ABSTRACT

WALDEN-SCHREINER, CHELSEY ALYSSA. Integrating Geospatial Resource Use and Environmental Data at Varying Scales for Protected Area Conservation (Under the direction of Dr. Yu-Fai Leung).

Biodiversity loss threatens the stability and function of ecosystems, yet is occurring at rates far greater than natural processes largely due to human activities. Protected areas (PAs) are a key strategy to conserve species and ecosystems, and represent global commitments to reduce future losses. Protected area objectives also encompass community development and ecosystem services, including nature-based recreation and tourism. Globally, recreation and tourism provide vital economic and political support to the existence and management of protected areas, yet also represent a potential threat to conserved resources already contending with pressures like climate change. Research on human (i.e., visitor)-environment interactions is vital to understanding relationships between use and natural resource impacts, thus informing effective management.

Methodological challenges and limited resources to collect detailed use data are often credited as main obstacles to integrated analysis of the effects of recreation and tourism on natural resources. The goal of this dissertation is to leverage recent advances in geospatial approaches to advance use monitoring and to foster integrated analysis regarding use-environment interactions in PAs at multiple scales. Specifically, both formal and crowdsourced geospatial data were analyzed using methods and analytical techniques applied in sub-disciplines of ecology, but only recently employed to monitor and understand visitor and visitor-related activities in relation to environmental conditions in PAs.

Implemented as three case studies, each study employed a specific geospatial method, spatio-temporal scale, or analytical approach. The first case study integrated observational techniques with global positioning system (GPS) tracking and behavior change point analysis to explore the behavior of domestic animals (i.e., pack stock) used to transport people and materials in PAs when released to graze during overnight periods. The complementary approaches identified multiple bouts of different behaviors during overnight release periods, including the separation of intense, localized grazing from ambulatory grazing relevant to future behavior and movement modelling parameters and management to reduce impacts to meadow habitats.

The second two case studies, which applied Maximum Entropy modelling to both formal and crowdsourced geospatial data, explored relationships between use, infrastructure, and natural environment. The second case study focused on pack stock use patterns at specific locations during 12-hour release periods, while the third case study examined park-level visitor use patterns at seasonal and annual temporal scales. Findings from both studies identified the importance of infrastructure (i.e., roads, trails) and management tactics (i.e., temporary fencing, grazing release locations of domestic animals) aimed at influencing distribution patterns of use in PAs.

Implications of findings from the three case studies highlight the significance of planning when implementing new tactics or building new infrastructure, because the resulting spatial change in use patterns will directly affect the location, timing, and types of impacts on natural resources. Results also provide empirical evidence about the level of efficiency in the methods employed and their ability to enhance collection and analysis of

spatial visitor use and environmental data. Limitations of each approach and modelling decisions are discussed, illustrating how findings can contribute to future research about other visitor activities, ecological impacts, and landscape types.

Integrating Geospatial Resource Use and Environmental Data at Varying Scales for Protected Area Conservation

by
Chelsey Alyssa Walden-Schreiner

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Forestry and Environmental Resources

Raleigh, North Carolina

2017

APPROVED BY:

Dr. Yu-Fai Leung Committee Chair	Dr. Stacy A. C. Nelson
Dr. Laura Tateosian	Dr. Thomas R. Wentworth

DEDICATION

I dedicate this dissertation to my family. To my parents for always encouraging my journey. To my husband, Scott, for his steadfast support and willingness to assist with data collection in any weather. Finally, to my son, Finn, whose enthusiasm for nature and delight upon seeing redwoods for the first time motivates me daily.

BIOGRAPHY

Chelsey Walden-Schreiner grew up in Madison, Wisconsin, where she developed an interest in conservation and communications. Combining her interests, she has a B.A. in Journalism and Mass Communication from Iowa State University and an M.S. in Natural Resources from North Carolina State University. Her research emphases include human-environment interactions and geographic information science pertaining to conservation and sustainable visitor use in global protected areas.

ACKNOWLEDGMENTS

The following dissertation was possible due to the collaboration, support, and guidance of several individuals. First, my sincerest thanks to my advisor and committee chair, Dr. Yu-Fai Leung. I am deeply grateful for the years of advice, encouragement, and energizing conversations. Dr. Leung's mentorship helped me grow as a researcher and person, and I cannot thank him enough for the opportunities he encouraged and supported. It has also been a privilege to learn from and work with my committee members, Drs. Stacy Nelson, Laura Tateosian, and Thomas Wentworth. I appreciate their expertise, guidance, constructive feedback, and the illuminating discussions throughout this process.

I thank the Department of Forestry and Environmental Resources for helping me reach my educational and professional goals throughout my doctoral program, including the opportunity and funding to conduct research internationally; the Department of Parks, Recreation, & Tourism Management for the opportunity to collaborate on multidisciplinary projects as a Research Associate; and the Center for Geospatial Analytics for opportunities to engage with cutting-edge geographic information science research, faculty, and technology. I also thank Bill Slocumb and Dr. Gary Howell of NC State University for their assistance with the high performance computing capabilities required for Chapter 2.

I also offer my sincere thanks to Todd Newburger, Tim Kuhn, and the Resources Management and Science staff at Yosemite National Park for the opportunity to collaborate on the pack stock animal study. I am very grateful for their expertise, time, and the gigabytes of data shared, as well as the funding from the National Park Service supporting the research.

Thank you also to Dr. Wei-Lun Tsai for her efforts on the data management and initial analyses, which helped inform the subsequent analyses in Chapters 2 and 3.

Thank you also to Dr. Rhonda Loh and Hawaii Volcanoes National Park for the opportunity to conduct research in the park, and to the Pacific Island Network Inventory and Monitoring program for the GIS data included in Chapter 4.

I'd also like to thank Dr. Catherine Pickering of Griffith University for her guidance, generous hospitality, and assistance with conducting research in Kosciuszko National Park, Australia during my doctoral program. The research in Australia was also made possible through funding from the Laarman International Gift Fund, and supported by conversations with and hospitality provided by several individuals from the New South Wales Office of Environment and Heritage (NSW OEH), specifically Pamela O'Brien, Anthony Evans, Dave Woods, Wil Allen, Andrew Smith, and Will Dorrington. Thank you also to Dr. Isabelle Wolf from the NSW OEH for her feedback on the study and site knowledge.

Finally, I thank my family, friends, and current and past graduate students who provided vital support and inspiration along the way.

TABLE OF CONTENTS

LIST OF TABLES

APPENDICES

LIST OF FIGURES

CHAPTER 1 – General Introduction

CHAPTER 2 – Integrating Direct Observation and GPS Tracking to Classify Animal Behavior in Natural Areas

CHAPTER 3 – Environmental and Managerial Factors Associated with Pack Stock Distribution in High Elevation Meadows: Case Study from Yosemite National Park

CHAPTER 4 – Digital Footprints in Protected Areas: Incorporating Crowdsourced Geographic Information to Support Resource Protection

APPENDICES

CHAPTER 1

General Introduction

> "Only within the moment of time represented by the present century has one species – man – acquired significant power to alter the nature of the world."
>
> – Rachel Carson, *A Silent Spring*, 1962

Humans have modified and affected landscapes throughout history, yet the rate and scale accelerated more rapidly in the last century than any other time, resulting in species extinction rates 100 to 1,000 times greater than natural rates (Millenium Ecosystem Assessment, 2005; Pimm et al., 2014; Steffen, Broadgate, Deutsch, Gaffney, & Ludwig, 2015). Research has demonstrated that the loss of biodiversity, whether from habitat conversion, deforestation, or unsustainable use of natural resources, reduces the stability of ecosystem structure and function, and availability of ecosystem services like clean water, timber, carbon storage, medicine, and food (Chapin et al., 2000). The dramatic acceleration of human activity and documented impact on ecosystems catalyzed a movement to name a new geologic epoch, the Anthropocene, and underscored the importance of research on conservation effectiveness, sustainable resource management, and human-environment interactions (Crutzen, 2002; Steffen et al., 2011, 2015).

Protected areas (PAs) represent a critical strategy to reduce the loss of biodiversity globally (Chape, Harrison, Spalding, & Lysenko, 2005; Heller & Zavaleta, 2009; Thomas et

al., 2012). According to the International Union for Conservation of Nature (IUCN), PAs are "clearly defined geographical space, recognized, dedicated, and managed, through legal or other effective means, to achieve the long term conservation of nature with associated ecosystem services and cultural values" (Dudley, 2008, p. 8). While the concept of PAs is not new, objectives have increasingly diversified from protecting iconic landscapes and charismatic species to also encompassing biodiversity conservation, ecosystem services, and community development (Castro et al., 2015; Watson, Dudley, Segan, & Hockings, 2014).

To fulfill their role in conservation, PAs require accurate and timely information about ecological conditions as well as human (i.e., visitor) use (DeFries et al., 2010; Hadwen, Hill, & Pickering, 2008; Monz & Leung, 2006). However, logistical and funding constraints commonly limit the capacity to monitor visitor use and impacts, and subsequently human (visitor)-environment interaction research. Monitoring provides empirical evidence to justify and prioritize research, and it is a key aspect of science-based and adaptive management paradigms. This dissertation contributes to integrated research on the management of visitors and their potential impacts in PAs by exploring how geospatial approaches can efficiently collect and analyze environmental and visitor data at varying scales to balance visitor use and conservation.

Biodiversity Conservation and Visitor Use in PAs

In 2010, the Strategic Plan for Biodiversity established the Aichi Biodiversity Targets as international goals to help reduce biodiversity losses by 2020 (Convention on Biological

Diversity, 2017). Within the five strategic goals focused on reducing human impact, promoting sustainable resource use, and ensuring equitable benefit sharing, visitors to PAs have the capacity to contribute to several of the 20 Targets. For example, Targets 1 and 19 aim to increase public awareness about the values of biodiversity and encourage knowledge transfer, while Targets 4, 6, and 7 promote the sustainable use of and mitigating impacts to natural resources. PAs represent opportunities to share information with visitors, showcase strategies to balance human use with conservation, and engage the public in conservation efforts through volunteerism and citizen science (Hvenegaard, Halpenny, & McCool, 2012). Target 11 establishes a goal of protecting 17% of terrestrial and 10% of marine areas by 2020 (Chape et al., 2005; Watson et al, 2014). As of December 2016, national and global arrangements protected approximately 14.8% and 13% of the world's terrestrial and marine areas, respectively (UNEP-WCMC and IUCN, 2016). The growth in marine protected areas was especially dramatic given the protection of 3.4% in 2014.

Although critical gains have been made in recent years toward achieving these targets, including Target 11, analyses indicate gaps still exist in conserving all areas ecologically representative of the world's ecosystems and species (Akasaka, Kadoya, Ishihama, Fujita, & Fuller, 2016; Joppa, Visconti, Jenkins, & Pimm, 2013), and area-based targets must still be properly located and managed to fulfill conservation objectives (Watson et al., 2016). Achieving Aichi Targets and maintaining conservation goals requires continued social, economic, and political support to ensure effective and equitable management.

Nature-based recreation and tourism serve as avenues of support, and governing body mandates often require PA managing organizations to protect ecological integrity while providing for compatible visitor activities. A study of in-country expenditures associated with the estimated eight billion tourist visits to PAs worldwide generated roughly $600 billion a year (Balmford et al., 2015). Increasing focus on the economic implications of nature-based tourism, as well as visitor satisfaction and the role of PAs in promoting health outcomes, further emphasizes the tourism industry as important PA stakeholders vital to achieving multiple Aichi Targets (Keniger, Gaston, Irvine, & Fuller, 2013; Puhakka, Pitkänen, & Siikamäki, 2016; Spenceley, 2015; Stolton, 2010).

Visitor Use Impacts in Protected Areas

Along with the many benefits of and to visitors in PAs, there are also concerns regarding ecological pressures attributed to nature-based recreation and tourism that, if unmanaged, can compromise PA objectives and are considered among the top threats to PA ecosystems (Cole & Landres, 1996; Marion, Leung, Eagleston, & Burroughs, 2016). Research has demonstrated that visitors can create negative impacts to soil, vegetation, wildlife, water, and air that are locally severe (e.g., erosion, vegetation trampling) as well as spatially dispersed (e.g., wildlife displacement) in highly significant ecological and cultural sites (Buckley, 2013; Hammitt, Cole, & Monz, 2015). Tourism and nature-based recreation threaten 42% of vascular plant species identified as critically endangered, endangered, or vulnerable by the IUCN Red List in Europe (Ballantyne & Pickering, 2013) and Australia

(Rankin, Ballantyne, & Pickering, 2015), and 35% of rare plants in the United States (Hernández-Yáñez et al., 2016). A global review of research on the impacts of nature-based recreation on wildlife also found 93% of the studies documented at least one impact on animals, with 59% of those impacts identified as negative (Larson et al., 2016). Scarce data on visitor use intensity, distribution, and seasonality are a common limitation to explorations of how the environment influences and response to visitor patterns and impacts, which informs management strategies.

Integrating Visitor Use and Environmental Data

In light of these threats and challenges, approaches are needed to gather and analyze data efficiently. Historically, monitoring in PAs has focused on specific locations, species, or impacts, with less emphasis on the complex and potentially interacting elements within the system of interest (Burton et al., 2014). As our understanding of natural and human systems expands, so too has the call for coupled human and natural systems (CHANS) research and exploration of the challenges associated with the integrated analysis of multiple datasets (McConnell et al., 2011; Sullivan & Meigh, 2007). CHANS are systems in which humans and the natural environment interact, with a research focus on the linking patterns and processes, integrated assessments, and accounting for interactions and feedbacks at multiple scales (Liu et al., 2007).

Limited data, mismatches in scale, resolution, spatial and temporal overlap and data type complicate analyses (Herr, 2007; McClintock, Johnson, Hooten, Ver Hoef, & Morales,

2014). An active area of research explores integration opportunities and methods, including addressing the spectrum of scales involved in multiple datasets and ecological and social relationships (Zvoleff & An, 2014). A workshop hosted by the US National Academy of Sciences identified geospatial data collection techniques and interdisciplinary data integration, including data derived from participatory methods, as a key research area for the National Geospatial Intelligence Agency (NRC, 2010). Advances in mobile technologies, data sharing, and low cost sensors are rapidly shifting data generation paradigms within Geographic Information Science (Elwood, Goodchild, & Sui, 2012; Elwood, 2009). This shift offers opportunities to increase efficiency and affordability to address data needs and analytical challenges, promoting innovative data visualization and dissemination to both researchers and the public, and applying integrative analytics to identify correlations in human-landscape interactions.

Methodological and analytical tools from geographic information science (GIScience) can contribute to linking natural and social science to answer questions about human-environment interactions (Zvoleff & An, 2014). Specific to visitor use within PAs, spatial data can incorporate biophysical and social data to quantify and proactively monitor visitor use and environmental conditions for conservation goals. Thus, monitoring and research can be prioritized to critical periods (Hadwen, Arthington, Boon, Taylor, & Fellows, 2011), locations (Barros, Pickering, & Gudes, 2014), or identify changes in visitation patterns potentially related to dynamic ecological or landscape characteristics (Buckley & Foushee, 2012).

Research Objectives

This dissertation follows the conceptual framework illustrated in Figure 1. To contribute to the goal of biodiversity conservation through PAs, where human (i.e., visitor) use generates support as well as resource pressures, understanding human-environment interactions and promoting sustainable use-conservation balance relies on multi- and interdisciplinary information. Increasingly, geospatial tools bridge and integrate natural and social science theories, methods, and analytics, offering opportunities to reduce data collection barriers, and build a strong foundation to answer enduring and evolving research questions salient to balancing use and biodiversity conservation.

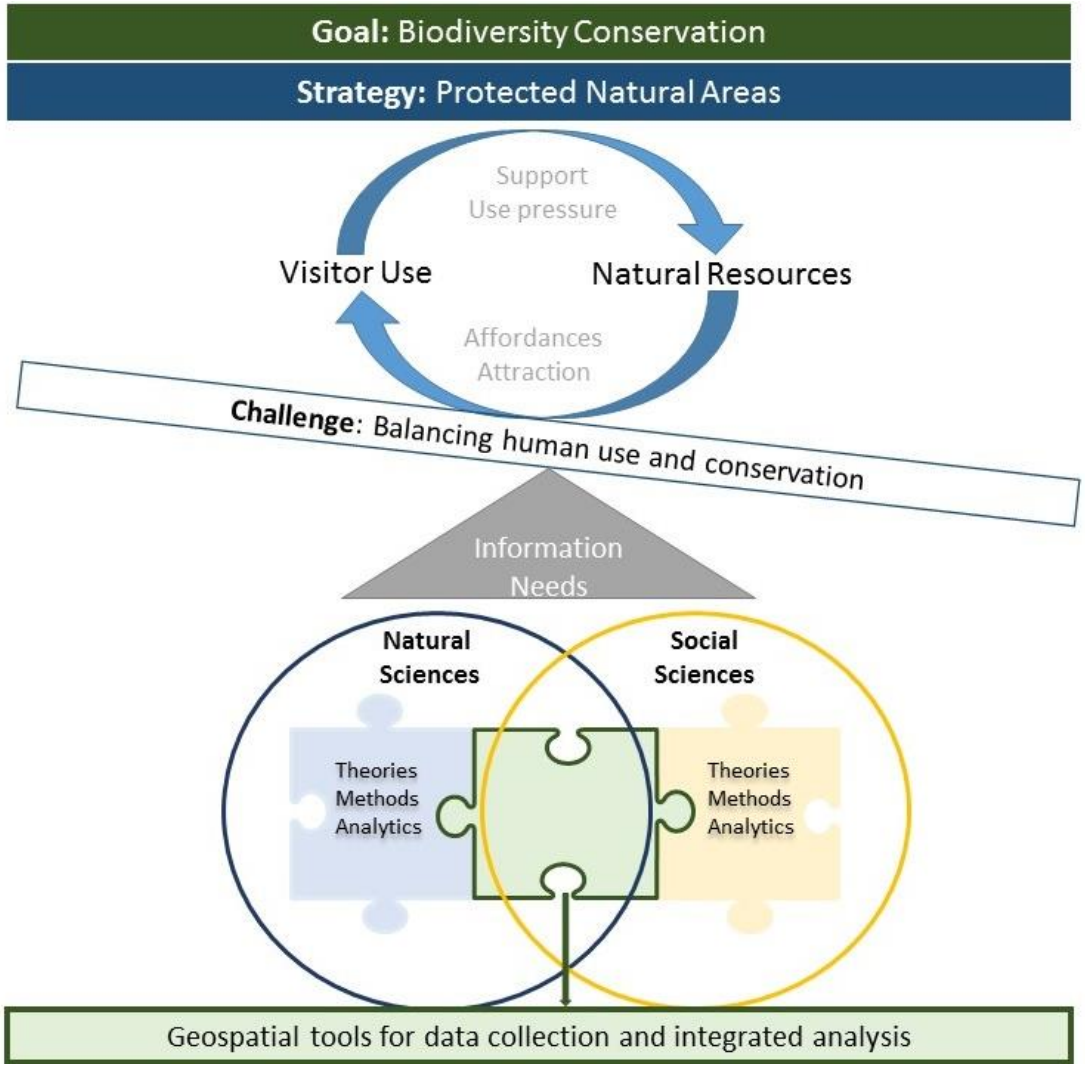


Figure 1 Conceptual diagram illustrating the use of geospatial approaches to integrate natural and social science methods and analytics for sustainable human use and biodiversity conservation within protected areas.

Specifically, this research:

1. Integrates geospatial methods to generate and analyze data about use patterns and behavior;
2. Explores how infrastructure and natural environmental factors influence use patterns at different spatial and temporal scales; and
3. Leverages crowdsourced geographic information to minimize barriers to visitor distribution modelling.

Dissertation Organization

Written as manuscripts prepared for peer-reviewed publication, the following three chapters provide case studies addressing identified objectives, followed by a general conclusion. Each manuscript conforms to the organization, style, and citation format of the target journal.

Chapter 2 addresses questions about pack stock behavior in PA wilderness. Many PAs allow for pack stock, which support private, commercial, and administrative operations. Research on pack stock grazing in PAs has identified myriad plant, soil, and water impacts. However, systematic assessments of pack stock behavior, which can influence the spatial patterns and intensity of impact, are rare. Therefore, the first paper couples global positioning system (GPS) technology, direct observation, and an exploratory movement modelling approach to identify behaviors of pack stock when released overnight to graze with limited

human oversight. Results highlight the applicability of the coupled approach to identify potentially impactful behaviors and inform successful impact mitigation strategies.

Building on the preceding chapter, Chapter 3 explores how environmental and managerial factors influence the distribution of pack stock animals during overnight grazing periods. Integrating geospatial data about environmental characteristics and management tactics (e.g., release point locations, fencing), this paper applies a distribution modelling approach appropriate for spatially and temporally autocorrelated movement data to identify which factors most influenced pack stock distribution patterns. Findings contribute to empirical assessments of how pack stock respond to managerial and environmental factors and inform management tactics designed to mitigate impacts associated with pack stock use.

Chapter 4 examines crowdsourced volunteered geographic information as a means to minimize data collection barriers at a park scale. Distribution models created using geotagged photos accessible on the site Flickr explore seasonal and annual spatial patterns of use and identify which environmental and managerial features most influence visitor distributions. Outputs of this study provide insights on how changes in the environment and infrastructure may influence visitor use and can serve as an inexpensive monitoring approach complementary to other methods.

The final chapter provides general conclusions gleaned from the research findings and summarizes implications for PAs. It also enumerates how the three studies collectively contribute to identified research gaps and next steps for future research.

References

Akasaka, M., Kadoya, T., Ishihama, F., Fujita, T., & Fuller, R. A. (2016). Smart protected area placement decelerates biodiversity loss: A representation-extinction feedback leads rare species to extinction. *Conservation Letters*, 1–8. Advance online publication. doi: 10.1111/conl.12302.

Balmford, A., Green, J. M. H., Anderson, M., Beresford, J., Huang, C., Naidoo, R., … Manica, A. (2015). Walk on the wild side: Estimating the global magnitude of visits to protected areas. *PLoS Biology*, *13*(2), e1002074.

Barros, A., Pickering, C., & Gudes, O. (2014). Desktop analysis of potential impacts of visitor use: A case study for the highest park in the Southern Hemisphere. *Journal of Environmental Management*, *150*, 179–195.

Buckley, L. B., & Foushee, M. S. (2012). Footprints of climate change in US national park visitation. *International Journal of Biometeorology*, *56*, 1173–1177.

Burton, A. C., Huggard, D., Bayne, E., Schieck, J., Sólymos, P., Muhly, T., … Boutin, S. (2014). A framework for adaptive monitoring of the cumulative effects of human footprint on biodiversity. *Environmental Monitoring and Assessment*, *186*, 3605–3617.

Butchart, S. H. M., Walpole, M., Collen, B., van Strien, A., Scharlemann, J. P. W., Almond, R. E. A, … Watson, R. (2010). Global biodiversity: Indicators of recent declines. *Science*, *328*(5982), 1164–1168.

Castro, A. J., Martín-López, B., López, E., Plieninger, T., Alcaraz-Segura, D., Vaughn, C. C., & Cabello, J. (2015). Do protected areas networks ensure the supply of ecosystem services? Spatial patterns of two nature reserve systems in semi-arid Spain. *Applied Geography*, *60*, 1–9.

Chape, S., Harrison, J., Spalding, M., & Lysenko, I. (2005). Measuring the extent and effectiveness of protected areas as an indicator for meeting global biodiversity targets. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *360*(1454), 443–455.

Chapin, F. S., Zavaleta, E. S., Eviner, V. T., Naylor, R. L., Vitousek, P. M., Reynolds, H. L., … Díaz, S. (2000). Consequences of changing biodiversity. *Nature*, *405*, 234–242.

Convention on Biological Diversity (2017). Aichi Biodiversity Targets. Retrieved April 10, 2017 from https://www.cbd.int/sp/targets/.

Crutzen, P. J. (2002). Geology of mankind. *Nature*, *415*, 23.

DeFries, R., Rovero, F., Wright, P., Ahumada, J., Andelman, S., Brandon, K., … Liu, J. (2010). From plot to landscape scale: Linking tropical biodiversity measurements across spatial scales. *Frontiers in Ecology and the Environment*, *8*(3), 153–160.

Dudley, N. (Ed.). (2008). *Guidelines for applying protected area management categories*. Gland, Switzerland: IUCN.

Elwood, S. (2009). Geographic information science: Emerging research on the societal implications of the geospatial web. *Progress in Human Geography*, *34*, 349–357.

Elwood, S., Goodchild, M. F., & Sui, D. Z. (2012). Researching volunteered geographic information: Spatial data, geographic research, and new social practice. *Annals of Association of American Geographers*, *102*, 571–590.

Hadwen, W. L., Arthington, A. H., Boon, P. I., Taylor, B., & Fellows, C. S. (2011). Do climatic or institutional factors drive seasonal patterns of tourism visitation to protected areas across diverse climate zones in eastern Australia? *Tourism Geographies*, *13*(2), 187–208.

Hadwen, W. L., Hill, W., & Pickering, C. M. (2008). Linking visitor impact research to visitor impact monitoring in protected areas. *Journal of Ecotourism*, *7*(1), 87–93.

Hammitt, W. E., Cole, D. N., & Monz, C. (2015). *Wildland recreation: Ecology and management* (3rd ed.). Chichester, UK: John Wiley & Sons, Ltd.

Heller, N. E., & Zavaleta, E. S. (2009). Biodiversity management in the face of climate change: A review of 22 years of recommendations. *Biological Conservation*, *142*, 14–32.

Herr, A. (2007). Data integration issues in research supporting sustainable natural resource management. *Geographical Research*, *45*, 376–386.

Joppa, L. N., Visconti, P., Jenkins, C. N., & Pimm, S. L. (2013). Achieving the Convention on Biological Diversity's goals for plant conservation. *Science*, *341*, 1100–1103.

Keniger, L., Gaston, K., Irvine, K., & Fuller, R. (2013). What are the benefits of interacting with nature? *International Journal of Environmental Research and Public Health*, *10*, 913–935.

Marion, A. J., Leung, Y.-F., Eagleston, H., & Burroughs, K. (2016). A review and synthesis of recreation ecology research findings on visitor impacts to wilderness and protected

natural areas. *Journal of Forestry*, *114*, 1–17.

McClintock, B. T., Johnson, D. S., Hooten, M. B., Ver Hoef, J. M., & Morales, J. M. (2014). When to be discrete: The importance of time formulation in understanding animal movement. *Movement Ecology*, *2*, 1–14.

McConnell, W. J., Millington, J. D. A., Reo, N. J., Alberti, M., Asbjornsen, H., Baker, L. A., … Pontius, R. G. (2011). Research on coupled human and natural systems (CHANS): Approach, challenges, and strategies. *Bulletin of the Ecological Society of America*, *92*, 218–228.

Millenium Ecosystem Assessment. (2005). *Ecosystems and human well-being: Biodiversity synthesis*. Washington, DC: World Resources Institute.

Monz, C., & Leung, Y.-F. (2006). Meaningful measures: Developing indicators of visitor impact in the National Park Service Inventory and Monitoring program. *George Wright Forum*, *23*(2), 17–27.

National Research Council. (2010). *Understanding the changing planet: Strategic directions for the geographical sciences*. Washington, DC: The National Academies Press.

Pimm, S. L., Jenkins, C. N., Abell, R., Brooks, T. M., Gittleman, J. L., Joppa, L. N., … Sexton, J. O. (2014). The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, *344*, 1246752-1 - 1246752-10.

Puhakka, R., Pitkänen, K., & Siikamäki, P. (2016). The health and well-being impacts of protected areas in Finland. *Journal of Sustainable Tourism.* Advance online publication. doi: 10.1080/09669582.2016.1243696.

Spenceley, A. (2015). Tourism and the IUCN World Parks Congress 2014. *Journal of Sustainable Tourism*, *23*, 1114–1116.

Steffen, W., Broadgate, W., Deutsch, L., Gaffney, O., & Ludwig, C. (2015). The trajectory of the Anthropocene: The Great Acceleration. *The Anthropocene Review*, *2*, 81–98.

Steffen, W., Persson, Å., Deutsch, L., Zalasiewicz, J., Williams, M., Richardson, K., … Svedin, U. (2011). The Anthropocene: From global change to planetary stewardship. *Ambio*, *40*, 739–761.

Stolton, S. (2010). *Arguments for protected areas: Multiple benefits for conservation and use*. London, UK: Earthscan.

Sullivan, C. A., & Meigh, J. (2007). Integration of the biophysical and social sciences using

an indicator approach: Addressing water problems at different scales. *Water Resources Management*, *21*, 111–128.

Thomas, C. D., Gillingham, P. K., Bradbury, R. B., Roy, D. B., Anderson, B. J., Baxter, J. M., … Hill, J. K. (2012). Protected areas facilitate species' range expansions. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 14063–14068.

UNEP-WCMC and IUCN. (2016). *Protected planet report 2016*. Cambridge, UK and Gland, Switzerland: UNEP-WCMC and IUCN.

Watson, J. E. M., Darling, E. S., Venter, O., Maron, M., Walston, J., Possingham, H. P., … Brooks, T. M. (2016). Bolder science needed now for protected areas. *Conservation Biology*, *30*, 243–248.

Watson, J. E. M., Dudley, N., Segan, D. B., & Hockings, M. (2014). The performance and potential of protected areas. *Nature*, *515*(7525), 67–73.

Zvoleff, A., & An, L. (2014). Analyzing human-landscape interactions: Tools that integrate. *Environmental Management*, *53*, 94–111.

CHAPTER 2

Integrating direct observation and GPS tracking to classify animal behavior in natural areas

Abstract

1. Understanding the behavior of pack animals in protected areas informs management of the associated negative impacts. However, systematic assessments of pack animal behavior are uncommon due to methodological and logistical constraints.
2. This study integrates behavior mapping and GPS tracking, and it applies behavior change point analysis to explore behaviors of pack animals at two temporal scales during overnight periods of free roaming.
3. The integrated approach identified multiple grazing patterns not feasible through a single methodology alone. The grazing patterns may influence impacts and associated management interventions. Specifically, locally intense grazing was differentiated from ambulatory grazing.
4. Results also indicated pack animals exhibited some level of activity throughout much of the overnight release period, which confirms past research indicating that free ranging equids spend upwards of 80% of their time foraging. This finding suggests potential impacts to the environment occurring during overnight hours most likely result from foraging and movement behaviors rather than passive behaviors.

5. *Synthesis and applications*. Identifying behaviors and corresponding environmental conditions, specifically the different patterns of grazing behavior, aids managers in implementing tactics designed to mitigate impacts associated with pack animals in natural areas. The integrated approach reduced time and logistical constraints of each method individually, and highlighted how multiple management tactics could reduce impacts to sensitive habitats.

Keywords

Behavior change point analysis (BCPA), grazing impacts, horse, mule, Yosemite National Park

Introduction

Understanding the spatial and temporal scales of ecological phenomena is fundamental to both basic and applied ecological research, (Chave, 2013; Levin, 1992). In animal and movement ecology, for example, technological advances increasing data resolution and computing capacity have facilitated rigorous explorations of the spatial and temporal scales of animal movement (Avgar, Mosser, Brown, & Fryxell, 2013; Fryxell et al., 2008; Nathan et al., 2008). Animal movement underlies many ecological phenomena and consists of complex, multidimensional behavioral structure associated with a wide range of

stimuli, constraints, and pressures operating at multiple scales (Cagnacci, Boitani, Powell, & Boyce, 2010; Gurarie & Ovaskainen, 2011). Increased observational capacity also enables examinations of how spatial and temporal sampling and analytical decisions influence the scope of animal movements observed (Laube & Purves, 2011; Postlethwaite & Dennis, 2013; Wiens, 1989). Understanding both the ecological and analytical dynamics of scale sheds light on behaviors like resource selection, intra- and interspecies interactions, migration, dispersal, and anthropogenic pressure effects important to many sub-disciplines of ecology (Berger-Tal et al., 2016; Foster-Smith & Evans, 2003; Levin, 1992).

Effective management of protected areas (PAs) requires knowledge of applied ecology, including animal behavior of wildlife and domestic animals. PAs are the main strategy for conserving biodiversity, supporting several Convention on Biological Diversity Aichi Targets (Cardinale et al., 2012; Chape, Harrison, Spalding, & Lysenko, 2005; Thomas et al., 2012; Watson et al., 2014). PAs also provide natural and cultural ecosystem services, including opportunities for outdoor recreation and tourism (Balmford et al., 2015; Hoehner et al., 2010). In remote or mountainous PAs around the world, the use of domestic animals (e.g., horses, donkeys, mules, llamas, yaks) as pack stock is common (Barros & Pickering, 2015; Byers, 2009; Geneletti & Dawa, 2009; McClaran & Cole, 1993; Torn, Tolbanen, Norokorpi, Tervo, & Siikamaki, 2009). Pack animals support tourism operations and other needs of managing agencies, representing established forms of recreation part of the cultural heritage. However, pack animal use in PAs can be controversial because of the potential negative environmental impacts to resources and therefore compatibility with PA values and

mission (Ostoja et al., 2014). Impacts can include soil compaction, accelerated erosion, vegetation trampling, changes in plant species composition, increases in bare ground, reductions in streambank stability, and facilitating the spread of invasive species (Barros, Monz, & Pickering, 2015; Cole & Monz, 2004; McClaran & Cole, 1993). As with other types of tourism and recreational activities, impacts vary based on environmental conditions and behavior (Hammitt et al., 2015).

The location and type of behavior in which pack animals engage has both resource and management implications. For example, horses and mules can graze while stationary or moving (Olson-Rutz, Marlow, Hansen, Gagnon, & Rossi, 1996; Ransom & Cade, 2009), resulting in different spatial patterns of impacts. The duration of behaviors is also of interest because sustained grazing could result in different intensities of resource impacts than episodic grazing punctuated by stationary or resting periods (Olson-Rutz et al., 1996). Yet, systematic assessments of stock behavior, especially when released to graze overnight when observation is not feasible, are rare (Ostoja et al., 2014). The purpose of this study is to present an integrated approach for exploring the behavioral structure of pack animal movement. We integrate a direct observation method adapted for natural areas with global positioning system (GPS) tracking during periods when pack animals roamed freely. This study defines behavior broadly as animal's activity (i.e., foraging, resting) within an environment rather than behavioral traits or specific levels of physical activity. Results explore the influence of temporal scale on behavior segmentation and inform decisions for managing pack animal use and impacts in natural areas.

INTEGRATING DIRECT OBSERVATION AND GPS TRACKING

Direct observation collects fine-scale information on animal behavior and resource utilization (Rutter, 2007) and can inform behavior classification models (Homburger, Schneider, Hilfiker, & Lüscher, 2014), yet the time required, amount of daylight, and terrain can limit applicability and scope. These constraints highlight the need to integrate observation with another data source. Studies employing GPS tracking to study animal movements and distributions over longer periods and distances have greatly increased in the last decade, driven largely by improvements in spatial and temporal data resolution, reduced device size, and increased battery longevity (Kays, Crofoot, Jetz, & Wikelski, 2015). GPS tracking obtains high frequency, accurate position data on animal locations within habitats over greater spatial and temporal scales than potentially feasible for field observers. Detailed movement data have been collected on a variety of terrestrial and marine animals, ranging in size from small birds to large mammals such as grizzly bears (*Ursos arctos*) and elk (*Cervus canadensis*) (Hebblewhite & Haydon, 2010; Kays et al., 2015).

