**DEADLINE 13th FEB**

**Title:**

The bias of crime statistics: Assessing the impact of data bias on police analysis and crime mapping

**Authors:**

David Buil-Gil. University of Manchester

Angelo Moretti. Manchester Metropolitan University

**Email:**

[david.builgil@manchester.ac.uk](mailto:david.builgil@manchester.ac.uk)

[A.Moretti@mmu.ac.uk](mailto:A.Moretti@mmu.ac.uk)

**Abstract (500 words):**

Police-recorded crimes are the main source of information used by police forces to analyse crime patterns, investigate the spatial concentration of crime, and design spatially targeted strategies. Police statistics are used to design and evaluate crime prevention policies and to develop theories of crime. Nevertheless, crimes known to police are affected by biases and unreliability driven by unequal crime reporting rates across social groups and geographical areas. The measures of error that affect the reliability of crime statistics is an issue that merits deeper scrutiny, since it affects police everyday practices, criminal policies and citizens’ everyday lives. Yet it is an understudied issue, and the implications of data biases for crime mapping are unknown. Moreover, police analyses are moving towards the study of smaller levels of geography than ever before, such as street segments with highly homogeneous communities. Maps produced from police records are used to foreground the micro places where rates of recorded crimes are larger. This paper presents a simulation study and an application to analyse the impact of data biases on crime maps produced from police records at the different spatial scales. It assesses whether micro-level maps are affected by a larger risk of bias than maps produced at larger scales.

Based on parameters obtained from the UK Census 2011, we simulated a synthetic population consistent with the social-demographic and spatial characteristics of Manchester, England. Then, based on model parameters derived from the Crime Survey for England and Wales 2011/12, we simulated the number and type of crimes suffered by individuals across social groups and areas, and predicted the likelihood of these crimes to be known to police. It allowed us to compare the relative difference between all crimes and police-recorded incidents at the different scales: (a) 1,530 Output Areas with an average of 328.8 residents, (b) 282 LSOAs, (c) 57 MSOAs, and (d) 32 wards. While the average relative difference between all crimes and those known to police is close to 62% for all geographical scales, the measures of dispersion of the relative difference between all crimes and police records are much larger when crime incidents are aggregated at the levels of small geographies. In other words, when producing maps at the scales of medium-level geographies, the percentage of crimes known to police is similar in all areas, and thus the risk that police statistics underestimate or overestimate crime rates in some areas more than others is small; whereas the percentage of unknown crimes varies widely across micro places. This has important implications for policing, policy making and research: police strategies, criminal policies and crime theories drawn from police records aggregated at the scales of small communities are likely to be affected by large biases that underestimate the prevalence of crime in certain places while overestimating its prevalence in others. We also provide an application that shows the large impact that data biases have on micro-level maps produced from crimes registered by Greater Manchester Police in 2011.

**Do we want to be considered for the special issue at the *Journal of the Royal Statistical Society: Series A*?**

YES