

Utah State University

DigitalCommons@USU

Undergraduate Honors Capstone Projects

Honors Program

5-2016

Biodiversity Prioritization: A Comparison of Data Types

Kari Norman

Utah State University

Follow this and additional works at: <https://digitalcommons.usu.edu/honors>



Part of the [Natural Resources and Conservation Commons](#)

Recommended Citation

Norman, Kari, "Biodiversity Prioritization: A Comparison of Data Types" (2016). *Undergraduate Honors Capstone Projects*. 608.

<https://digitalcommons.usu.edu/honors/608>

This Thesis is brought to you for free and open access by the Honors Program at DigitalCommons@USU. It has been accepted for inclusion in Undergraduate Honors Capstone Projects by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.



BIODIVERSITY PRIORITIZATION: A COMPARISON OF DATA TYPES

by

Kari Norman

**Thesis submitted in partial fulfillment
of the requirements for the degree**

of

DEPARTMENTAL HONORS

in

**Conservation and Restoration Ecology
in the Department of Wildland Resources**

Approved:

Thesis/Project Advisor
Dr. Ethan White

Departmental Honors Advisor
Dr. David Koons

Director of Honors Program
Dr. Kristine Miller

UTAH STATE UNIVERSITY
Logan, UT

Spring 2016

© 2016 Kari Norman
All Rights Reserved

ABSTRACT

The identification of important areas for biodiversity is essential for effective allocation of limited conservation resources. Prioritizing regions for conservation based on biodiversity is typically done using global biodiversity maps created using range map data for one or more taxa. While the use of range maps makes pragmatic sense since large-scale survey data is rarely available, it is important to understand the sensitivity of the results to the use of range map data. We studied how prioritizations may change between data types using the North American Breeding Bird survey (BBS) and BirdLife International range maps as a comparison case study. Diversity maps were generated using the North American Breeding Bird survey data for both total species richness and the richness of rare species. Rarity was defined as species present at less than the median number of sites. To account for spatial sampling bias in the location of BBS routes, maps were created based on a subsampling of sites within 100 square kilometer grid cells. For comparison, similar maps were generated using range maps for equivalent species.

Analyses of the Breeding Bird Survey data and range map data show that for species richness there is only 12% - 15% overlap in hotspots at different scales. Hotspots for rare species have 56% - 57% overlap. Multiple regions, such as the southern and eastern states have high biodiversity for one data type and low biodiversity for another. Maps for rare species are generally more congruent, with hotspot concentrations along the southern border of the United States. Biodiversity patterns for species richness vary greatly between data types.

ACKNOWLEDGEMENTS

We thank Bird Life and the USGS Breeding Bird Survey for providing the data that made this research possible. We especially thank members of the Weecology lab for their productive comments on early drafts of this manuscript. This research was supported by an NSF CAREER award (DEB-0953694) and by the Gordon and Betty Moore Foundation's Data-Driven Discovery Initiative (Grant GBMF4563).

TABLE OF CONTENTS

Introduction	1
Methods	3
<i>Data Sources</i>	3
<i>Methods</i>	4
Results	5
Discussion	7
Literature Cited	11
Figure 1	13
Figure 2	14
Figure 3	15
Reflective Writing	16
Author Biography	19

INTRODUCTION

Identifying important areas for conservation is essential for addressing threats to biodiversity such as growing human populations, rising global temperatures, and widespread land use change. Approaches for systematically identifying the most important areas to conserve and efficiently allocating limited conservation resources are broadly referred to as conservation prioritization¹.

Conserving biodiversity is widely accepted to be both a goal of conservation and a metric associated with other positive conservation outcomes. As a result, many conservation prioritization analyses focus on maximizing the number of species in a given area, or an area's "species richness". Other desirable conservation criteria include endemism (the number of species occurring only in a particular area), vulnerability (species designation as threatened or endangered) , and level of threat (likelihood of future habitat loss).