GPS-derived position data alone, however, are limited in determining specific behavior beyond coarse segmentation (Löttker et al., 2009) or actual resource consumption. This is especially salient for grazing animals who change behaviors frequently within a small area and can exhibit complex and slow movement patterns (Homburger et al., 2014). Cagnacci and others (2010) suggest that fine-scale behavior details are only possible through additional methods. Researchers engaged in sensor-derived animal movement analysis still

find benefit in direct observation techniques to validate sensor-derived behavior classification (Purves, Laube, Buchin, & Speckmann, 2014). Hebblewhite and Haydon (2010) suggest "technology should not replace field biology" (p. 2306). Additionally, direct observation contributes and can verify data when canopy cover or landscape features cause GPS data errors or loss (Williams, Dechen Quinn, & Porter, 2012).

MODELLING ANIMAL MOVEMENT AND BEHAVIOR

The rapid increase in availability, quality, and complexity of movement data has led to a variety of analytical approaches employed to examine and classify behavior with no clear rules as to which approach to use (Börger, 2016; Gurarie et al., 2016; Postlethwaite, Brown, & Dennis, 2013). Movement data are correlated in space and time, multidimensional, and subject to sampling error (e.g., missed GPS fixes), complicating statistical analysis. Statistical models using direct behavior observation for training have included variants of random walk, time-series (Gurarie, Andrews, & Laidre, 2009), state-space (McClintock et al., 2014), and hidden Markov models (Patterson, Basson, Bravington, & Gunn, 2009).

Specific to grazing, Homburger and others (2014) evaluated eight models and machine learning algorithms. Comparisons included random forest classification, linear discriminant analysis, state-space models (SSM), and a kernel-based vector machine to classify GPS data of cattle into walking, grazing, and resting behaviors using observation data for training (Homburger et al., 2014). Common to the models is classifying the GPS track into movement metrics including step-lengths (i.e., Euclidean distance between two

successive fixes), turning angles, and in some instances, speed, to segment bouts of behavior when compared to an observed behavior using time as the connecting variable. The complementary benefits and limitations of direct observation and GPS tracking, along with advances in statistical analysis of autocorrelated movement data, offer an opportunity to build on research integrating the two methods using the shared attribute of time to address questions of overnight pack animal behavior.

Gurarie and others (2016) highlight that exploratory analysis of the behavioral structure is a critical analytical component, prior to any predictive movement modelling, to address questions about animal behavior. For highly correlated and irregularly sampled data, behavioral change point analysis (BCPA) often outperformed other methods, in both simulated and real animal tracks when compared to methods representing different analytical traditions: correlated random walk, first passage time, and Bayesian partitioning of Markov models (Gurarie et al., 2016). BCPA identifies changes in movement parameters (i.e., velocity, persistence velocity) by sweeping a window over a GPS track to identify significant change points in the structure of the movement. BCPA can be computational efficient and does not require a priori assumptions about movement or pre-determined number of behavior states as compared in other methods (Gurarie et al., 2009, 2016; Zhang, O'Reilly, Perry, Taylor, & Dennis, 2015). Additionally, BCPA provides an estimate of the significance of the changes and, because it is a parametric approach, distributional assumptions can be tested by examining residuals (Edelhoff, Signer, & Balkenhol, 2016). Subtle behavior transitions, however, challenge BCPA segmentation if combinations of derived movement parameters

obscure differences in others, or if parameters remain similar yet environmental factors influence behavior change (Gurarie et al., 2016). BCPA also does not directly categorize individual data points into a behavior state, but rather identifies bouts of behavior. To assign each data point to a behavior state requires additional analysis like k-means clustering (Zhang et al., 2015).

Materials and Methods

STUDY AREA

Yosemite National Park (YNP) in California, a UNESCO World Heritage site, is home to over 400 species of vertebrates, including two potentially endemic species, and five biotic zones ranging from 600 to nearly 4,000 m in elevation. In 2015, YNP attracted 4.1 million recreational visits (USDI NPS, 2016a). Ninety-five percent of the 3,027-km^2 park is federally-designated wilderness, intended to preserve natural conditions, prohibit motorized vehicles and equipment, and provide unconfined recreation opportunities. Pack animals are allowed in YNP, including wilderness areas, for private, commercial, and NPS administrative purposes, with an average of 1,571 stock nights per year reported between 2004 and 2012 (Ostoja et al., 2014).

Pack animals are often released overnight in wilderness meadows to graze. Meadows account for 3% of total park area, yet support disproportionate amounts of plant, animal, and invertebrate biodiversity, and attract visitors due to their aesthetic and recreational appeal (Ballenger et al., 2011; Viers et al., 2013). Assessment of pack animal impact, specifically in

wilderness meadows, is part of the park's ongoing monitoring efforts associated with several management plans (Ballenger et al., 2011; Kuhn, Ballenger, Scherer, & Williams, 2015).

DATA COLLECTION

To support wilderness management and pack animal use monitoring in YNP, North Carolina State University researchers and NPS staff collaboratively developed integrative data collection procedures. Following pilot testing, NPS staff collected geospatial data from GPS tracking and behavior mapping of administrative pack animals in montane (i.e., above 900 m a.s.l) and sub-alpine (i.e., above 2450 m a.s.l) meadows within designated wilderness areas of the park between June and September 2014. A purposive sampling strategy identified meadows experiencing all types of pack animal use allowed in the park (i.e., administrative, commercial, private); representing all biotic zones in wilderness areas of the park that receive pack animal use; and complementing existing pack animal-related research, travel logistics, and accessibility.

When released to graze, pack animals wore custom, professional-grade GPS receivers, developed by Oregon State University, placed inside a protective Pelican 1010 case attached to the animal's halter. The combined apparatus weighed less than 1 kg, well under the threshold of less than 5% of the animal's body weight advocated by animal movement researchers to minimize the receiver's effect on animal welfare and behavior (Kays et al., 2015). GPS receivers collected positional fixes every second for the duration of the release period, which ranged from four to 13 hours. GPS positions were differentially

corrected in real time to improve accuracy. The one second interval facilitated integration with behavior observation data, provided the temporal granularity to differentiate behaviors into at least slow and fast states (Postlethwaite & Dennis, 2013), and afforded the opportunity to explore different temporal resolutions on results.

NPS staff concurrently conducted one-hour behavior mapping sessions during daylight hours of pack animal release. Behavior mapping is a direct observation method designed to capture the actual use of space by pairing a spatial location with an observed individual's behavior. This study adapted procedures from behavior mapping applied in open, natural settings (Walden-Schreiner & Leung, 2013). Observations occurred at least 25 m away from the nearest animal so as not to influence behavior. The 25-m threshold was identified through discussions with NPS mounted Wilderness Patrol Rangers and other staff familiar with the animals (T. Kuhn, personal communication). Observers selected locations offering maximum visibility of all animals meeting conditions outlined above. Several observation points throughout the meadows mitigated instances of obscured views (Figure 1).

Figure 1 Observers collecting behavior mapping data on pack animal behavior. Animals are wearing GPS receivers (yellow) attached to their halter. Photo courtesy of T. Kuhn.

Observers employed a scan sampling approach that repeated continuously for the duration of the one-hour behavior mapping session. Scan sampling records the animal's behavior state at the moment of observation and was selected to maximize the number of records within the limited window for observations and number of animals observed. Behavior categories (Table1), identified during pilot testing and informed by equid behavior studies conducted by the U.S. Geological Survey (USGS), included grazing, loafing, moving with intent, drinking, rolling, and unknown (Ransom & Cade, 2009).

Table 1 Behavior definitions for direct observation of pack animal (adapted from Ransom & Cade, 2009).

Behavior Category	Definition
Grazing	Animal is actively consuming vegetation (e.g., biting and ingesting). May be moving while grazing.
Loafing	Animal is stationary or traveling at very slow rate of speed (e.g., meandering, lack of attention and relaxed); may be standing or laying down; resting
Moving with Intent	Animal is traveling at a speed fast or consistent enough to indicate motivation beyond foraging (e.g., travel between patches or to water). May include walking, trotting, cantering, galloping, jumping, or swimming.
Drinking	Animal is actively consuming water
Rolling	Animal is on the ground or in the water and actively rolling
Unknown	Behavior not certain (i.e., body partially obscured) or does not fit within other categories

DATA ANALYSIS

The temporal resolution of behavior and GPS data can determine appropriate modelling parameters and can influence results and computational capacity (Laube & Purves, 2006). As Gurarie and others (2016) indicate, although no firm rules exist due to the diversity of objectives, establishing parameters for exploratory behavior analysis should be informed by biology and research objectives. Applied to pack animals, BCPA models were fit to GPS data from receivers obtaining a fix rate success of 80% or greater following GPS data filtering. Data were filtered using techniques available in MoveBank (www.movebank.org), an online database designed to manage and analyze movement data. Specifically, filtering excluded points that indicated the animal was traveling at an unlikely speed. Informed by research on horse gait analysis, a threshold of 13 meters per second (mps), the approximate

mid-point of speeds achieved by a galloping horse, was used to remove outliers (Huang et al., 2013). . Fix rate success was calculated using the number of records obtained divided by number of records possible after data filtering. This threshold ensured robust sample sizes for analysis and subsampling, less subject to bias resulting from habitat-related data loss (Laube & Purves, 2011).

From the eligible receivers, tracks from two locations were chosen because they represented an overnight release of at least 12 hours, in which at least three receivers achieved the 80% fix rate success threshold, and included at least an hour of direct behavior observations. The two meadows were of similar size and represented different environmental conditions, therefore potentially affording different behaviors. Hook Lake is a 3.36 ha subalpine meadow at 2,860 m a.s.l. and Lake Vernon is a 3.84 ha mixed conifer forest at 2,002 m a.s.l. Animal activity and grazing levels prior to release were similar for both locations.

Using code written for the statistical software R (version 3.1.2) provided by Gurarie and others (2016), BCPA was applied to the first hour of the release when animals were being observed directly as well as to the full GPS track (i.e., between 12 and 13 hours) to examine how changing the temporal analysis scale influenced behavior bouts identified. Averaging data over two minutes smoothed the GPS track because the occasional limited mobility of the animals could make small errors significant (Gurarie et al., 2016). The two-minute window was based on the average amount of time between successive behavior observations and previous analysis using First Passage Time (FPT) to identify clear visual

separation between behavior bouts without data loss. The same model parameters were applied to all tracks and based on recommendations in the literature (Gurarie et al., 2009, 2016; Zhang et al., 2015). Specifically, the BCPA modeled the persistence velocity (PV) of the animals (i.e., scalar velocity multiplied by the cosine of the turning angle) using a continuous-time autoregressive process across the time series, and obtained a likelihood for three parameters: mean PV, standard deviation, and temporal autocorrelation.

A window of a fixed number of records, in this instance 30, swept over the data at one step increments to identify locations of significant changes in any of the three parameters using a modified Bayesian Information Criterion (BIC). A 30-record window was selected to meet minimum sample size requirements for model selection and identify fine-scale behavior changes. The modified BIC chooses which parameters best model the change in the three parameters to locate the most likely change points within each window. The sensitivity threshold, the modification to the BIC equation, allows adjustments to the penalty associated with extra parameters. For this study, the sensitivity parameter was reduced from a default value of 2 to 1, encouraging more conservative analysis, selection of simpler models, and compensating for the more sensitive window size. Outputs of the BCPA were modelled using a flat approach, meaning the window sweep sought the most commonly identified change points and estimated the parameter values between two change points (Gurarie et al., 2016). Our analysis clustered change points within three records to filter minor changes and minimize noise. Outputs of the first hour after releasing the animals and the full release

period BCPAs were integrated with behavior mapping results based on time stamp to characterize behavior phases.

Results

Pack animals wore active GPS receivers for 13 hours and 24 minutes (i.e., 18:52 on 27 Aug to 08:15 on 28 Aug) in Hook Lake and 11 hours and 55 minutes (i.e., 18:45 on 12 Aug to 06:40 on 13 Aug) in Lake Vernon, resulting in 116,680 and 88,164 total records for the three selected units, respectively. In Hook Lake, GPS fix rate success ranged from 26.95 to 97.1%. The three receivers selected exhibited fix rate success of 92.5% and above. In Lake Vernon, GPS fix rate success ranged from 29.91% to 97.3%, with the three receivers selected achieving fix rate success of 84.7% and above. Direct observation results from the three animals wearing collars with fix rate successes greater than 80% indicates animals spent over 75% of the time observed grazing, exhibiting few other behaviors (Table 2).

Table 2 Behaviors observed for the three animals wearing GPS receivers during the first hour after release

	Grazing	Moving with Intent	Rolling	Loafing	Drinking	Unknown
Hook Lake (n=60)	75%	18%	3%	3%	0%	0%
Lake Vernon (n=23*)	83%	13%	4%	0%	0%	0%

*Difference in sample size between the two locations is to the result of observers needing to frequently move to new observation locations in Lake Vernon because of environmental conditions obstructing views.

FIRST HOUR FOLLOWING ANIMAL RELEASE

The complexity and individuality of integrated behavior-BCPA outputs are presented visually for comparison (Figures 2-5), with diagnostic plots and persistence velocity details included in the Appendix A. BCPA outputs for the first hour of release identified multiple changes in mean persistence velocity, standard deviation, and/or time-scale of temporal autocorrelation, indicating frequent switching between behavior patterns for all animals in both locations. Vertical purple lines indicate identified change points, with the width of the lines related to the number of times a change point was deemed significant. The horizontal black lines indicate the mean PV for a phase, with the horizontal red lines providing a 95% confidence interval (i.e., plus or minus two standard deviations). The color of the data points visualizes the time scale of temporal autocorrelation (i.e., τ), with smaller values represented by cooler colors (i.e., blue) and greater values represented by warmer (i.e., orange) colors.

In Lake Vernon, mean PV across the three tracks ranged from 0.00 mps to 0.58 mps, with the number of change points ranging from 6 to 18 (Figure 2). For all units, greater values of τ occurred early in the release period compared with the end of the first hour, indicating less tortuous movements. A similar pattern emerged in Hook Lake, with a range of 0.01 to 0.61 mps and 2 to 27 change points (Figure 3). However, early movements were more tortuous for all animals in Hook Lake as compared to later in the release period.

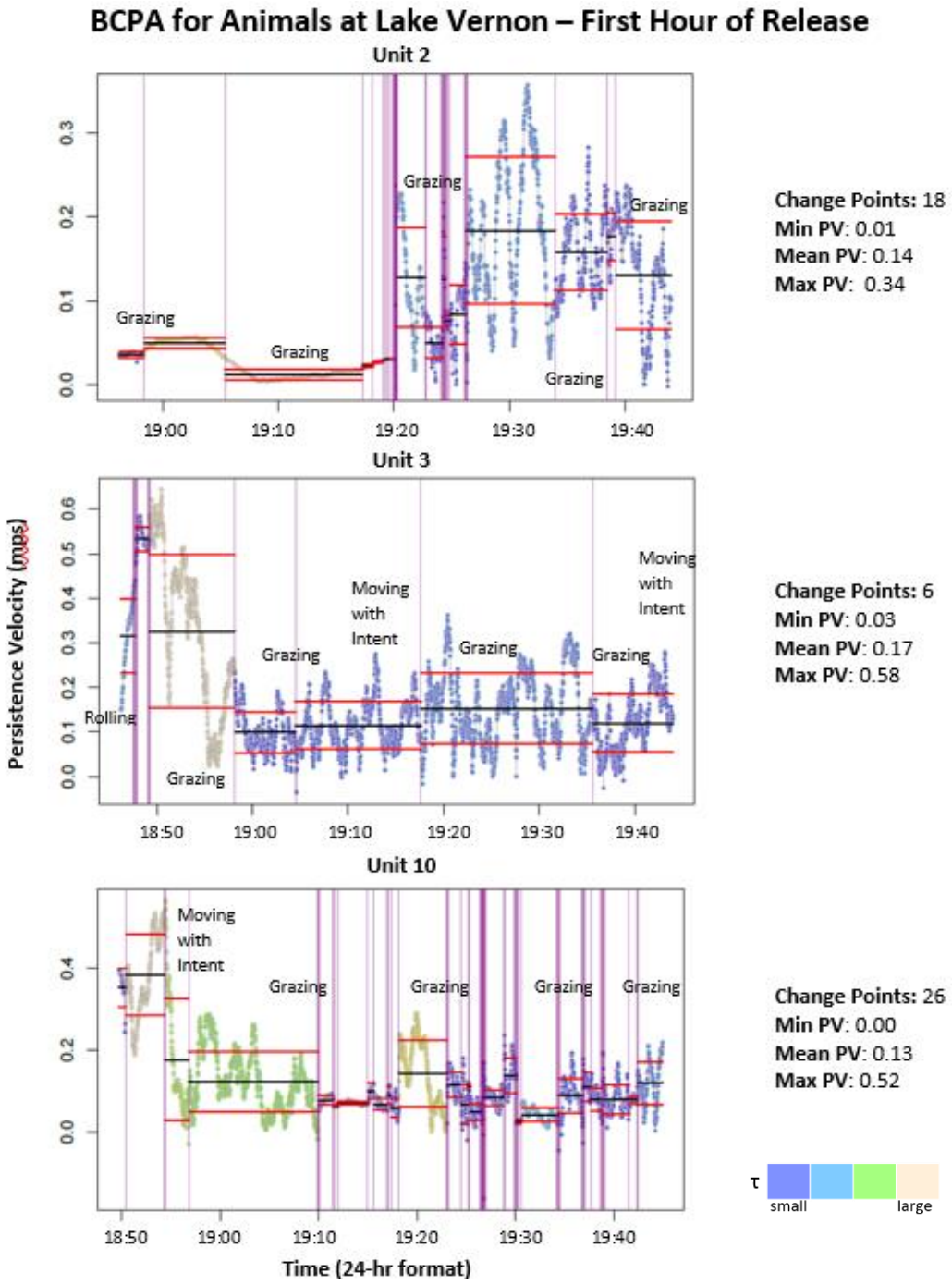


Figure 2 Flat behavior change point analysis (BCPA) output for three animals at Lake Vernon between 18:45 and 19:45 on August 12. Purple lines indicate behavior change points. Black lines indicate average persistence velocity for the behavior phase, with a 95% prediction interval indicated in red. Color indicates the time scale of temporal autocorrelation (i.e., τ), with cooler (i.e., blue) and warmer (i.e., orange) colors indicating smaller and larger values, respectively.

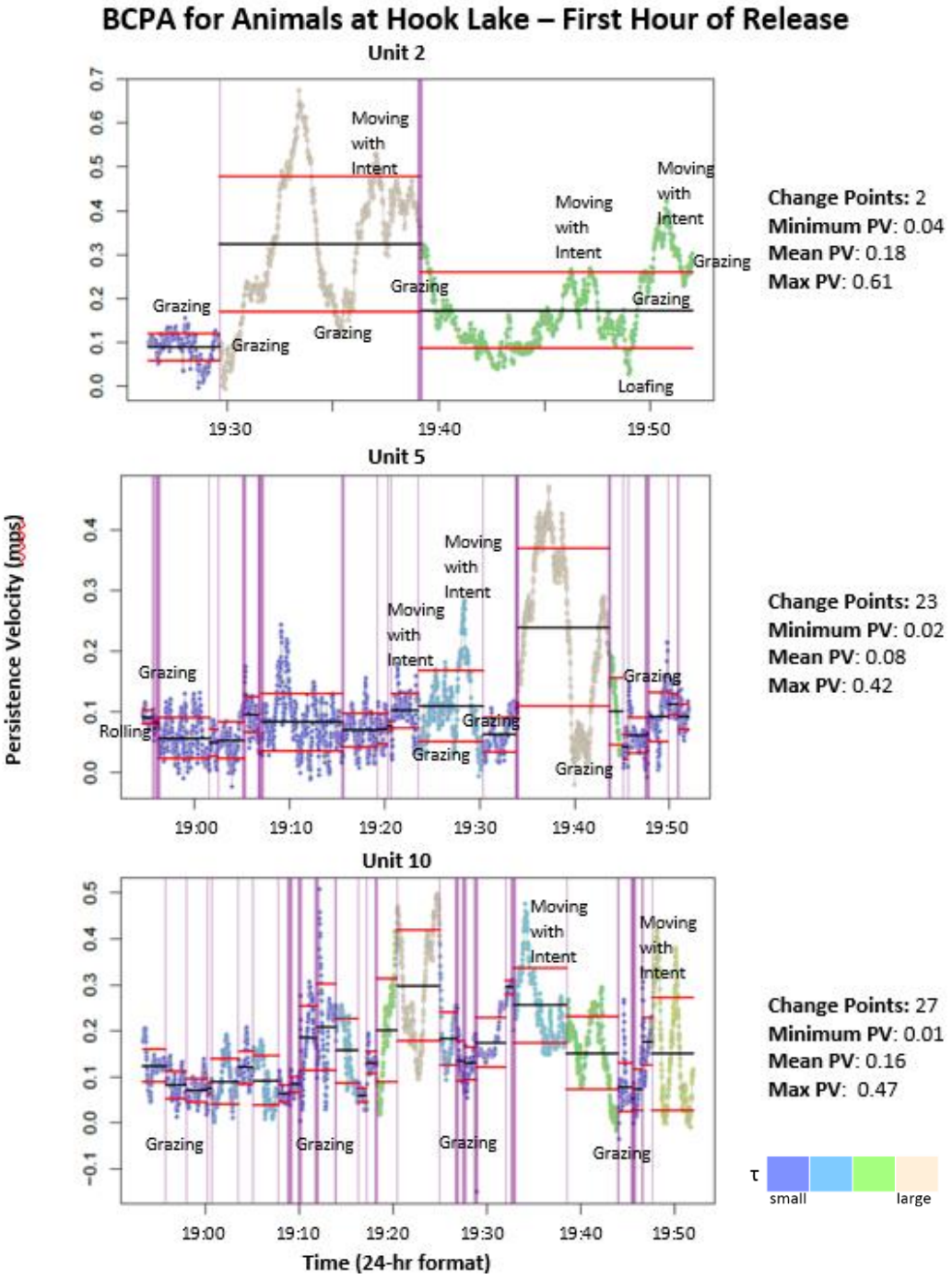


Figure 3 Flat behavior change point analysis (BCPA) output for three animals from 18:52 to 19:52 on August 27. Note, unit 2 was missing data from 18:45 to 19:20. Purple lines indicate behavior change points. Black lines indicate average persistence velocity for the behavior phase, with a 95% prediction interval indicated in red. Color indicates the time scale of temporal autocorrelation (i.e., τ), with cooler (i.e., blue) and warmer (i.e., orange) colors indicating smaller and larger values, respectively.

Although grazing was the predominant behavior witnessed during behavior observations, multiple grazing patterns emerged for all animals when integrating behavior data with BCPA outputs. In Lake Vernon, the animal wearing unit 2 was classified as grazing during all observations, however the BCPA shows an intense grazing period with low PV and variance, yet more autocorrelated from 18:45 to 19:20, followed by a period of grazing characterized by faster, more varied and less autocorrelated movements. This suggests more stationary or concentrated grazing was differentiated from other grazing patterns. The scan sampling of observations documented behavior, on average, every two minutes, thus potentially missing changes in behavior between observations.

In Hook Lake grazing again separated into multiple patterns noted in the integrated BCPA output, with all three animals beginning the release period with slower, more tortuous grazing patterns. Intermediate grazing patterns also emerged toward the end of the first hour for all three animals, as well as peaks of high PV values also identified periods when animals were observed moving with intent. The intermediate grazing patterns interspersed with times identified as moving with intent, though the BCPA did not always separate the two behaviors. For example, the animal wearing unit 5 was observed as moving with intent at 19:21, 19:24 and 19:26. Although the exact start and duration is not known from the behavior observations themselves, the first two observations fall within the same phase, separated from two phases of grazing. The third moving with intent observation was not separated from a grazing period, perhaps in part due to the use of the flat BCPA analysis, which assumes homogeneous behavior between change points. Few, if any, phases identified by the

integrated behavior-BCPA outputs suggested the animals were resting (i.e., loafing) during the first hour after release, characterized by very low PV and standard deviation. Collectively, in Lake Vernon, animal behavior for the three animals indicated faster, more linear movements immediately after release, followed by less temporally autocorrelated movements with lower PV and standard deviation values. This is in contrast to Hook Lake where the reverse pattern was observed during the first hour.

FULL RELEASE PERIOD

Modelling the time-series of the entire release period for both locations also illustrated frequent changes in movement as identified by the BCPA and inferred from the integrated behavior mapping and BCPA outputs from the first hour of release, with persistence velocities ranging from 0.00 mps to 2.00 mps in both meadows. Similar numbers of change points emerged for the full release period in comparison to the first hour, ranging from 12 to 15 in Lake Vernon and 7 to 17 in Hook Lake. For both meadows, BCPA identified phases of tortuous movements, with the exception of the animal wearing unit 5 in Hook Lake.

In comparison with the movement patterns classified with the assistance of the behavior observations in the first hour, the full release period was less straightforward given the overlap in PV values for the different behaviors and increased PV range. The maximum PV observed increased from 0.61 mps in the first hour to 2.00 mps during the full release, an increase of 227%. The increased temporal scale of analysis obscured finer changes in

movement behavior detectable in the first hour analysis, specifically the different types of grazing identified through the integration of the behavior observations, though instances of potential resting periods emerged.

In Lake Vernon, the flat BCPA peaks of PV for two of the three animals between 22:00 and 23:00, corresponded with the period when the animals (wearing units 2 and 3) moved outside the management-defined boundaries of the Lake Vernon area. Coupled with increased PV values for behaviors identified as moving with intent, these phases suggest motivation for more linear, faster movements (Figure 4). The animal wearing unit 2 appears to rest (i.e., very low PV and time scale of temporal autocorrelation) between 01:00 and 03:30, while the animal wearing unit 3 continues to exhibit moderate PV with lower time scale of temporal autocorrelation (i.e., more tortuous movements).

In Hook Lake, peaks in PV were observed around 03:00 for all three animals and again shortly before 06:00 for two animals (wearing units 2 and 10; Figure 5). The animal wearing unit 2 is missing data from 20:00 until approximately 23:30 and the animal wearing until 5 is missing data for the 06:00 hour, and neither appear to engage in extended periods of resting behavior throughout the entire release period. In general, for both locations, the animals exhibited slow movements with low temporal autocorrelation for the duration of the release period.

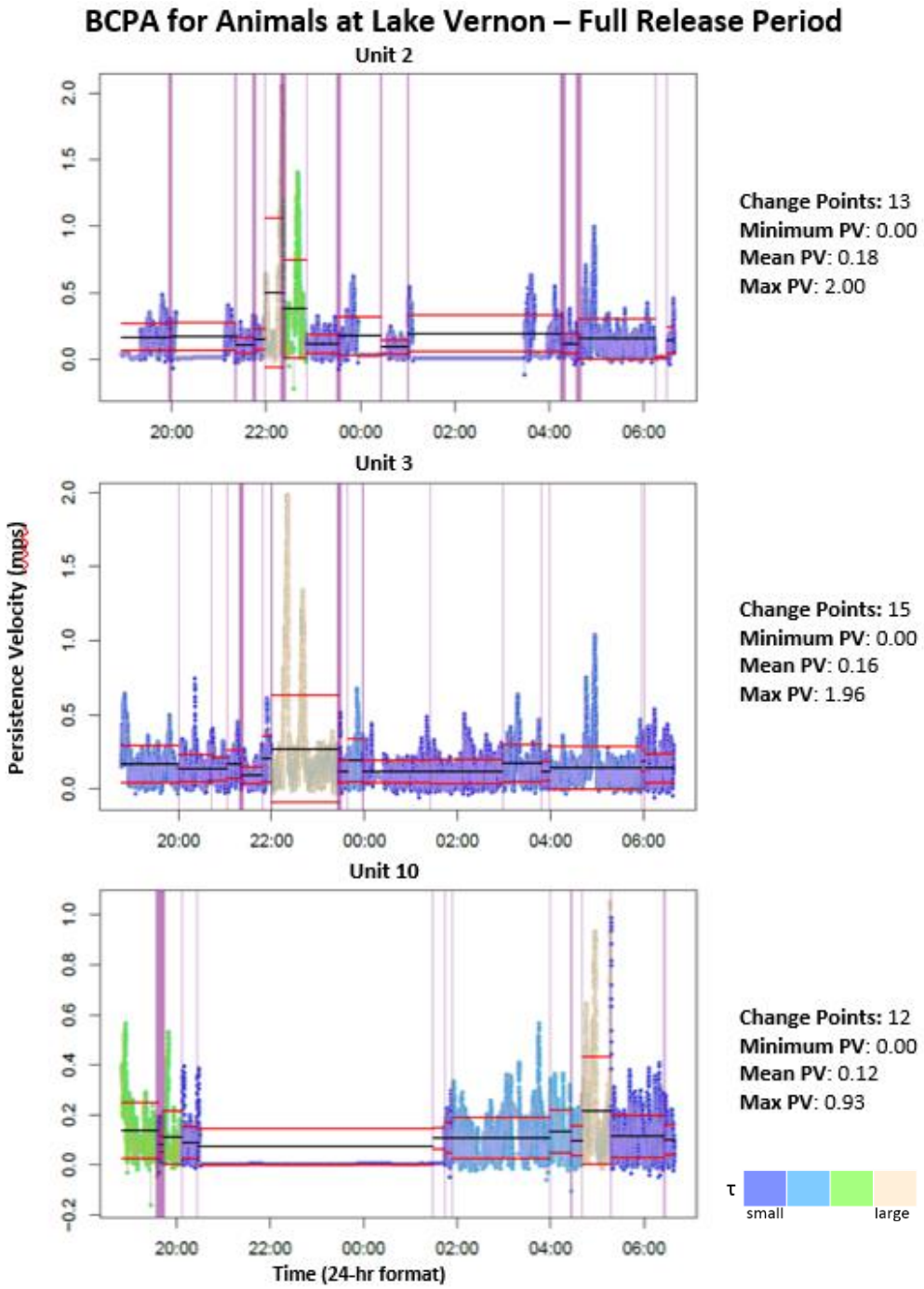


Figure 4 Flat behavior change point analysis (BCPA) output for three animals released at 18:45 on August 12 at Lake Vernon and retrieved at 06:40 on August 13. Purple lines indicate behavior change points. Black lines indicate average persistence velocity for the behavior phase, with a 95% prediction interval indicated in red. Color indicates the time scale of temporal autocorrelation (i.e., τ), with cooler (i.e., blue) and warmer (i.e., orange) colors indicating smaller and larger values, respectively.

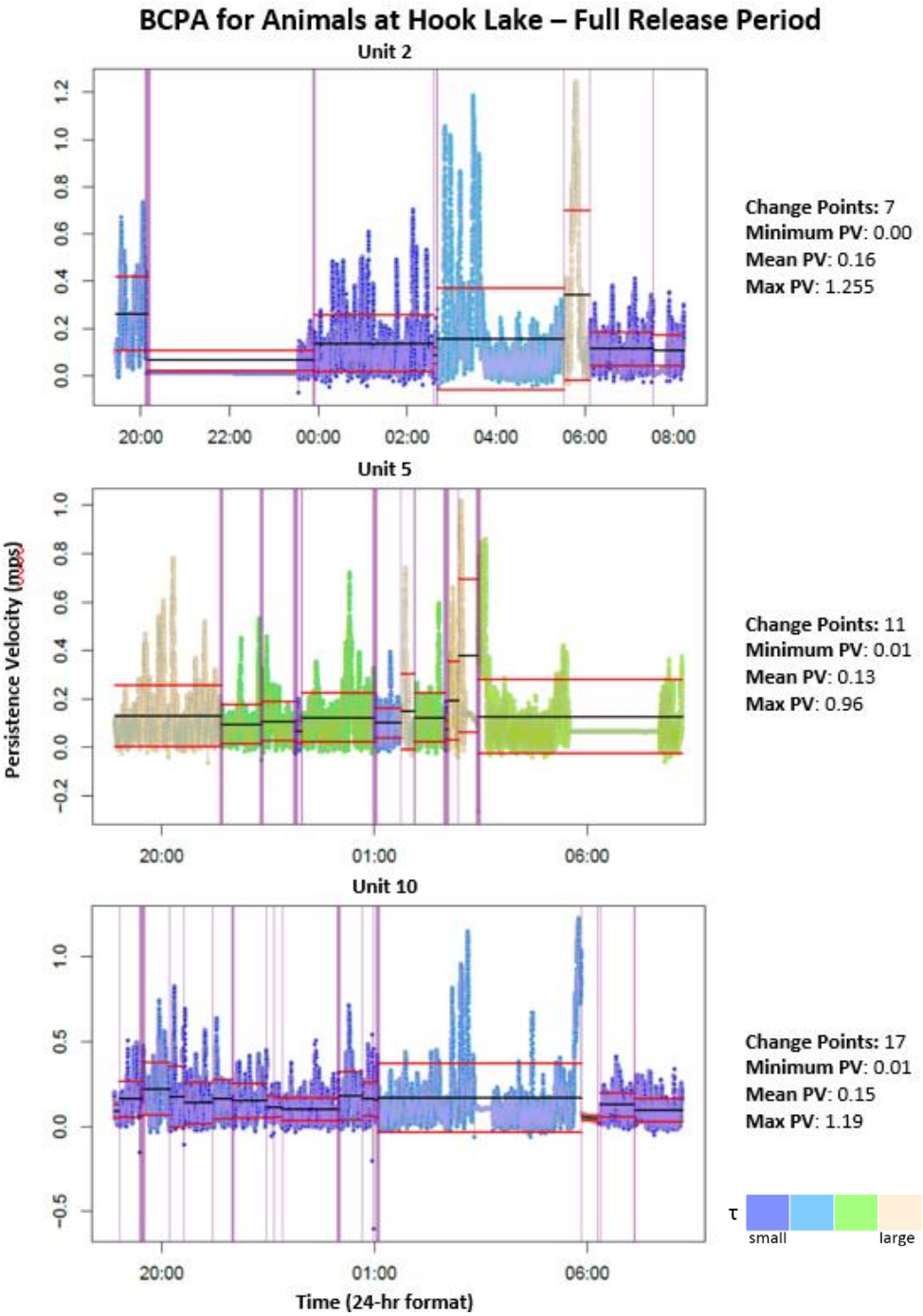


Figure 5 Flat behavior change point analysis (BCPA) output for three animals released at 18:52 on August 27 at Hook Lake and retrieved at 8:15 on August 28. Purple lines indicate behavior change points. Black lines indicate average persistence velocity for the behavior phase, with a 95% prediction interval indicated in red. Color indicates the time scale of temporal autocorrelation (i.e., τ), with cooler (i.e., blue) and warmer (i.e., orange) colors indicating smaller and larger values, respectively.

