53

Integrating datasets to improve estimates of species distribution and abundance using non-linear models and a mean-dispersion parameterization of the Beta distribution

The increasing number and availability of ecological datasets motivates the development of methods for integrating datasets from different sources over the same spatial or temporal domain. In theory, combining the information in different datasets can improve the accuracy and precision of model estimates and predictions, but differences in format, scale, and sampling methodology must be appropriately accounted for in the modeling process. In this work, we develop an integrated distribution and abundance model for a vegetation community in Yellowstone National Park, USA, using vegetation cover class data from the National Park Service (NPS) Inventory Monitoring Program and percent cover data from the National Ecological Observation Network (NEON). The spatial extent of the NEON data is contained within the spatial extent of the NPS data, is more densely sampled, and contains information at a finer grain (percent cover) than the NPS data (cover class). We develop a calibration model in which we use the NEON data to adjust the estimates of distribution and mean abundance generated by the NPS dataset. After calibrating the data, we use a joint-likelihood approach to estimate spatial distribution and abundance in the context of a generalized additive model optimized using MCMC methods. Because the joint likelihood is in the form of a mean-dispersion parameterization of the Beta distribution, we directly model the response surface for the mean, and derive precision estimates from Bayesian posteriors. Results from our study can be used to inform management practices in the national park, and the technique we develop can be broadly applied to integrate two common forms of vegetation data: ordinal cover class and percent cover.