

speciesgeocodeR and *sampbias*: new tools for detecting erroneous coordinates and accessibility bias in biological collection data

Alexander Zizka, University of Gothenburg

The public availability of large-scale species distribution data has increased dramatically over the last ten years. In particular, the digitalization of collections from museums and herbaria and the aggregation of information in public databases, such as the Global Biodiversity Information Facility (GBIF) have contributed significantly to this development. However, the use of these data for ecological and biogeographic research as well as conservation planning is often hampered by issues regarding data quality. Two major issues are I) erroneous geographical information due to wrong or overly imprecise geographic coordinates, and II) undocumented sampling effort leading to uncertainty on sampling biases, as for example caused by differences in geographical accessibility. Here we present two new software tools: *speciesgeocodeR* to identify potentially errors in geographic coordinates common to biological collections and *sampbias* to quantify the influence of geographic structures related to accessibility, such as roads or rivers on sampling intensity. Furthermore, we illustrate the use of these tools on GBIF data. The software can be obtained from <https://github.com/azizka> and www.antonelli-lab.net.