

sampbias, a method for quantifying geographic sampling biases in  
species distribution data

Appendix S1 - Supplementary Figure

Alexander Zizka, Alexandre Antonelli, Daniele Silvestro

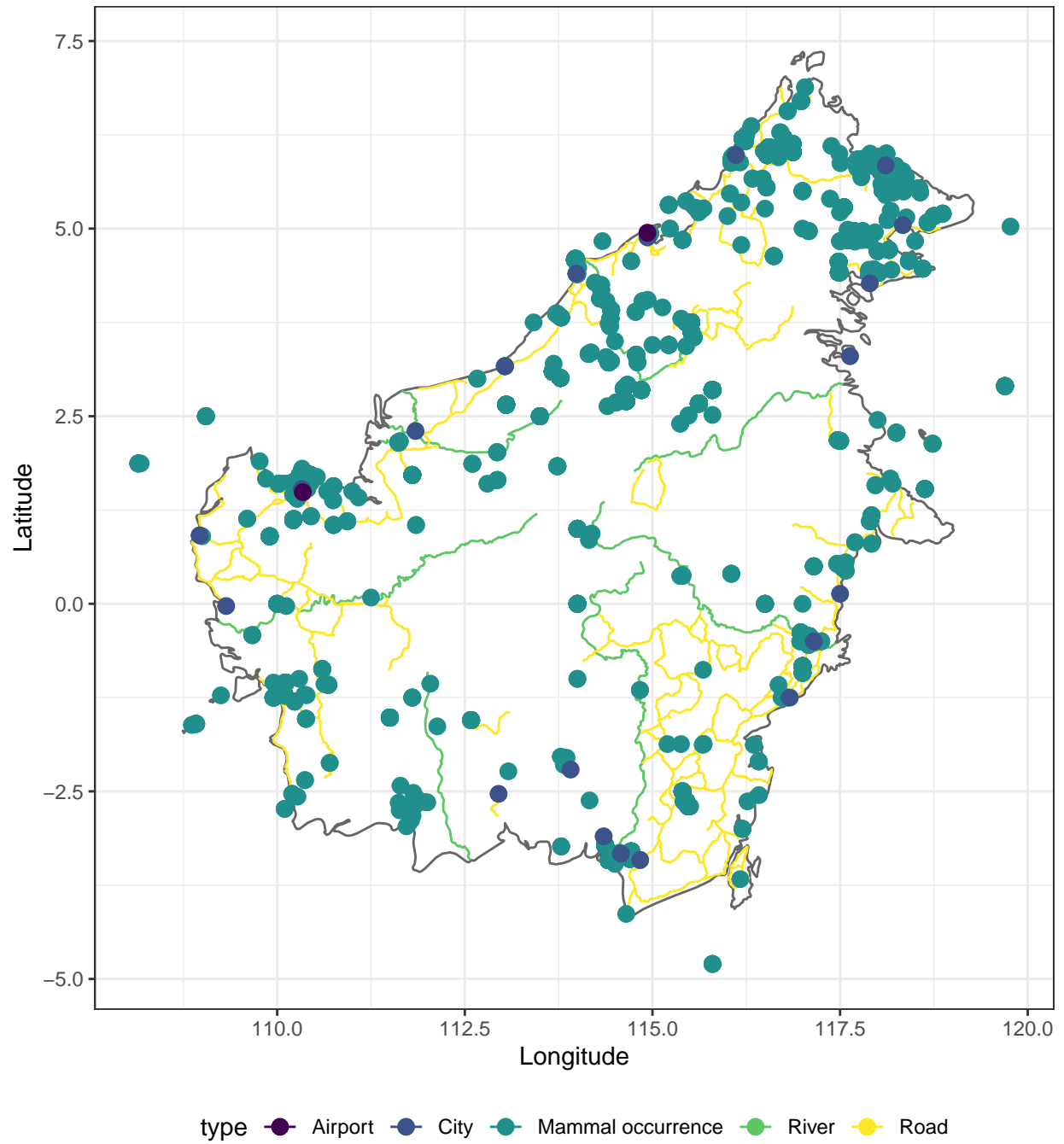


Figure S1: The example dataset of mammal occurrences from the island of Borneo, as downloaded from [www.gbif.org](https://www.gbif.org) ( $n = 6,262$ , 2016, <https://doi.org/10.15468/dl.7fg4zx>), and the geographic gazettters of main cities, roads, rivers and airports used for the sampbias analysis. Note that the records in the sea around Borneo were excluded by with the help of the user-provided study area.

