Practicum on Point Referenced Data Analysis

Alexandra M. Schmidt and Carlo Zaccardi

03-06, June 2025

Question 1

Analysis of the Parana rainfall data. Check the separate `pdf` file `RainfallParana.pdf`.

Question 2

Fit a geostatistical model to predict fine particulate matter PM2.5 in the USA using the INLA and SPDE approaches. See file `Example-Moraga.r`.

Question 3 - *Model-based kriging with binary responses*

You have available data on presence/absence of malaria on village resident children in the Gambia. This dataset was analysed by Diggle *et al.* (2002)[1] The malaria prevalence data set consists of measurements of the presence of malarial parasites in blood samples obtained from children in 65 villages in the Gambia. Other child- and village-level indicators include age, bed net use, whether the bed net is treated, whether or not the village belonged to the primary health care structure, and a measure of 'greenness' using a vegetation index.

x: x coordinate of the village (UTM),

y: y coordinate of the village (UTM),

pos: presence (1) or absence (0) of malaria in a blood sample taken from the child,

age: age of the child in days,

netuse: indicator variable denoting whether the child regularly sleeps under a bed net,

treated: indicator variable denoting whether the bed net is treated,

green: satellite-derived measure of the greenness of vegetation in the vicinity of the village,

phc: indicator variable denoting the presence or absence of a health center in the village.

id_area: indicator of the village area

The main goal is to investigate if treated nets are associated with a higher probability of presence of malaria. Propose a spatial generalized mixed model that accounts for the structure of the data.

[1] Diggle, P., Moyeed, R., Rowlingson, B., Thomson, M. (2002). Childhood Malaria in the Gambia: A Case-Study in Model-Based Geostatistics, Journal of the Royal Statistical Society Series C: Applied Statistics, Volume 51, Issue 4, Pages 493–506.