
From passive monitoring to the mapping of migration routes: accounting for identification errors when building species abundance models and inferring seasonal connectivity

Charlotte Roemer*¹, Yves Bas², and Christian Kerbiriou²

¹Centre d'Ecologie et des Sciences de la COnservation – Museum National d'Histoire Naturelle, Sorbonne Universite, Centre National de la Recherche Scientifique – France

²Centre d'Ecologie et des Sciences de la COnservation – Muséum National d'Histoire Naturelle (MNHN) – France

Abstract

Accurate spatial information on species distribution and connectivity is urgently needed for conservation purposes, notably for migrating species, particularly vulnerable to anthropogenic sources of mortality. Abundance data collected with standardised citizen science is one of the best sources of information for this purpose. Yet, the quantity of data requires an automatic process for the identification of species, which is associated with identification errors. In addition, when connectivity is inferred based on species distribution models (SDM), it is unknown how the identification errors inherited from the automatic process impact connectivity models.

We based our method on data from the passive acoustic monitoring of bat populations collected in France between 2014 and 2022 (26,001 sampled points, 115,375 full nights). These recordings were automatically identified thanks to a random forest species classifier. We then compared four possible methods for accounting for identification errors: (1) using the original dataset without any filter or weighting, (2) sorting out all data with a probability of classification below 50 %, (3) sorting out all data with a probability of classification below 90 %, and (4) using the probability of classification as a weight. We then modelled species distribution for each of these methods. For this, we trained random forest models using environmental predictors. To model connectivity, we used the results of the SDM to define origin and goal locations and to build a map of movement costs, and we calculated the randomized shortest paths. We then compared the results obtained after the four different methods for accounting for identification errors.

The maps of SDM and connectivity displayed very similar results, although sorting out all data with a probability of classification below 90 % was the most different from the other methods. The maps confirm general patterns deduced from multiple complementary sources in the literature about species distribution. They also shine a light on ecological patterns that were never documented before, such as the likely altitudinal migration of several species. Our study is the first methodological workflow for producing connectivity maps based on results from automated identification of species with abundance data. This study should be

*Speaker

reproduced with other types of datasets and to other scales to assess whether the manner to account for the species identification errors systematically leads to very similar results or not. This method could also be applied to other monitoring data, such as the automatic identification of species on camera trap.

Keywords: SDM, SAM, bats, bioacoustics, vigie, chiro, AI, machine learning