Documenting distributions of hyperdiverse groups (terrestrial invertebrates) through grid-based spatial sampling, high-throughput sequencing, and occupancy modeling

## Author Information

**Matthew L. Bowser,\* John M. Morton**

**M.L. Bowser** USFWS Kenai National Wildlife Refuge, Soldotna, Alaska 99669

**J.M. Morton** USFWS Kenai National Wildlife Refuge, Soldotna, Alaska 99669 (retired)

\*Corresponding author: matt\_bowser@fws.gov

## Introduction

Alaska National Wildlife Refuges have been given mandates to conserve natural biodiversity, habitats, biological integrity, and environmental health (96th Congress [1980](#ref-96th_Congress_1980); 105th Congress [1997](#ref-105th_Congress_1997)). Woodward and Beever ([2010](#ref-Woodward_Beever_2010)) noted that, under these broad purposes of National Wildlife Refuges in Alaska to protect natural landscapes and entire ecosystems, they must develop methods that can successfully monitor biodiversity.

Morton et al. ([2009](#ref-Morton_et_al_2009)) determined that the best way to measure biodiversity at the scale of Alaska Refuges was to document occurrences and co-occurrences of many species on a set of points distributed widely across the landscape. This kind of sampling scheme would yield species distributions and assemblages and, if repeated over time, could be used to assess whether Alaska Refuges were fulfilling their conservation purposes. Magness et al. ([2008](#ref-Magness_et_al_2008)) demonstrated the utility of this sampling design for generating species distribution data that could serve this purpose.

A key deficiency identified by Morton et al. ([2009](#ref-Morton_et_al_2009)) was an inability to feasibly monitor most of biodiversity due to the high cost and lengthy time required to obtain identifications for species in hyperdiverse groups. Since that time, high-throughput sequencing (HTS) methods have emerged as a means to obtain identfications for mixed samples and multiple studies (Bowser et al. [2017](#ref-Bowser_et_al_2017), [2020](#ref-Bowser_et_al_2020)) demonstrated that assemblages of diverse taxa can be efficiently identified using these methods.

The second deficiency recognized by Morton et al. ([2009](#ref-Morton_et_al_2009)) and Magness et al. ([2008](#ref-Magness_et_al_2008)) was the inability to account for imperfect detection. When a species is present but there is a chance that it will not be detected, this may lead to false negatives (i.e., recording an absence where a species is present), leading to bias in the resulting estimates of species distributions. Mackenzie et al. ([2002](#ref-MacKenzie_et_al_2002), [2006](#ref-MacKenzie_et_al_2006)) presented occupancy modeling methods that can yield unbiased estimates of species occurrence and distributions. These models explicitly account for imperfect detection through spatially or temporally repeated sampling.

The biological inventory work of Bowser et al. ([2020](#ref-Bowser_et_al_2020)) was designed in part to provide the kind of occurrence data required for monitoring distributions of multiple species (i.e., monitoring assemblages of species) and accounting for imperfect detection by spatial subsampling. Our intent in this paper is to test the usefulness of the methods of Bowser et al. ([2020](#ref-Bowser_et_al_2020)) for delivering unbiased estimates of species occurrence and distributions of hyperdiverse taxa.

## Methods

A summary of all occurrences by metabarcoding from sweep net samples obtained by Bowser et al. ([2020](#ref-Bowser_et_al_2020)) were downloaded from Arctos on 12 November 2020 (saved search URI: <https://arctos.database.museum/saved/2020-11-12-1157_Slikok_metabarcoding_taxa_list>), yielding a list of 975 unique identifications. For unidentified molecular operational taxonomic units (MOTUS) where 10 or more occurrences were observed, we attempted to improve identifications by submitting the corresponding sequences to BOLD’s Identification Engince and NCBI BLASTn. We obtained no new identifications.

The 2,375 occurrences were downloaded on 12 November 2020 (saved search URI: <https://arctos.database.museum/saved/2020-11-12-1400_Slikok_project_metabarcoding_occurrences>).

## References

105th Congress. 1997. 111 Stat. 1252 - National Wildlife Refuge System Improvement Act of 1997. *In* United States Statutes at Large, Volume 111, 105th Congress, 1st Session. pp. 1252–1260. Available from <https://www.govinfo.gov/app/details/STATUTE-111/STATUTE-111-Pg1252>.

96th Congress. 1980. Alaska National Interest Lands Conservation Act. *In* United States Statutes at Large, Volume 94, 96th Congress, 2nd Session. U.S. Government Printing Office. Available from <https://www.govinfo.gov/app/details/STATUTE-94/STATUTE-94-Pg2371>.

Bowser, M.L., Brassfield, R., Dziergowski, A., Eskelin, T., Hester, J., Magness, D.R., McInnis, M., Melvin, T., Morton, J.M., and Stone, J. 2020. Towards conserving natural diversity: A biotic inventory by observations, specimens, DNA barcoding and high-throughput sequencing methods. Biodiversity Data Journal **8**: e50124. doi:[10.3897/BDJ.8.e50124](https://doi.org/10.3897/BDJ.8.e50124).

Bowser, M.L., Morton, J., Hanson, J., Magness, D., and Okuly, M. 2017. Arthropod and oligochaete assemblages from grasslands of the southern Kenai Peninsula, Alaska. Biodiversity Data Journal **5**: e10792. doi:[10.3897/BDJ.5.e10792](https://doi.org/10.3897/BDJ.5.e10792).

MacKenzie, D.I., Nichols, J.D., Lachman, G.B., Droege, S., Royle, J.A., and Langtimm, C.A. 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology **83**(8): 2248–2255. doi:[10.1890/0012-9658(2002)083[2248:ESORWD]2.0.CO;2](https://doi.org/10.1890/0012-9658(2002)083%5B2248:ESORWD%5D2.0.CO;2).

MacKenzie, D.I., Nichols, J.D., Royle, J.A., Pollock, K.H., Bailey, L.L., and Hines, J.E. 2006. Occupancy Estimation and Modeling. Elsevier, New York.

Magness, D.R., Huettmann, F., and Morton, J.M. 2008. Using Random Forests to provide predicted species distribution maps as a metric for ecological inventory & monitoring programs. *In* Applications of Computational Intelligence in Biology: Current Trends and Open Problems. *Edited by* T.G. Smolinski, M.G. Milanova, and A.-E. Hassanien. Springer-Verlag, Berlin. pp. 209–229. doi:[10.1007/978-3-540-78534-7\_9](https://doi.org/10.1007/978-3-540-78534-7_9).

Morton, J.M., Bowser, M.L., Berg, E., Magness, D., and Eskelin, T. 2009. Long Term Ecological Monitoring Program on the Kenai National Wildlife Refuge, Alaska: An FIA adjunct inventory. *In* 2008 Forest Inventory and Analysis (FIA) Symposium; October 21-23, 2008; Park City, UT. *Edited by* W. McWilliams, G. Moisen, and R. Czaplewski. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, Colorado. pp. 1–17. Available from <http://www.treesearch.fs.fed.us/pubs/33332>.

Woodward, A., and Beever, E.A. 2010. Framework for ecological monitoring on lands of Alaska National Wildlife Refuges and their partners. Report, Reston, Virginia. doi:[10.3133/ofr20101300](https://doi.org/10.3133/ofr20101300).