Discussion

This analysis integrated behavior mapping, GPS tracking, and BCPA to explore behaviors of pack animals within a limited area for impact monitoring and management. Outputs provided an efficient method to explore and visualize animal movements and behavior, including periods when direct observation was not feasible. Multiple movement patterns associated with grazing, a key finding, could be identified only by integrating the two methods. This finding suggests implementing multiple management tactics, like enclosure or exclosure fencing and rotational release points, to minimize impacts associated with different grazing behaviors.

Alone, BCPA outputs indicated animals moved at relatively low PVs, engaging in some movement through much of the full release period. Equids can remain alert for extended periods of time, though will rest, often in shifts with at least one animal remaining alert, if in a familiar, non-threatening location (McDonnell, 2003). Less familiar surroundings may contribute to less resting identified in the GPS tracks, as well as the staggered periods of rest with one animal remaining alert as seen in Lake Vernon. Additionally, limited variation in behavior was noted during direct observations, which limited the comparison of parameter values to several behaviors. However, results are consistent with research suggesting natural forage grazing equids spend the majority of their time feeding (Mills & McDonnell, 2005), accounting for 60 to 80% of their time (McDonnell, 2003).

Temporal resolution helped differentiate behaviors of slower moving species, specifically the high sampling frequency of GPS data and segmenting the BCPA outputs to examine the first hour separately. Outputs at this resolution provide managers with approaches to explore behavior changes under different environmental conditions. For example, comparing patterns from years with high precipitation totals to years with drought would explore the influence of soil moisture and forage quality, especially in semi-arid grasslands where plant growth is already limited by precipitation.

The narrow range of PV challenged the identification of some behaviors within the full release BCPA for which direct observations were not feasible. Overlapping movement patterns and parameter values were identified for the three main behaviors observed – loafing, grazing, and moving with intent. Yet, the fine scale enabled visual separation into different behavior states. This finding parallels those of other studies of grazing animal movement, where 20-second intervals provided granularity capable of visual segmentation (Homburger et al., 2014). In this study, the estimated duration of behaviors ranged between 14 and 33 seconds for the first hour and 8 to 12 seconds for the full release in both locations. As suggested by Postlethwaite and colleagues (2013), approximate behavior time scales matched the sampling interval. As a result, we propose maintaining a high sampling frequency.

Overlapping patterns are also supported by the fact that horses and mules graze both while stationary or in motion, and rest either in a standing position or laying down, further complicating behavior segmentation. Identifying when an animal is stationary is also

complex given GPS accuracy (Laube & Purves, 2011). To minimize noise in the data, GPS tracks were smoothed prior to BCPA, and BCPA outputs averaged parameter values within a window sweep, clustering more minor change points. Future research should examine how the inclusion of additional sensors, like a mercury switch or accelerometer, could help further differentiate behavioral structure for animals exhibiting a limited range of PV, as well as the influence of animal characteristics (e.g., age, sex) and number of pack animals released in the same area. Infrared camera traps may also help increase the number of behaviors observed by extending the observation period into twilight hours. Specifically, capturing a series of photos when animals are present would provide a time-series for behavior classification.

MANAGEMENT IMPLICATIONS

Implications for resource management related to pack animals center on the finding that the animals remained active to varying degrees throughout the release period and exhibited multiple grazing patterns, suggesting areas of intense space use indicate concentrated grazing during the overnight hours. Focused grazing could lead to increased negative impact to resources within that area (Olson-Rutz et al., 1996a), influenced by quality and abundance of forage, and affect management tactics designed to influence distribution, or social dynamics. This is especially important in the riparian areas of arid and semi-arid landscapes where research suggests that excluding grazing during late summer months may contribute more to mitigating impacts than rotating grazing (Bailey & Brown, 2011).

Examining behavior before and after management action would further explore the applicability of using BCPA as a means to evaluate the effect and efficacy of management designed to reduce impact. Example tactics relevant to handling pack stock in wilderness include providing supplementary feed, adjusting the timing of grazing releases, and the use of portable electric fences to restrict the movement of the lead animal in the group and strategically releasing animals in locations near preferred use areas. Research suggests the use of enclosure fencing and release point placement can influence pack animal distribution (Walden-Schreiner, Leung, Kuhn, Newburger, & Tsai, 2017). Future applications should also map behavior to environmental conditions to inform how multifaceted management approaches could mitigate impacts informed by pack animal behavior.

Conclusion

This study highlighted the importance of temporal scale and resolution to behavior analysis, uncovering multiple grazing patterns in the first hour of release obscured when applying BCPA to the full release period. The different grazing patterns may imply varying levels of impact, especially when the environmental context is considered. Prolonged, stationary grazing in sensitive habitats may affect the conservation of other plant and animal species and therefore favor management tactics like exclosure fencing over changing the location at which animals are released to graze. However, identifying behaviors during the full release period through the integrated approach also highlighted the dynamic nature of pack animal behavior during periods when observation was not possible.

Exploring fine-scale, high temporal resolution and longer-term behavior patterns through integrated approaches affords opportunities to explore not only the influence of scale on animal behavior (Fryxell et al., 2008) but also to uncover multiple behavioral responses (Bejder, Samuels, Whitehead, & Gales, 2006). A focus on scale and data resolution is increasingly important as environmental and anthropogenic factors contribute to plant and animal range shifts (Chen, Hill, Ohlemüller, Roy, & Thomas, 2011; Lindström, Brown, Sisson, Phillips, & Shine, 2013; Moritz et al., 2008), influencing behavior and impacting communities and corresponding human activities. The integrated approach presented strives to minimize time and resource burdens for managers while providing information about behavior applicable to myriad species-impact concerns and highlights the importance of multiple scales of analysis.

Acknowledgements

This work was supported by the USDI National Park Service under the Cooperative Ecosystem Studies Unit Task Agreement P14AC01493. All applicable international, national, and/or institutional guidelines for the care and use of animals were followed.

The authors thank collaborators who provided time and expertise to this project, including NPS Resources Management and Science staff L. Austin, E. Babich, J. Baccei, M. Booher, A. Dickenson, L. Jones, M. Marschall, L. Mazzu, and A. Puchkoff and for efforts in the field and data management; NPS Wranglers T. Bernacchi, R. Colflesh, and R. Fredy for managing pack animals incorporated in this study; S. Ostoja from the USGS/USFS for

helpful feedback throughout the project; and D. Johnson, G. Alexander, and A. Day from Oregon State University for assembly and testing of custom GPS receivers.

References

Avgar, T., Mosser, A., Brown, G.S. & Fryxell, J.M. (2013) Environmental and individual drivers of animal movement patterns across a wide geographical gradient. *The Journal of Animal Ecology*, **82**, 96–106.

Bailey, D.W. & Brown, J.R. (2011) Rotational grazing systems and livestock grazing behavior in shrub-dominated semi-arid and arid rangelands. *Rangeland Ecology & Management*, **64**, 1–9.

Ballenger, L., Wilkin, K., Acree, L., Baccei, J., Whittaker, T. & Babich, E. (2011) *2010 Assessment of Meadows in the Merced River Corridor, Yosemite National Park*. USDI National Park Service, El Portal, CA.

Balmford, A., Green, J.M.H., Anderson, M., Beresford, J., Huang, C., Naidoo, R., Walpole, M. & Manica, A. (2015) Walk on the wild side: Estimating the global magnitude of visits to protected areas. *PLoS Biology*, **13**, e1002074.

Barros, A., Monz, C. & Pickering, C. (2015) Is tourism damaging ecosystems in the Andes? Current knowledge and an agenda for future research. *Ambio*, **44**, 82–98.

Barros, A. & Pickering, C.M. (2015) Impacts of experimental trampling by hikers and pack animals on a high-altitude alpine sedge meadow in the Andes. *Plant Ecology & Diversity*, **8**, 265–276.

Bejder, L., Samuels, A., Whitehead, H. & Gales, N. (2006) Interpreting short-term behavioural responses to disturbance within a longitudinal perspective. *Animal Behaviour*, **72**, 1149–1158.

Berger-Tal, O., Blumstein, D.T., Carroll, S., Fisher, R.N., Mesnick, S.L., Owen, M.A., Saltz, D., St. Claire, C.C. & Swaisgood, R.R. (2016) A systematic survey of the integration of animal behavior into conservation. *Conservation Biology*, **30**, 744–753.

Börger, L. (2016). Stuck in motion? Reconnecting questions and tools in movement ecology. *Journal of Animal Ecology*, **85**, 5–10.

Byers, A.C. (2009) A comparative study of tourism impacts on alpine ecosystems in the Sagarmatha (Mt. Everest) National Park, Nepal and the Huascaran National Park, Peru. *Ecotourism and Environmental Sustainability: Principles and Practice* (eds J. Hill & T. Gale), pp. 51–72. Routledge, London, UK.

Cagnacci, F., Boitani, L., Powell, R.A. & Boyce, M.S. (2010) Animal ecology meets GPS-

based radiotelemetry: A perfect storm of opportunities and challenges. *Philosophical Transactions of the Royal Society: Series B, Biological Sciences*, **365**, 2157–2162.

Cardinale, B.J., Duffy, J.E., Gonzalez, A., Hooper, D.U., Perrings, C., Venail, P., Narwani, A., Mace, G.M., Tilman, D., A.Wardle, D., Kinzig, A.P., Daily, G.C., Loreau, M., Grace, J.B., Larigauderie, A., Srivastava, D.S. & Naeem, S. (2012) Biodiversity loss and its impact on humanity. *Nature*, **489**, 326–326.

Chape, S., Harrison, J., Spalding, M. & Lysenko, I. (2005) Measuring the extent and effectiveness of protected areas as an indicator for meeting global biodiversity targets. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **360**, 443–455.

Chave, J. (2013) The problem of pattern and scale in ecology: What have we learned in 20 years? *Ecology Letters*, **16**, 4–16.

Chen, I.-C., Hill, J.K., Ohlemüller, R., Roy, D.B. & Thomas, C.D. (2011) Rapid range shifts of species associated with high levels of climate warming. *Science*, **333**, 1024–1026.

Cole, D.N. & Monz, C.A. (2004) Impacts of camping on vegetation: Response and recovery following acute and chronic disturbance. *Environmental Management*, **32**, 693–705.

Edelhoff, H., Signer, J. & Balkenhol, N. (2016) Path segmentation for beginners: An overview of current methods for detecting changes in animal movement patterns. *Movement Ecology*, **4**, 1–21.

Foster-Smith, J. & Evans, S.M. (2003) The value of marine ecological data collected by volunteers. *Biological Conservation*, **113**, 199–213.

Frair, J.L., Fieberg, J., Hebblewhite, M., Cagnacci, F., DeCesare, N.J. & Pedrotti, L. (2010) Resolving issues of imprecise and habitat-biased locations in ecological analyses using GPS telemetry data. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **365**, 2187–2200.

Fryxell, J.M., Hazell, M., Börger, L., Dalziel, B.D., Haydon, D.T., Morales, J.M., McIntosh, T. & Rosatte, R.C. (2008) Multiple movement modes by large herbivores at multiple spatiotemporal scales. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 19114–19119.

Geneletti, D. & Dawa, D. (2009) Environmental impact assessment of mountain tourism in developing regions: A study in Ladakh, Indian Himalaya. *Environmental Impact Assessment Review*, **29**, 229–242.

Gurarie, E., Andrews, R.D. & Laidre, K.L. (2009) A novel method for identifying behavioural changes in animal movement data. *Ecology Letters*, **12**, 395–408.

Gurarie, E., Bracis, C., Delgado, M., Meckley, T.D., Kojola, I. & Wagner, C.M. (2016) What is the animal doing? Tools for exploring behavioural structure in animal movements. *The Journal of Animal Ecology*, **85**, 69–84.

Gurarie, E. & Ovaskainen, O. (2011) Characteristic Spatial and temporal scales unify models of animal movement. *The American Naturalist*, **178**, 113-123.

Hammitt, W.E., Cole, D.N. & Monz, C. (2015) *Wildland Recreation: Ecology and Management*, 3rd ed. Wiley Blackwell, Chichester, UK.

Hebblewhite, M. & Haydon, D.T. (2010) Distinguishing technology from biology: A critical review of the use of GPS telemetry data in ecology. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **365**, 2303–2312.

Hoehner, C.M., Brownson, R.C., Allen, D., Gramann, J., Behrens, T.K., Floyd, M.F., Leahy, J., Liddle, J.B., Smaldone, D., Spain, D.D., Tardona, D.R., Ruthmann, N.P., Seiler, R.L. & Yount, B.W. (2010) Parks promoting physical activity: Synthesis of findings from interventions in seven national parks. *Journal of Physical Activity and Health*, **7**, S67–S81.

Homburger, H., Schneider, M.K., Hilfiker, S. & Lüscher, A. (2014) Inferring behavioral states of grazing livestock from high-frequency position data alone. *PloS One*, **9**, e114522.

Kays, R., Crofoot, M.C., Jetz, W. & Wikelski, M. (2015) Terrestrial animal tracking as an eye on life and planet. *Science*, **348**, aaa2478-1 - aaa2478-9.

Kuhn, T., Ballenger, L., Scherer, R. & Williams, J.N. (2015) *Data Analysis and Assessment of High Elevation Wilderness Meadows Surveyed from 2008 to 2011, Yosemite National Park Resources Management and Science*. USDI National Park Service, Fort Collins, CO.

Laube, P. & Purves, R.S. (2006) An approach to evaluating motion pattern detection techniques in spatio-temporal data. *Computers, Environment and Urban Systems*, **30**, 347–374.

Laube, P. & Purves, R.S. (2011) How fast is a cow? Cross-scale analysis of movement data. *Transactions in GIS*, **15**, 401–418.

Levin, S.A. (1992) The problem of pattern and scale in ecology: The Robert H. MacArthur

Award Lecture. *Ecology*, **73**, 1943–1967.

Lindström, T., Brown, G.P., Sisson, S.A., Phillips, B.L. & Shine, R. (2013) Rapid shifts in dispersal behavior on an expanding range edge. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 13452–13456.

Löttker, P., Rummel, A., Traube, M., Stache, A., Šustr, P., Müller, J. & Heurich, M. (2009) New possibilities of observing animal behaviour from a distance using activity sensors in gps-collars: An attempt to calibrate remotely collected activity data with direct behavioural observations in red deer *Cervus elaphus*. *Wildlife Biology*, **15**, 425–434.

McClaran, M.P. & Cole, D.N. (1993) *Packstock in Wilderness: Use, Impacts, Monitoring, and Management*. USDA Forest Service, Ogden, UT.

McClintock, B.T., Johnson, D.S., Hooten, M.B., Ver Hoef, J.M. & Morales, J.M. (2014) When to be discrete: The importance of time formulation in understanding animal movement. *Movement Ecology*, **2**, 1–14.

McDonnell, S.M. (2003) *The Equid Ethogram: A Practical Field Guide to Horse Behavior*. Eclipse Press, Lexington, KY.

Mills, D.S., & McDonald, S. (2005) *The Domestic Horse: The Origins, Development and Management of its Behavior*. Cambridge University Press, Cambridge, UK.

Moritz, C., Patton, J.L., Conroy, C.J., Parra, J.L., White, G.C. & Beissinger, S.R. (2008) Impact of a century of climate change on small-mammal communities in Yosemite National Park, USA. *Science*, **322**, 261–264.

Nathan, R., Getz, W.M., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D. & Smouse, P.E. (2008) A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 19052–19059.

Olson-Rutz, K.M., Marlow, C.B., Hansen, K., Gagnon, L.C. & Rossi, R.J. (1996) Packhorse grazing behavior and immediate impact on a timberline meadow. *Journal of Range Management*, **49**, 546–550.

Ostoja, S.M., Brooks, M.L., Moore, P.E., Berlow, E.., Blank, R., Roche, J., Chase, J. & Haultain, S. (2014) Potential environmental effects of pack stock on meadow ecosystems of the Sierra Nevada, USA. *The Rangeland Journal*, **36**, 411–427.

Patterson, T.A., Basson, M., Bravington, M. V & Gunn, J.S. (2009) Classifying movement behaviour in relation to environmental conditions using hidden Markov models. *The*

Journal of Animal Ecology, **78**, 1113–1123.

Postlethwaite, C.M., Brown, P. & Dennis, T.E. (2013) A new multi-scale measure for analysing animal movement data. *Journal of Theoretical Biology*, **317**, 175–185.

Postlethwaite, C.M. & Dennis, T.E. (2013) Effects of temporal resolution on an inferential model of animal movement. *PLoS ONE*, **8**, 1–12.

Purves, R.S., Laube, P., Buchin, M. & Speckmann, B. (2014) Moving beyond the point: An agenda for research in movement analysis with real data. *Computers, Environment and Urban Systems*, **47**, 1–4.

Ransom, J.I. & Cade, B.S. (2009) *Quantifying Equid Behavior: A Research Ethogram for Free-Roaming Feral Horses*. US Geological Survey, Reston, VA.

Rutter, S.M. (2007) The integration of GPS, vegetation mapping and GIS in ecological and behavioral studies. *Revista Brasileira de Zootecnia*, **36**, 63–70.

Thomas, C.D., Gillingham, P.K., Bradbury, R.B., Roy, D.B., Anderson, B.J., Baxter, J.M., Bourn, N.A.D., Crick, H.Q.P., Findon, R.A., Fox, R., Hodgson, J.A., Holt, A.R., Morecroft, M.D., O'Hanlon, N.J., Oliver, T.H., Pearce-Higgins, J.W., Procter, D.A., Thomas, J.A., Walker, K.J., Walmsley, C.A., Wilson, R.J. & Hill, J.K. (2012) Protected areas facilitate species' range expansions. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 14063–14068.

Torn, A., Tolbanen, A., Norokorpi, Y., Tervo, R. & Siikamaki, P. (2009) Comparing the impacts of hiking, skiing and horse riding on trail and vegetation in different types of forest. *Journal of Environmental Management*, **90**, 1427–1434.

USDI NPS. (2016) National Park Service Visitor Use Statistics - Yosemite NP, https://irma.nps.gov/Stats/SSRSReports/Park Specific Reports/Annual Park Recreation Visitation (1904 - Last Calendar Year)?Park=YOSE

Viers, J.H., Purdy, S.E., Peek, R. A, Fryjoff-Hung, A., Santos, N.R., Katz, J.V.E., Emmons, J.D., Dolan, D. V & Yarnell, S.M. (2013) *Montane Meadows in the Sierra Nevada: Changing Hydroclimatic Conditions and Concepts for Vulnerability Assessment*. UC Davis, Center for Watershed Science, Davis, CA.

Walden-Schreiner, C. & Leung, Y.-F. (2013) Spatially characterizing visitor use and its association with informal trails in Yosemite Valley Meadows. *Environmental Management*, **52**, 163–178.

Walden-Schreiner, C., Leung, Y.-F., Kuhn, T., Newburger, T. & Tsai, W.-L. (2017)

Environmental and managerial factors associated with pack stock distribution in high elevation meadows: Case study from Yosemite National Park. *Journal of Environmental Management*, **193**, 52-63.

Watson, J.E.M., Dudley, N., Segan, D.B. & Hockings, M. (2014) The performance and potential of protected areas. *Nature*, **515**, 67–73.

Wiens, J.A. (1989) Spatial scaling in ecology. *Functional Ecology*, **3**, 385–397.

Williams, D.M., Dechen Quinn, A. & Porter, W.F. (2012) Impact of habitat-specific GPS positional error on detection of movement scales by first-passage time analysis. *PloS one*, **7**, e48439.

Zhang, J., O'Reilly, K.M., Perry, G.L.W., Taylor, G.A. & Dennis, T.E. (2015) Extending the functionality of behavioural change-point analysis with k-means clustering: A case study with the little penguin (*Eudyptula minor*). *PloS one*, **10**, e0122811.

CHAPTER 3

Environmental and Managerial Factors Associated with Pack Stock Distribution in High Elevation Meadows: Case Study from Yosemite National Park[1]

Abstract

Parks and protected areas are integral strategies for biological diversity conservation, and their management often involves balancing visitor use with resource protection. Effectively balancing these objectives requires data about how use is distributed within areas of interest and how management strategies and environmental conditions interact to minimize negative impacts. This study examined which environmental and managerial factors most influenced the distribution of domestic pack stock animals, a common visitor use-related activity, when released to graze in high elevation meadows. Using a species distribution modelling approach, MaxEnt, managerial factors were found to be among the top contributors to models. Pack stock animals concentrated use near the locations where they were released and near portable enclosure fencing confining the lead animal, even though the remainder were allowed to roam freely. Elevation was the highest contributing environmental factor, with animals remaining at elevations similar to that of the meadow, even if moving into nearby understory. Results highlight the importance of release point and fence locations to overall

[1] Published as: Walden-Schreiner, C., Leung, Y.-F., Kuhn, T., Newburger, T., & Tsai, W.-L. 2017. Environmental and managerial factors associated with pack stock distribution in high elevation meadows: Case study from Yosemite National Park. Journal of Environmental Management, 193, 52-63.

pack stock animal distribution and show that rotational or strategic placement can be a tactic for mitigating impacts to sensitive habitats.

Keywords

Horse; mule; maximum entropy; meadow; GPS tracking

Introduction

Biodiversity underpins healthy ecosystem function, provides ecosystem services, and supports human well-being. Parks and protected areas (PAs) represent a core component of biodiversity conservation strategies (Chape et al., 2005), with global agreements and efforts striving to increase the number and size of PAs to meet biodiversity conservation goals (Watson et al., 2014). In addition to conserving the resources within their boundaries, PA managing agencies may also have mandates to provide for compatible recreational opportunities. Balancing conservation and recreation needs can be challenging in light of recent estimates indicating terrestrial PAs around the world attract eight billion visits each year collectively (Balmford et al., 2015). Proactive management needs timely and accurate data about recreational use and environmental conditions to assess the magnitude, trends, and significance of recreation-associated disturbance. The purpose of this study is to identify environmental and managerial factors influencing the distribution of one type of visitor-

related activity, use of pack stock animals (stock), to assess the influence of management tactics designed to minimize negative impacts.

In many PAs around the world, stock use is an established form of recreation, supports tourism operations, and is often part of the cultural heritage (Barros and Pickering 2015; Byers, 2009; Cole et al., 2004; Geneletti and Dawa, 2009; Newsome et al., 2002). The managing organization may also use stock to support administrative efforts. Stock use, which can include horses, mules, donkeys, yaks, or llamas, is one example of a visitor-related activity that can affect the ecological integrity of preserved ecosystems (Barros et al., 2015; Byers, 2009; Cole et al., 2004; Geneletti and Dawa, 2009). Impacts attributed to stock presence in alpine and mountainous landscapes include soil compaction, altering of hydrologic patterns, increases in erosion, and trampling of vegetation (Barros et al., 2014; Byers, 2009; Kuhn et al., 2015; Ostoja et al., 2014). Research demonstrates horse trampling occurs at greater intensities than that of hikers (Weaver and Dale, 1978), and, in some cases, other hooved animals (Cole and Spildie, 1998). Additionally, transporting forage for stock may not always be feasible due to load sizes, geographic location, or invasive species concerns. Instead, stock may be released into meadows and adjacent areas to graze. Estimates indicate that an individual horse or mule consumes an average 14.8 kg of forage per night (Jacoby and The Society for Range Management, 1989).

Meadows are vital ecosystems supporting disproportionately high levels of biodiversity. In the western U.S., meadows also represent one of the most at-risk landscapes (Viers et al., 2013). Several studies reveal relationships between stock grazing and impacts to meadow

ecosystem function. Overgrazing can decrease vegetation cover and meadow productivity, shift plant species composition through selective grazing of palatable species, change local soil chemistry and micro-climate conditions, increase bare ground cover, and facilitate the encroachment of invasive or woody species (Cole et al., 2004; Kuhn et al., 2015; McClaran and Cole, 1993; Moore et al., 2000; Olson-Rutz et al., 1996a, 1996b; Ostoja et al., 2014). Mitigating impacts from stock in PAs, while providing access for stock users, relies on understanding the interaction between stock and the environment (Moore et al., 2000). Research on the resistance and resiliency of grass, forb, and shrub-dominated habitats indicates proactive management is essential in limiting the propagation and proliferation of negative impacts (Cole et al., 2004; McClaran and Cole, 1993; Olson-Rutz et al., 1996b). Studies of grazing impacts in PAs are often based on inventories made before and after an animal has left the area, employing plots with varying levels of grazing treatments to measure ecosystem response, or observational studies of impacts to a variety of ecosystem components (Alzerreca et al., 2006; Cole et al., 2004; Holmquist et al., 2014; Kuhn et al., 2015; Ostoja et al., 2014).

How grazing intensity and environmental variables vary, however, has been identified as a research need within recreation ecology (Monz et al., 2010). Specifically, researchers and managers highlight a need to understand patterns of stock distribution and use during potential disturbance and how it is associated with environmental (e.g., vegetation composition, access to water) and managerial (e.g., fencing, stock release locations) factors, including overnight releases when direct supervision is less feasible (McClaran, 1989; Ostoja

et al., 2014). Studies of stock resource use at the site level offer an opportunity to combine advances in data collection instruments, computing capacity, and statistical modelling methods from several disciplines interested in use-environment interactions to address this research need (Brown et al., 2013; Gaylord and Sanchez, 2014; Jeltsch et al., 2013; Kays et al., 2015; Meijles et al., 2014; Rutter, 2007). The convergence of these advances have facilitated greater integration of animal distribution and movement data with environmental covariates (Kays et al., 2015).

Research objectives

The purpose of this study is to integrate data collection and analytical techniques from sub-disciplines of ecology to assess environmental and managerial variables influencing fine scale stock distribution and resource use. Specifically, this study:

(1) Identifies which environmental and managerial factors contributed most to models of free-roaming stock distribution in meadows and adjacent areas during overnight release periods; and

(2) Evaluates the accuracy of the models used to characterize stock distribution in relation to environmental and managerial factors.

Background Literature

To address the research objectives, three components were required: accurate spatio-temporal stock movement data, high-resolution environmental and managerial data, and a modelling method capable of addressing sampling limitations and quantifying model accuracy. This section reviews related work surrounding spatial use data and modelling use-environment interactions to identify which best address the research objectives.

Spatial use data

In the fields of recreation ecology, animal ecology, and movement ecology, global positioning system (GPS) data tracking visitor and animal movement are increasing in spatial and temporal accuracy, enhanced by expanding data collection capacity and environmental data captured by mobile and remote sensors (Handcock et al., 2009; Kays et al., 2015). Accuracy and capacity improvements are driven by reductions in GPS unit size, increases in battery longevity, and improved data retrieval options, rapidly increasing the amount and quality of tracking data available for analysis (Michelot et al., 2016). The development and application of statistical methods capable of garnering information from complex GPS datasets represents a growing research focus, because data collection capacity has challenged computational capacity and traditional statistical assumptions about independence, and highlighted the need for user-friendly software (Long and Nelson, 2013; Michelot et al., 2016).

Modelling use-environment relationships

Distribution and movement models have also been employed to draw inference from species occurrence records (Edren et al., 2010). Species distribution models (SDMs) examine occurrence (or presence) records with surrounding environmental variables to determine habitat suitability. The underlying concept of SDMs is that species are more likely to be present in locations with conditions that suit their needs. Distribution models have used environmental envelopes, generalized linear models (GLM), generalized additive models (GAM), support vector machines, and boosted regression trees (BRT), to name a few, and have been applied to a diverse range of terrestrial and marine species to help inform conservation and management (Smith, 2013).

Key assumptions of SDMs, however, are that the presence records are independent from one another or representative of the species' distribution. When leveraging species data derived from GPS or telemetry methods, these assumptions are challenged and no systematic assessment of absences (i.e., where a species is not found) is inherent. Incorporating absence data is ideal for SDM and enables presence probability to be predicted using methods like logistic regression. Without systematic absence data, model outputs should be interpreted as relative rates of occurrence (Merow et al., 2013). Lack of absence data is shared in SDMs derived from museum collections or opportunistic sampling of species locations common in plant and animal ecology and has received considerable attention in the last decade,

especially following the development of the modelling tool, MaxEnt (Edren et al., 2010; Elith et al., 2006; Guillera-Arroita et al., 2015; Phillips et al., 2006).

MaxEnt is a machine-learning algorithm built on the principle of maximum entropy (Jaynes, 1957, 1982). The model compares covariate conditions at presence locations to the conditions at random locations in the greater area of potential occurrence (Elith et al., 2011; Phillips et al., 2006), and has been applied to GPS tracking data (Edren et al., 2010; Poor et al., 2012; van Gils and Kayijamahe, 2010; Warton and Aarts, 2013). MaxEnt models examine patterns in the available data to make predictions based on the assumption of incomplete information (Jaynes, 1982). They are also less sensitive to position errors and limited sample sizes (Elith et al., 2006). Research demonstrates that MaxEnt models are successful in modelling species distributions based on presence-only data when compared to other methods, including GLM and GAM (Elith et al., 2006). While often applied at regional scales, MaxEnt has also been successfully applied to finer scales (Smith et al., 2012).

Environmental predictors, termed features in MaxEnt, can be linear, quadratic, or cubic terms, or based on hinge or threshold functions. Coefficients are used to weight features, with iterative changes in the model determining the point at which coefficients converge on a probability distribution that maximizes the likelihood of use or probability of species occurrence (Braunisch et al., 2011). Benefits of MaxEnt models include the ability to integrate presence-only data with environmental data, incorporate continuous and categorical variables, employ methods to minimize overfitting, and evaluate each covariate’s influence.

As with any model or method, MaxEnt is not without limitations. Critiques of MaxEnt discuss the potential for extreme values of environmental conditions not observed in the study area due to the use of an exponential probability model and sensitivity to sampling efforts (Elith et al., 2011; Phillips et al., 2006). These critiques recommend caution when interpreting model outputs and metrics (Yackulic et al., 2013). To address this concern, background selection should be ecologically informed and representative of a species distribution (Phillips and Elith, 2016). It is important to note that MaxEnt outputs are indices of habitat suitability (i.e., relative probability of presence) and reflective of potential biases in sampling (Royle et al., 2012). Options to minimize sampling biases include spatially filtering the number of occurrence records in oversampled locations, which can reduce the sample sizes needed for modelling, or corrective measures like choosing background data that reflect the same sampling bias as the occurrence data or bias correction files (Kramer-Schadt et al., 2013).

Additionally, Renner and Warton (2013) explored the mathematical similarities between Poisson regression, and therefore Poisson point process models (PPM), to MaxEnt. They found equivalencies between the two methods depending on how model parameters are configured and analyzed. Renner and others (2015) demonstrated how MaxEnt could be used as a means for fitting a Poisson PPM, namely by retaining duplicate presence and examining the effect of the number of background points on the regularized training gain. For the purposes of this analysis, which examines distributions within a small scale to determine the influence of factors on observed distributions rather than predicting distributions in other

environments, duplicate presence points were retained and the MaxEnt algorithm allowed to choose the best feature type (i.e., linear, quadratic, product, threshold, or hinge).

Use of MaxEnt with spatially and temporally autocorrelated GPS-derived position data also represents a challenge to modelling assumptions (Aarts et al., 2008; Baldwin, 2009; Gurarie et al., 2009). Solutions proposed to address this include modelling each species individually, using the home range as the background for model fitting, and comparing the percent contribution of each covariate for the individual model to determine which were most important (Baldwin, 2009). Subsampling also reduces autocorrelation and ensures no species contributes more positions to the model than another (Edren et al., 2010; Gschweng et al., 2012).

To evaluate model performance and address research objective 2, MaxEnt models assess the area under the curve (AUC) of the receiver operating characteristic (ROC), with values closer to one indicating better model performance. The ROC plots the percentage of correctly classified occurrence points against the percentage of correctly classified pseudo absence points, derived from available background points. Unlike model performance assessments when true absences are included, the AUC from a MaxEnt output should be interpreted as the model's ability to correctly identify areas of high probability of occurrence from random background pixels, rather than suitable or unsuitable locations (Jalali et al., 2015; Yackulic et al., 2013).

Material and Methods

Study area

Yosemite National Park (YNP) is located in the Sierra Nevada mountain range of California. In 1890, it was one of the first National Parks established in the U.S. and was designated a United Nations Educational, Scientific, and Cultural Organization (UNESCO) World Heritage Site in 1984. The park, managed by the U.S. National Park Service (NPS), is known globally for its biodiversity, glacially created landforms, waterfalls, sequoia groves, and recreational opportunities. The park provides diverse habitats for over 400 species of vertebrates, including two potentially endemic species, in five biotic zones ranging from 600 to 4,000 m in elevation (UNESCO, 2016), and attracted over 4.1 million recreational visits in 2015 (USDI NPS, 2016a). Approximately 95% of the 3,027 sq. km park is federally designated wilderness, and two federally designated Wild and Scenic Rivers, the Merced and Tuolumne, flow within its boundaries.

Stock use is allowed in YNP, limited to 25 animals or less for NPS administrative functions, commercial trips, and private parties (Abbe and Ballenger, 2011). Commercial stock trips account for half of all overnight stock use in the park, with current management relying on permits and zoning to direct trail, camp, and meadow usage based on the number of stock use nights (e.g., number of animals at a site per night). Ostoja and others (2014) reported an average of 1,571 stock animal nights per year between 2004 and 2012. Meadows, some of which are located within a quarter mile of an NPS system trail, may be used for

overnight grazing by stock animals, account for 3% of the total park area, and yet support disproportionate amounts of biodiversity and aesthetic and recreational opportunities for visitors (Viers et al., 2013).

Concerns about overuse impacts related to stock in YNP are documented as early as the 1930s (Sumner, 1935), with a call for monitoring in 1969 (Moore et al., 2000). Recent monitoring efforts of potential stock and visitor use impacts have been associated with management plan development for the Merced and Tuolumne Rivers, both of which require formal, dedicated management plans due to their status as Wild and Scenic Rivers (Ballenger et al., 2010; Ballenger et al., 2011; USDI NPS, 2014a, 2014b). These monitoring efforts also carried forward to inform the park's Wilderness Stewardship Plan (USDI NPS, 2016b).

As part of a greater NPS study, purposive sampling of stock meadow use in designated wilderness areas of the park was conducted during the summer months of 2014 by NPS staff using methods developed in collaboration with the authors. This paper compares two contrasting, graminoid-dominated meadow sites where similar number of animals were released for similar durations and managed with two common techniques. Techniques employed by management included changing the location at which the animals were released for free-range grazing (release point), and the use of a portable electric fencing (fence). When fences were used, the lead horse or horses were enclosed inside the fence and the remaining animals were allowed to roam freely. The two meadows included Castle Camp (CACA, UTM Zone 11N centroid at 293047.9, 4210568.4) and Upper Kerrick-North (UPKE, UTM Zone 11N centroid at 282509.2, 4221912.0). Animals were released for approximately 12

hours (i.e., release period) with differences in management techniques (Figure 1, Table 1). Both sites are subalpine meadow habitat, dominated by native grasses, with perennial stream channels and wetland complexes, surrounded by lodgepole pine (*Pinus contorta*) forest.