Myers' seminal paper in 2000 is arguably the first example of biodiversity-based conservation prioritization, and certainly the first global assessment of conservation need². The analysis by Myers et al. yielded locations for 25 global biodiversity hotspots based on vascular plant endemism and threat. Building on this work, hotspot prioritizations have now been created for a number of different taxa³⁻⁵, compared to current reserve networks to evaluate their effectiveness in protecting biodiversity^{6,7,8}, and used to assess the scope of human impact on biodiversity centers⁹.

With its beginnings in conceptually simple but geographically comprehensive hotspot analyses, conservation prioritization has now been expanded for

application to a wide range of questions. Recent development of software such as Zonation and Marxan provide an algorithmic approach to prioritization that incorporates greater levels of ecological and human complexity¹⁰. Such advances have transformed conservation prioritization from a technique exclusively used in global-scale categorizations to a viable tool for local managers. Managers now use conservation prioritization to inform decisions such as the allocation of conservation money and expansion of local reserve networks¹.

As use of conservation prioritization proliferates throughout the conservation community, a largely unacknowledged methodological divide has emerged between studies using two distinct kinds of data. Global-scale hotspot analyses rely almost exclusively on geographic range map data^{2-4,6,11}. This data is heavily informed by expert opinion and potential habitat¹². The relatively low cost of that information means range maps are accessible for a wide array of taxa at continental to global scales. However, range map data have two potential weaknesses: 1) they are typically temporally static; and 2) they reflect biodiversity at spatial scales of nearly 2x2 degrees (~40,000 km²)¹². In contrast, managers working at smaller scales typically use survey data. Though costly to collect, these data provide direct observations of species richness or abundance in a particular region. These regions are often much smaller than the 2x2 degree grid cells approximated using range map data.

Despite this clear dichotomy between range map and survey based approaches, there have been no analyses to examine how differences in data type influence the regions prioritized for conservation. Comparisons of range map and

survey data for biodiversity patterns and models more generally show significant disparities between range map and survey based estimates¹³, but it is unclear how that discrepancy may translate to prioritization results.

We seek to address the question of sensitivity through comparison of hotspot analyses created based on survey and range map data of landbird species. Reflective of typical hotspot analyses, we created maps for overall species richness and the richness of rare species. Comparisons were made between data types at two levels of aggregation, the site and cell level, and geographic discrepancies between hotspot locations were quantitatively and qualitatively assessed.

METHODS

Data Sources

We compared biodiversity patterns of North American breeding landbird species based on survey and range map data. Digital breeding range maps were obtained from the Birdlife International^{4,18} Survey data were from 2,769 routes of the North American Breeding Bird Survey (BBS), collected in 2009¹⁷. Each route is 24.5-miles long and surveyed annually in June. Three-minute point counts are made along the route every 0.5 miles, in which every bird seen or heard within 0.25 miles is recorded. The 2009 surveys included data on 1,819,908 individuals from 347 species. Further information on this subset of BBS data can be found in White et al.¹⁴.

Methods

The first step in comparing prioritizations based on survey and range map data was to ensure a fair comparison between inherently different data types and methodological approaches. Survey data are discrete point estimates of richness at each site. In contrast, range maps data are typically aggregated into cells and richness patterns analyzed across those cells. We compared survey and range map based prioritizations at both site and cell levels for species richness and the richness of rare species.

Point richness estimates for both data types were made at the starting position of each BBS survey route. For survey data, estimates were the number of species counted on each route. Range map estimates were calculated by counting the number of individual species ranges intersecting with each point. Cell level richness was calculated for 100 km² cells across North America¹⁵. This is a typical cell area that accounts for range map resolution¹² while still producing a large number of cells for analysis. Cell richness values for survey and range map data were averages of the point richness estimates for sites within each cell. Sites and cells were then mapped for the North American extent on a color gradient, with lighter colored sites and cells corresponding to lower richness and vice versa.

Maps for rare species were created using the same methods for site and cell level richness estimates, but for the subset of species considered rare. We classified a species as rare when it occurred at a proportion of sites less than the median proportion of site occurrence⁴. The rarity proportion is sensitive to variation in the intensity of spatial sampling, because a species could be considered rare simple

because there happened to be few sampling locations within its range. Therefore richness estimates for rare species were made for a subset of sites adjusted to have consistent sampling intensity across the study area. This subset was made up of three randomly selected sites from within 100 km² cells across North America; cells containing less than three sites were treated as empty. This combination of cell size and number of samples was sufficient to address the bias while retaining a sufficient percentage of the data for a meaningful analysis.

