, Dae-Jin Lee, Paul H.C. Eilers

PII: S2211-6753(16)30031-8
DOI: http://dx.doi.org/10.1016/j.spasta.2016.06.007
Reference: SPASTA 177

To appear in: *Spatial Statistics*

Received date: 17 November 2015
Accepted date: 6 June 2016

Please cite this article as: Ayma, D., Durbán, M., Lee, D.-J., Eilers, P.H.C., Penalized composite link models for aggregated spatial count data: A mixed model approach. *Spatial Statistics* (2016), http://dx.doi.org/10.1016/j.spasta.2016.06.007

Penalized composite link models for aggregated spatial count data: a mixed model approach

Diego Ayma[a], María Durbán[a], Dae-Jin Lee[b,*], Paul H. C. Eilers[c]

[a] *Department of Statistics, Universidad Carlos III de Madrid, 28911 Leganés, Spain.*
[b] *Basque Center for Applied Mathematics, 48009 Bilbao, Spain.*
[c] *Department of Biostatistics, Erasmus University Medical Center, 3015GE Rotterdam, The Netherlands.*

Abstract

Mortality data provide valuable information for the study of the spatial distribution of mortality risk, in disciplines such as spatial epidemiology and public health. However, they are frequently available in an aggregated form over irregular geographical units, hindering the visualization of the underlying mortality risk. Also, it can be of interest to obtain mortality risk estimates on a finer spatial resolution, such that they can be linked to potential risk factors that are usually measured in a different spatial resolution. In this paper, we propose the use of the penalized composite link model and its mixed model representation. This model considers the nature of mortality rates by incorporating the population size at the finest resolution, and allows the creation of mortality maps at a finer scale, thus reducing the visual bias resulting from the spatial aggregation within original units. We also extend the model by considering individual random effects at the aggregated scale, in order to take into account the overdispersion. We illustrate our novel proposal using two datasets: female deaths by lung cancer in Indiana, USA, and male lip cancer incidence in Scotland counties. We also compare the performance of our proposal with the area-to-point Poisson kriging approach.

Keywords: Penalized composite link models, Mixed models, Mortality rates, Disease mapping

1. Introduction

Disease mapping studies commonly consider public health data that are only available in an aggregated form over irregular geographical units, like counties, districts, and municipalities. Epidemiologists, health care practitioners, and

*Corresponding author
Email addresses: dayma@est-econ.uc3m.es (Diego Ayma), mdurban@est-econ.uc3m.es (María Durbán), dlee@bcamath.org (Dae-Jin Lee), p.eilers@erasmusmc.nl (Paul H. C. Eilers)