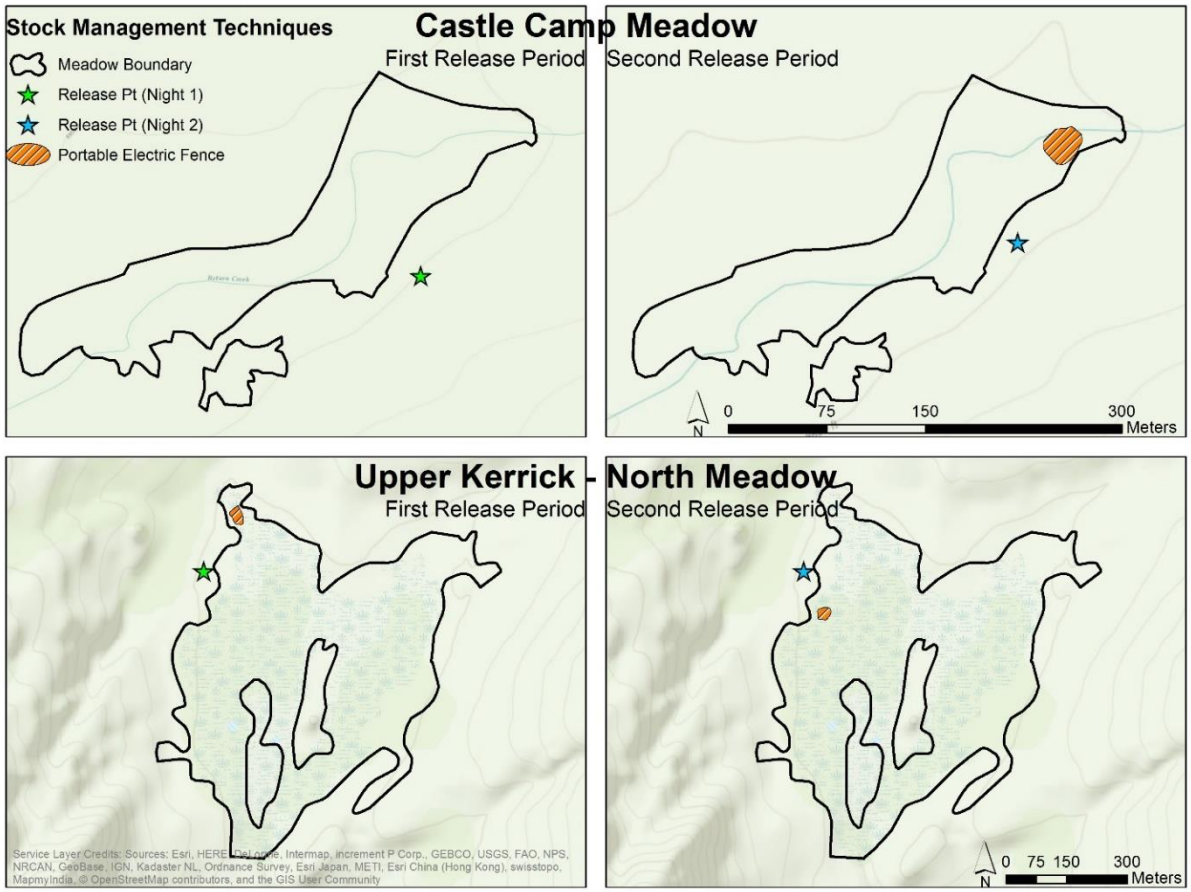


Figure 1 Location of release points and portable electric fences, when used, for both release periods in Castle Camp and Upper Kerrick-North meadows. Note: the release point in Upper Kerrick-North was the same for both release periods.

Table 1 General characteristics of the two study sites and stock management techniques.

Site Name (Abbreviation)	Area (ha)	Elevation (m)	Dominant Vegetation Associations[1]	Start Date (Time, PST) End Date (Time, PST)	Stock Management
Castle Camp (CACA)	2.9	2,677	DESCES, CALMUI	8/25/2014 (19:30) 8/26/2014 (07:35)	**Release Point** : Animals released outside of meadow along the eastern boundary **Fence**: No fence used; All animals allowed to graze freely
				8/26/2014 (19:30) 8/27/2014 (08:00)	**Release Point:** Animals released outside of meadow approximately 50m closer to meadow boundary to the east **Fence**: Two horses enclosed in 643 sq. m. fence located in the NE of meadow
Upper Kerrick – North (UPKE)	21.4	2,858	STIKIN, CALMUI / OREALP/ VACCAE, DESCES, ELYTRA, HORBRA	9/22/2014 (18:15) 9/23/2014 (07:48)	**Release Point**: Animals released at NW edge of meadow **Fence**: One horse enclosed inside 642 sq. m. fence located at north end of meadow approximately 150m NE of release point
				9/24/2014 (18:20) 9/25/2014 (07:45)	**Release Point**: Animals released from same location as previous release **Fence**: One horse enclosed inside 509 sq. m. fence located approximately 100m to the SE of the release point

[1]Only the top five dominant vegetation associations are listed for general information on the plant communities present at each site (see Kuhn et al. 2015 for full inventory). Abbreviations include *Deschampcia cespitosa* (DESCES); *Calamagrostis muriana* (CALMUI); *Stipa kingii* (STIKIN); *Oreostemma alpigenum* (OREALP); *Vaccinium cespitosum* (VACCAE); *Elymus trachycaulus* (ELYTRA); and *Hordeum brachyanthemum* (HORBRA).

Data for each night were analyzed separately to examine animal distribution under differing techniques, with outputs labeled by the meadow abbreviation and release number (e.g., CACA1 for release period starting the evening of August 25 and CACA2 for release period starting the evening of August 26). The trial at CACA included three horses and four mules, and the trial at UPKE included one horse and four mules.

Data collection and filtering

Stock performing administrative duties for the NPS were fitted with custom-built, professional-grade GPS units attached to the animal's halter when released to graze (Figure 2). Developed by Oregon State University, the GPS units used in this study contained the u-blox Max-7c module. Combined with a Pelican 1010 case for protection, the units weighed less than 1 kg to minimize influence or impact on the animal's comfort or behavior. During early pilot trials, one animal of the herd was not equipped with a unit so researchers could observe whether its presence affected behavior. For each pilot trial, the non-collared animal randomly varied. The presence of a unit did not appear to have an effect on animal behavior. For subsequent trials, including those at CACA and UPKE, all animals were equipped with units. Calibration testing of the units, conducted by Oregon State University under conditions similar to those of meadows, documented a circular error probability (CEP) of less than 6 m (D. Johnson and T. Kuhn, personal communication). CEP is the radius of a circle, calculated using the standard deviations of x and y coordinates, in which 50% of the values occur, and

is a tested method for assessing the accuracy of GPS units (Frair et al., 2010; Graves and Waller, 2006; Lewis et al., 2007).

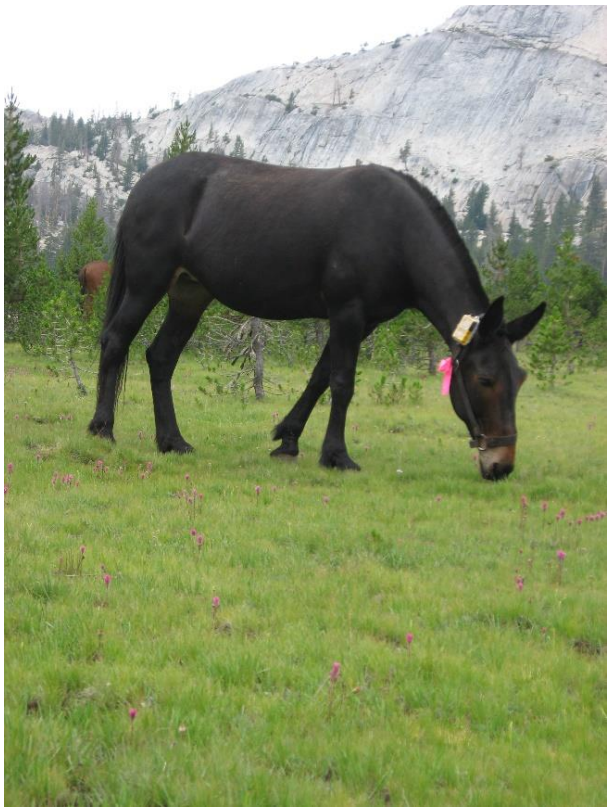


Figure 2 Stock animal wearing a custom GPS unit attached to its halter. Photo courtesy of T. Kuhn.

The GPS units recorded the animal's position every second for the duration of the release period, and differentially corrected in real-time to improve positional accuracy. NPS staff retrieved each unit at the conclusion of a release period and downloaded the data for later conversion to Environmental Systems Research Institute (ESRI) shapefiles for analysis.

GPS data were examined for spatial and temporal errors and processed in Microsoft Excel to remove duplicate records and spatial or temporal outliers. Data filters available in MoveBank (www.movebank.org), an online database designed to manage and analyze movement data, confirmed outliers prior to their removal. Filtering excluded any points that indicated the animal was traveling at an improbable speed. Informed by research on horse gait analysis, a threshold of 13 meters per second (mps), the approximate mid-point of speeds achieved by a galloping horse, was used to remove outliers (Huang et al., 2013).

Studies of grazing animals indicate distributions can be influenced by environmental conditions like slope, distance to water, and forage quantity and quality (Bailey et al., 1996; Duncan, 1983). Therefore, to address research objective 1, data sources relevant to this analysis included the U.S. Geological Survey (USGS) 3-meter resolution National Elevation Dataset ([NED] USGS, 2016), color infrared 1-meter resolution imagery from the U.S. Department of Agriculture (USDA) National Agricultural Imagery Program ([NAIP] USDA, 2016), USDA CALVEG vegetation survey data (USDA Forest Service, 2008), and a hydrology layer containing rivers and waterbodies obtained from the NPS and derived from 1992 USGS 1:24,000 topographic maps (USDI NPS and USGS, 2011). The NED was used to identify elevation at stock locations and to derive the slope (i.e., percent rise using a 3 x 3-pixel window) using the slope tool in ESRI ArcGIS 10.3.1 software.

NAIP images were cloud-free and color balanced using the mosaic color balance tool in ArcGIS 10.3.1 and used to calculate a Normalized Vegetation Difference Index (NDVI). NDVI is an indicator of surface greenness calculated by using two bands from spectral

imagery, red and infrared. NDVI values range from -1 to 1, with greater positive values associated with more green vegetation. NAIP images used in the analysis for CACA were taken on August 15, 2014, 10 days prior to stock data collection in that area, and July 24, 2014 for UPKE, two months prior to stock data collection in that area. From the CALVEG vegetation survey, the regional dominance type was selected to classify land cover and indicates the primary (dominant) vegetation alliance in an area, as well as the secondary understory hardwood species in mixed cover types, based on procedures developed by the U.S. Forest Service (USDA Forest Service, 2008). Trimble Pro-XRT GPS receivers recorded the locations of the release point and fences, when used, and data were differentially corrected prior to inclusion in the model. Distance-to release point, fencing, and hydrology raster layers were constructed using the near tool in ArcGIS.

Data analysis

We identified the potential area of stock use (i.e., background area of interest) to be the greatest distance to the release point for each animal by any animal based on GPS data during any release to the site. The overall maximum value for the site was rounded to the nearest 100 m to create the area of interest. The nearest 100 m was chosen as it would allow for consideration of GPS (in)accuracy and account for potentially visible yet not visited locations. The release point was then radially buffered by this distance, and converted to a minimum-bounding polygon as the study area of interest (AOI) for analysis. AOI polygons included the meadow, as well as upland habitat, water bodies, and forested areas. This aligns

with recommendations that the area from which unvisited sample points are taken (i.e., background points) reflect environmental conditions that are accessible to the species in question or have conditions of interest (Merow et al., 2013; Radosavljevic and Anderson, 2014).

The AOI is also directly relevant to management planning efforts focusing on the meadow and immediately adjacent area, including the area stock could reasonably visit during an overnight release. Additionally, this accounts for sampling bias as both presence and background samples represent the same conditions relevant to the research question, which is interested in a specific area's use rather than predicting to new locations (Yackulic et al., 2013). All data were clipped to the AOI using the extract by mask tool in ArcGIS at the coarsest resolution (3m).

Fourcade and colleagues (2014) found systematic subsampling of presence records to be among the top performing bias reduction methods for MaxEnt modelling when compared to three additional types of bias correction (i.e., bias files, restricted background, cluster sampling, and splitting datasets over large geogrpahic regions). Although they selected their subsample to be distributed in a large region, they do note that this could underestimate the contribution of areas where species are found in high densities. Therefore, due to the small size of the AOIs in this study when compared to other MaxEnt studies, a systematic random subsample was selected following similar procedures outlined in Edren and others (2010) and Gschweng and others (2012). Gschweng and others (2012) identified the lowest number of locations for an individual animal as a threshold and subsequently randomly sampled other

units to that threshold. Similarly, the total number of successful fixes for each GPS unit for all stock were examined. The fixed number for random sampling was determined by taking 30% of the unit with the least number of fixes and randomly sampling the remaining units to have the same number of fixes for the final sample.

The Maximum Entropy Species Distribution Modelling software, version 3.3.3k, fit MaxEnt models (http://www.cs.princeton.edu/~schapire/maxent/). Simplified to an n by two matrix, GPS tracking data contained X and Y UTM coordinates for data from all units. All predictors were included in model fitting to allow the machine learning algorithm to determine which were important, through regularization (Phillips et al., 2006). Following a machine learning approach, high collinearity among predictors is less of an issue when compared with statistical methods (Elith et al., 2011), so long as results are interpreted as the predictive accuracy of the presences (Merow et al., 2013).

To fit the model, 30% of the data were randomly selected to be used as test points for cross-validation of the model, with 10 replicate runs using random seed subsampling (i.e., a different, randomly selected partition was selected for each run for the test/training partition and background sample). The regularization multiplier and maximum number of background points remained at the defaults of 1 and 10,000, respectively. The maximum number of iterations was set at 500, with a convergence threshold of 0.00001 (Phillips et al., 2006; Young et al., 2011). The software also created response curves and a jackknifing procedure was used to measure variable importance. The jackknife procedure examines the model,

using AUC gain, without the specified variable and with only the specified variable, and compares it to the model with all variables included.

Model outputs ranked each predictor variable based on its contribution to the overall model. To evaluate the method and address research objective 2, predictive accuracy was assessed using AUC (Jalali et al., 2015). AUC scores above 0.5 indicate the model performs better than random and above 0.75 are considered acceptable for predictive accuracy (Phillips and Dudík, 2008; Richards and Friess, 2015), with >0.8 considered good and 0.9 considered excellent (Coppes and Braunisch, 2013).

Results

GPS tracking

A total of 267,538 and 250,825 location fixes were obtained from seven animals in CACA and five animals in UPKE, for both release periods. Of those records, 236,419 and 202,142 locations remained for CACA and UPKE, respectively, after removing points collected for periods of confinement if applicable. The remaining data for free-ranging animals were analyzed for duplicates and outliers, with 189,012 (80%) and 156,352 (77%) locations retained for CACA and UPKE, respectively. Successful fixes from the remaining sample ranged from 5,830 to 36,849 locations per unit in CACA and 7,174 to 46,199 per unit in UPKE. To minimize the influence of varied fix rate success, a subsample of data was randomly selected from each unit, using a threshold of 30% of the total number of locations

from the unit with the least number of fixes (i.e., 5,830). Thus, 1,750 records for each animal in each meadow were used for analysis.

AOI characteristics

Using the filtered full data sets (i.e., 189,012 in CACA and 156,352 in UPKE), the AOI for CACA was identified as 1100 m, because the farthest an animal traveled away from the release point during either release period was 1085 m. For UPKE, the AOI was 2100 m from the release point, because the farthest any animal traveled was 2044 m. The regional dominant vegetation alliances for both CACA and UPKE AOIs were predominately comprised of conifer species, with wet meadows representing only 0.66% and 6.48%, respectively (see Appendix B for details on all vegetation alliances in the AOIs).

MaxEnt results

A MaxEnt model was fit for each of the four release periods, CACA1, CACA2, UPKE1, and UPKE2 (Table 2). The relative contributions of each variable to the MaxEnt model were examined in two ways. The first, percent contribution, is estimated by adding (or subtracting) the increase (or decrease) in regularized gain for each iteration of the training algorithm to rank variable importance. These values depend on the path taken to converge on a solution, with subsequent runs potentially yielding different values, and may indicate greater importance for highly correlated variables (Baldwin, 2009). Four additional iterations of the models with the same data yielded similar values, with the order of the top variables

remaining constant. As an alternative, permutation importance examines variable importance by reevaluating the model following a randomization of values for each variable separately and measuring the change in AUC as a normalized percentage. Unlike the percent contribution, the values are not dependent on the path taken to identify the optimal solution, but rather the values from the final model. Predictive accuracy as evaluated by the AUC indicates high accuracy (i.e., AUC > 0.9) for all four models.

Table 2 Percent contribution and permutation importance for variables by release period and meadow. The number 1 following the meadow abbreviation designates the first release period, with the second release period designated with the number 2. Bolded values indicate the variable is one of the top two factors contributing to the model for that area of interest.

	CACA1		**CACA2**		**UPKE1**		**UPKE2**	
	Percent Contribution	**Permutation Importance (%)**	**Percent Contribution**	**Permutation Importance (%)**	**Percent Contribution**	**Permutation Importance (%)**	**Percent Contribution**	**Permutation Importance (%)**
Distance to Fence (meters)	n/a	n/a	**72.6**	**93.8**	1.8	2.8	**79.6**	**96.1**
Distance to Release Point (meters)	**82.4**	**93**	**24.4**	1.7	**67.7**	**80.1**	1.4	0.3
Elevation (meters above sea level)	**12.3**	**3.6**	0.8	1.4	**25.7**	**11.3**	**16.7**	**2.8**
Distance to waterbodies (meters)	5.0	1.6	2.0	0.5	4.4	2.9	1.3	0.6
Regional dominant veg cover	0.1	0.8	0.1	0.4	0.3	1.1	0.2	0.1
Slope (percent rise)	0.1	1.0	0.2	**2.1**	0.1	1.9	0.8	0.2
NDVI	0.0	0.0	0.0	0.0	0.0	0.0	0.0.	0.0
Average Test AUC (AUC Standard Deviation)	0.918 (0.0027)		0.947 (0.0022)		0.935 (0.0024)		0.961 (0.0020)	
Average Training AUC	0.9198		0.9465		0.9346		0.9617	

Response curves (Figure 3) illustrated how each variable influences the logistic prediction, with upward trends indicating positive associations. Response curves representing a separate MaxEnt model created with just the variable of interest are discussed below and are ‘clamped’ at the minimum and maximum values in the training data by the software to avoid errors due to extrapolation. The red line represents the MaxEnt run, with the blue area indicating plus or minus one standard deviation.

Distance to the release point was one of the top two variables in three of the four models based on percent contribution, contributing between 24% to 82%. The response curves illustrate dependence of the predicted occurrence at varying values of each factor. Greater values along the y-axis of a response curve indicates higher logistic probability of animal presence than lower values. For example, the model suggests a high probability of stock presence within 300 m for CACA and 500 m for UPKE of the release point, with negative slope indicating declines beyond those values (Figure 3). A second peak near 450 m from the release point during the first release in CACA represents a cluster of points occurring in the SW of the meadow near the river.

Distance to fence was the variable contributing the most to the models for the second releases in both meadows, contributing over 70% of the model for each (Figure 3). A fence was used only during the second release period in CACA, and for both releases in UPKE, but placed in different locations of the meadow (Table 1, Figure 1). Stock presence is more likely within 200 m of the fence used during the second release period in CACA, and between 600 m and 400m for the first and second release periods in UPKE, respectively. The fence was

not among the top variables during the first release period in UPKE, with elevation instead becoming a top contributing factor.

Elevation contributed as one of the top two variables in three of the four models, though its contribution was less than 25%. The response curves for elevation indicate the relative rate of occurrence increases at and around the elevation range of the meadow, approximately 2677 m for CACA and 2858 m for UPKE. Distance to a water body was often in the middle of the variable ranking and the regional dominant species also did not contribute much to the models. When examining the response curves independently from other variables, however, two specific dominant species did indicate greater relative probability of presence (i.e., logistic probability of presence greater than 0.4). Specifically, of the 12 dominant vegetation alliances identified in the AOI for CACA and 16 dominant vegetation alliances identified in the AOI for UPKE (see Appendix B), results suggested an increased relative probability of presence (greater than 0.4 for all models) in lodgepole pine (*Pinus contorta*) and wet meadow areas. This also corresponds to the response curves from the NDVI variable, which although not contributing to the MaxEnt models independently, indicated higher relative probability of presence between values of 0.2 and 0.6. These NDVI values correspond to values associated with shrub, grasslands, and increasingly dense green vegetation (Figure 3).

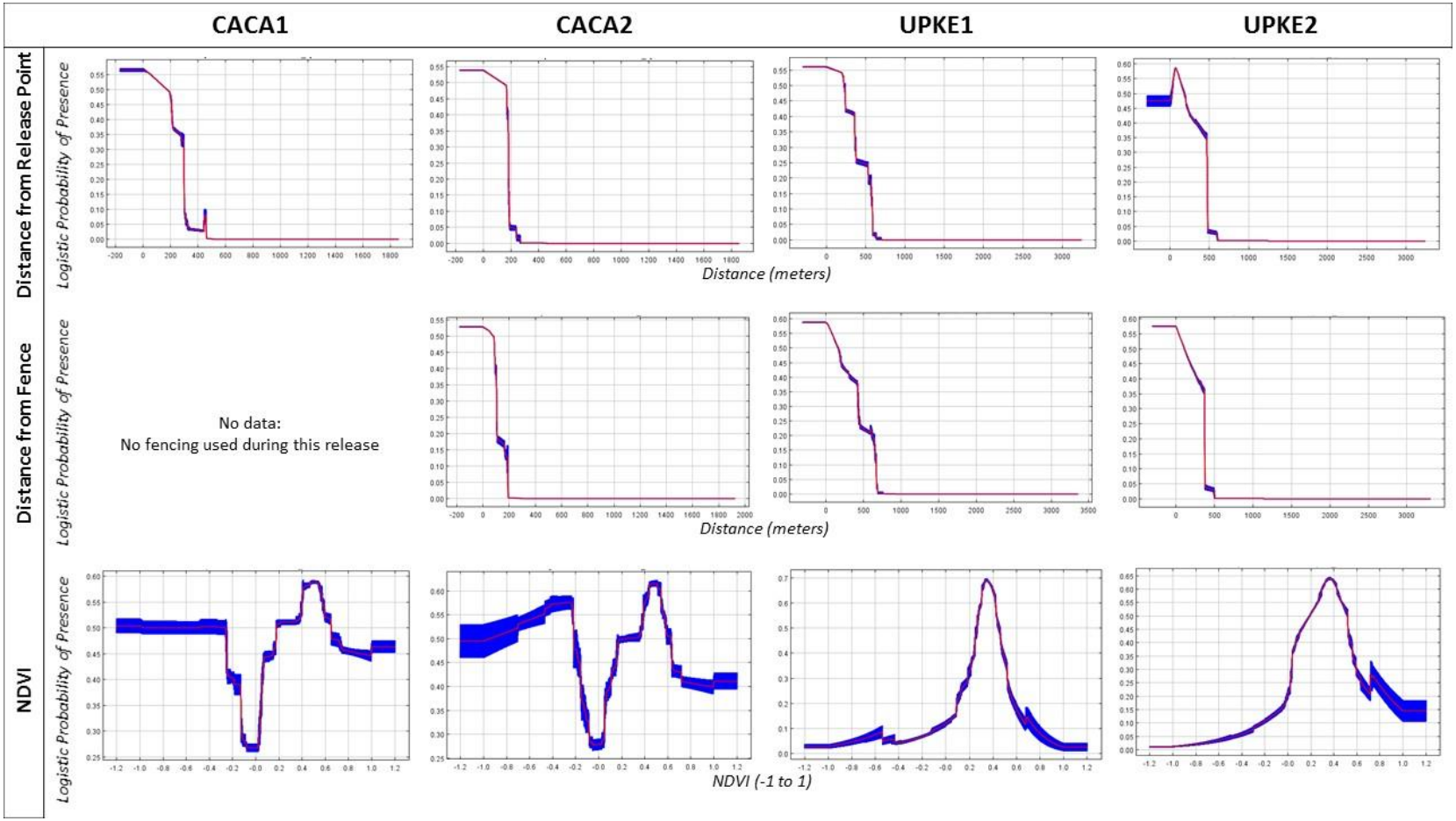


Figure 3 Response curves for distance to release point, distance to fence and NDVI values for CACA1, CACA2, UPKE1, and UPKE2. The red line indicates the mean response of the 10 replicate MaxEnt runs and the blue line is plus and minus one standard deviation. Note the different scales between CACA and UPKE due to site extent differences.

Jackknife values compared a MaxEnt model without the specified variable in teal (i.e., what would the MaxEnt AUC be without the specified variable), a model created with only the specified variable in blue, and the full model with all variables in red (i.e., AUC values reported in Table 2). The length of the teal bar indicates how much is lost in AUC by excluding the variable. The length of the blue bar indicates the amount to which the specified variable contributes to the overall AUC by itself. For CACA meadow (Figure 4a), on the first night of release, the distance to release point, when modeled in isolation, contained the most useful information by itself (blue). Distance to the release point also decreased the gain the most when omitted from the model, indicating it contained information not present in the other variables. NDVI, by itself, achieved almost no gain, so the variable in isolation is not as useful for determining stock distribution. Similar results were found for the second night of release in CACA, although distance to fence contains the most useful information as well as information not contained by other variables for predicting distribution of occurrence data when measuring using AUC (Figure 4b).

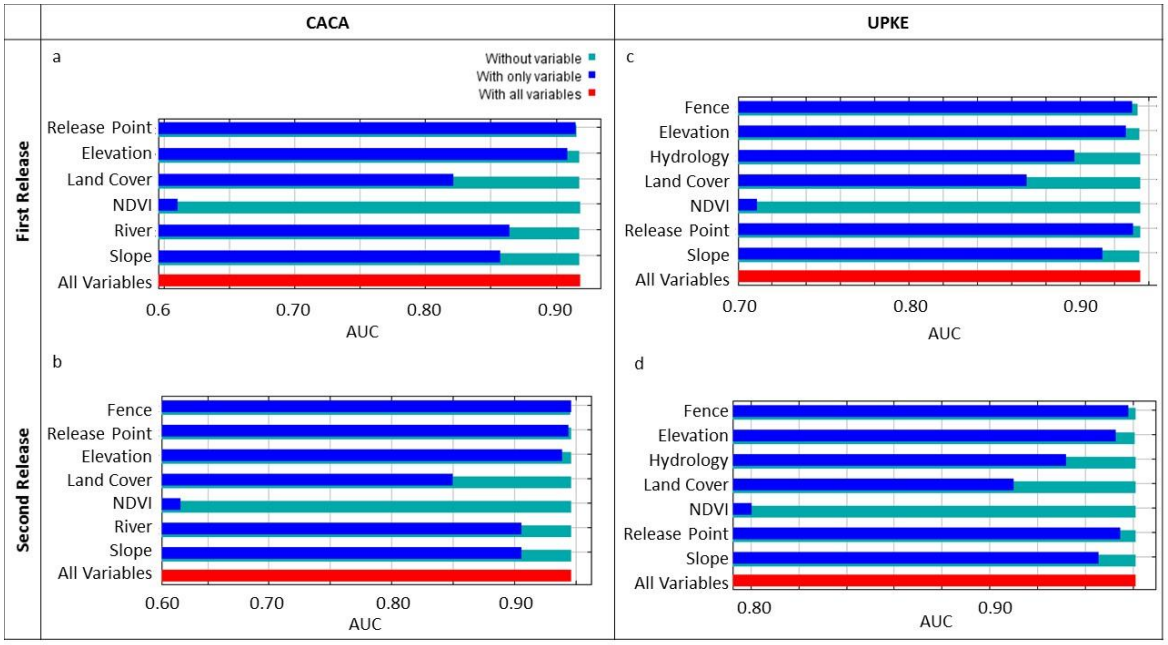


Figure 4 Jackknife plots for AUC for (a) CACA1, (b) CACA2, (c) UPKE1, and (d) UPKE2. Fence, release point, and river/hydrology are distance-to variables.

For UPKE, the jackknife plots of AUC for the first night of release (Figure 4c) indicate that the release point distance contributes the highest gain when modeled by itself and therefore contains the most useful information in isolation, whereas elevation decreases the gain the most when omitted and therefore contains information that is not present in other variables. For the second night of release, UPKE2 (Figure 4d), distance to the fence

contributes the highest gain and therefore most useful information by itself, and elevation again decreases the gain the most.

Discussion

The MaxEnt model results from the four release periods indicate the importance of managerial factors on the overall distribution of stock within and near meadows. For all release periods, a managerial variable was the top contributor to the model in terms of percent contribution, permutation importance, and jackknife variable importance when examined in isolation (i.e., contained the most useful information for the model as compared to the other variables). This included the release point, which was a top contributor in three of the four releases, with animals more likely to be present near the location of release. This is consistent with findings from an NPS pilot study incorporating 14 additional sites using the data collection procedures described in this paper and examining intensively used areas with kernel density estimates (Walden-Schreiner et al., 2015). Moving the location at which stock are released is a feasible tactic in landscapes where terrain is not a limiting factor. Further research on how this tactic influences distribution, especially in arid environments or locations where increasing numbers of tourists are leveraging stock to transport supplies in limited spaces, is needed.

The only instance where release point was not one of the top two variables in terms of percent contribution was during the second release in UPKE when the fence was 50 m further from the release point and became the top contributor to the model, with a permutated

importance of 96.1%. The high permutation importance and model AUC gain by the fence variable as illustrated by the jackknife plot suggests that the release point did not contain additional information not already captured by other variables. This suggests releasing animals away from sensitive habitats or in rotating locations may encourage grazing near the area of release and limit impacts.

Additionally, the use of an electric fence was also often an important factor, though the distance animals traveled varied between releases. Findings suggests additional conditions contributed to the effectiveness of the fence at limiting the distribution of the group and warrants further investigation of meadow size, fence visibility, and the social dynamics of the group. This was most noticeable during the first release in UPKE, as the fence was not among the top contributors to the model. Instead, the release point, which was the same location on both nights, became the top contributor to the model, with elevation and distance to a water body also ranking higher than distance to fencing. During the first night, the fence was approximately 50m closer to the release point and 200 m further north in the meadow than the second night. A potential explanation could include the fence being closer to the release point and not contributing as much new information to the model. However, the distance changes are relatively small at 50 m, and social dynamics within the group may also have contributed to differences. If directing stock grazing to smaller areas without overgrazing a particular location, the use of a fence to confine the lead animal may limit the dispersal of the other animals, but that social dynamics of the herd may also be important

factors of animal dispersal. This finding is relevant for tourism operations that rely on the same animals who may have prior knowledge of grazing locations.

With respect to environmental variables, elevation was one of the top contributors to three of the four models. When considered in conjunction with land cover, animals were more often present at similar elevations as the meadow, even when outside of the meadow boundary. This may result from similar foraging opportunities associated with plant communities within the elevation gradient, or stock seeking resting locations not requiring traversing steep slopes. The relative logistic probability of presence dropped sharply to near zero at a slope of approximately five degrees (10% rise) or less. This parallels findings of studies of domestic grazing livestock (i.e., cattle, sheep) that also avoid steep slopes (see Bailey and Brown, 2011).

NDVI and dominant regional vegetation alliance consistently ranked as not contributing much new information, by themselves, to the models across all release periods. This is in slight contrast to Poor and others (2012) who found NDVI to be among the top two variables (and fencing among the least) contributing to a MaxEnt model of GPS-tracked pronghorn migration and suitable habitat in Montana, USA and Saskatchewan, Canada. However, the area of interest in this study is at the site level, as opposed to regional scale, and examines the distribution of domestic grazing animals within a natural environment. Thus, spatial and temporal scale could limit the opportunity for stock to experience a range of NDVI values. Additionally, while NDVI may not have contributed much new information to the model not contained in other variables, it is important to note that NDVI, by itself, achieved AUC

values between 0.62 and 0.80 (i.e., better than random to fair model performance) as displayed in the jackknife plots. Response curves of the variable only, which enables interpretation free of correlations among predictor variables, indicate at which values of the predictor variable increased probability of presence was likely. In this site-scale study, lodgepole pine (*Pinus contorta*) and wet meadow regional vegetation alliances had a higher probability of stock presence than others. Further exploration of NDVI as a predictor variable in stock distribution is needed, especially in arid environments where stark contrasts may exist between meadows and adjacent areas.

Management Implications

Results highlight the importance of managerial factors in the distribution of stock. Managers tasked with mitigating impacts associated with stock use when animals are released to graze can consider the strategic placement of release points and portable fences. Encouraging release points in locations near quality forage and away from sensitive habitats, if available, may promote a concentration strategy, with stock remaining near the release location and away from areas more prone to impact. Conversely, rotating release points can promote a dispersal strategy at a broader spatial scale, spreading out grazing impacts over the course of a growing season and grazing area. How to implement strategies, concentration or dispersal, should be informed by area size, vegetation and hydrologic composition, number of stock use nights, and context of use.

Additionally, strategies can be used independently or together at different scales to accomplish resource protection goals. For example, a smaller site that is heavily used may choose a concentration strategy to minimize the area impacted. However, concentration relies on ample forage to support grazing in the area, otherwise animals will move further away to graze, reducing the influence of the release location. Stock in this study also concentrated use near portable electric fencing, when used during a release period. This tactic may also support both concentration and confinement strategies, depending on whether the fencing is moved systematically throughout the area, and may be best implemented in areas allowing for commercial or administrative stock use.

Limitations and Future Research

Due to the dynamic nature of pack stock trips, only a limited number of releases at the same site under similar conditions were available for analysis. Future research should select additional sites from which to develop a randomized study design, incorporating repeat visits that cover the same length of release, similar stock handling techniques, and pre-release activity and grazing. Additionally, incorporating more habitat heterogeneity and geographically diverse release points and fence locations may further explore distance-decay relationships related to stock distribution and managerial factors. Although the NDVI was not among the top contributors to the model, fine-scale vegetation mapping and phenology coupled with NDVI would examine fine-scale relationships between stock habitat use and vegetation. This would further elucidate how moving release points interacts with vegetation,

as previous research in high elevation meadows in the Sierra Nevada indicated horses selected graminoids over forbs until foraging became limited within the area (Olson-Rutz et al., 1996a).