Biodiversity hotspots were prioritized based on the highest richness values. The 5% most biodiverse points or cells were considered hotspots, and plotted in red^{4,12}. Hotspot locations based on survey data were compared to those based on range map data for general richness and rare species richness for point and cell level analyses. The percent of hotspots that were shared between range map and survey based approaches was calculated by direct site-to-site and cell-to-cell comparisons.

RESULTS

Maps of hotspot locations for species richness at site and cell level (Fig 1) differed notably between data types. Range map based site-level maps show heavily concentrated priority areas primarily in the northern Rockies, the area around Lake Winnipeg, and the Great Lakes Region. Smaller hotspots are also found in the mountains of northern California and desert areas of the western United States. At the cell level, hotspots in the Great Lakes Region disappeared, with concentrations

remaining around Lake Winnipeg, the northern Rockies, the northern Californian mountains, and the desert of the Western United States. Richness maps based on survey data at the site level give generally more dispersed priorities. Hotspots are located across the northeastern United States to the Great Lakes region up into Lake Winnipeg, and stretching down to southern states. Smaller hotspot contingents are also seen in the Canadian Rockies, across the mountain regions of Colorado, and in southern Arizona. Locations are generally the same at the cell level, but with a disappearance of hotspots in southern states, and the appearance of prioritized areas in northern California. At both levels, discrepancies are obvious. At the site level the biggest differences are in southern states and the Northeast, which are prioritized throughout for survey based map, but are some of the least biodiverse areas on the range map based map. Another clear difference is seen in the Canadian Rockies, where the reverse is true: a heavy concentration of hotspots for range map data, and relatively low biodiversity for survey data. Discrepancies remained in the Northeast and extended into the Great Lakes region for cell level maps.

Richness patterns of rare species show more congruence across level and data type (Fig 2). All four prioritizations give hotspots along the United States-Mexico border, and into southern California.

Direct site-to-site or cell-to-cell comparisons for maps of different data types reinforce the qualitative assessments of hotspot differences. For species richness there is only 12% overlap in hotspots at the site level and 15% at cell level (Fig 3). Higher similarity at the cell level is likely a function of greater aggregation. Hotspots for rare species had 56% overlap at the site level and 57% at the cell level (Fig 3).

Quantitatively there is a meaningful difference between data types, which is especially prominent in prioritizations of species richness.

DISCUSSION

Clear differences exist between the locations of conservation priority areas based on survey and range map data. Given these differences it is important to understand the reasons for discrepancies, and their implications for the appropriate situations in which to use each data type to ensure the most informed conservation decisions.

The biggest discrepancies between hotspot locations were seen in maps for species richness. Though at the site level some hotspots were in qualitatively similar regions for both data types, such as the Great Lakes and Lake Winnipeg, direct site comparison showed only a 12% overlap (Fig 3). Aggregation to the cell level revealed very little regional overlap qualitatively, and a similarly low 15% overlap from direct comparison (Fig 3). Discrepancies in the overall biodiversity pattern are emphasized by multiple regions for which there was high biodiversity for one data type and low for the other, such as the southern states, the Northeast, and the Canadian Rockies.

Priority areas for rare species were not nearly as incongruent, with 56% and 57% overlap for site and cell level comparisons respectively (Fig 3). Hotspot regions were qualitatively consistent across data type and level, with hotspots along the southern border of the United States and into southern California. Greater overlap in

hotspots for rare species compared to overall richness could be attributed to the different drivers each pattern responds too. Species richness patterns are largely based on species with wide ranges and respond to drivers such as area, habitat heterogeneity, productivity, and geometric constraints¹⁶. The most important driver for rare species is topographic heterogeneity¹⁶. Geometric constraints, or the constraints imposed by boundaries, appear to play a role in discrepancies between data types for richness hotspots. Geometric boundary issues lead to larger concentrations of species in the middle of a bounded region, a pattern we see reflected in richness maps based on range map data¹⁶. Further assessment of potential explanations for both consistencies and inconsistencies between data types, and their relationship to richness drivers, is key to using both data types more affectively.

