

Abstract

Model-based data integration provides a promising framework for fitting species distribution models using citizen science data together with structured survey data, but a common challenge is how to properly include biased citizen science data in an integrated model.

I implement an integrated species distribution model using two data sets of freshwater fish in Norway: one which is a structured survey data set and one which is a citizen science data set. For the underlying distribution, I use a log-Gaussian Cox-process. Together with this, I assume separate observation processes for each data set, but with shared environmental covariates and a shared spatial field. In addition, the observation process for the citizen science data is given a separate spatial field which is estimated only from the citizen science data, referred to as the effort spatial field. This allows us to estimate the spatial bias of these observations.

By comparing the estimated separate spatial field across four different species of freshwater fish, we see that even in fish with very different distributions, the effort spatial field is very similar. When comparing variations of integrated models to a survey-only model, the integrated models perform consistently better than the single-dataset model.

The integrated nested Laplace approximation (INLA) methodology is used to fit all models, and gives great flexibility as well as very efficient computation.

1 Introduction

2 Method

3 Results

4 Discussion

5 Conclusion