Finally, an important next step is to examine and incorporate stock behavior and group dynamics to expand the analysis of factors influencing distribution of stock when released to graze. Coupling direct observation with GPS tracking provides context for metrics (e.g., speed, turning angle) derived from GPS data that underpin the modelling of spatio-temporal behavior (Gurarie et al., 2016). Modelling group movement, which may help account for social dynamics, is also a growing area of research in movement ecology (Langrock et al., 2014). Modelling both behavior and group movement from data collected as part of this study may further explain why top management variables identified in three models were not among the top variables in the fourth model and afford the opportunity to examine differences between administrative, commercial, and private stock use and behavior.

Conclusions

This study explored the importance of two management tactics (i.e., release locations and the use of portable electric fences) and environmental factors (i.e., elevation, slope, distance to water, dominant vegetation alliance cover, and NDVI) on stock distribution in two high elevation meadows using GPS tracking and MaxEnt modelling. Evaluation of the model's predictive success using AUC values indicated the models performed well (AUC >0.9), suggesting MaxEnt offered an accessible method to generate distribution models for pack

stock with respect to continuous and categorical data. The findings from this study, as well as the application of this modelling method to GPS tracking data of domestic animals when released to graze in natural areas, can help managers and planners in determining which tactics may be more effective in minimizing and mitigating impacts associated with free-range grazing in sensitive habitats.

Acknowledgements

We thank all of the collaborators who provided their time and expertise to this project, including NPS wranglers R. Colflesh, R. Fredy, and T. Bernacchi for managing the stock involved in this study; NPS Resources Management and Science staff L. Austin, J. Baccei, A. Dickenson, M. Marschall, E. Babich, M. Booher, A. Puchkoff, L. Mazzu, and L. Jones for their dedicated efforts in the field and compiling the data; S. Ostoja from the USGS/USFS for his helpful feedback throughout the project; and to D. Johnson, G. Alexander, and A. Day from Oregon State University for the assembly and testing of the custom GPS units used in this project.

Funding: This project was funded by the USDI National Park Service under the Cooperative Ecosystem Studies Unit Task Agreement P14AC01493.

References

Aarts, G., MacKenzie, M., McConnell, B., Fedak, M., Matthiopoulos, J., 2008. Estimating space-use and habitat preference from wildlife telemetry data. Ecography. 31, 140–160.

Abbe, J.D., Ballenger, L., 2011. Managing overnight stock use at Yosemite National Park: A science-based approach. Park Sci. 28, 99–102.

Alzerreca, H., Laura, J., Loza, F., Luna, D., Ortega, J. 2006. Importance of carrying capacity in sustainable management of key high-Andean Puna rangelands (bofedales) in Ulla Ulla, Bolivia. In: Spehn, E., Liberman, M., Korner, C. (Eds.), Land Use Change and Mountain Biodiversity. Taylor & Francis: Boca Raton, FL, pp. 167-186.

Bailey, D.W., Brown, J.R., 2011. Rotational grazing systems and livestock grazing behavior in shrub-dominated semi-arid and arid rangelands. Rangel. Ecol. Manag. 64, 1–9.

Bailey, D.W., Gross, J.E., Laca, E.A., Rittenhouse, L.R., Coughenour, M.B., Swift, D.M., Sims, P.L., 1996. Mechanisms that result in large herbivore grazing distribution patterns. J. Range Manag. 49, 386–400.

Baldwin, R.A., 2009. Use of maximum entropy modeling in wildlife research. Entropy 11, 854–866.

Ballenger, E., Baccei, J., Acree, L., 2010. 2008 Pack Stock Use Assessment in Subalpine Meadows of the Tuolumne River Watershed. USDI National Park Service. El Portal, CA. Retrieved from https://www.nps.gov/yose/learn/management/upload/2008_Stock-Use-Report-Final_two_sided_12-02-10.pdf

Ballenger, L., Wilkin, K., Acree, L., Baccei, J., Whittaker, T., Babich, E., 2011. 2010 Assessment of Meadows in the Merced River Corridor, Yosemite National Park. USDI National Park Service. El Portal, CA. Retrieved from https://www.nps.gov/yose/learn/management/upload/MRP-Meadows_Report_20120424.pdf

Balmford, A., Green, J.M.H., Anderson, M., Beresford, J., Huang, C., Naidoo, R., Walpole, M., Manica, A., 2015. Walk on the wide side: Estimating the global magnitude of visits to protected areas. PLOS Biol. 13.

Barros, A., Monz, C., Pickering, C., 2015. Is tourism damaging ecosystems in the Andes? Current knowledge and an agenda for future research. Ambio 44, 82-98.

Barros, A., Pickering, C.M., 2015. Impacts of experimental trampling by hikers and pack animals on a high-altitude alpine sedge meadow in the Andes. Plant Ecol. Divers. 8, 265-276.

Barros, A., Pickering, C.M., Renison, D., 2014. Short-term effects of pack animal grazing exclusion from Andean alpine meadows. Arctic, Antarct. Alp. Res. 46, 333–343.

Braunisch, V., Patthey, P., Arlettaz, R., 2011. Spatially explicit modeling of conflict zones between wildlife and snow sports: Prioritizing areas for winter refuges. Ecol. Appl. 21, 955–967.

Brown, D.D., Kays, R., Wikelski, M., Wilson, R., Klimley, A., 2013. Observing the unwatchable through acceleration logging of animal behavior. Anim. Biotelemetry 1, 1–16, e1002074.

Byers, A.C., 2009. A comparative study of tourism impacts on alpine ecosystems in the Sagarmatha (Mt. Everest) National Park, Nepal and the Huascaran National Park, Peru., in: Hill, J., Gale, T. (Eds.), Ecotourism and Environmental Sustainability: Principles and Practice. Routledge, London, UK, pp. 51–72.

Chape, S., Harrison, J., Spalding, M., Lysenko, I., 2005. Measuring the extent and effectiveness of protected areas as an indicator for meeting global biodiversity targets. *Philo. Trans. R. Soc. Lond. B. Biol. Sci.* **360**, 443–55.

Cole, D.N., Spildie, D.R., 1998. Hiker, horse, and llama trampling effects on native vegetation in Montana, USA. J. Environ. Manag. 53, 61–71.

Cole, D.N., van Wagtendonk, J.W., McClaran, M.P., Moore, P.E., McDougald, N.K., 2004. Response of mountain meadows to grazing by recreational pack stock. J. Range Manage. 57, 153–160.

Coppes, J., Braunisch, V., 2013. Managing visitors in nature areas: Where do they leave the trails? A spatial model. Wildlife Biol. 19, 1–11.

Duncan, P., 1983. Determinants of the use of habitat by horses in a Mediterranean wetland. J. Anim. Ecol. 52, 93–109.

Edren, S.M.C., Wisz, M.S., Teilmann, J., Dietz, R., Soderkvist, J., 2010. Modelling spatial patterns in harbour porpoise satellite telemetry data using maximum entropy. Ecography 33, 698–708.

Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A,

Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J., Townsend Peterson, A., Phillips, S.J., Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberón, J., Williams, S., Wisz, M.S., Zimmermann, N.E., 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29, 129–151.

Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E., Yates, C.J., 2011. A statistical explanation of MaxEnt for ecologists. Divers. Distrib. 17, 43–57.

Fourcade, Y., Engler, J.O., Rödder, D., Secondi, J., 2014. Mapping species distributions with MAXENT using a geographically biased sample of presence data: A performance assessment of methods for correcting sampling bias. PLoS One 9, e97122.

Frair, J.L., Fieberg, J., Hebblewhite, M., Cagnacci, F., DeCesare, N.J., Pedrotti, L., 2010. Resolving issues of imprecise and habitat-biased locations in ecological analyses using GPS telemetry data. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 365, 2187–2200.

Gaylord, A.J., Sanchez, D.M., 2014. Causes and consequences of timing errors associated with global positioning system collar accelerometer activity monitors. J. Fish Wildl. Manag. 5, 372–379.

Geneletti, D., Dawa, D., 2009. Environmental impact assessment of mountain tourism in developing regions: A study in Ladakh, Indian Himalaya. Environ. Impact Assess. Rev. 29, 229–242.

Graves, T.A., Waller, J.S., 2006. Understanding the causes of missed global positioning system telemetry fixes. J. Wildl. Manage. 70, 844–851.

Gschweng, M., Kalko, E.K. V., Berthold, P., Fiedler, W., Fahr, J., 2012. Multi-temporal distribution modelling with satellite tracking data: Predicting responses of a long-distance migrant to changing environmental conditions. J. Appl. Ecol. 49, 803–813.

Guillera-Arroita, G., Lahoz-Monfort, J.J., Elith, J., Gordon, A., Kujala, H., Lentini, P.E., Mccarthy, M.A., Tingley, R., Wintle, B.A., 2015. Is my species distribution model fit for purpose? Matching data and models to applications. Glob. Ecol. Biogeogr. 24, 276–292.

Gurarie, E., Andrews, R.D., Laidre, K.L., 2009. A novel method for identifying behavioural changes in animal movement data. Ecol. Lett. 12, 395–408.

Gurarie, E., Bracis, C., Delgado, M., Meckley, T.D., Kojola, I. Wagner, C.M. 2016. What is the animal doing? Tools for exploring behavioural structure in animal movements. J. Ani. Ecol. 85, 69-84.

Handcock, R.N., Swain, D.L., Bishop-Hurley, G.J., Patison, K.P., Wark, T., Valencia, P., Corke, P., O'Neill, C.J., 2009. Monitoring animal behaviour and environmental interactions using wireless sensor networks, GPS collars and satellite remote sensing. Sensors 9, 3586–3603.

Holmquist, J.G., Schmidt-Gengenbach, J., Ballenger, E.A., 2014. Patch-scale effects of equine disturbance on arthropod assemblages and vegetation structure in subalpine wetlands. Environ. Manag. 53, 1109-1118.

Huang, T.-C., Huang, Y.-J., Lin, W.-C., 2013. Real-time horse gait synthesis. Comput. Animat. Virtual Worlds 24, 87–95.

Jacoby, P.W., The Society for Range Management, 1989. A Glossary of Terms Used in Range Management: A Definition of Terms Commonly Used in Range Management, third ed. Society for Range Management, Denver, CO.

Jalali, M.A., Ierodiaconou, D., Monk, J., Gorfine, H., Rattray, A., 2015. Predictive mapping of abalone fishing grounds using remotely-sensed LiDAR and commercial catch data. Fish. Res. 169, 26–36.

Jaynes, E.T., 1957. Information theory and statistical mechanics. Phys. Rev. 106, 620–630. doi:10.1103/PhysRev.106.620

Jaynes, E.T., 1982. On the rationale of maximum-entropy methods. Proc. IEEE 70, 939–952.

Jeltsch, F., Bonte, D., Pe'er, G., Reineking, B., Leimgruber, P., Balkenhol, N., Schröder, B., Buchmann, C.M., Mueller, T., Blaum, N., Zurell, D., Böhning-Gaese, K., Wiegand, T., Eccard, J.A, Hofer, H., Reeg, J., Eggers, U., Bauer, S., 2013. Integrating movement ecology with biodiversity research - exploring new avenues to address spatiotemporal biodiversity dynamics. Mov. Ecol. 1, 1–13.

Kays, R., Crofoot, M.C., Jetz, W., Wikelski, M., 2015. Terrestrial animal tracking as an eye on life and planet. Science 348, aaa2478-1—aaa2478-9. 1222–1223.

Kramer-Schadt, S., Niedballa, J., Pilgrim, J.D., Schröder, B., Lindenborn, J., Reinfelder, V., Stillfried, M., Heckmann, I., Scharf, A.K., Augeri, D.M., Cheyne, S.M., Hearn, A.J., Ross, J., Macdonald, D.W., Mathai, J., Eaton, J., Marshall, A.J., Semiadi, G., Rustam, R., Bernard, H., Alfred, R., Samejima, H., Duckworth, J.W., Breitenmoser-Wuersten, C., Belant, J.L., Hofer, H., Wilting, A., 2013. The importance of correcting for sampling bias in MaxEnt species distribution models. Divers. Distrib. 19, 1366–1379.

Kuhn, T., Ballenger, L., Scherer, R., Williams, J.N., 2015. Data Analysis and Assessment of High Elevation Wilderness Meadows Surveyed from 2008 to 2011, Yosemite

National Park. Natural Resource Report NPS/YOSE/NRR-2014/923. Fort Collins, CO. Retrieved from https://www.nps.gov/yose/learn/nature/upload/Data-Analysis-and-Assessment-of-High-Elevation-Wildereness-Meadows-Surveyed-from-2008-to-2011-Yosemite-National-Park.pdf.

Langrock, R., Hopcraft, J.G.C., Blackwell, P.G., Goodall, V., King, R., Niu, M., Patterson, T.A., Pedersen, M.W., Skarin, A., Schick, R.S., 2014. Modelling group dynamic animal movement. Methods Ecol. Evol. 5, 190–199.

Lewis, J.S., Rachlow, J.L., Garton, E.O., Vierling, L.A., 2007. Effects of habitat on GPS collar performance: Using data screening to reduce location error. J. Appl. Ecol. 44, 663–671.

Long, J.A., Nelson, T.A., 2013. A review of quantitative methods for movement data. Int. J. Geogr. Inf. Sci. 27, 292–318.

McClaran, M.P., 1989. Recreational pack stock management in Sequoia and Kings Canyon National Parks. Rangelands 11, 3–8.

McClaran, M.P., Cole, D.N., 1993. Packstock in wilderness: Use, impacts, monitoring, and management. General Technical Report INT-301. USDA Forest Service, Intermountain Research Station, Ogden, UT.

Meijles, E.W., de Bakker, M., Groote, P.D., Barske, R., 2014. Analysing hiker movement patterns using GPS data: Implications for park management. Comput. Environ. Urban Syst. 47, 44–57.

Merow, C., Smith, M.J., Silander, J.A., 2013. A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. Ecography 36, 1058–1069.

Michelot, T., Langrock, R., Patterson, T.A., 2016. moveHMM: An R package for the statistical modelling of animal movement data using hidden Markov models. Methods Ecol. Evol. 7, 1308-1315.

Monz, C.A., Cole, D.N., Leung, Y.-F., Marion, J.L., 2010. Sustaining visitor use in protected areas: Future opportunities in recreation ecology research based on the USA experience. Environ. Manag. 45, 551–562.

Moore, P.E., Cole, D.N., van Wagtendonk, J.W., McClaran, M.P., McDougald, N., McCool, S.F., Borrie, W.T., O'Loughlin, J., 2000. Meadow response to packstock grazing in the Yosemite wilderness: Integrating research and management. In: Cole, D.N., McCool, S.F., Borrie, W.T., O'Loughlin, J., (Eds.), Wilderness Science in a Time of

Change Conference: Wilderness Ecosystems, Threats, and Management; Proceedings RMRS-P-15-VOL-5. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Ogden, UT, pp. 160–164.

Olson-Rutz, K.M., Marlow, C.B., Hansen, K., Gagnon, L.C., Rossi, R.J., 1996a. Recovery of a high elevation plant community after packhorse grazing. J. Range Manag. 49, 541–545.

Olson-Rutz, K.M., Marlow, C.B., Hansen, K., Gagnon, L.C., Rossi, R.J., 1996b. Packhorse grazing behavior and immediate impact on a timberline meadow. J. Range Manag. 49, 546–550.

Ostoja, S.M., Brooks, M.L., Moore, P.E., Berlow, E.., Blank, R., Roche, J., Chase, J., Haultain, S., 2014. Potential environmental effects of pack stock on meadow ecosystems of the Sierra Nevada, USA. Rangel. J. 36, 411–427.

Phillips, S., Elith, J., 2016. On estimating probability of presence from use — availability or presence — background data. Ecology 94, 1409–1419.

Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. Ecol. Modell. 190, 231–259.

Phillips, S.J., Dudík, M., 2008. Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. Ecography 31, 161–175.

Poor, E.E., Loucks, C., Jakes, A., Urban, D.L., 2012. Comparing habitat suitability and connectivity modeling methods for conserving pronghorn migrations. PLoS One 7, e49390.

Radosavljevic, A., Anderson, R.P., 2014. Making better Maxent models of species distributions: Complexity, overfitting and evaluation. J. Biogeogr. 41, 629–643.

Renner, I.W., Warton, D.I., 2013. Equivalence of MAXENT and Poisson Point Process models for species distribution modeling in ecology. Biometrics 69, 274–281.

Renner, I.W., Elith, J., Baddeley, A., Fithian, W., Hastie, T., Phillips, S.J., Popovic, G., Warton, D.I. 2015. Point process models for presence-only analysis. Methods Ecol. and Evol. 6, 366–379.

Richards, D.R., Friess, D.A., 2015. A rapid indicator of cultural ecosystem service usage at a fine spatial scale: Content analysis of social media photographs. Ecol. Indic. 53, 187–195.

Royle, J.A., Chandler, R.B., Yackulic, C., Nichols, J.D., 2012. Likelihood analysis of species occurrence probability from presence-only data for modelling species distributions. Methods Ecol. Evol. 3, 545–554.

Rutter, S.M., 2007. The integration of GPS, vegetation mapping and GIS in ecological and behavioural studies. Rev. Bras. Zootec. 36, 63–70.

Smith, A., Page, B., Duffy, K., Slotow, R., 2012. Using Maximum Entropy modeling to predict the potential distributions of large trees for conservation planning. Ecosphere 3, 1–21.

Smith, A.B., 2013. On evaluating species distribution models with random background sites in place of absences when test presences disproportionately sample suitable habitat. Divers. Distrib. 19, 867–872.

Sumner, E.J., 1935. Report on the Yosemite Saddle and Pack Stock Grazing Problem. USDI National Park Service; Yosemite National Park, CA.

UNESCO, 2016. Yosemite National Park. http://whc.unesco.org/en/list/308 (accessed 01.06.16).

USDA, 2016. NAIP Imagery. Aer. Imag. URL https://www.fsa.usda.gov/programs-and-services/aerial-photography/imagery-programs/naip-imagery/ (accessed 28.03.16).

USDA Forest Service, 2008. CALVEG Vegetation Classification and Mapping. URL http://www.fs.usda.gov/detail/r5/landmanagement/resourcemanagement/?cid=stelprdb5347192.

USDI NPS, 2014a. Merced Wild and Scenic River Final Comprehensive Management Plan and Environmental Impact Statement. http://www.nps.gov/yose/parkmgmt/mrp_finalplan.htm (accessed 15.02.15).

USDI NPS, 2014b. Tuolumne Wild and Scenic River Final Comprehensive Management Plan and Environmental Impact Statement. http://www.nps.gov/yose/parkmgmt/trp_feis.htm (accessed 15.02.15).

USDI NPS, 2016a. Annual visitation summary report for 2015. https://irma.nps.gov/Stats/SSRSReports/National Reports/Annual Visitation Summary Report (1979 - Last Calendar Year). (accessed 01.06.16).

USDI NPS, 2016b. Wilderness Stewardship Plan - Yosemite National Park. https://www.nps.gov/yose/getinvolved/wsp.htm (accessed 10.08.16).

USDI NPS, USGS, 2011. Hydrology geodatabase for Yosemite National park. https://irma.nps.gov/DataStore/Reference/Profile/2170437 (accessed 28.03.16).

USGS, 2016. The National Map: Elevation. http://nationalmap.gov/elevation.html (accessed 28.03.16).

van Gils, H., Kayijamahe, E., 2010. Sharing natural resources: Mountain gorillas and people in the Parc National des Volcans, Rwanda. Afr. J. Ecol. 48, 621–627.

Viers, J.H., Purdy, S.E., Peek, R.A, Fryjoff-Hung, A., Santos, N.R., Katz, J.V.E., Emmons, J.D., Dolan, D. V, Yarnell, S.M., 2013. Montane Meadows in the Sierra Nevada: Changing Hydroclimatic Conditions and Concepts for Vulnerability Assessment. Technical Report CWS-2013-01. University of California, Davis, Davis, CA.

Walden-Schreiner, C., Tsai, W.-L., Leung, Y.-F., 2015. Geospatial analysis of pack stock movement and grazing patterns pilot study in subalpine meadows at Yosemite National Park. Unpublished Final Technical Report submitted for CESU Task Agreement P14AC01493. North Carolina State University, Raleigh, NC.

Warton, D., Aarts, G., 2013. Advancing our thinking in presence-only and used-available analysis. J. Anim. Ecol. 82, 1125–1134.

Watson, J.E.M., Dudley, N., Segan, D.B. Hockings, M., 2014. The performance and potential of protected areas. Nature, 515, 67–73.

Weaver, T., Dale, D., 1978. Trampling effects of hikers, motorcycles and horses in meadows and forests. J. Appl. Ecol. 15, 451–457.

Yackulic, C.B., Chandler, R., Zipkin, E.F., Royle, J.A., Nichols, J.D., Campbell Grant, E.H., Veran, S., 2013. Presence-only modelling using MAXENT: When can we trust the inferences? Methods Ecol. Evol. 4, 236–243.

Young, N., Carter, L., Evangelista, P., Jarnevich, C., 2011. A MaxEnt Model v3.3.3e Tutorial (ArcGIS v10).

CHAPTER 4

Digital Footprints in Protected Areas:
Incorporating Crowdsourced Geographic Information to Support Resource Protection

Abstract

Biodiversity loss driven by anthropogenic pressures highlights the importance of conservation efforts in protected areas globally. Protected areas are also locations providing myriad ecosystem services, including recreation and tourism. Advancements in mobile and web technologies have expanded the capabilities and accessibility of user generated spatial content increasingly leveraged for research. This study explores the use of crowdsourced geographic information to model, at varying temporal scales, spatial patterns of visitor use and identify factors contributing to distribution patterns in a dynamic landscape, Hawaii Volcanoes National Park, Hawaii, USA. Specifically, this study integrated geotagged photos publicly shared on Flickr with raster data about infrastructure and natural environmental using MaxEnt modelling. Infrastructure designated for visitor use (i.e., roads, trails) contributed most to models of visitor distribution for all years and seasons. During the spring months, elevation was also a top contributing variable to the model. Results suggest crowdsourced data provided empirical assessments of covariates influencing visitor distributions, informing how changes in infrastructure and environmental factors may

influence visitor use, and therefore resource pressures, to help researchers, managers, and planners with efforts to mitigate negative impacts.

Keywords: Geotagged photos, Information communication technology (ICT), MaxEnt modelling, User generated spatial content

Introduction

Environmental changes and biodiversity loss attributed to human activities increased more rapidly in the last century than during any other time in recorded history (Millenium Ecosystem Assessment, 2005; Pimm et al., 2014). This unprecedented rate and scale of human impact and ecological footprint ushered proposals to name a new geological epoch, the Anthropocene (Crutzen, 2002; Steffen et al., 2011). Geographic Information Science (GIScience) is central to exploring, understanding, and modelling the human footprint (Sanderson et al., 2002) and human-environment interactions (Zvoleff & An, 2014), coupling social and ecological datasets through geographic location for integrated analysis. Advances in mobile and remote sensing technologies have increased the amount, access, and accuracy of spatial data available for integrated analysis, fostering a paradigm shift in GIScience. Specifically, how spatial data are produced and consumed have rapidly expanded to new users and platforms (Goodchild, Aubrecht, & Bhaduri, 2017; Sui, Elwood, & Goodchild, 2013). As of 2010, the amount of geocoded data grew an average of one exabyte (i.e., one

million terabytes) per day, with personal location data from mobile devices comprising approximately one petabyte (i.e., 1,000 terabytes) of data each year (Sui et al., 2013).

The ubiquitous and accessible nature of big (geo)data facilitated a proliferation and examination of non-traditional spatial data products and innovative data collection methods in research exploring human-environment interactions (Longley & Adnan, 2016; See et al., 2016; Senaratne, Mobasheri, Ali, Capineri, & Haklay, 2017). Portable electronic devices, including mobile phones and tablets, enable individuals to act as sensors of their environments capable of generating and sharing geospatial data on big data scales. While important considerations about the digital divide and accessibility persist, a 2016 survey of global trends found a median 43% of adults own a smartphone, with over 70% ownership observed in the U.S., Australia, Spain, Israel, and Korea (Pew Research Center, 2016). Digital footprints, or the "digital traces each of us leaves behind as we conduct our lives," (Weaver & Gahegan, 2007, p. 324), are a prime example of VGI. Digital footprints captured by mobile devices and shared through social media platforms, like geotagged photos, offer an opportunity to study human-environment interactions central to geography at high spatial and temporal resolutions in near real-time (Wang, Guo, Fu, & Li, 2014). This research explores VGI originating outside of formal research endeavors to quantify human presence and model infrastructure and environment interactions in areas designated for biodiversity conservation.

Protected areas (PAs) are an especially relevant context in which to explore the capacity of VGI in addressing human-environment interactions and associated impacts. PAs are designated terrestrial and marine landscapes formally managed to conserve nature,

ecosystem services, and cultural values. PAs are core components in the strategy to minimize biodiversity loss globally, and human (i.e., visitor) data are often limited due to myriad competing research needs and challenges associated with visitor data collection. Part of the barrier to detailed visitor data is that monitoring visitor use in natural areas can be difficult to sample and time consuming (Cessford & Muhar, 2003), especially in large or topographically complex PAs. Additionally, staffing or financial constraints may necessitate prioritizing other objectives or concentrating visitor monitoring to select locations.

The dearth of visitor data in PAs is credited as a main limitation in implementing proactive management strategies to minimize visitor impact on resources (Hadwen, Hill, & Pickering, 2008) and to measure ecosystem services (Schägner, Brander, Maes, Paracchini, & Hartje, 2016). PA management strives to conserve natural, cultural, and wildlife resources within its boundaries, while allowing visitor use deemed appropriate based on the PA classification. Encroaching development or resource extraction near PA boundaries, climate change, and limited budgets can challenge efforts to balance use with biodiversity. To understand the spatial implications of environmental changes on visitor use patterns and support sustainable and effective management of PAs, managers require timely data on resource conditions and visitor use at site specific and landscape scales, as well as analytical tools to conduct integrated analyses to understand human-environment (i.e., visitor-environment) interactions.

Research objectives

This study advances the spatio-temporal understanding of human visitation in PAs by leveraging their digital footprints, manifested as publicly shared geotagged photos, as an efficient spatial data source to:

(1) Model the spatial distribution of visitor use to identify spatial patterns of visitation at multiple temporal scales and;

(2) Assess relationships between infrastructure, environmental factors, and visitor distribution patterns.

Review of Methods

Spatial visitor use monitoring

Visitor use and impact monitoring within PAs focuses on use levels, locations, and associated environmental and social impacts to identify trends and alert management to departures from desired conditions (Cessford & Burns, 2008; Leung & Monz, 2006; Tomczyk & Ewertowski, 2011). The number of visitors to an area represents the most basic yet vital information for PAs, important for planning facility design and capacity, aligning with PA mission objectives, and identifying potential for negative impacts (Cessford & Muhar, 2003). A variety of counting methods exist, ranging from observation studies, registries, permits, and on-site surveys to the use of electronic traffic counters, sensors, and

photographic or videographic methods (Arnberger, Haider, & Brandenburg, 2005; Marvin et al., 2016; O'Connor, Zerger, & Itami, 2005; Xia & Arrowsmith, 2008).

Global positioning system (GPS) tracking and computer modelling represent increasingly leveraged methods for understanding spatial use patterns and provide high-resolution data to quantify use intensities, explore changes in spatio-temporal patterns, and understand resource impacts (Beeco & Brown, 2013; Hallo et al., 2012; Meijles et al., 2014; Van Kirk et al., 2014). Collectively, traditional and emerging approaches provide necessary and important information for PA conservation. However cost, efficiency, or environment can still hinder wider or repeat implementation (Eagles, 2014). Additionally, data collection is often the responsibility of the research partner or managing organization and subject to staffing and budget constraints.

Prioritizing sampling strategy to high use areas or to coincide with peak periods of visitation is vital in protecting ecological and site integrity at well-known destinations and ensuring adequate sample sizes, but this approach may miss emerging trends or seasonal variations. Emerging trends can lead to unintended negative impacts. Recreation ecology literature has shown that impacts can happen quickly, with small amounts of use causing disproportionate amounts of impact (Hammitt et al., 2015; Monz et al., 2010). Furthermore, although changing climate patterns vary geographically in intensity and manifestations, changes in the location or quality of resources may affect utilization and visitor-environment interaction patterns (Jones & Scott, 2006). Recent studies demonstrate that visitors are already changing their decisions about the timing, frequency, or duration of PA visits

(Buckley & Foushee, 2012; Loomis & Richardson, 2006; McEvoy, Cavan, Handley, McMorrow, & Lindley, 2008; Richardson & Loomis, 2005; Scott, Jones, & Konopek, 2007), further complicating the sampling schemes of visitor data collection.

The focus on temporal changes in visitation patterns, coupled with changing environmental conditions, also suggests the possibility of spatial shifts in visitor-environment interactions. Changes in average precipitation and temperature have altered some species distributions (Chen et al., 2011) and phenology (Lesica & Kittelson, 2010; Menzel et al., 2006), together with changing seasonal access in PAs, affords opportunities for temporal and spatial changes in use patterns. These changing patterns may not be congruent with existing infrastructure (i.e., managed trails) or management strategies, resulting in new or increasing negative impacts. Quantifying and visualizing visitor distributions is especially relevant when the landscape itself is also dynamic or unique, as patterns may change quickly or unexpectedly. Therefore, methods are also needed to collect near real-time visitor use data at a landscape scale without unduly burdening data collection and analysis efforts.

Research applications of geotagged photographs

VGI has helped address data collection constraints, with increasing work exploring concerns regarding accuracy, reliability, and credibility (Connors et al. 2012; Flanagin & Metzger 2008; Levin et al. 2017; Senaratne et al. 2017). Including the public in helping to collect scientific data (citizen science) continues as an active area of research in GIScience, though less attention has focused on crowdsourced data arising outside of a formal research

collaboration or citizen science effort (Connors et al., 2012). These platforms offer a potential solution to informing monitoring efforts of emerging trends or time-sensitive spatial data needs beyond the capacity of traditional monitoring.

Photo sharing sites like Flickr, Panoramio, Instagram, and others allow for cloud storage of user photos and map-based visualization of geotagged photo locations. In accordance with a site's privacy and use policy, researchers can query photo metadata using the site's application programming interface (API). APIs allow requests to be made of a server, app, or software, which then responds with data. APIs are also used to create mashups, websites or applications that conceptualize or integrate data from another source to create something new (Elwood et al., 2012). The use of APIs to extract human movement patterns from geotagged media has progressively been documented in tourism and recreation literature (see Zheng, Zha, & Chua, 2010), with Flickr offering a popular data mining platform due to the number of photos and accessible API (Levin et al., 2017). By 2007, one year after the inclusion of geotagging functionality, Flickr hosted over 20 million geotagged photos (Zheng et al., 2010). That number does not take into account the large volume of photos without geotags, for which researchers are also trying to determine how to retroactively assign a spatial attribute (Kalogerakis, Vesselova, Hays, Efros, & Hertzmann, 2009; Liu, Yuan, Cong, & Xu, 2014).

Flickr users have the option of manually geotagging a photo by placing it on a map if the camera or device does not have internal GPS capabilities, or leveraging an external GPS device to record the geographic location (Senaratne et al., 2017). However, of the photos

uploaded to Flickr in 2014, the most common devices included popular smart phone brands (Dove, 2015). Mobile phones with GPS are capable of achieving horizontal accuracy similar to recreational grade GPS units, often within 10 m of true position (Zandbergen & Barbeau, 2011).

The volume of geotagged photos suggests the impact of incorrectly tagged photos can be minimal, and crowdsourced correction capacity exists on sharing sites. Zielstra and Hochmair (2013) examined the spatial accuracy of geotagged photos through a random selection of geotagged photos from Flickr and Panoramio. Their results indicated median positional errors of 15 m and 46 m in North America for Panoramio and Flickr images, respectively, when compared to manually corrected position based on image content. Antoniou, Morley, and Haklay (2010) found geotagged photos in the UK were usually also uploaded within a few weeks following capture. In their sample, only 8.4% of Flickr photos were uploaded more than a year later, suggesting that data from Flickr can be both accurate and timely enough for some applications.

Applications of geotagged photo data relevant to visitors in PAs specifically are also garnering research attention. Wood and others (2013) used data from Flickr to estimate visitation rates at 836 natural and cultural recreation sites in 31 countries and found the number of uploaded photos was positively correlated with empirical visitation counts. They argue that geotagged photos can serve as reliable proxy for PA visitation rates and provide a statistical foundation from which to base further exploration. Sonter and others (2016) also found correlations between photo numbers and survey visits to Vermont, USA, PAs and

estimated economic contribution of nature-based tourism and landscape characteristics influencing the spatio-temporal patterns observed using Flickr photos.

Beyond simple location counts, Flickr photos also have the potential to document visitor movements and distribution within a PA. Addressing concerns about the relatively low density of geotagged photos in natural areas, Orsi and Geneletti (2013) employed a gravity model to estimate visitor flows for the Dolomites UNESCO World Heritage Site, a nine-unit site in Italy comprised of 231,170 ha of core and buffer areas, based on popular locations determined through high densities of geotagged photos validated with on-site counts. Results highlight how geotagged photos can provide visitor use estimates for a large area in a quick and inexpensive way.

Counts and estimated volumes are critical first steps toward understanding potential visitor pressure on natural resources within a PA, with research also needed to address what variables are influencing such distributions of use. Richards and Friess (2015) classified the content of geotagged photos to examine cultural ecosystem services in natural areas of Singapore. They found using geotagged photos provided a rapid assessment of what visitors derived from different habitats, as well as broad activity categories. Keeler and others (2015) used geotagged photos as a proxy for recreational visits to Midwestern lakes and found increased visits corresponded with increased water clarity. The accuracy, timeliness, and availability of geotagged photos represents a complement to traditional visitor use monitoring methods to assist in gathering park-scale data to inform focused monitoring efforts and possible resource impacts. The ability of geotagged photos to visualize visitor

distributions and potentially identify emerging and near real-time human-environment interactions difficult to capture with more traditional methods is especially relevant when the landscape itself is also dynamic.

Methods

Study area

Hawaii Volcanoes National Park (HVNP) is an example of a frequently changing landscape in which data on human-environment interactions offer a unique opportunity to explore spatio-temporal changes in visitor use patterns. HVNP is located on Island of Hawaii, the southernmost island in the Hawaiian Archipelago (Figure 1), and managed by the U.S. National Park Service (NPS). Established as a national park in 1916, HVNP was also designated a United Nations Educational, Scientific and Cultural Organization (UNESCO) Biosphere Reserve in 1980 (UNESCO, 2016), followed closely by its inscription as a UNESCO World Heritage site in 1986. Both designations make HVNP an important component of global conservation efforts, geared toward innovative and integrative strategies to protect outstanding natural and cultural values.