In choosing a data type for future prioritizations, one clear criterion is how the scale of the data type reflects the scale at which it is attempting to inform decision. Our cell level analyses were performed at a resolution of approximately 1 degree (100 km² cells), a resolution typical for range map based prioritizations. Yet, this is a far coarser grain than would ever be used for local management decisions. As described in Jenkins et al.⁴, an area of that size in some parts of the world contains multiple mountain ranges and punctuating valleys. Still, attempts to simply analyze range map data at a scale more appropriate for conservation (e.g., Jenkins et al.⁴) may be misleading, especially in light of estimates of range maps' true resolution. Assessments show that analyses of range maps at resolutions less than 2

degrees result in overestimates of biodiversity and distorted spatial patterns, suggesting that range map data represent empirical patterns at this 2 degree scale¹².

The mismatch between the resolution of range map data and the scale of conservation questions leaves survey data as a natural replacement. Its local, explicit scale ensures biodiversity patterns will be representative of reality at the scales important for conservation. The increased availability of survey data through citizen science and large-scale government efforts means that using survey data in place of range map data is also realistic for many taxa, with more data becoming available in the future. Survey data's potential to improve the accuracy of biodiversity-based conservation indicates that its further availability is a worthwhile investment for the conservation community.

Despite the many benefits of survey data, conservation decisions cannot wait for its comprehensive availability. It is therefore also important to explore improvements for range map data use, including methods for downscaling range map data to more useful scales, updating old maps to reflect changes such as range shifts and land use changes, and new approaches for addressing range map porosity¹³. Establishing relationships between richness drivers and congruence between data types, as previously described, will play an important role in making methods for accurate range map data use possible.

Our findings underline the importance of understanding the implications of data type when prioritizing areas for conservation. Discrepancies in hotspot location and overall biodiversity patterns between data types give evidence of the current tradeoff between accuracy and availability in data type selection. Further

exploration of biodiversity analyses' sensitivity to data type is essential for partitioning the appropriate roles of each data type, and ensuring the most effective conservation planning into the future.

1. Lehtomäki, J., Tomppo, E., Kuokkanen, P., Hanski, I. & Moilanen, A. Applying spatial conservation prioritization software and high-resolution GIS data to a national-scale study in forest conservation. *For. Ecol. Manag.* **258**, 2439–2449 (2009).
2. Myers, N. Biodiversity Hotspots for Conservation Priorities. *Nature*
3. Brooks, T. M. Global Biodiversity Conservation Priorities. *Science* **313**, 58–61 (2006).
4. Jenkins, C. N., Pimm, S. L. & Joppa, L. N. Global patterns of terrestrial vertebrate diversity and conservation. *Proc. Natl. Acad. Sci.* **110**, E2602–E2610 (2013).
5. Ceballos, G. Global Mammal Conservation: What Must We Manage? *Science* **309**, 603–607 (2005).
6. Jenkins, C. N., Van Houtan, K. S., Pimm, S. L. & Sexton, J. O. US protected lands mismatch biodiversity priorities. *Proc. Natl. Acad. Sci.* **112**, 5081–5086 (2015).
7. Brooks, T. M. *et al.* Coverage Provided by the Global Protected-Area System: Is It Enough? *BioScience* **54**, 1081 (2004).
8. Rodrigues, A. S. L. *et al.* Effectiveness of the global protected area network in representing species diversity. *Nature* **428**, 640–643 (2004).
9. Gaston, K. J., Blackburn, T. M. & Goldewijk, K. K. Habitat conversion and global avian biodiversity loss. *Proc. R. Soc. B Biol. Sci.* **270**, 1293–1300 (2003).
10. Moilanen, A. & Wilson, K. A. *Spatial conservation prioritization : quantitative methods and computational tools.* (Oxford University Press, 2009).
11. Brooks, T. M. *et al.* Habitat Loss and Extinction in the Hotspots of Biodiversity. *Conserv. Biol.* **16**, 909–923 (2002).

