

## Reducing bias in multivariate analyses due to the modifiable areal unit problem.

Tuson, M<sup>1</sup>, Kok, MR<sup>1</sup>, Yap, M<sup>1</sup>, Vickery, A<sup>1</sup>, Boruff, B<sup>1</sup>, Murray, K<sup>1</sup>, Turlach, B<sup>1</sup>, and Whyatt, D<sup>1</sup>

<sup>1</sup>University of Western Australia

### Introduction

The Modifiable Areal Unit Problem (MAUP) arises from the aggregation of data organized by spatially defined boundaries. Aggregated values are influenced by the shape (zone effect) and scale of the aggregated units. Aggregations of the same data using different zones or scales can give different analytical results, none reliable.

### Objectives and Approach

Using population-level administrative health data in Western Australia, the objectives were to: accurately measure the association between health service utilization and demographic, socio-economic, and service accessibility variables; and develop models to accurately forecast areas of high health service utilization into the future.

Multiple zone designs and aggregation scales were used to examine the impact of MAUP in association studies. These zone designs and scales were then used in all-subset model selection processes, combined with repeated k-fold cross-validation, to generate forecast maps of areas having high future rates of health service utilization.

### Results

The impact of the MAUP and methods to reduce this bias in association studies will be presented, for both simple and complex model designs. Maps indicating gradients of predicted probabilities of high rate of health service demand in the future can be used to optimize the placement of services, through the use of catchment areas based on road-network travel distance and population distributions.

### Conclusion/Implications

The impact of the MAUP on the analysis of spatially-aggregated data has been considered intractable. However, methods to reduce the impact of the MAUP can improve policy and planning decisions based on such studies.