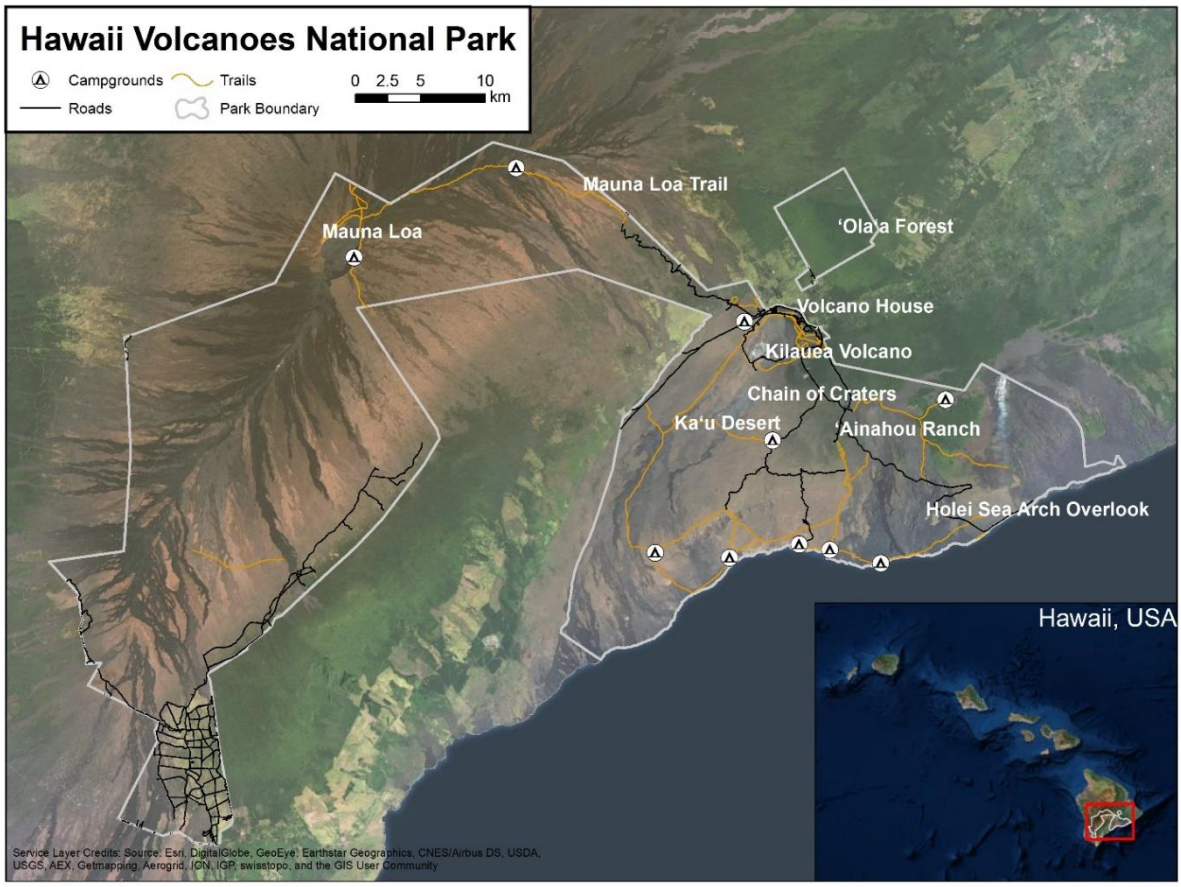


Figure 1 Map of Hawaii Volcanoes National Park

Due to the geographically isolated location of the archipelago, HVNP conserves habitat for a multitude of endemic plant and animal species. The NPS actively protects 54 threatened or endangered plant and animal species within seven contiguous biotic zones in the park's nearly 1300 km^2 (IUCN, 1987). Mauna Loa, considered the largest volcano in the world, is one of two active shield volcanoes in the park. Kīlauea is the second volcano in the

park, erupting continuously since 1983 and adding over 200 ha of new land and covering 14 km of the Chain of Craters Road along the southern boundary of the park.

The actively changing landscape and unique volcanic activity, geologic features, and cultural importance attracted over 1.8 million visits in 2015 and 2016 (NPS, 2016), and was ranked as the third most visited attraction, natural or built, in the Hawaiian Islands in 2015 (Hawaii Tourism Authority, 2015). The level of visitation and rapidly changing landscape of HVNP underscores the need for efficient snapshots of visitor use distributions. The park concluded the visitor comment period for the draft general management plan in June 2015, in which the NPS proposed plans to disperse visitors within the park through different methods described in each alternative and to increase access through proposed land acquisition (NPS, 2015). The potential for new visitor use patterns and locations, especially relevant in an actively changing landscape, could be supported through an understanding of emerging trends in the spatial distribution of visitors available from novel data sources like geotagged photos.

Geotagged photo acquisition

Geotagged photos were queried from the Flickr API using the statistical software program R (version 3.2.5), employing the code in Appendix C. We chose Flickr because it has been included as a data source in GIScience, geography, and tourism literature, offers an accessible API, and has experienced continued growth (Alivand & Hochmair, 2017). The boundary coordinates of HVNP established a bounding box (bbox) in which to census all

photos taken between January 1, 2011 and December 31, 2015 containing a geotag. Attributes for each photo record included latitude, longitude, owner, date taken, date uploaded, number of views, textual tags, photo title, accuracy rating, image URL, and capture device. The owner attribute is an auto-generated alpha-numeric identifier that attributes the photo record to a registered Flickr user. Records were written to comma-separated values files for statistical analysis in R and converted to shapefiles for geospatial analysis using ESRI ArcGIS (10.3.1).

As different resource management questions focus on multiple levels and locations of use, each photo was considered an incident of use and not weighted for multiple uploads by the same individual. If the same individual uploaded multiple photos from different locations within the park, each new spatial location is an indicator of use. Similarly, if an individual uploaded multiple photos from the same location taken over the course of several hours, the length of stay has resource use and impact implications. For example, visitor presence may displace wildlife from that location, indicating that the duration, and therefore each incidence, is important. To address the research objectives of this study focused on the geographical distribution of visitors, only the metadata from each photo were examined.

Geotagged photo data coordinates were assumed to be based on the WGS 1984 datum as it is the spatial reference for many aerial imagery base maps used in online mapping applications including DigitalGlobe which provides base maps for Flickr and Google Earth if users manually geotag their photos within the application. In converting photo records to shapefiles, the data were projected to UTM Zone 5N using the NAD 1983 datum to

correspond to environmental data projections and data formatting requirements for the distribution modelling software. Resulting shapefiles were clipped to park boundaries.

Visitor spatial distribution modelling

Distribution modelling approaches draw inference from presence-absence records and comparing the corresponding covariate conditions at each location type. The model assumes the phenomenon of interest will occur in locations with suitable conditions, a systematic assessment of absence has been conducted, and that observations are independent. Geotagged photos violate these assumptions, as do many plant and animal species distribution data sources. Maximum entropy (MaxEnt) models have been successfully applied using incomplete and correlated information (Elith et al., 2006, 2011). Used widely in species distribution modelling in ecology (Phillips et al., 2006), MaxEnt has only recently been applied to visitor use distributions in tourism and recreation research (Westcott & Andrew, 2015). MaxEnt models can integrate continuous and categorical predictor variables, minimize overfitting, and evaluate each covariate's influence. While other distribution modelling approaches can achieve similar objectives, MaxEnt was chosen because it has been shown to perform well, if not better, in comparison to other approaches when absence data are unavailable (Braunisch & Suchant, 2010; Elith et al., 2006) and offers an accessible graphical user interface. The use of MaxEnt for PA management of visitor impacts on natural resources has included human-wildlife interactions (Braunisch et al., 2011), prediction of off-

trail recreation behavior (Coppes & Braunisch, 2013; Westcott & Andrew, 2015), and cultural ecosystem services (Richards & Friess 2015).

Recent critiques of MaxEnt have included explorations of mathematical similarities to Poisson point process models (PPM) (Renner & Warton, 2013). Equivalencies between the two methods depend on how model parameters are configured and analyzed. Renner and others (2015) demonstrated how MaxEnt could be used as a means for fitting a Poisson PPM, namely by retaining duplicate presence records and examining the effect of the number of background points on the regularized training gain. For the purposes of this analysis, we focused on distributions within a park rather than predicting patterns in new environments. Duplicate presence points were retained and the MaxEnt algorithm allowed to choose the best feature type (i.e., linear, quadratic, product, threshold, or hinge). In machine learning, features are transformations of original covariate data, similar to basis functions in regression approaches, used to model complex, non-linear responses (Elith et al., 2011). Models presented in this paper were based on 10 replications for each temporal segment of interest, using random seed subsampling and withholding 30% of the data for training. The maximum number of background points was set at 10,000 with a convergence threshold of 0.00001 (Merow et al., 2013; Phillips et al., 2006; Poor et al., 2012; Young et al., 2011).

Independent variable data sources

Visitor infrastructure and environmental characteristics served as independent variables in the MaxEnt model (Table 1). If the original format was vector, the distance to each feature

was calculated and resulting values converted to raster. All independent variables either were originally, or resampled to, the coarsest resolution (i.e. 30 m). Response curves for each predictor were created and the variables were ranked based on their overall contribution to the model. Response curves illustrate how values of the predictor variable influence the logistic probability of presence, with upward trends indicating positive associations. The curves are separate MaxEnt models with only the predictor variable of interest and are clamped at the minimum and maximum values identified in the training data to minimize errors associated with extrapolation.

Predictive accuracy was evaluated using the area under the curve (AUC) (Jalali et al., 2015). Scores greater than 0.75 are considered acceptable for predictive accuracy (Phillips & Dudík 2008; Richards & Friess 2015), with AUC scores greater than 0.9 considered excellent (Coppes & Braunisch, 2013). The influence of each predictor variable was examined through permutation importance because the values are not dependent on the path taken by the model to arrive at a solution. Permutation importance reevaluates the MaxEnt model by randomizing the values for each variable separately to measure the decrease in AUC.

Table 1 Variables included in MaxEnt model

Variable	**Metric**	**Data Sources[1]**
Visitor use (dependent)	nx2 matrix of UTM coordinates of geotagged photos	Flickr
Buildings	Distance-to raster surface created from point features designated as buildings and structures	NPS
Trails	Distance-to raster surface created from line features designated as formal trails	NPS
Parking	Distance-to raster surface created from polygon features designated as formal parking	NPS
Roads	Distance-to raster surface created from line features designated as visitor accessible roads	NPS; HI DOT
Campgrounds	Distance-to raster surface created from point features identified as formal campgrounds	NPS
Slope	Percent slope derived from digital elevation model	USGS NED
Land cover	Categorical raster surface created from classified polygons	NPS
Water	Distance-to raster surface created from line (e.g., steam) and polygon (e.g., ocean) features	USGS

[1]Data Source Abbreviations: U.S. National Park Service (NPS), Hawaii Department of Transportation (HI DOT), U.S. Geological Survey (USGS), National Elevation Dataset (NED).

Photo clustering and hot spot identification

To complement park-level models of visitor distribution, spatial statistics quantified and characterized statistically significant hot spots of visitor use (García-Palomares et al., 2015). The Optimized Getis-Ord Gi* statistic examines hot spots based on the number of points in an area, and within the context of neighboring points, instead of numeric measurements at a specific location, and corrects for multiple testing and spatial dependence using the False Discovery Rate (FDR) method (Alivand & Hochmair, 2017). As a local indicator of spatial association (LISA), the Getis-Ord Gi* statistic provides an additional

level of analysis beyond density maps and distribution models to quantify and visualize locations of high or low density, reducing subjectivity in assessing the resulting visualization by identifying only locations achieving statistical significance.

Results

Geotagged photo sample characteristics

A query of public, geotagged photos on Flickr taken in the park between January 1, 2011 and December 31, 2015 resulted in 12,048 records. Photos were uploaded by 897 Flickr users, with a median of four photo uploads per user (interquartile range: 1 to 12 photos per user). Photos were uploaded to Flickr a median 12 days after capture (interquartile range: 2.75 to 35 days), with 40% of the sample uploaded within a week and over 70% within 30 days. Devices used to capture photos included film (0.18%) and digital cameras (66.4%), as well as mobile devices (19.3%) and video cameras (0.03%). Although digital cameras comprised more than two-thirds of the total 5-year sample (66.4%), recent years indicated a shift away from digital cameras and toward mobile devices (Figure 2). Of the mobile devices, which included mobile phones and tablet computers, the most common models included Apple iPhone (76.7%) and Samsung Galaxy S lines (11.5%).

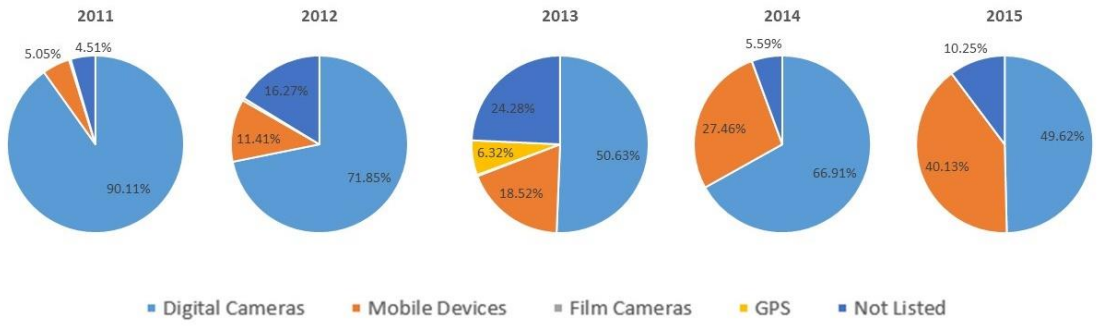


Figure 2 Type of device used to capture image shared to Flickr. Percentages based on the total number of records per year.

Records were grouped by year and season to investigate annual and seasonal variation in the models, respectively (Table 2). Years 2013 and 2014 had the greatest number of photos with over 2700 taken each year. A smaller sample size of approximately 1600 was obtained for 2015. Aggregated by season, the highest number of photos taken and shared varied by year, with the fall (i.e., September through November) having the highest number of photos in 2011 and 2012, and shifting to the summer months (i.e., June through August) in 2013 and 2014. In 2015, winter months (i.e., December through February) witnessed the largest sample size. Histograms in appendix D illustrate the distribution of photo records with respect to each environmental and infrastructure variable.

Table 2 Number of geotagged photos posted to Flickr within park boundary by season and year

	Winter *(Dec – Feb)*	**Spring** *(Mar – May)*	**Summer** *(Jun – Aug)*	**Fall** *(Sep – Nov)*	**All Seasons**
2011	428	680	649	703	2460
2012	539	454	414	930	2337
2013	474	747	1324	326	2871
2014	639	412	1227	492	2770
2015	521	370	393	326	1610
Total	**2601**	**2663**	**4007**	**2777**	**12048**

MaxEnt models

Annual models

Visualizations of model output across years show similar locations with high logistic probabilities of visitor presence, spatially coinciding with roads, trails, and other visitor infrastructure (Figure 1, Figure 3). Few differences were visually identified between the years, with the exception of 2014. In 2014, increased probability of presence was observed along the northern portion of the park (i.e., Mauna Loa trail and summit) compared to other years. Corroborating the visual assessment, we used zonal statistics to examine MaxEnt values within a 1 km radius of the Mauna Loa trailhead for each year. The 2014 model indicated the highest logistic probability of presence values when compared to the same area across years (Table 3).

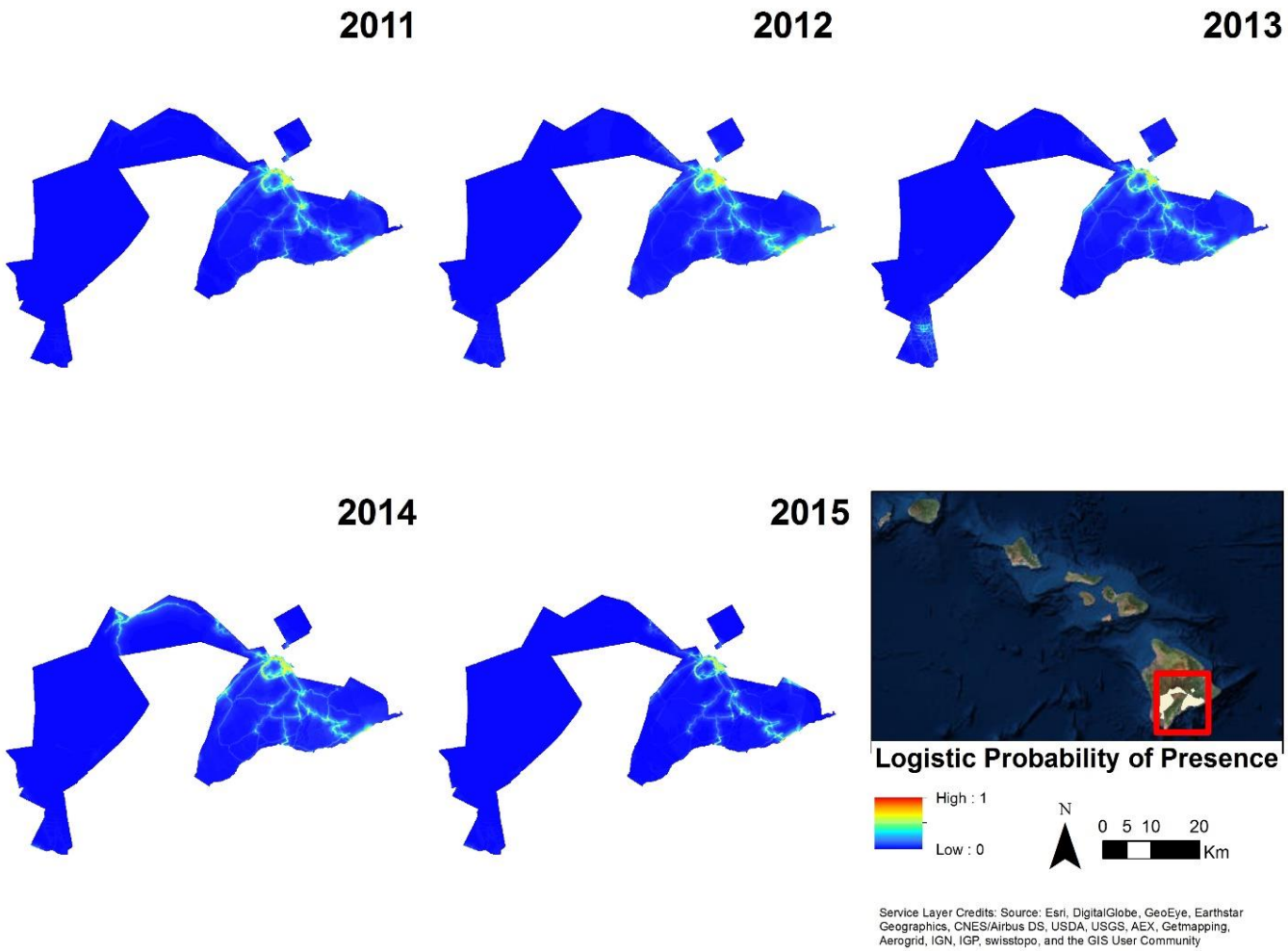


Figure 3 MaxEnt models by year. Warmer colors indicate higher logistic probabilities of presence.

Table 3 Zonal statistics calculated on MaxEnt logistic probability of presence values within a 1 km radius of known features. The minimum (min), maximum (max), range, and mean are reported for each year.

Mauna Loa Trail	**2011**	**2012**	**2013**	**2014**	**2015**
Min	0.0003	0.0010	0.0009	0.0076	0.0002
Max	0.0056	0.0064	0.0125	0.1344	0.0048
Range	0.0053	0.0053	0.0116	0.1268	0.0047
Mean	0.0014	0.0021	0.0022	0.0220	0.0006
Volcano House	**2011**	**2012**	**2013**	**2014**	**2015**
Min	0.1619	0.2094	0.1315	0.1395	0.0662
Max	0.8556	0.8355	0.8639	0.8724	0.8683
Range	0.6936	0.6261	0.7324	0.7329	0.8021
Mean	0.4864	0.5282	0.4719	0.4841	0.4313

Consistent across years was an increased logistic probability of presence in the northeastern part of the park, coincident with the roads and trails providing access to the park's visitor center, hotel (i.e., Volcano House), and the Kīlauea caldera. Maximum logistic probability of presence values were above 0.8 for all years within a 1 km radius of the park's hotel (Table 3). Examining the concentrated area of high probability of use near the northeastern part of the park (Figure 4), similar hot spots (i.e., orange to red) persist, with less to the west of the caldera in 2012 and 2015. In 2012, higher probabilities of presence cover a broader extent than other years where higher probability of presence is more concentrated to roads and trails, indicated by visual outputs of the model and the highest coefficient of variation (Table 3).

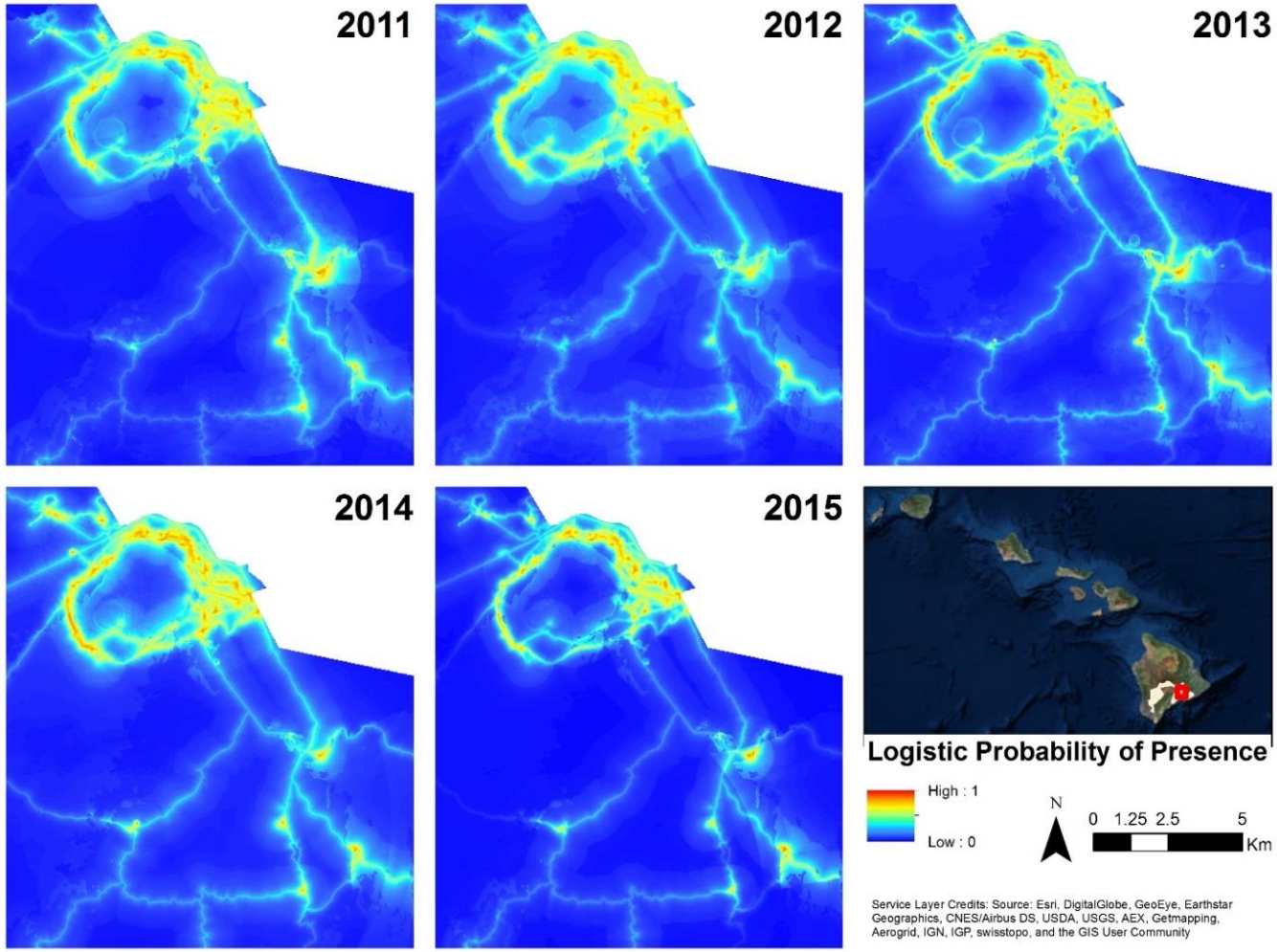


Figure 4 MaxEnt model outputs by year near park visitor center and Kīlauea caldera

For all years, distance to a road was a top contributing variable, representing at least 30% of AUC training gain (Table 4). It was the top contributing variable in all models with the exception of 2014, when distance to trail was the top contributing variable at 43.2%. Distance to trail also contributed at least 20% to the overall model in 2014, 2015, and the full model (i.e., all years). In 2012 and 2013, elevation also contributed at least 20% of the AUC training gain. For all six models, the predictive accuracy was considered high with AUC values greater than 0.9.

Table 4 Permutated importance for environmental and managerial variables by year. Bolded values indicate the variable contributed at least 20% to the AUC training gain.

	Permutated Importance (%)					
	2011	**2012**	**2013**	**2014**	**2015**	**All Years**
Distance to Road	**44.3**	**41.7**	**41.2**	**30.4**	**57.1**	**37.8**
Distance to Parking	10.9	15.1	1.9	13.4	3.2	2.8
Distance to Trail	11.1	5.1	8.4	**43.2**	**21**	16
Distance to Building	12.1	1.6	10.4	2	6.1	12
Elevation	18.1	**32.7**	**34.4**	4.8	9.9	**24.9**
Distance to Campsite	1.3	2.1	1.8	1.9	0.4	3.1
Land Cover	1.7	1.2	0.9	1.3	1.1	1.1
Distance to Water	0.5	0.3	0.8	2.9	0.9	2.1
Slope	0.1	0.1	0.2	0.1	0.2	0.2
Test AUC	0.960	0.960	0.958	0.957	0.973	0.912
AUC St. Dev.	0.005	0.003	0.003	0.003	0.004	0.003

To exercised caution when interpreting variable contribution due to correlation among predictor variables, variables were also examined as individual response curves. Distance to roads and trails indicated high logistic probability of presence on or within a few

hundred meters of the infrastructure (Figure 5). Similar response curve profiles resulted for distance to road, with the exception of 2014. In 2014, an logistic increasing probability of presence was also found as distance increased above 17,500 m from the road. This corresponds to increased logistic probability of presence along the trail to the caldera and summit of Mauna Loa (see Figure 3), which is over 27 km from where the park road ends at the trailhead. Higher logistic probability of visitor presence was also observed on or very near trails, dropping sharply after a few hundred meters.

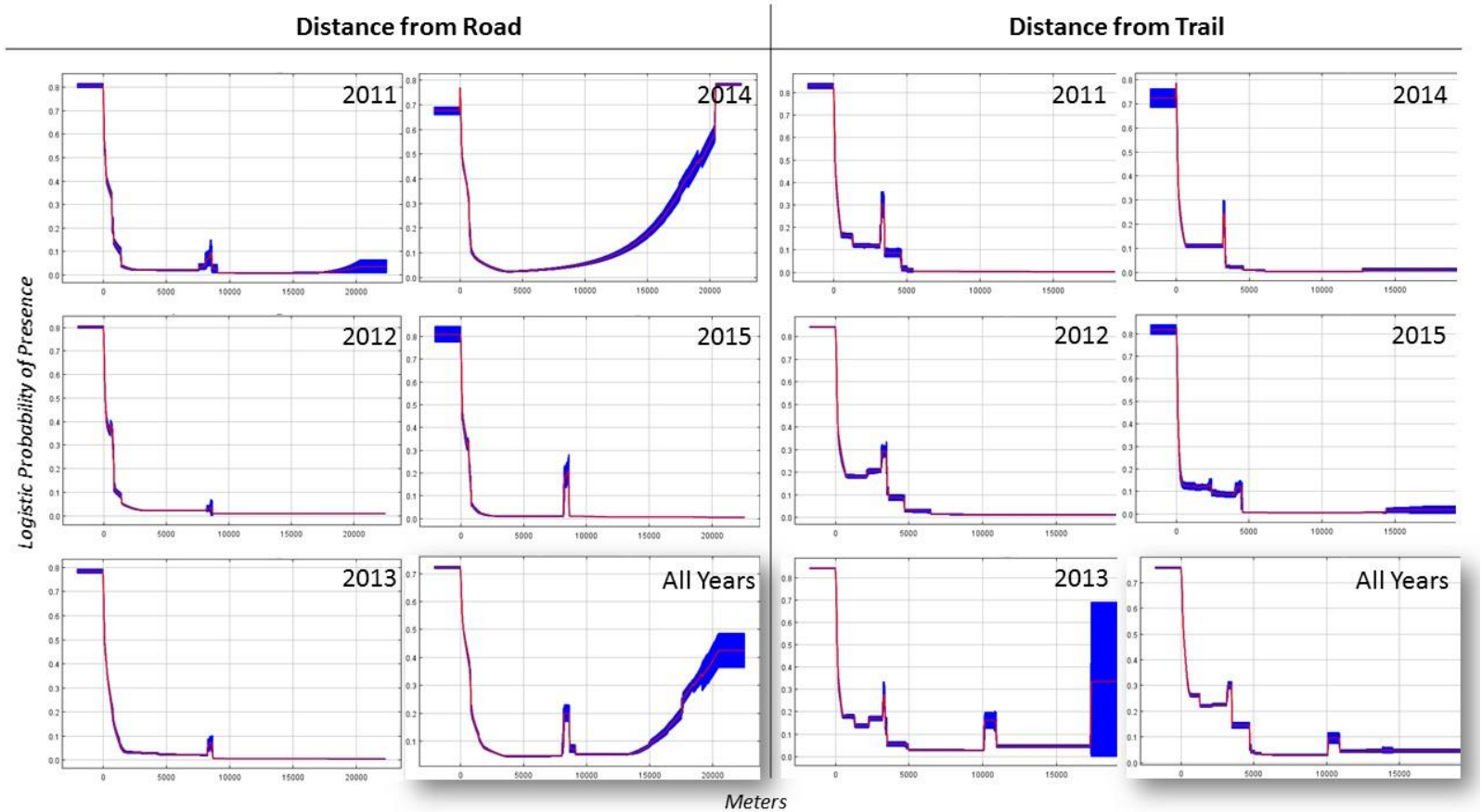


Figure 5 MaxEnt response curves for distance to roads and trails by year. Red line represents the logistic probability of presence, with the blue providing a plus and minus one standard deviation from the 10 replicate runs.

Seasonal models

Results from seasonal models also highlight increased logistic probability of presence near the visitor center and Kīlauea caldera, as well as along Chain of Craters Road (Figure 6). More noticeable in the seasonal models are differences in visitor presence along Mauna Loa trail in the summer months, with greatest logistic probability of presence values within 1 km of the trailhead (Table 5), and lower logistic probability of presence on trails in the Kua Desert in the winter and fall as compared to the spring and summer months (Figure 1, Figure 6). Lower logistic probability of presence along trails was also noted during the fall, greater values concentrated along the roads in the southern portion of the park.

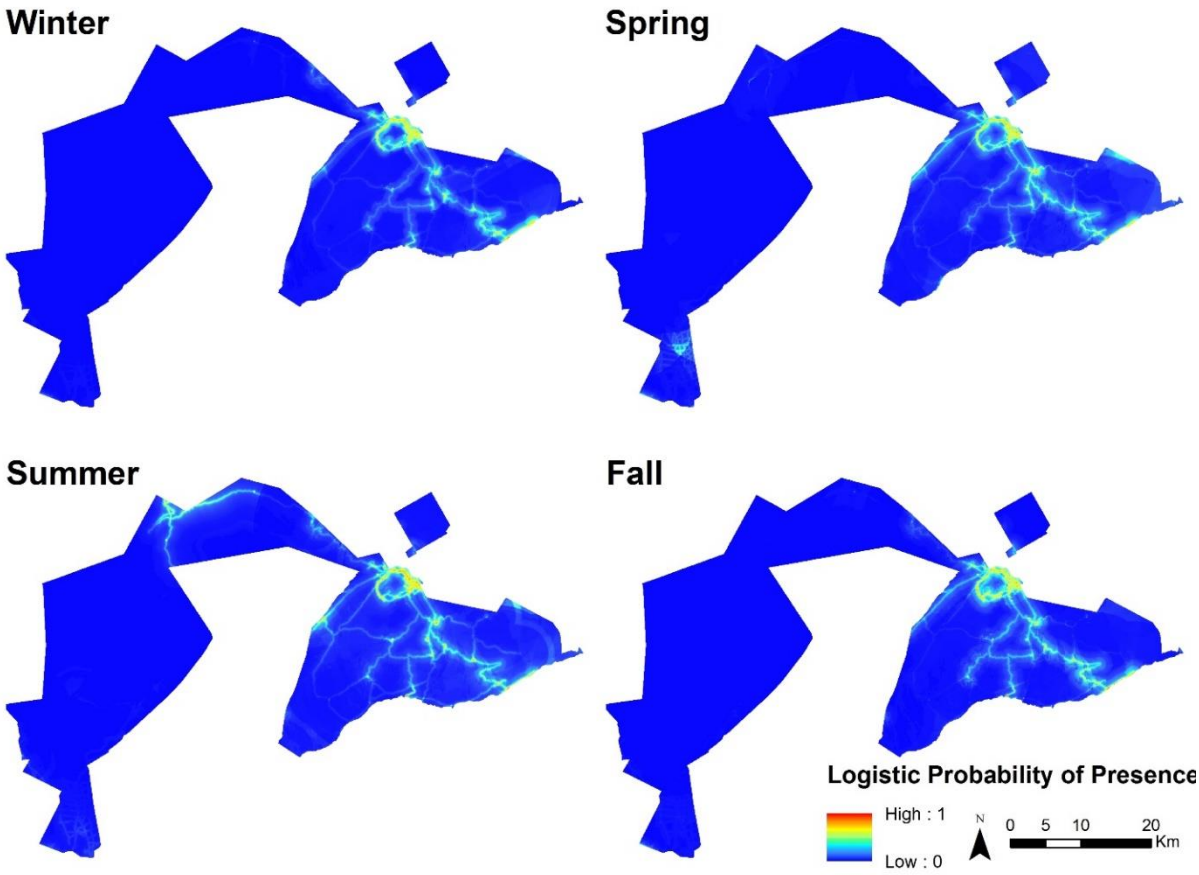


Figure 6 MaxEnt models by season. Warmer colors indicate higher logistic probability of visitor presence as modeled using geotagged photo records.

Table 5 Zonal statistics calculated for seasonal MaxEnt logistic probability of presence values within 1 km of known features.

Mauna Loa Trail	Winter	Spring	Summer	Fall
Min	0.0011	0.0003	0.0046	0.0009
Max	0.0307	0.0043	0.0528	0.0090
Range	0.0296	0.0040	0.0482	0.0080
Mean	0.0043	0.0007	0.0127	0.0022
Volcano House	**Winter**	**Spring**	**Summer**	**Fall**
Min	0.1072	0.1670	0.1591	0.1744
Max	0.8320	0.8388	0.8504	0.8452
Range	0.7248	0.6718	0.6912	0.6708
Mean	0.4846	0.4750	0.5101	0.4895

Model visualizations near the visitor center and Kīlauea caldera indicate more areas of very high logistic probability of presence (i.e., orange and red) during the summer months along the eastern edge of the caldera and along the western edge during the winter (Figure 7). Although maximum MaxEnt values are similar across the four seasons within 1 km of the park hotel (i.e. Volcano House), the summer had the greatest mean and standard deviation values, suggesting high and clustered visitor presence (Table 5).