12. Hurlbert, A. H. & Jetz, W. Species richness, hotspots, and the scale dependence of range maps in ecology and conservation. *Proc. Natl. Acad. Sci.* **104**, 13384–13389 (2007).
13. Hurlbert, A. H. & White, E. P. Disparity between range map- and survey-based analyses of species richness: patterns, processes and implications: Range map- vs. survey-based species richness. *Ecol. Lett.* **8**, 319–327 (2005).
14. White, E. P., Thibault, K. M. & Xiao, X. Characterizing species abundance distributions across taxa and ecosystems using a simple maximum entropy model. *Ecology* **93**, 1772–1778 (2012).
15. Orme, C. D. L. *et al.* Global hotspots of species richness are not congruent with endemism or threat. *Nature* **436**, 1016–1019 (2005).
16. Jetz, W. Geographic Range Size and Determinants of Avian Species Richness. *Science* **297**, 1548–1551 (2002).
17. Sauer, J. R., J. E. Hines, J. E. Fallon, K. L. Pardieck, D. J. Ziolkowski, Jr., and W. A. Link. 2014. The North American Breeding Bird Survey, Results and Analysis 1966 - 2013. Version 01.30.2015 USGS Patuxent Wildlife Research Center, Laurel, MD
18. BirdLife International and NatureServe (2014) Bird species distribution maps of the world. BirdLife International, Cambridge, UK and NatureServe, Arlington, USA.

