Adaptive Geostatistical Designs (AGD)

Michael G. Chipeta*
College of Medicine, Blantyre, Malawi/University of Liverpool, UK -Michael.Chipeta@liverpool.ac.uk

Dianne J. Terlouw
Liverpool School of Tropical Medicine, UK/Malawi Liverpool Wellcome Trust, Blantyre, Malawi - Anja.Terlouw@lstmed.ac.uk

Kamija Phiri
College of Medicine, Blantyre, Malawi - kamijaphiri@gmail.com

Peter J. Diggle
University of Lancaster, Lancaster, UK/University of Liverpool, UK - p.diggle@lancaster.ac.uk

Non-adaptive geostatistical designs (NAGD) offer standard ways of collecting and analysing geostatistical data in which sampling locations x_i are fixed in advance of any data collection and cannot be changed. The two most common types are: *random* sampling, which is relatively efficient for parameter estimation, and *regular* sampling, which is generally more efficient for spatial prediction when model parameters are known. In contrast, Adaptive geostatistical designs (AGD) allow collection of exposure and outcome data over time to depend on information obtained from previously collected data so as to optimise data collection towards the analysis objective. Sampling locations are thus customized to ensure they optimize understanding of spatial variation in the area of interest. AGDs are becoming more important in spatial mapping studies, particularly in poor resource settings where uniformly precise mapping may be unrealistically costly and the priority is often to identify critical areas where interventions can have the most health impact. AGDs offer several constructions, which can be classified initially as *singleton adaptive sampling* and *batch adaptive sampling*. In singleton sampling, locations x_i are chosen sequentially and at each stage, x_{k+1} can depend on data obtained at locations $x_1, \ldots, x_k$. In batch sampling, locations are chosen in batches of size $b > 1$, allowing a new batch, $\{x_{(kb+1)}, \ldots, x_{(k+1)b}\}$, to depend on data obtained at locations $x_1, \ldots, x_{kb}$. Batch adaptive sampling cannot be more efficient theoretically than singleton adaptive sampling, but is almost always more realistic in practice. Using simulated data, we evaluated some specific batch adaptive sampling designs and assessed their efficiency relative to their singleton adaptive and non-adaptive counterparts by comparing their average prediction variance. We will present the simulation results and show how we apply these findings to inform an adaptive geostatistical sampling design of a 1-year rolling (monthly) cross-sectional Malaria Indicator Survey (rMIS) that is part of a large scale, five-year malaria transmission reduction project in Malawi. In this 1-year rMIS survey, the first objective is to describe the local variation in malaria infection to identify *hotspots* that could guide more targeted disease control efforts. The secondary objective is to describe the association of the observed malaria infection patterns with malaria control coverage, and environmental, epidemiological and other risk-factors, based on both collected data during the survey and publicly remote sensed climate and environmental information. rMIS results will inform the geostatistical sampling design of a subsequent randomised community-level malaria intervention trial.

Keywords: Sampling strategies; Disease mapping; Malaria; Spatial statistics; .