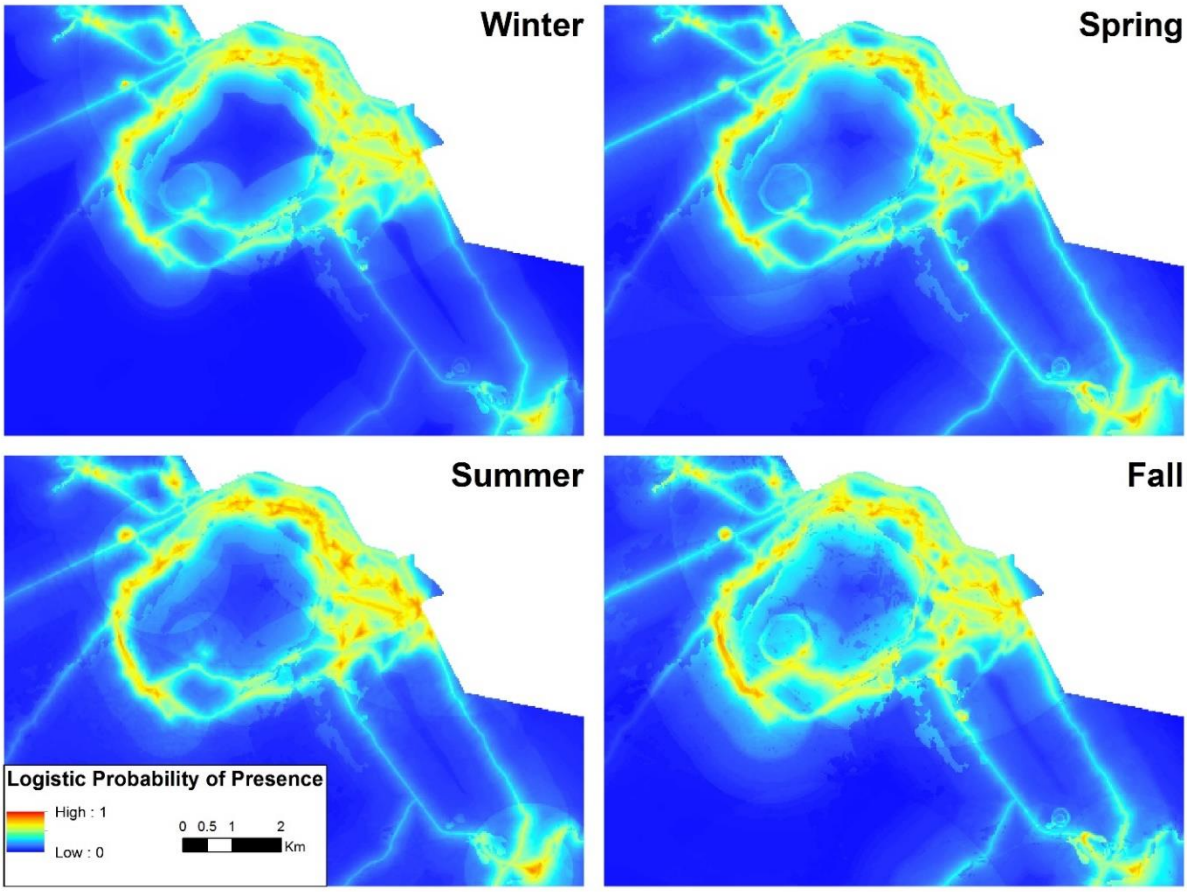


Figure 7 MaxEnt model outputs by season near park visitor center and Kīlauea caldera

As in the annual models, the top contributing predictor variables to the model included distance to roads for all seasons (Table 6). Distance to trail contributed more than 37% in the summer months, and elevation contributed more than 54% in the spring. Distance to trail did not contribute as much in the seasonal models, with the exception of the summer, as compared to annual models. This finding is also evident in the model visualizations, with visitor probability of presence concentrated on roads and less on trails in the fall and winter

(Figure 6). Predictive accuracy for the four seasonal models was high, with AUC values greater than 0.95.

Table 6 Permutated importance for environmental and managerial variables by season. Bolded values represent variables contributing to at least 20% of the AUC gain.

	Permutated Importance (%)				
	Winter	**Spring**	**Summer**	**Fall**	**All Seasons**
Distance to Road	**46**	**29.4**	**33.3**	**58.7**	**37.8**
Distance to Parking	13.1	1.1	7.3	13.2	2.8
Distance to Trail	14.6	5.7	**37.8**	10.4	16
Distance to Building	9	6.3	8.6	2.9	12
Elevation	12.5	**54.2**	8.9	10.9	**24.9**
Distance to Campsite	1.8	1.4	2.6	0.9	3.1
Land Cover	1.1	0.9	0.7	2	1.1
Distance to Water	2	0.8	0.7	0.6	2.1
Slope	0.1	0.1	0.2	0.4	0.2
Test AUC	0.962	0.953	0.952	0.959	0.912
AUC St. Dev.	0.002	0.004	0.004	0.003	0.003

Response curves for the seasonal models reflected findings in the annual models, with visitor presence very likely on or within a few hundred meters of a road or trail, with an increase in the summer months to approximately 15,000 m, related to the increased logistic probability of presence along the Mauna Loa Trail and summit. The response curves for elevation indicate high logistic probability of presence at sea level and 1,000-1,250 m above sea level (a.s.l.), the latter corresponding to the elevations of the visitor center and hotel near Kīlauea caldera (Figure 8). Increased logistic probability of presence above 3,000 m a.s.l.

during the summer months, also related to increased probability of presence along Mauna Loa Trail, which starts at approximately 2,000 m a.s.l and ends just above 4,000 m a.s.l.

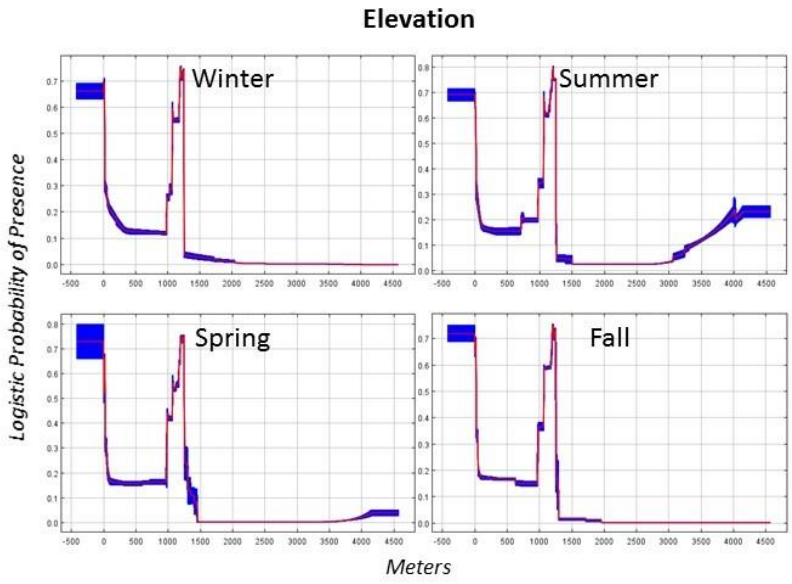


Figure 8 Response curve for elevation for each seasonal model

Hot Spots

The MaxEnt models suggested high probabilities of visitor clustering in specific areas of the park and within close proximity to visitor infrastructure, and average nearest neighbor ratios confirmed visitor clustering was statistically significant ($p < 0.000$) across all years (Appendix E). Complementing MaxEnt outputs, Getis-Ord Gi* statistics identified the

location of statistically significant visitor hot spots concentrated along roads and trails near the park's visitor center, hotel, and main volcanic craters (Figure 9, Figure 10). Considered in the context of each sample rather than across years, 2015 witnessed several hot spots occurring along the Chain of Craters Road in addition to the expected hot spots near the park's visitor center and hotel.

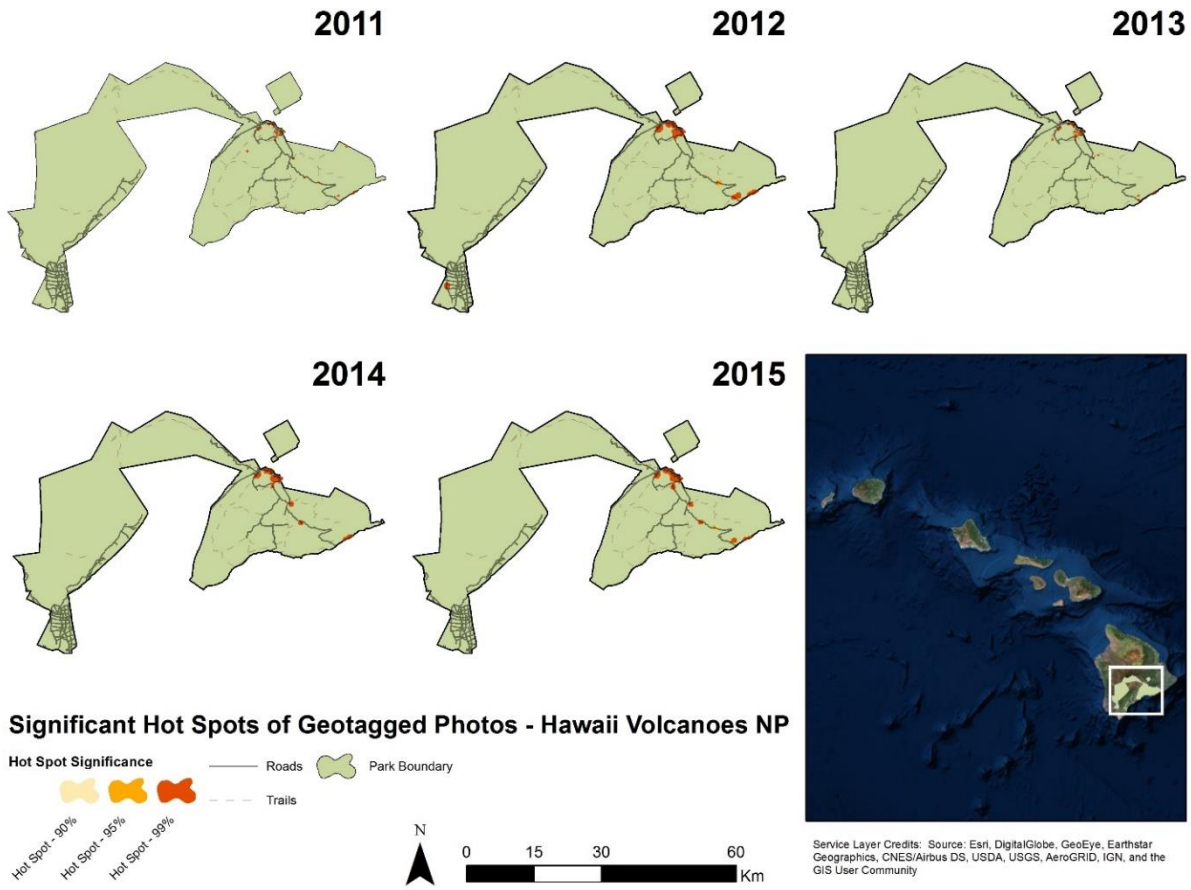


Figure 9 Hot spots of visitor activity by year

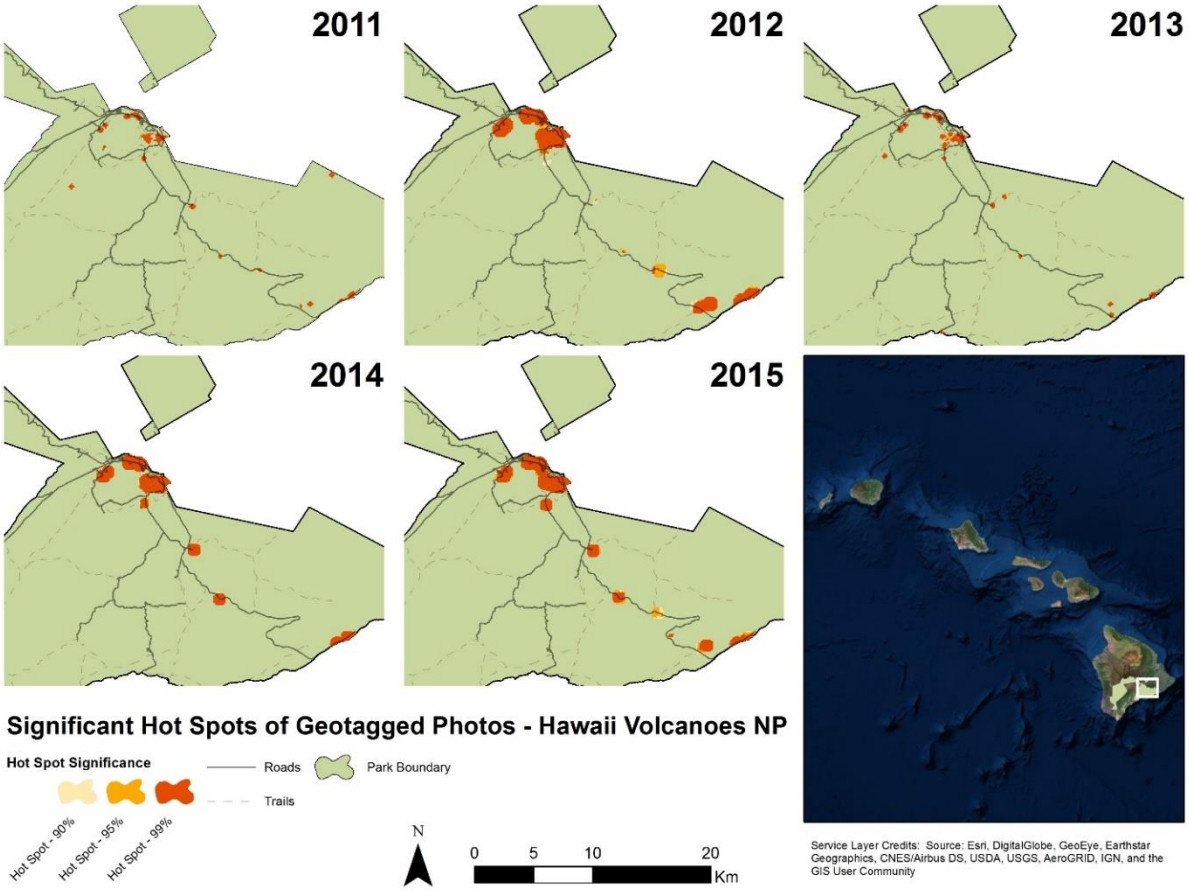


Figure 10 Details of visitor hot spots near Kilauea crater and main park roads by year

Seasonal hot spots also mirrored MaxEnt outputs, with similar locations indicative of high logistic probability of presence and statistically significant hot spots of visitor digital footprints (Figure 11). Hot spots during the summer months occurred near the park's visitor center and Kilauea crater, as well as along the southern border of the park. Photos taken during the fall and winter months concentrated in small clusters at several locations along and near visitor infrastructure.

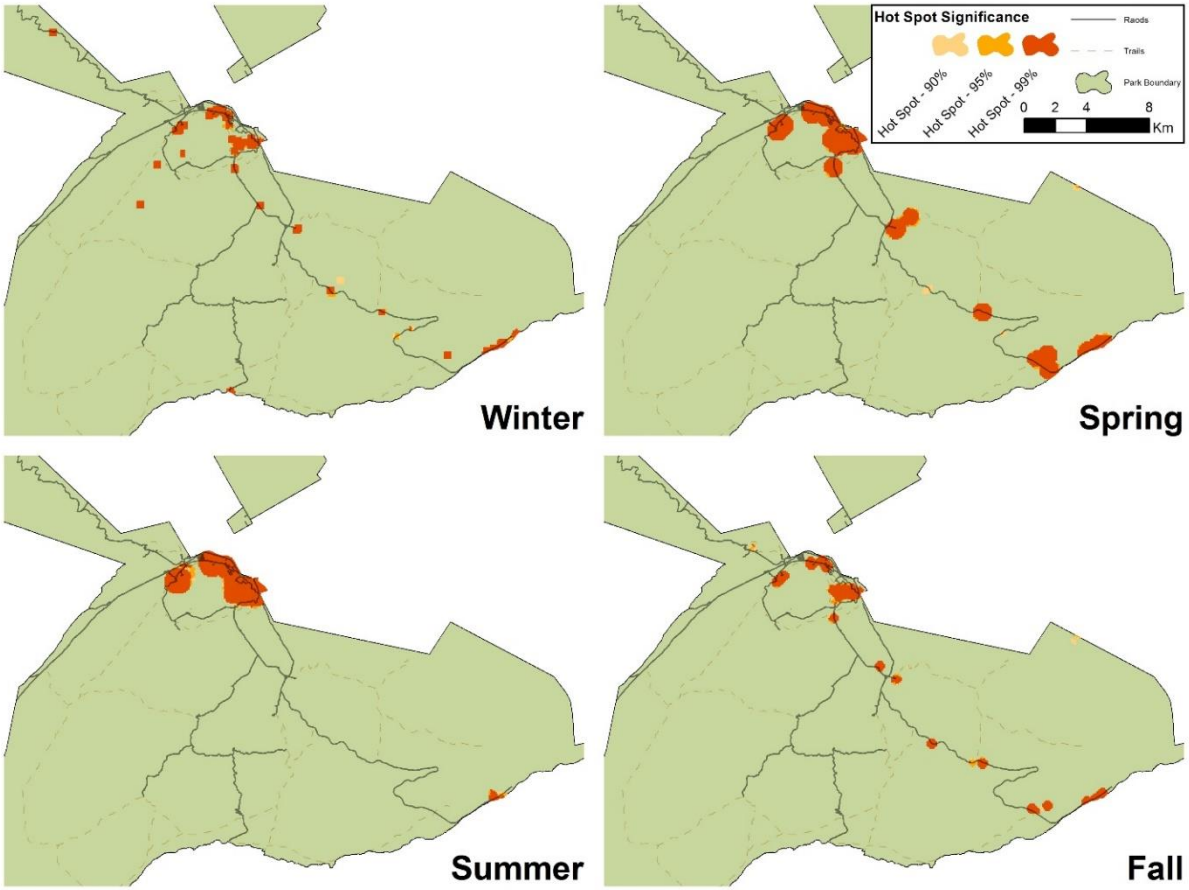


Figure 11 Details of visitor hot spots near Kilauea crater and park roads by season

Discussion and Conclusion

Flickr provided adequate sample sizes to achieve the research objectives outlined in this study using the proposed methods. Data were uploaded within approximately two weeks of their capture, similar to the timeliness noted by Antoniou and colleagues (2010), and at similar numbers of photos per user (Richards & Friess 2015; Sonter et al. 2016). The shift from digital cameras toward mobile devices of the devices reported between 2011 and 2015

has implications for spatial accuracy assumptions in future modelling and research. Many of the digital cameras reported do not have integrated GPS, though from the data in this study, we could not distinguish the method for geotagging the resulting image (i.e., external GPS for camera or manual map placement). If location services are turned on for a mobile device, the image will contain the geotag from the phone's internal GPS and potentially reduce opportunities for errors introduced during manual geotagging. However, this also represents a challenge as data sets contain devices capable of varying resolution (Girardin, Calabrese, Fiore, Ratti, & Blat, 2008). Spatial accuracy assessment could improve parameter settings for models and spatial statistics, including the ones in this study, which are sensitive to distance metrics for calculations.

MaxEnt models highlighted the importance of infrastructure in the overall distribution of visitor photos, similar to MaxEnt models of winter recreation impacts on wildlife studied by Coppes and Braunisch (2013). The high logistic probability of presence along roads is an important empirical finding for mitigating roadside impacts to vegetation as research has found they can be greater and more extensive than trailside impacts (Wolf & Croft, 2014). It also highlights the importance of infrastructure and transportation planning given the potential impact to flora and fauna (Hammitt et al., 2015) when, where, and how areas are made accessible to visitors. This pattern was observed in all annual and seasonal models, though the peak in probability of visitor presence approximately 17.5 km from a road in 2014 (Figure 5) highlights the importance of considering the spatial location in context with how they correspond to other variables and potential visitation patterns. For example, although

there are other areas in the park where visitors could venture considerable distances from a road, this peak corresponded with increased digital footprints along the Mauna Loa trail and summit area.

Consideration of how values of variables correspond to logistic probability of presence both individually as well as spatially with other variables was important for understanding distribution patterns. Elevation, for example, contributed over 20% to some of the models with the response curves indicating increase logistic probability of presence at elevations similar to that of park infrastructure and the trail along Mauna Loa. The importance of elevation to overall spatial patterns also represents an important attribute for further consideration in areas where visitor infrastructure traverses more elevation gradients. Specifically, explorations of areas with documented shifts in species presence, climate patterns, and fire regimes would advance our understanding of if and how visitors are changing their spatial patterns of use (McEvoy et al., 2008).

The ranking of predictor variables is especially important for human-environment interaction research and visitor use management in dynamic landscapes like HVNP to identify which factors most greatly influence distribution and how rapid change in the landscape may affect subsequent visitor use patterns. Top environmental factors inform how changes in the landscape influence use patterns. Top infrastructure factors indicate where impacts may concentrate and perpetuate, as well as which management tactics or infrastructure are most influential on visitor distribution and at what times of the year. Interestingly, while less inter-annual variation was noted in this study, the MaxEnt models

did suggest seasonal variation in infrastructure importance. For example, lower logistic probability of presence along the Mauna Loa trail during the fall and winter months may be weather related, as the summit area does experience snow and low temperatures, potentially limiting the number of visitors attempting the backcountry hike. The summit is open year round, with occasional closures due to weather or high winds.

Surprising in the photo records and MaxEnt models was the high logistic probability of presence along the western part of Kīlauea caldera (Figure 4, Figure 7) as parts of this area closed in 2008 due to hazardous conditions. Random inspection of photos (approximately 15%) georeferenced to the area indicated many were aerial photographs, suggesting visitors were present in the area, however not participating in ground-based activity. Instead, visitors were viewing the site from the air (e.g., helicopter tour). Few appeared to be the result of incorrect geotags, which has been identified as a source of error in other studies (Orsi & Geneletti, 2013; Zielstra & Hochmair, 2013).

The finding that some photos were taken from aircraft is important to other studies assessing impacts based on visitor presence with geotagged photos, given that photos may not represent ground-based activity. These records are still an important indicator of impact as helicopter tours can disturb wildlife (Stockwell, Bateman, & Berger, 1991) and contribute to noise pollution (Pickering, Harrington, & Worboys, 2003). This is also true of cable cars, chair lifts, or other air-based transportation and tourism operations, where the photo location has an altitude component. We recommend further research continue exploring distribution

models following a content analysis of the photos (see Richards & Friess 2015), segmented by visitor activity, to model activity zones as the impacts to natural resources vary.

A major limitation associated with geotagged photos, in addition to potentially incorrect geotags, are that they represent a sample of those visitors taking and sharing geotagged media, and therefore potentially biased to digital footprints. Growing work on the quality, accuracy, and contributors of VGI have found more educated individuals (Crampton et al., 2013), and those employed in management, business, science, and arts were more likely to create geotagged media on Flickr and Twitter (Li, Goodchild, & Xu, 2013). The implication for spatial distribution modelling is that leveraging only one platform may miss certain populations, use types, and result in different distributions or growth patterns (Alivand & Hochmair, 2017).

More research is also needed to explore user demographics for different social media platforms and how experience, values, accessibility, and environmental characteristics influence if and at which locations visitors take and share photos to inform future modelling efforts (Levin et al., 2017). An example includes the globally-popular Instagram, with real-time uploads and an estimated 28% of individuals with internet access in the U.S. contributing content (Pew Research Center, 2016; van Zanten et al., 2016). Additionally, pairing modelling results with a media content analysis at finer temporal scales may also illuminate how events, both natural and programmatic, influence the number, location, and timing of visitor photos. For example, the NPS along with media outlets published several stories throughout 2014 about the increased seismic activity from Mauna Loa similar to its

last eruption in 1984 (Klemetti, 2014; USGS, 2014). Although a formal assessment of these publications was beyond the scope of this study, it may relate to the greater number of photos shared in 2014, especially the increased digital footprint along the Mauna Loa trail.

Crowdsourced geographic information can reduce barriers to obtaining spatially explicit data about visitation patterns in PAs to answer long-standing and emerging questions on human-environment interactions and, subsequently, inform planning and management efforts. Geotagged photos from Flickr illustrated temporal changes in the spatial patterns of visitor use and how infrastructure and natural environmental characteristics were associated with distribution patterns, though patterns derived from MaxEnt models and hot spot analyses indicated a level of predictability in visitor use. This predictability is useful if existing empirical evidence can inform impact, interactions, and spatial analyses.

However, within the context of a changing climate and dynamic responses of ecosystems, this predictability may only hold at certain spatial and temporal scales and warrant further investigation of the role of temporal scale. An example includes analysis of visitor patterns coinciding with media coverage of events like visibility of migratory wildlife (i.e., whale watching, birding) or geological events (i.e., volcanic activity). This is possible with the availability of free and timely crowdsourced data, capable of high spatial resolution, and from sources experiencing growth in the amount and extent of content shared (Alivand & Hochmair, 2017). Continued exploration of these sources and their capacity for addressing conservation and PA management challenges represent a rich area of research in GIScience, as the supporting technologies and possibilities continue advancing.

Acknowledgements

Thank you to the NPS Pacific Islands Inventory and Monitoring Program for providing the GIS data used in the analysis and Hawaii Volcanoes NP for permission to conduct research within the park. Thank you also to L. Tateosian, S. Nelson, and T. Wentworth for their feedback on the study design and implementation of this research.

References

Alivand, M. and Hochmair, H.H., 2017. Spatiotemporal analysis of photo contribution patterns to Panoramio and Flickr. *Cartography and Geographic Information Science*, 44 (2), 170–184.

Antoniou, V., Morley, J. and Haklay, M., 2010. Web 2.0 geotagged photos: assessing the spatial dimension of the phenomenon. *Geomatica*, 64 (1), 99–110.

Arnberger, A., Haider, W. and Brandenburg, C., 2005. Evaluating visitor-monitoring techniques: a comparison of counting and video observation data. *Environmental Management*, 36 (2), 317–327.

Beeco, J.A. and Brown, G., 2013. Integrating space, spatial tools, and spatial analysis into the human dimensions of parks and outdoor recreation. *Applied Geography*, 38, 76–85.

Braunisch, V., Patthey, P. and Arlettaz, R., 2011. Spatially explicit modeling of conflict zones between wildlife and snow sports: prioritizing areas for winter refuges. *Ecological Applications*, 21 (3), 955–967.

Braunisch, V. and Suchant, R., 2010. Predicting species distributions based on incomplete survey data: the trade-off between precision and scale. *Ecography*, 33 (5), 826–840.

Buckley, L.B. and Foushee, M.S., 2012. Footprints of climate change in US national park visitation. *International Journal of Biometeorology*, 56 (6), 1173–1177.

Cessford, G. and Burns, R., 2008. *Monitoring visitor numbers in New Zealand national parks and protected areas: a literature review and development summary*. Wellington, New Zealand: New Zealand Department of Conservation.

Cessford, G.R. and Muhar, A., 2003. Monitoring options for visitor numbers in national parks and natural areas. *Journal for Nature Conservation*, 11, 240–250.

Chen, I.-C. et al., 2011. Rapid range shifts of species associated with high levels of climate warming. *Science*, 333 (6045), 1024–1026.

Connors, J.P., Lei, S. and Kelly, M., 2012. Citizen science in the age of neogeography: utilizing volunteered geographic information for environmental monitoring. *Annals of the Association of American Geographers*, 102 (6), 1267–1289.

Coppes, J. and Braunisch, V., 2013. Managing visitors in nature areas: where do they leave the trails? A spatial model. *Wildlife Biology*, 19 (1), 1–11.

Crampton, J.W. et al., 2013. Beyond the geotag: situating “big data” and leveraging the potential of the geoweb. *Cartography and Geographic Information Science*, 40 (2), 130–139.

Crutzen, P.J., 2002. Geology of mankind. *Nature*, 415 (6867), 23–23.

Dove, J., 2015. *Canon and Apple were the top Flickr cameras in 2014* [online]. The Next Web. Available from: http://thenextweb.com/apps/2015/01/10/canon-dslrs-and-iphone-5-dominated-photo-posting-on-flickr-in-2014/ [Accessed 4 August 2015].

Eagles, P.F.J., 2014. Research priorities in park tourism. *Journal of Sustainable Tourism*, 22 (4), 528–549.

Elith, J. et al., 2011. A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, 17 (1), 43–57.

Elith, J. et al., 2006. Novel methods improve prediction of species’ distributions from occurrence data. *Ecography*, 29 (2), 129–151.

Elwood, S., Goodchild, M.F. and Sui, D.Z., 2012. Researching volunteered geographic information: spatial data, geographic research, and new social practice. *Annals of Association of American Geographers*, 102 (3), 571–590.

Flanagin, A.J. and Metzger, M.J., 2008. The credibility of volunteered geographic information. *GeoJournal*, 72 (3–4), 137–148.

Girardin, F. et al., 2008. Digital footprinting: uncovering tourists with user-generated content. *IEEE Pervasive Computing*, 7 (4), 36–43.

Goodchild, M.F., Aubrecht, C. and Bhaduri, B., 2017. New questions and a changing focus in advanced VGI research. *Transactions in GIS*, 21, 189–190.

Hadwen, W.L., Hill, W. and Pickering, C.M., 2008. Linking visitor impact research to visitor impact monitoring in protected areas. *Journal of Ecotourism*, 7 (1), 87–93.

Hallo, J.C. et al., 2012. GPS as a method for assessing spatial and temporal use distributions of nature-based tourists. *Journal of Travel Research*, 51 (5), 591–606.

Hammitt, W.E., Cole, D.N., and Monz, C., 2015. *Wildland recreation: ecology and management.* 3rd ed., Chichester, UK: Wiley Blackwell.

Hawaii Tourism Authority, 2015. *Hawaii Volcanoes National Park, Hawaii Island* [online]. Available from: http://www.gohawaii.com/en/big-island/regions-

neighborhoods/kau/hawaii-volcanoes-national-park/ [Accessed 9 September, 2015].

Jalali, M.A. et al., 2015. Predictive mapping of abalone fishing grounds using remotely-sensed LiDAR and commercial catch data. *Fisheries Research*, 169, 26–36.

Jones, B. and Scott, D., 2006. Climate change, seasonality and visitation to Canada's national parks. *Journal of Park and Recreation Administration*, 24 (2), 42–62.

Kalogerakis, E. et al., 2009. Image sequence geolocation with human travel priors. *2009 IEEE 12th International Conference on Computer Vision*, 29 September–2 October 2009 Kyoto, Japan. Washington, D.C.: IEEE, 253–260.

Keeler, B.L. et al., 2015. Recreational demand for clean water: evidence from geotagged photographs by visitors to lakes. *Frontiers in Ecology and the Environment*, 13 (2), 76–81.

Klemetti, E. 2014. Earthquake swarms at Mauna Loa betray magma moving deep below. *Wired Magazine* [online]. Available from: https://www.wired.com/2014/06/earthquake-swarms-at-mauna-loa-betray-magma-moving-deep-below/ [Accessed 1 February, 2017].

Lesica, P. and Kittelson, P.M., 2010. Precipitation and temperature are associated with advanced flowering phenology in a semi-arid grassland. *Journal of Arid Environments*, 74 (9), 1013–1017.

Leung, Y.-F. and Monz, C., 2006. Visitor impact monitoring: old issues, new challenges – an introduction to this special issue. *George Wright Forum*, 23 (2), 7–10.

Levin, N., Lechner, A.M. and Brown, G., 2017. An evaluation of crowdsourced information for assessing the visitation and perceived importance of protected areas. *Applied Geography*, 79, 115–126.

Li, L., Goodchild, M.F. and Xu, B., 2013. Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. *Cartography and Geographic Information Science*, 40 (2), 61–77.

Liu, B. et al., 2014. Where your photo is taken: geolocation prediction for social images. *Journal of the Association for Information Science and Technology*, 65 (6), 1232–1243.

Longley, P.A. and Adnan, M., 2016. Geo-temporal Twitter demographics. *International Journal of Geographical Information Science*, 30 (2), 369–389.

Loomis, J.B. and Richardson, R.B., 2006. An external validity test of intended behavior: comparing revealed preference and intended visitation in response to climate change. *Journal of Environmental Planning and Management*, 49 (4), 621–630.

Marvin, D.C. et al., 2016. Integrating technologies for scalable ecology and conservation. *Global Ecology and Conservation*, 7, 262–275.

McEvoy, D. et al., 2008. Changes to climate and visitor behaviour: implications for vulnerable landscapes in the north west region of England. *Journal of Sustainable Tourism*, 16 (1), 101–121.

Meijles, E.W. et al., 2013. Analysing hiker movement patterns using GPS data: implications for park management. *Computers, Environment and Urban Systems*, 47, 44–57.

Menzel, A. et al., 2006. European phenological response to climate change matches the warming pattern. *Global Change Biology*, 12 (10), 1969–1976.

Merow, C., Smith, M.J. and Silander, J.A., 2013. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography*, 36 (10), 1058–1069.

Millenium Ecosystem Assessment, 2005. *Ecosystems and human well-being: biodiversity synthesis*. Washington, D.C.: World Resources Institute.

Monz, C.A. et al., 2010. Sustaining visitor use in protected areas: future opportunities in recreation ecology research based on the USA Experience. *Environmental Management*, 45 (3), 551–562.

NPS, 2015. *Hawaii Volcanoes National Park draft general management plan, wilderness study, and environmental impact statement*. Seattle, WA: USDI National Park Service.

NPS, 2016. *Hawaii Volcanoes NP - annual park recreation visititation report* [online], Available from: https://irma.nps.gov/Stats/SSRSReports/Park Specific Reports/Annual Park Recreation Visitation (1904 - Last Calendar Year)?Park=HAVO [Accessed 4 January, 2017].

O'Connor, A., Zerger, A. and Itami, B., 2005. Geo-temporal tracking and analysis of tourist movement. *Mathematics and Computers in Simulation*, 69, 135–150.

Orsi, F. and Geneletti, D., 2013. Using geotagged photographs and GIS analysis to estimate visitor flows in natural areas. *Journal for Nature Conservation*, 21 (5), 359–368.

Pew Research Center, 2016. *Smartphone ownership and internet usage continues to climb in emerging economies* [online], Available from: http://www.pewglobal.org/2016/02/22/smartphone-ownership-and-internet-usage-continues-to-climb-in-emerging-economies/ [Accessed 23 February, 2017].