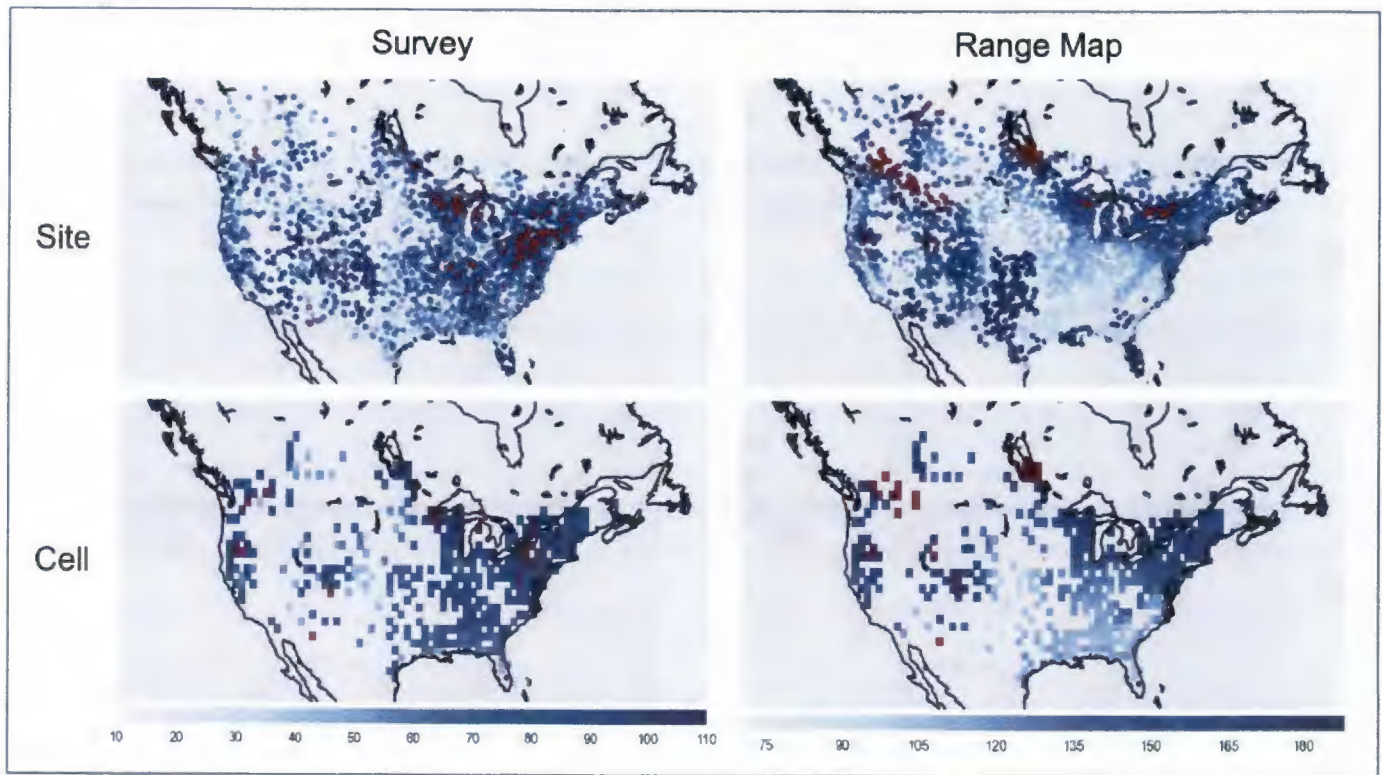


Figure 1. Maps of species richness based on survey and range map data at the site and cell level. Darker regions correspond to higher biodiversity, with hotspots (richest 5%) marked in red.

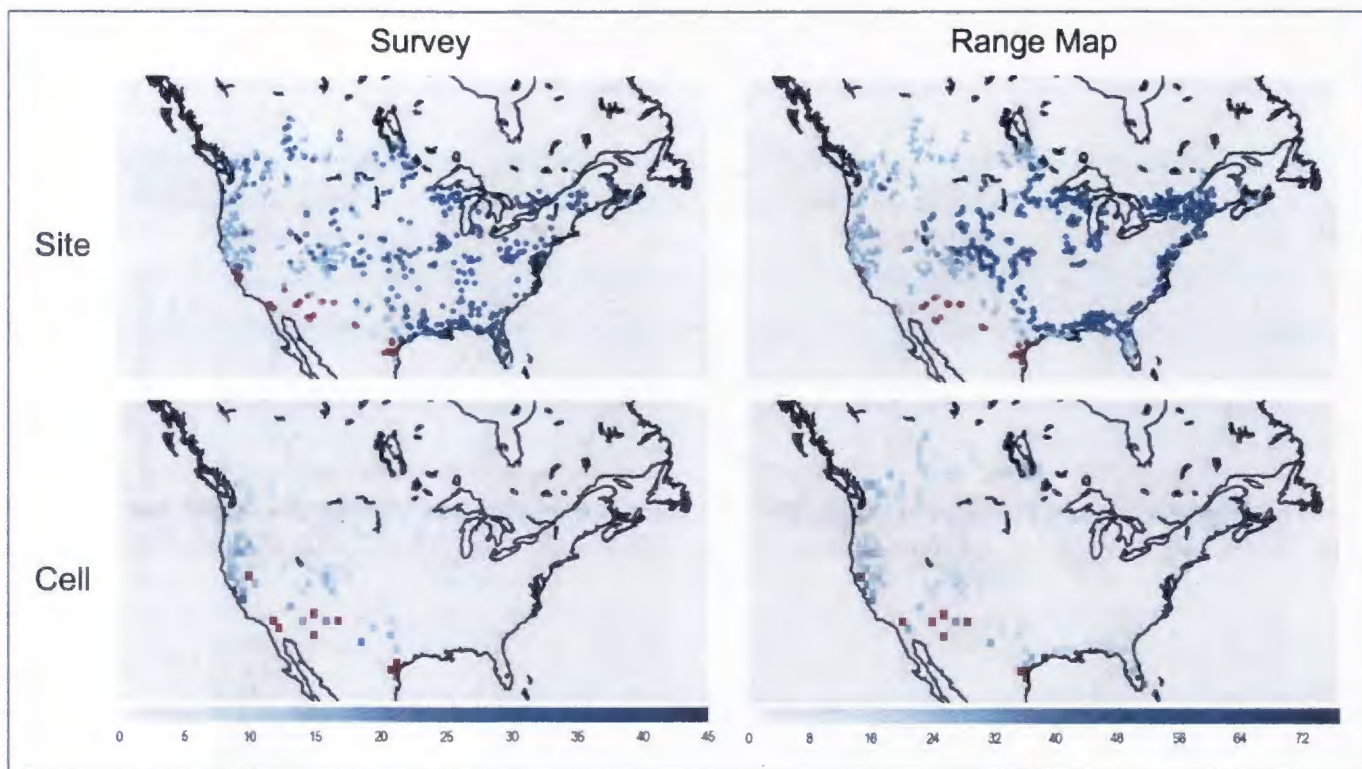


Figure 2. Maps of the richness of rare species based on survey and range map data at the site and cell level. Darker regions correspond with more rare species, with hotspots (richest 5%) marked in red.