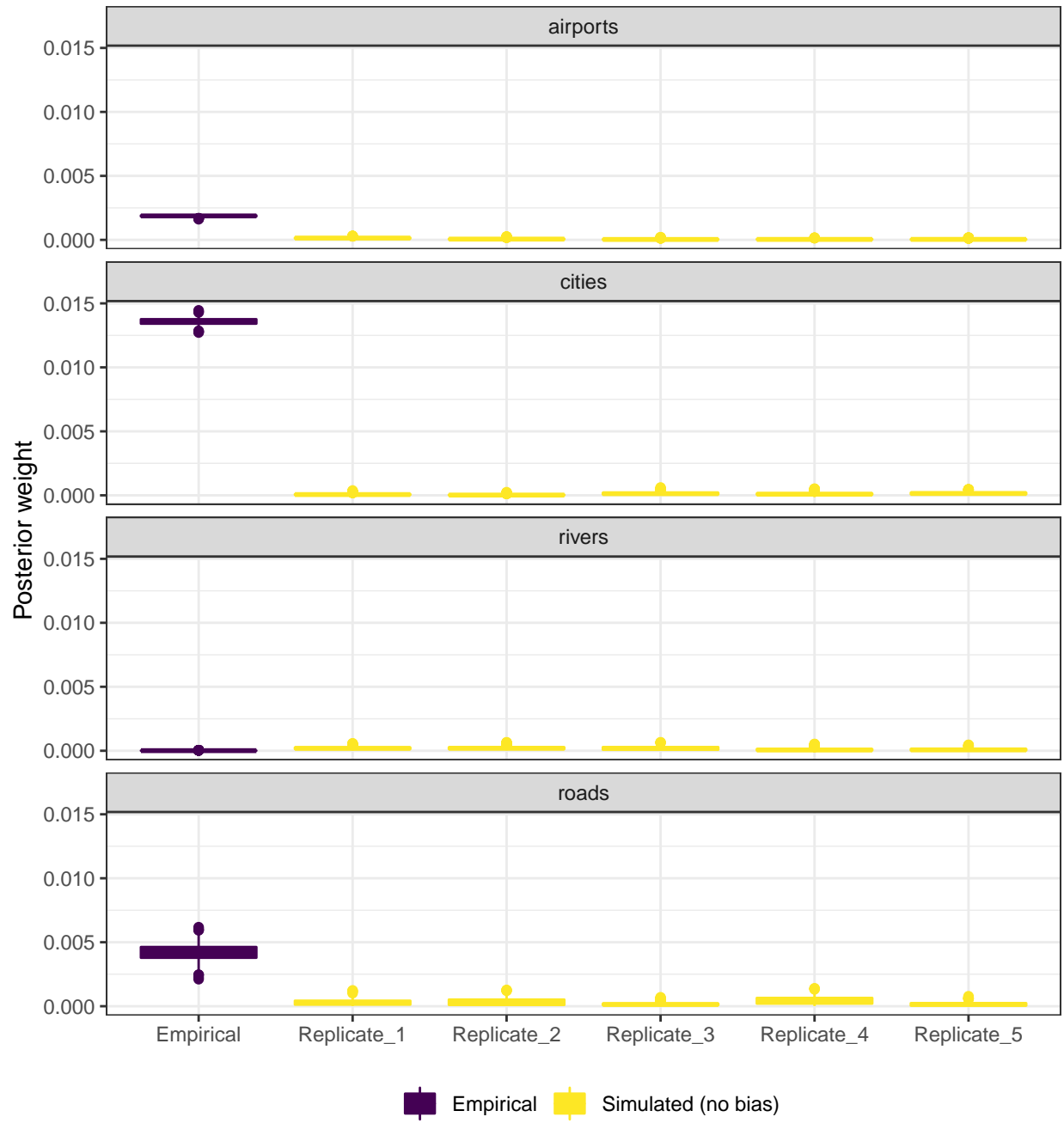


Figure S2: The bias weights ( $w$ ) of the empirical analyses of records from mammals from Borneo (6262 records, [www.gbif.org](http://www.gbif.org), 2016, <https://doi.org/10.15468/dl.7fg4zx>) in comparison to ten simulated datasets of the same size with randomly picked locations across Borneo (no bias).

