245

Estimating Population Change with Citizen Science Data and Causal Machine Learning

Citizen and community-science (CS) projects have unrealized potential to provide population monitoring across broad spatial extents, given the large volumes of high-resolution data collected globally every year. Yet, the flexible protocols that enable many CS projects to collect large volumes of data often lack the structure necessary to control for confounding sources of interannual variation when individual participants decide when, where, and how to conduct their surveys. Here we discuss how causal machine learning models can be used to estimate spatially-explicit population trends using semi-structured CS data. These models use a two-stage regression to account for selection bias and interannual confounding while estimating population change. Machine Learning is used to fit each regression, making it possible to exploit large sets of informative covariates to estimate complex patterns of population change in the face of complex selection biases.

To demonstrate this approach, we estimated interannual trends from 2007-2019 for 451 breeding bird species in North America using species data from the CS project eBird with environmental data derived from remote sensing. We estimated population trends and associated uncertainty within each 27km × 27km grid cell across each species' range. We demonstrate the reliability of these methods with species-specific simulations and comparisons with independently derived estimates from the North American Breeding Bird Survey. These high-resolution trends provide a novel, valuable source of ecological information to help understand causes of bird declines and plan conservation actions. These results add to a growing literature demonstrating how to unlock ecological information from citizen science data.