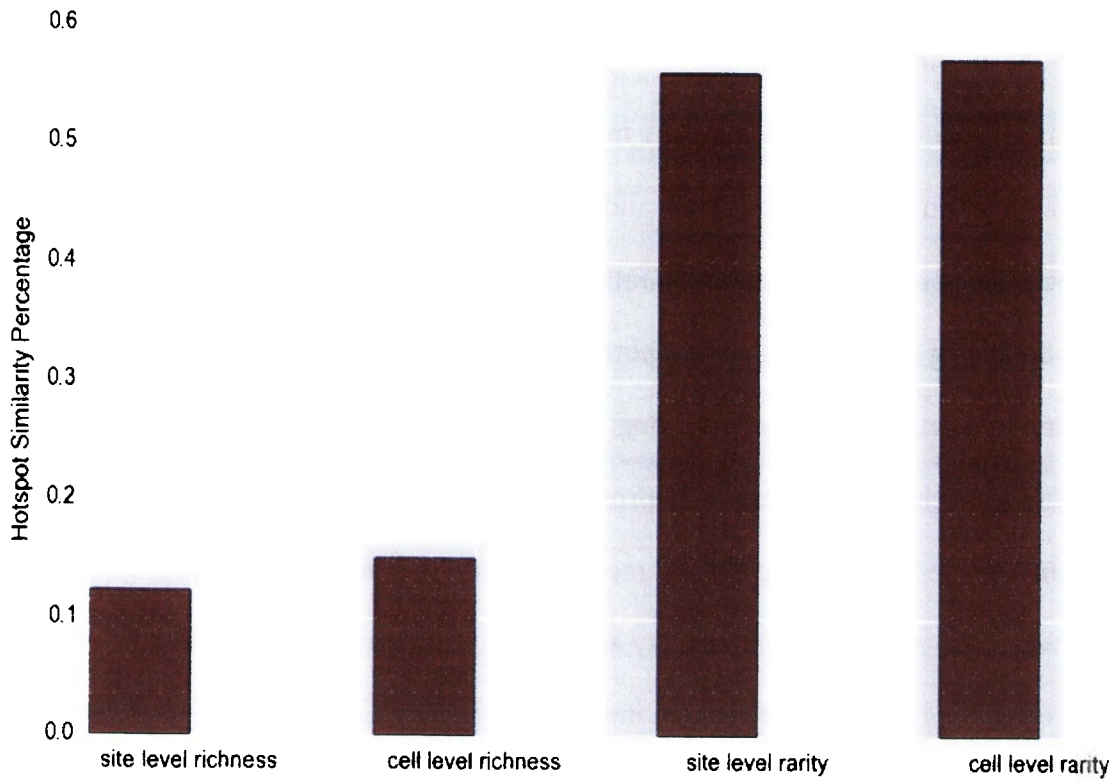


Figure 3. Percent hotspot similarity between range map and survey data types for richness and rare species at the site and cell level.

REFLECTIVE WRITING

As an Undergraduate Research Fellow I was involved in research from day one of my undergraduate career. Like most Freshman I had only a vague idea of where I wanted to take my research and career in general. I navigated to early projects based on keywords like “quantitative” and “modeling”, in the hopes I would find something I loved that married both my ecology and statistics backgrounds. This led to a pattern of project hopping, in which I gained a broad range of skills in everything from animal behavior to programming to watershed science. While I’m grateful for the broad ecological and methodological grounding those years gave me, they were also marked by uncertainty and anxiety about my research path. I enjoyed each project because it gave me the opportunity to learn new things and because of my love for research itself, but I worried I would never find something I was passionate enough about to work on for the rest of my life.