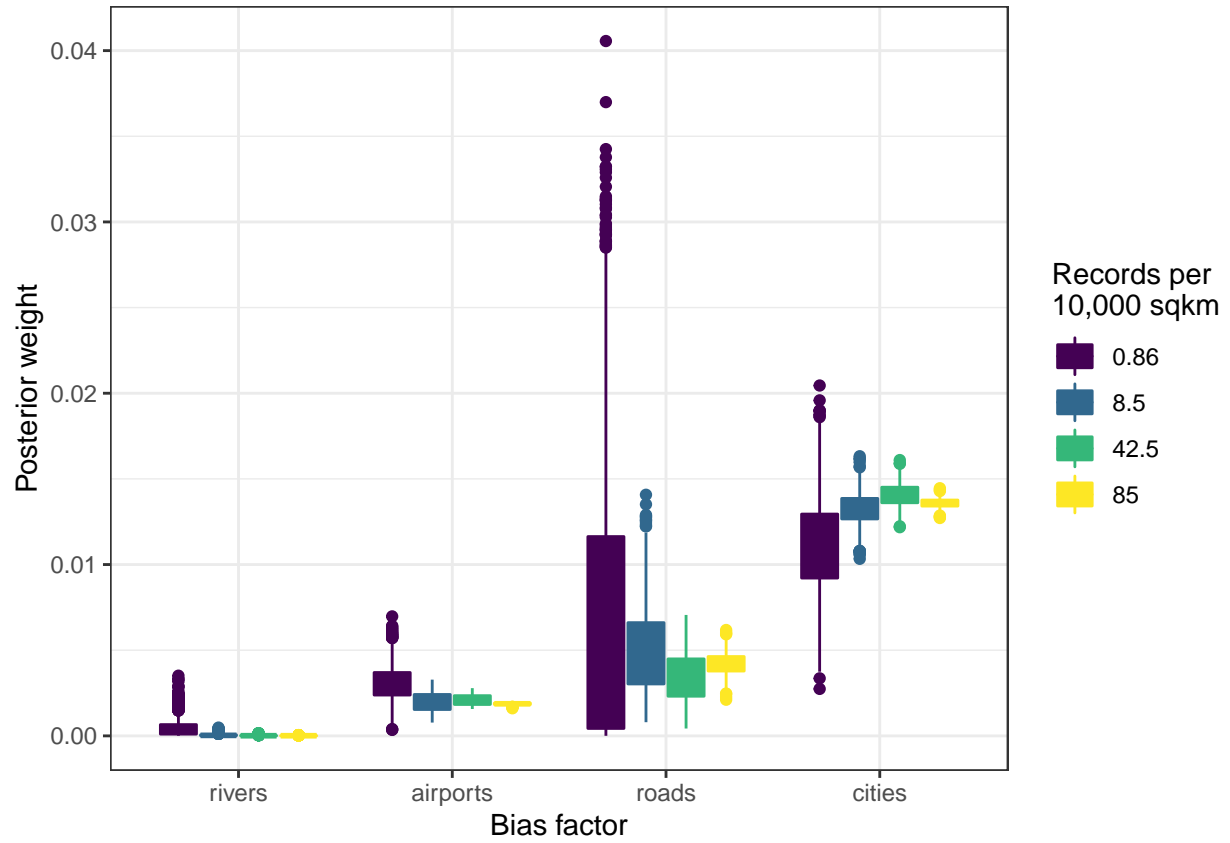


Figure S3: The bias weights ( $w$ ) quantifying the effects of each bias factor estimated from data sets with differing density of occurrence records across the study area. Datasets generated by down-sampling the empirical example dataset of mammal occurrences on Borneo ([www.gbif.org](http://www.gbif.org), 2016, <https://doi.org/10.15468/dl.7fg4zx>).

