

# Modelling the Heterogeneity within Citizen Science Data for Biodiversity Research

Diana E Bowler<sup>‡</sup>, Nick J B Isaac<sup>§</sup>, Aletta Bonn<sup>¶,‡</sup>

<sup>‡</sup> German Centre for Integrative Biodiversity Research (iDiv), Leipzig, Germany

<sup>§</sup> Centre for Ecology & Hydrology, Wallingford, United Kingdom

| Helmholtz-Center for Environmental Research – UFZ, Leipzig, Germany

<sup>¶</sup> Friedrich Schiller University Jena, Jena, Germany

Corresponding author: Diana E Bowler ([diana.e.bowler@gmail.com](mailto:diana.e.bowler@gmail.com))

## Abstract

Large amounts of species occurrence data are compiled by platforms such as the Global Biodiversity Information Facility (GBIF) but these data are collected by a diversity of methods and people. Statistical tools, such as occupancy-detection models, have been developed and tested as a way to analyze these heterogeneous data and extract information on species' population trends. However, these models make many assumptions that might not always be met. More detailed metadata associated with occurrence records would help better describe the observation/detection submodel within occupancy models and improve the accuracy/precision of species' trend estimates. Here, we present examples of occupancy-detection models applied to citizen science datasets, including dragonfly data in Germany, and typical approaches to account for variation in sampling effort and species detectability, including visit covariates, such as list length. Using results from a recent questionnaire in Germany asking citizen scientists about why and how they collect species occurrence data, we also characterize the different approaches that citizen scientists take to sample and report species observations. We use our findings to highlight examples of key metadata that are often missing (e.g., length of time spent searching, complete checklist or not) in data sharing platforms but would greatly aid modelling attempts of heterogeneous species occurrence data.

## Keywords

biodiversity monitoring, metadata, survey methods

## Presenting author

Diana E. Bowler

## **Presented at**

TDWG 2021

## **Conflicts of interest**