

Challenges analyzing avian distributions and abundance at scale with community science data: Perspectives from the eBird Status and Trends project

Daniel Fink
Cornell Lab of Ornithology
Ithaca, NY, USA
daniel.fink@cornell.edu

Abstract

Information on species' distributions and abundances and how they change over time are central to the study of wildlife populations and their conservation. For many taxa, this information is challenging to obtain at relevant spatiotemporal scales and extents, leading to large knowledge gaps. However, birds are conspicuous, occur in all habitats, and are enjoyed by tens of millions of individuals globally. The goal of the eBird Status and Trends project is to use data from eBird, a global citizen science bird monitoring project, to produce and freely disseminate a standardized source of up-to-date high resolution biodiversity information for the world's bird populations.

Community Science (CS) projects like eBird have emerged as attractive data sources for biodiversity studies because of their success collecting large volumes of species data continuously across large spatial extents. However, most of these observational data sets lack critical sampling structure making bias a concern. In this presentation, we discuss three challenges faced when using eBird data to estimate species' distributions, abundances at scale and how they change. First, we discuss a feature engineering approach to account for variation among participants in how they search for and detect birds. Next, we discuss the divide-and-recombine strategy we use to scale estimates of species distributions and abundance across large spatial and temporal extents dealing with associated challenges of non-stationarity, varying data density, and computational cost. Finally, we discuss how the eBird observation process has evolved over time, presenting a number of interrelated challenges for when studying year-to-year changes in bird populations, a central question in conservation.