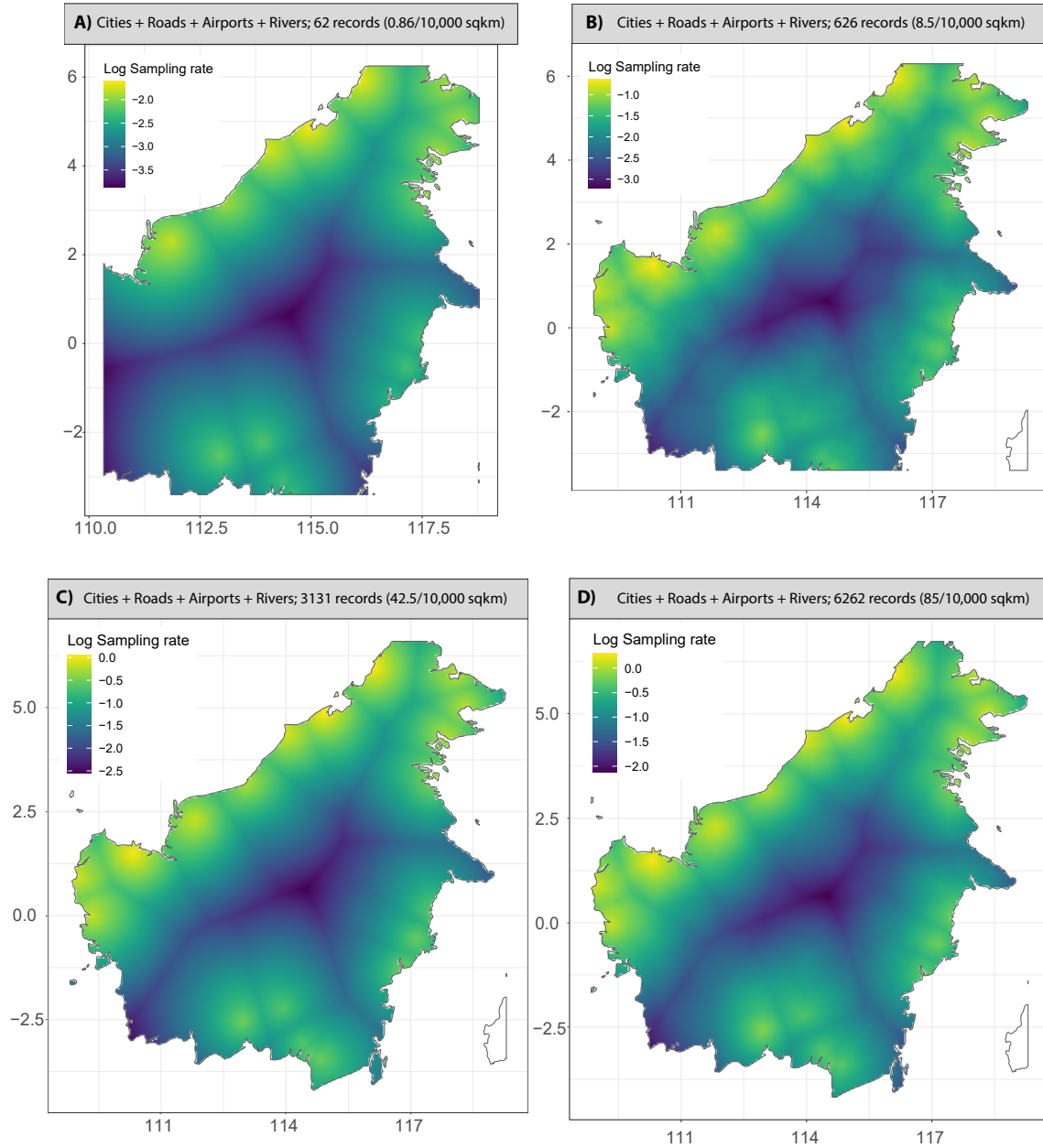


Figure S4: Projected bias surface for the distribution data of mammals from Borneo ([www.gbif.org](http://www.gbif.org), 2016, <https://doi.org/10.15468/dl.7fg4zx>), at different levels of downsampling, from a random replicate. The plotted extent is adapted to the study area, which may not cover the entire island for rarified datasets with few records (A) and B))