Just as I was sure from the beginning of my college career that research would be an important aspect of my time as an undergrad, I also planned to make the opportunity for studying abroad a priority. And so I found myself sitting in a course on Systematic Conservation Prioritization at the University of Helsinki, spring of my sophomore year. Throughout the course of that semester I came to discover a line of inquiry that incorporated many of the aspects of both ecology and statistics I found most interesting: comprehensive conservation solutions using advanced computational techniques. My passion for these approaches, and their further development, was much like an emergent property; I could not have predicted it based on my previous interest in its smaller components. With the full

intention of gaining a more in depth understanding of the field of conservation prioritization I did everything I could to get a clear vision of what that meant before leaving Helsinki. Though my excitement was unwavering, conversations with professors there and my own personal research proved disheartening. The kind of interdisciplinary software development happening at Helsinki was in many ways unique, and no one could point me to similar work being done in the States.

Upon returning home I had the singular goal of identifying and acquiring the additional skills I would need to pursue a career in large-scale conservation ecology, despite the fact that active pursuit of that career would be on hold until an international graduate degree. Included in those preparations was a course in programming for biologists taught by Dr. Ethan White, who was soon to become my research mentor. Little did I know that hidden away behind the microbiologists and physiologists in the Biology department lived one of the foremost Macroecologists in the country. In his class I quickly realized that, while he was not actively developing conservation software, his work using large ecological data sets to address global biodiversity patterns was the theoretical underpinning for the work that has so inspired me in Helsinki.

Ethan and I worked together, first over the course of the semester I was in his class, and then for the rest of my time at Utah State, to develop a project built on my experience in Helsinki, addressing an important question in the scientific community. I independently delved into the literature, bringing back to our meetings questions and connections and interesting tidbits, the product of which was the project that would eventually turn into this thesis.

While I consider the discovery and pursuit of a topic I'm truly passionate about to be a great triumph in and of itself, the execution of this project from development to manuscript has been one of the most fulfilling aspects of my time as an undergraduate. Much of my work required the manipulation of large datasets, and therefore the development of computational tools oftentimes unfamiliar to even my mentor. I spent many hours in trial and error with new software and programming languages. A few entire months were swallowed up in what Dr. Philip Guo of the University of Rochester terms "command line bullshittery", the phenomena of the software installation time suck. Yet despite setbacks, I remained committed and excited about the project and the potential it held for a career I loved.

If there were one message I could give to those coming after me, it would be to trust the things you love, even if you don't know where they are going to take you. I could not have possibly imagined on day one where I would be now, I didn't even know the field existed! But I did know I loved ecology and conservation, and the power statistics gives to address those topics. Had I not gone to Helsinki despite the fact it might set me back in course work, I would never have discovered the field of systematic conservation. Had I not continued to pursue that field even after it seemed all hope was lost, I would never have met the mentor that opened my eyes to the scope of questions and approaches possible. This project, my position as an Undergraduate Research Fellow, and the Honors program changed my life. They have defined my undergraduate career and future path, one that would not have been possible without these programs and the people and experiences at Utah State.

AUTHOR BIOGRAPHY

Kari Norman graduated Magna Cum Laude in spring of 2016 with degrees in Conservation and Restoration Ecology and Statistics, a minor in Biomathematics, and University and Departmental Honors. She was awarded the Outstanding Statistics Undergraduate her senior year. During her time at Utah State University she was a founding member of the Ecology club, a member of the Range Club Plant Identification Team, a member of the Sustainability Club, and served on the Natural Resources Student Council as the Wildland Resources Department representative. She has worked as a Fellow on the iUtah project examining watershed response to climate change, a DAAD Rise intern in Goettingen, Germany developing models for predicting forest biomass from LiDAR data, and as an Undergraduate Research and Creative Opportunities grant recipient, examining the impact of climate change on Uinta Ground Squirrels. She has presented current and past work at conferences such as the Utah Conference of Undergraduate Research, the National Conference of Undergraduate Research, Research on Capitol Hill, and the annual meeting of the Ecological Society of America. She will continue her work while pursuing a PhD at the University of California, Berkeley, where she also hopes to participate in her other passions: playing violin, cuddling with her cat, and reading for fun.