Phillips, S.J., Anderson, R.P. and Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190 (3–4), 231–259.

Phillips, S.J. and Dudík, M., 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, 31 (2), 161–175.

Pickering, C.M., Harrington, J. and Worboys, G., 2003. Environmental impacts of tourism on the Australia Alps protected areas: judgments of protected area managers. *Mountain Research and Development*, 23 (3), 247–254.

Pimm, S.L. et al., 2014. The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, 344, 988–997.

Poor, E.E. et al., 2012. Comparing habitat suitability and connectivity modeling methods for conserving pronghorn migrations. *PloS One*, 7 (11), e49390.

Renner, I.W. and Warton, D.I., 2013. Equivalence of MAXENT and Poisson point process models for species distribution modeling in ecology. *Biometrics*, 69 (1), 274–281.

Richards, D.R. and Friess, D.A., 2015. A rapid indicator of cultural ecosystem service usage at a fine spatial scale: content analysis of social media photographs. *Ecological Indicators*, 53, 187–195.

Richardson, R.B. and Loomis, J.B., 2005. Climate change and recreation benefits in an alpine national park. *Journal of Leisure Research*, 37 (3), 307–320.

Sanderson, E.W. et al., 2002. The human footprint and the last of the wild. *BioScience*, 52 (10), 891–904.

Schägner, J.P. et al., 2016. Mapping recreational visits and values of European national parks by combining statistical modelling and unit value transfer. *Journal for Nature Conservation*, 31, 71–84.

Scott, D., Jones, B. and Konopek, J., 2007. Implications of climate and environmental change for nature-based tourism in the Canadian Rocky Mountains: a case study of Waterton Lakes National Park. *Tourism Management*, 28 (2), 570–579.

See, L. et al., 2016. Crowdsourcing, citizen science or volunteered geographic information?

The current state of crowdsourced geographic information. *ISPRS International Journal of Geo-Information*, 5 (5), 55–78.

Senaratne, H. et al., 2017. A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science*, 31 (1), 139–167.

Sonter, L.J. et al., 2016. Spatial and temporal dynamics and value of nature-based recreation, estimated via social media *PLOS One*, 11 (9), e0162372.

Steffen, W. et al., 2011. The Anthropocene: from global change to planetary stewardship. *Ambio*, 40 (7), 739–761.

Stockwell, C.A., Bateman, G.C. and Berger, J., 1991. Conflicts in national parks: a case study of helicopters and bighorn sheep time budgets at the grand canyon. *Biological Conservation*, 56 (3), 317–328.

Sui, D., Elwood, S. and Goodchild, M. eds., 2013. *Crowdsouring geographic knowledge: volunteered geographic information (VGI) in theory and practice*. Dordrecht, The Netherlands: Springer.

Tenerelli, P., Demšar, U. and Luque, S., 2016. Crowdsourcing indicators for cultural ecosystem services: a geographically weighted approach for mountain landscapes. *Ecological Indicators*, 64, 237–248.

Tomczyk, A.M. and Ewertowski, M., 2011. Degradation of recreational trails, Gorce National Park, Poland. *Journal of Maps*, 7 (1), 509–518.

UNESCO. 2016. Biosphere Reserve information - Hawaiian Islands [online]. *MAB Biosphere Reserve Directory*. Available from: http://www.unesco.org/mabdb/br/brdir/directory/biores.asp?mode=all&code=USA+32 [Accessed 15 December, 2016].

USGS. 2014. Mauna Loa – a stiring giant? [online]. *Hawaiian Volcano Observatory*. Available from: https://hvo.wr.usgs.gov/volcanowatch/view.php?id=232 [Accessed 1 February, 2017].

Wang, S. et al., 2014. A study of the spatial patterns of tourist sightseeing based on volunteered geographic information: the case of Jiuzhai Valley. *Tourism Tribune*, 29 (2), 86–92.

Weaver, S.D. and Gahegan, M., 2007. Constructing, visualizing, and analyzing a digital footprint. *Geographical Review*, 97 (3), 324–350.

Westcott, F. and Andrew, M.E., 2015. Spatial and environmental patterns of off-road vehicle recreation in a semi-arid woodland. *Applied Geography*, 62, 97–106.

Wolf, I.D. and Croft, D.B., 2014. Impacts of tourism hotspots on vegetation communities show a higher potential for self-propagation along roads than hiking trails. *Journal of Environmental Management*, 143, 173–185.

Wood, S.A. et al., 2013. Using social media to quantify nature-based tourism and recreation. *Scientific Reports*, 3, 1-7.

Xia, J. and Arrowsmith, C.A., 2008. Techniques for counting and tracking the spatial and temporal movement of visitors. In R. Gimblett and H. Skov-Petersen, eds. *Monitoring, simulation and management of visitor landscape*. Tucson, AZ: The University of Arizona Press, 85–106.

Young, N. et al., 2011. A MaxEnt model v3.3.3e tutorial (ArcGIS v10). Available from: http://ibis.colostate.edu/WebContent/WS/ColoradoView/TutorialsDownloads/A_Maxent_Model_v8.pdf [Accessed 10 August, 2016].

Zandbergen, P. A. and Barbeau, S.J., 2011. Positional accuracy of assisted GPS data from high-sensitivity GPS-enabled mobile phones. *Journal of Navigation*, 64 (3), 381–399.

Van Kirk, R. et al., 2014. Computer simulation modeling to determine trailhead quotas for overnight wilderness visitor use. *Journal of Park and Recreation Administration*, 32 (3), 29–48.

van Zanten, B.T. et al., 2016. Continental-scale quantification of landscape values using social media data. *Proceedings of the National Academy of Sciences of the United States of America*, 113 (46), 12974–12979.

Zheng, Y.-T., Zha, Z.-J. and Chua, T.-S., 2010. Research and applications on georeferenced multimedia: a survey. *Multimedia Tools and Applications*, 51 (1), 77–98.

Zielstra, D. and Hochmair, H.H., 2013. Positional accuracy analysis of Flickr and Panoramio images for selected world regions. *Journal of Spatial Science*, 58 (2), 251–273.

Zvoleff, A. and An, L., 2014. Analyzing human-landscape interactions: tools that integrate. *Environmental Management*, 53 (1), 94–111.

CHAPTER 5

General Conclusions

Protected areas (PAs) are a key global strategy in conserving biological diversity, important to ecosystem function and the provisioning of ecosystem services. However, PAs contend with myriad ecological and anthropogenic pressures inside and beyond designated borders that can affect conservation efforts, with climate change representing a key example (Cole, Millar, & Stephenson, 2010). PAs are also social constructs, predicated on and sustained by political and economic support (Buckley, 2009). Nature-based recreation and tourism contribute to the support for PAs, with many areas mandated by regulating agencies to balance conservation and visitor use. Understanding how visitor use and visitation-related activities spatially and temporally relate to infrastructure and natural environmental conditions furthers the science of human-environment interactions as well as the efficacy of management strategies and tactics to balance conservation and use.

Synthesis of findings

This dissertation contributed to research exploring human-environment relationships in PAs by providing an integrated assessment of interactions and patterns at multiple spatial and temporal scales. Specifically, case studies described in Chapters 2 through 4 examined geospatial approaches to monitor movements and model distributions from designed research

approaches to crowdsourced geographic information, hourly to annual patterns, and site to park scales.

At the site level, Chapter 2 explored pack stock animal behavior and movement using high-resolution geospatial data. Results indicated animals spent much of the overnight release period engaged in some type of movement, with coupled behavior mapping and GPS tracking suggesting multiple movement patterns associated with grazing. Specifically, the integrated approach differentiated localized, intense grazing from more ambulatory grazing. Without integrating behavior mapping with the movement analysis, the limited range of variation in persistence velocity observed during release periods obscured different patterns of grazing.

This finding has implications for impacts to meadow habitat and can inform management tactics. For example, localized, intense grazing may negatively affect vegetation in a confined area whereas more ambulatory grazing may lessen the intensity of impacts to a particular location but increase the spatial extent. Future research examining pack stock animal movement behavior should explore additional environmental conditions, like semi-arid grasslands, and social behavior to extend our understanding of overnight grazing patterns. The incorporation of camera traps could also supplement behavior observation during twilight hours, complementing daytime behavior mapping to validate behavior classifications.

Because grazing can negatively affect meadow productivity, floristic composition, soil chemistry, microclimate conditions, and ground cover, while facilitating the encroachment of

invasive of woody species, management strives to balance user access afforded by stock use with resource protection (Barros & Pickering, 2015; Byers, 2009; Ostoja et al., 2014). The third chapter explored the influence of two common management tactics designed to confine free roaming stock use on animal distribution patterns. Animals congregated near release locations and areas fenced to confine a lead animal, resulting in greater use intensities in these settings. Strategically placing and rotating these settings can encourage grazing away from sensitive habitats or concentrate grazing in specific areas that may mitigate negative. Moving fences and release points may identify thresholds for points at which environmental factors like meadow size, access to water, and forage quality increase influence over subsequent distributions. Additionally, applying the methods and strategic placements and combinations of management tactics in additional environments and with different stock user groups (i.e., commercial outfitters, privately owned animals) can help identify if similar spatial associations between management tactics and animal location persist.

Scaling up from the site and hourly scales of the preceding chapters, the fourth chapter modelled visitor distributions in a PA at seasonal and annual scales, examining which infrastructure and environmental factors most contributed to patterns observed, using crowdsourced geographic information. For Hawaii Volcanoes National Park, a highly dynamic landscape resulting from ongoing volcanic activity, visitor use was most associated with visitor infrastructure (i.e., roads and trails), with the spatial locations suggestive of high probability of visitor presence expanding during the summer months to additional areas of the park. Crowdsourced geographic information also enabled descriptive statistics of visitor

clustering intensities, with results providing insight into the location and timing of continued and complementary visitor and resource monitoring methods.

The geospatial approaches incorporated in these case studies shed light on how visitors and visitation-related activities (i.e., use of pack stock) distribute and concentrate in natural landscapes at times or scales challenging to sample. Results demonstrated the capacity of a machine learning approach (i.e., MaxEnt) and a maximum-likelihood method (i.e., Behavior Change Point Analysis) in detecting distribution and behavior patterns, respectively, with data generated from sources experiencing increasing spatial and temporal resolution. This dissertation contributed to use-environment interaction research by providing empirical examples of how the selected geospatial approaches can be adapted, integrated, and leveraged at resolutions and scales relevant to interactions in PA landscapes and biodiversity conservation. The findings also provided assessments of how domestic pack animals and visitors interacted spatially with several environment factors and highlighted how integrating efficient and replicable methods inform efforts to reduce negative impacts to natural resources. Mitigating and minimizing impacts support PA missions to conserve biodiversity and provide for visitor experiences within the confines of environmental, logistical, and financial constraints.

More broadly, through addressing and minimizing challenges to obtaining and analyzing visitor and visitation-related use data, the geospatial approaches leveraged in this dissertation can generate data sets and analyses necessary for contributing to theory within natural and social science disciplines. New theories can emerge inductively from increasing observations

made possible through technology and GIScience (Blaschke & Merschdorf, 2014). Alternatively, increasing data provide opportunities to explore and enrich existing theories deductively. Either approach relies on accessible methods to generate data at scales or about phenomena not previously feasible. With data on interactions, feedbacks, factors, and conditions generated daily by authoritative and volunteered sources, exploring whether patterns and theories hold in new situations and environments is increasingly achievable across a diversity of questions and contexts. While methodological improvements are always possible, the approaches in this dissertation encourage strengthened collaboration across disciplines to improve our understanding of the world and address the grand challenges we face now and in the years ahead.

Implications and future directions

As existing pressures persist, new threats emerge, and visitor activities and patterns evolve, the need for efficient and accurate data on visitors remains critical to understanding dynamic interactions and supporting evidence-based and effective management (Monz & Leung, 2006). Visitor monitoring remains a top priority in park research, including an emphasis on developing methods that balance accuracy and efficiency and leverage electronic approaches (Eagles, 2014). Continued investigation contributing to these questions will also need to address growing analytical and privacy questions surrounding big data and of advancing possibilities of geospatial technologies. For example, are data posted publicly on social media sites truly public (Goodchild et al., 2017)? Should contributors of

volunteered geographic information be able to opt-out if researchers leverage their content for other purposes (Connors et al., 2012)? While the third case study leveraged only the geographic location of the shared media, the question of ethical limits for reuse is an important consideration for research analyzing photo contents, especially instances of social media documenting or facilitating illegal activities like vandalism and poaching.

Continued explorations of sampling bias inherent in crowdsourced geographic data are also critical for descriptive and statistical use of social media data. Increasing research on the quality and contributors of crowdsourced geographic have explored measures of accuracy, credibility, and how unequal access, demographics, and proximity can influence sampling bias (Antoniou et al., 2010; Brovelli, Minghini, & Zamboni, 2016; Brown, 2017; Brown, Kelly, & Whitall, 2013; Brown, Strickland-Munro, Kobryn, & Moore, 2017; Haklay, 2010; Li & Goodchild, 2012). Results often indicate the ability of volunteered geographic information comparable to authoritative sources (Goodchild et al., 2016). Deeper understanding of who posts what, where, when, and why help illuminate the capabilities, capacity, and limitations of using crowdsourced geographic information for research. This is also true of the spatial accuracy of data captured on mobile devices in locations without mobile or wireless services or areas prone to GPS error.

Increasing capacity of geospatial approaches contributes to advancing opportunities to make traditionally non-spatial data spatially explicit, a key objective outlined in the human-environment research agenda, and enhancing the capacity to integrate biological, physical, and social science disciplines (Moran, 2010). Collectively, this emphasis on a systems and

interdisciplinary approach affords new and innovative opportunities for furthering scientific inquiry across disciplines, engaging the public in both research efforts and educational opportunities, and contributing to biodiversity conservation and PA relevancy, value, and political support.

References

Antoniou, V., Morley, J., & Haklay, M. (2010). Web 2.0 geotagged photos: Assessing the spatial dimension of the phenomenon. *Geomatica*, *64*(1), 99–110.

Barros, A., & Pickering, C. M. (2015). Impacts of experimental trampling by hikers and pack animals on a high-altitude alpine sedge meadow in the Andes. *Plant Ecology & Diversity*, *8*(2), 265–276.

Blaschke, T., & Merschdorf, H. (2014). Geographic information science as a multidisciplinary and multiparadigmatic field. *Cartography and Geographic Information Science*, *41*(3), 196-213.

Brovelli, M. A., Minghini, M., & Zamboni, G. (2016). Public participation in GIS via mobile applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, *114*, 306–315.

Brown, G. (2017). A review of sampling effects and response bias in internet participatory mapping (PPGIS/PGIS/VGI). *Transactions in GIS*, *21*(1), 39–56.

Brown, G., Kelly, M., & Whitall, D. (2013). Which "public"? Sampling effects in public participation GIS (PPGIS) and volunteered geographic information (VGI) systems for public lands management. *Journal of Environmental Planning and Management*, *57*(2), 190–214.

Brown, G., Strickland-Munro, J., Kobryn, H., & Moore, S. A. (2017). Mixed methods participatory GIS: An evaluation of the validity of qualitative and quantitative mapping methods. *Applied Geography*, *79*, 153–166.

Buckley, R. (2009). Parks and tourism. *PLoS Biology*, *7*(6), 1–2.

Byers, A. C. (2009). A comparative study of tourism impacts on alpine ecosystems in the Sagarmatha (Mt. Everest) National Park, Nepal and the Huascaran National Park, Peru. In J. Hill & T. Gale (Eds.), *Ecotourism and environmental sustainability: Principles and practice* (pp. 51–72). London, UK: Routledge.

Cole, D. N., Millar, C., & Stephenson, N. L. (2010). Responding to climate change: A toolbox of management strategies. In D. N. Cole & L. Yung (Eds.), *Beyond naturalness: Rethinking park and wilderness stewardship in an era of rapid change* (pp. 179–196). Washington, D.C.: Island Press.

Connors, J. P., Lei, S., & Kelly, M. (2012). Citizen science in the age of neogeography: Utilizing volunteered geographic information for environmental monitoring. *Annals of*

the Association of American Geographers, *102*(6), 1267–1289.

Eagles, P. F. J. (2014). Research priorities in park tourism. *Journal of Sustainable Tourism*, *22*(4), 528–549.

Golder, S., & Macy, M. (2014). Digital footprints: Opportunities and challenges for online social research. *Annual Review of Sociology, 40*(1), 129-152.

Goodchild, M. F., Aubrecht, C., & Bhaduri, B. (2017). New questions and a changing focus in advanced VGI research. *Transactions in GIS*, 21, 189-190.

Haklay, M. (2010). How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environment and Planning B: Planning and Design*, *37*(4), 682–703.

Li, L., & Goodchild, M. F. (2012). Constructing places from spatial footprints. In *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Crowdsourced and Volunteered Geographic Information - GEOCROWD '12* (pp. 15–21). New York, NY: ACM Press.

Mayer, A.L., Donovan, R.P., & Pawlowski, C.W. (2014). Information and entropy theory for sustainability of coupled human and natural systems. *Ecology and Society, 19*(3), 11.

Monz, C., & Leung, Y.-F. (2006). Meaningful measures: Developing indicators of visitor impact in the National Park Service Inventory and Monitoring Program. *George Wright Forum*, *23*(2), 17–27.

Moran, E. F. (2010). *Environmental social science: Human-environment interactions and sustainability*. Malden, Mass.: Wiley-Blackwell.

Ostoja, S. M., Brooks, M. L., Moore, P. E., Berlow, E. ., Blank, R., Roche, J., … Haultain, S. (2014). Potential environmental effects of pack stock on meadow ecosystems of the Sierra Nevada , USA. *The Rangeland Journal*, *36*, 411–427.

APPENDICES

Appendix A: Supplement to Chapter 2

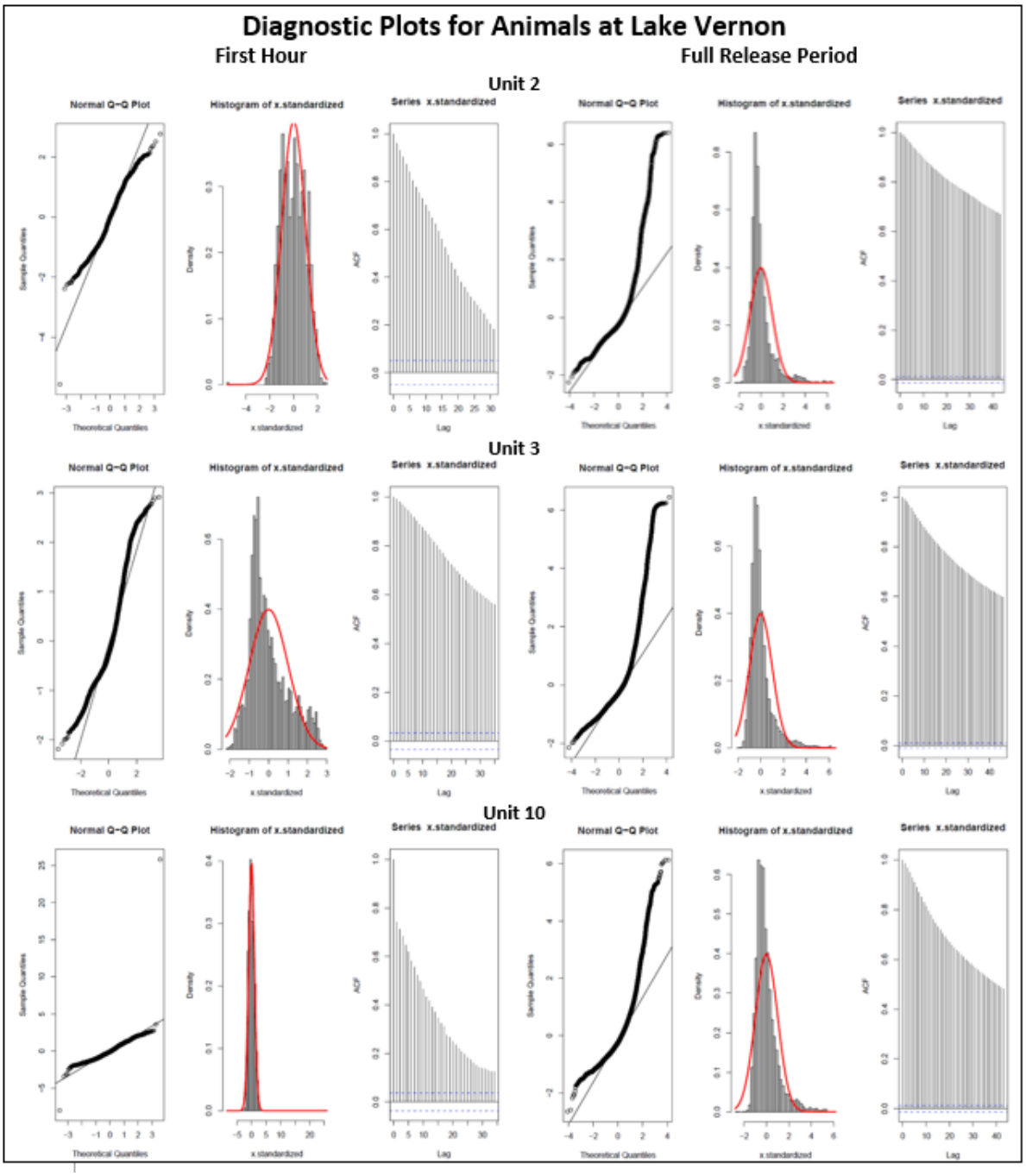


Figure A.1 Diagnostic plot of residuals for animals in Lake Vernon

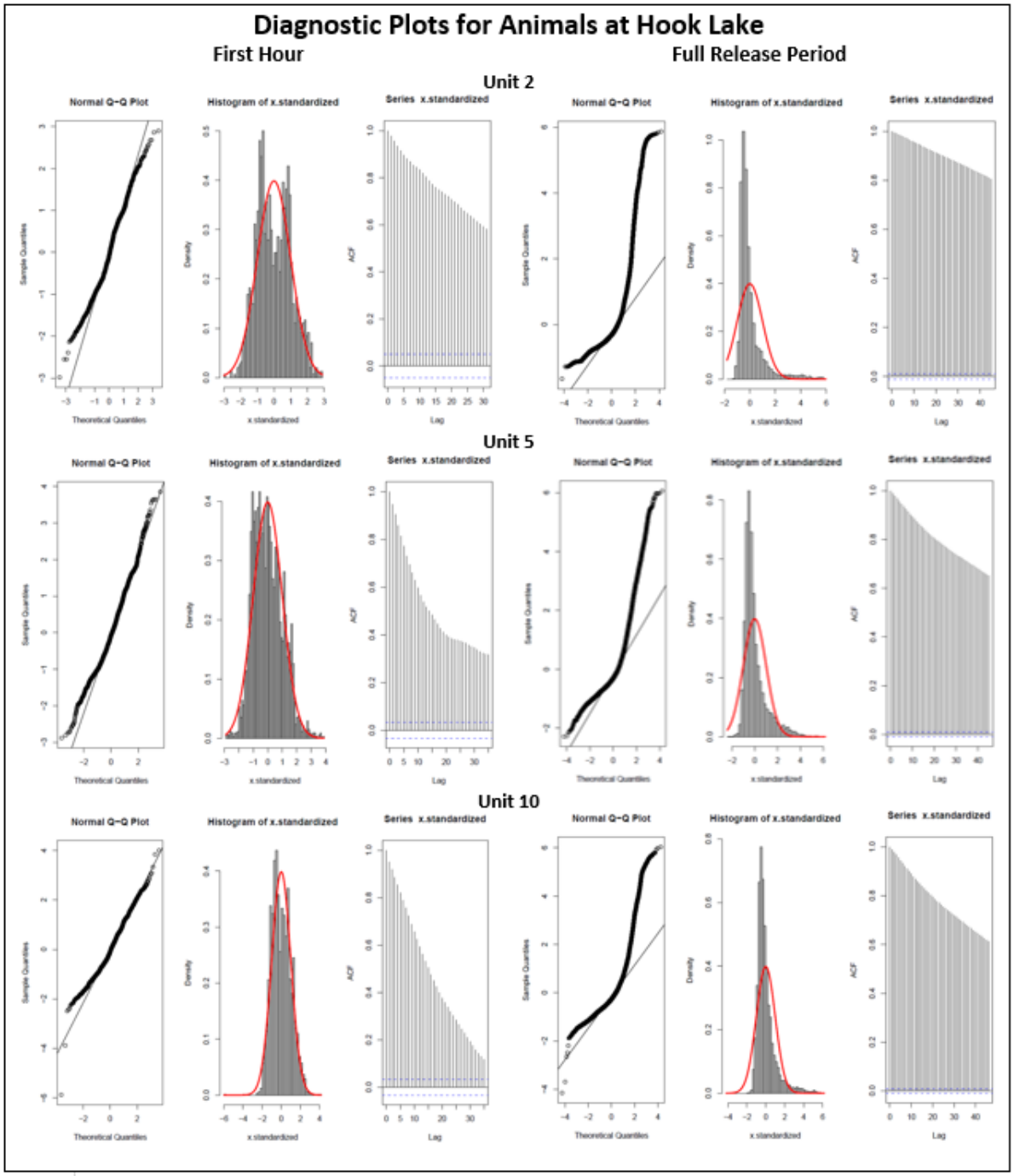


Figure A.2 Diagnostic plot of residuals for animals in Hook Lake

Table A.1 Summary of parameter values from BCPA in Lake Vernon. The average persistence velocity (in mps) observed between change points is denoted by $\hat{\mu}$, $\hat{\sigma}$ is the standard deviation, and $\hat{\rho}$ is the estimated temporal scale of autocorrelation (in seconds).

Lake Vernon														
		Full Release							First Hour					
		Minimum	**IQR1**	**Median**	**Mean**	**IQR3**	**Maximum**		**Minimum**	**IQR1**	**Median**	**Mean**	**IQR3**	**Maximum**
Unit 2	$\hat{\mu}$	0.0049	0.0775	0.1225	0.1819	0.2033	1.9979		0.0077	0.7896	0.1314	0.1368	0.1857	0.3356
	$\hat{\sigma}$	0.0000	0.0177	0.0273	0.0330	0.0423	0.2723		0.0010	0.1507	0.0233	0.0265	0.0354	0.0990
	$\hat{\rho}$	0.0006	6.1457	12.5977	26.9949	26.2862	2134.7495		0.0251	6.1223	14.8548	33.6424	28.2003	382.7495
		Minimum	**IQR1**	**Median**	**Mean**	**IQR3**	**Maximum**		**Minimum**	**IQR1**	**Median**	**Mean**	**IQR3**	**Maximum**
Unit 3	$\hat{\mu}$	0.0050	0.0771	0.1162	0.1576	0.1807	1.9574		0.0264	0.0821	0.1345	0.1686	0.2002	0.5789
	$\hat{\sigma}$	0.0000	0.0206	0.0293	0.3442	0.0424	0.2877		0.0034	0.0192	0.0262	0.0308	0.0377	0.1093
	$\hat{\rho}$	0.0006	4.8485	8.9731	12.3392	17.3440	219.4996		0.6189	5.4212	10.7068	14.1877	22.3554	50.2497
		Minimum	**IQR1**	**Median**	**Mean**	**IQR3**	**Maximum**		**Minimum**	**IQR1**	**Median**	**Mean**	**IQR3**	**Maximum**
Unit 10	$\hat{\mu}$	0.0045	0.6150	0.0931	0.1221	0.1488	0.9255		0.0045	0.0675	0.0952	0.1287	0.1515	0.5071
	$\hat{\sigma}$	0.0000	0.0167	0.0256	0.0304	0.0378	0.3928		0.0004	0.0150	0.0221	0.0264	0.0336	0.1169
	$\hat{\rho}$	0.0010	5.0980	10.1240	25.9240	20.5840	4711.2500		0.0374	6.7687	12.5257	15.5698	24.6275	95.7497

Table A.2 Parameter values from BCPA in Hook Lake. The average persistence velocity (in mps) observed between change points is denoted by $\hat{\mu}$, $\hat{\sigma}$ is the standard deviation, and $\hat{\rho}$ is the estimated temporal scale of autocorrelation (in seconds).

Hook Lake														
		Full Release							First Hour					
		Minimum	**IQR1**	**Median**	**Mean**	**IQR3**	**Maximum**		**Minimum**	**IQR1**	**Median**	**Mean**	**IQR3**	**Maximum**
Unit 2	$\hat{\mu}$	0.0008	0.0577	0.0949	0.1565	0.1710	1.2216		0.0377	0.1036	0.1788	0.2187	0.3105	0.6128
	$\hat{\sigma}$	0.0000	0.0205	0.0285	0.0332	0.0400	0.2309		0.0064	0.0183	0.0270	0.0297	0.0369	0.0841
	$\hat{\rho}$	0.0170	3.5740	6.6540	15.1410	13.5390	3096.2500		0.5914	4.0358	7.5674	10.8037	14.4056	29.9997
		Minimum	**IQR1**	**Median**	**Mean**	**IQR3**	**Maximum**		**Minimum**	**IQR1**	**Median**	**Mean**	**IQR3**	**Maximum**
Unit 5	$\hat{\mu}$	0.0087	0.0553	0.0842	0.1267	0.1372	0.9646		0.0160	0.0573	0.0796	0.1078	0.1176	0.4237
	$\hat{\sigma}$	0.0000	0.0174	0.0246	0.0283	0.0349	0.3065		0.0000	0.0153	0.0219	0.0229	0.0294	0.0805
	$\hat{\rho}$	0.0161	3.5876	6.6367	11.9829	13.0313	1916.4120		0.0161	3.2432	6.4483	8.8151	11.8829	29.9997
		Minimum	**IQR1**	**Median**	**Mean**	**IQR3**	**Maximum**		**Minimum**	**IQR1**	**Median**	**Mean**	**IQR3**	**Maximum**
Unit 10	$\hat{\mu}$	0.0059	0.6772	0.1082	0.1511	0.1797	1.9114		0.0122	0.0814	0.1392	0.1588	0.2151	0.4724
	$\hat{\sigma}$	0.0000	0.0188	0.0274	0.0322	0.0397	0.4979		0.0009	0.1556	0.0247	0.0298	0.0378	0.1281
	$\hat{\rho}$	0.0006	4.5067	8.9032	13.6954	18.9756	423.9997		0.0194	5.8114	13.2267	15.5499	24.7116	58.8036

Appendix B: Supplement to Chapter 3

Table A.3 Summary of vegetation alliances within the area of interest for the MaxEnt models. Vegetation alliance classifications are from the USDA CALVEG dataset.

Vegetation Alliance	Percent of CACA AOI	Percent of UPKE AOI
Alpine Grasses and Forbs		3
Alpine Mixed Scrub		0.7
Barren	3.76	15.92
Barren - Snow/Ice		0.06
Basin Sagebrush	2.24	
Great Basin - Mixed Chaparral Transition		0.59
Jeffrey Pine	0.72	
Lodgepole Pine	32.36	16.54
Low Sagebrush		1.69
Mixed Conifer - Fir	11.03	
Mountain Hemlock	18.02	
Mountain Sagebrush		0.16
Perennial Grasses and Forbs		2.34
Quaking Aspen	1.94	
Red Fir	0.16	
Subalpine Conifers	18.53	26.02
Upper Montane Mixed Chaparral		0.73
Water		1.77
Western White Pine	5.18	0.58
Wet Meadows	0.66	6.48
Whitebark Pine	3.85	17.86
Willow	1.37	5.56

Appendix C: Flickr API code for R

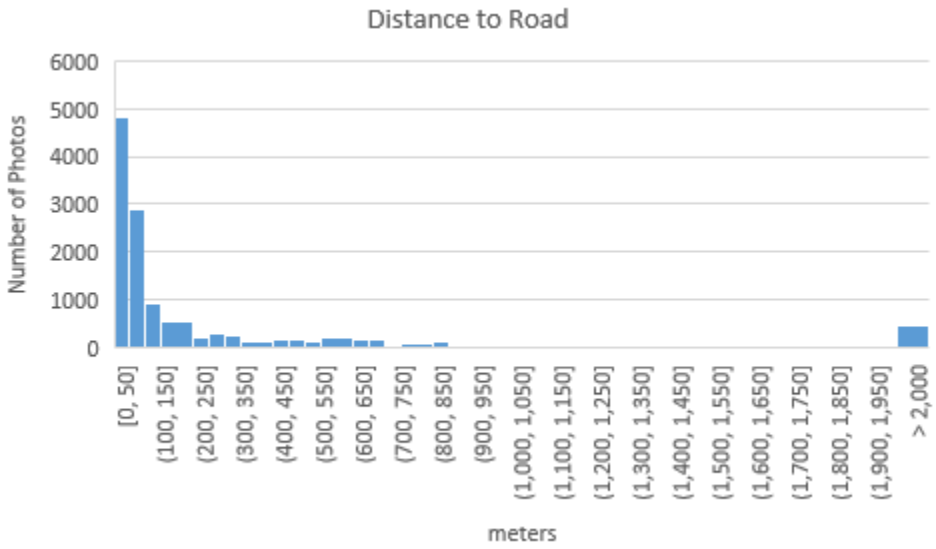
```
# Purpose: Create a spreadsheet of geotagged photo data from Flickr.
# Results based on search dates and bounding box (minimum latitude, minimum longitude,
#                                                 maximum latitude, maximum longitude)
# Functions adapted from Fredheim, R. (2015). Digital data collection course.
# Retrieved June 1 2015 from http://www.r-bloggers.com/digital-data-collection-course/.

# Install and require necessary R packages.
install.packages("RCurl")
install.packages("rjson")
install.packages("RJSONIO")

require(RCurl)
require (rjson)
require(RJSONIO)

# Set working directory with path name in quotation marks and / instead of \
setwd("C:/Users/Desktop")


# Flickr API access keys
# Key must have single quotations around it
# Example: 'A8MNR7997LLOP123'
api_key =
api_secret =

# Create function to return URL using flickr.photos.search
# More information at https://www.flickr.com/services/api/flickr.photos.search.html
getURL = function(api_key, minDate, maxDate, minLon, minLat, maxLon, maxLat, pageNum){
  root = 'https://api.flickr.com/services/rest/?method=flickr.photos.search&'
  u = paste0(root,"api_key=",api_key, "&min_taken_date=",  minDate,"&max_taken_date=",
                maxDate,"&bbox=", minLon,"%2C+", minLat, "%2C+", maxLon, "%2C+", maxLat,
               "&has_geo=1&extras=geo%2C+date_taken%2C+date_upload%2C+
                views%2C+tags%2c+url_o&per_page=250&page=",
                pageNum, "&format=json&nojsoncallback=1" )
  return(URLencode(u))
}

#Create function to return URL to get Exif data from image metadata
getURL2 = function(api_key, id, secret){
  root2 = 'https://api.flickr.com/services/rest/?method=flickr.photos.getExif&'
  u2 = paste0(root2,"api_key=",api_key,"&photo_id=",id,"&secret=",secret,"&format=json&nojsoncallback=1"
)
  return(URLencode(u2))
}
```

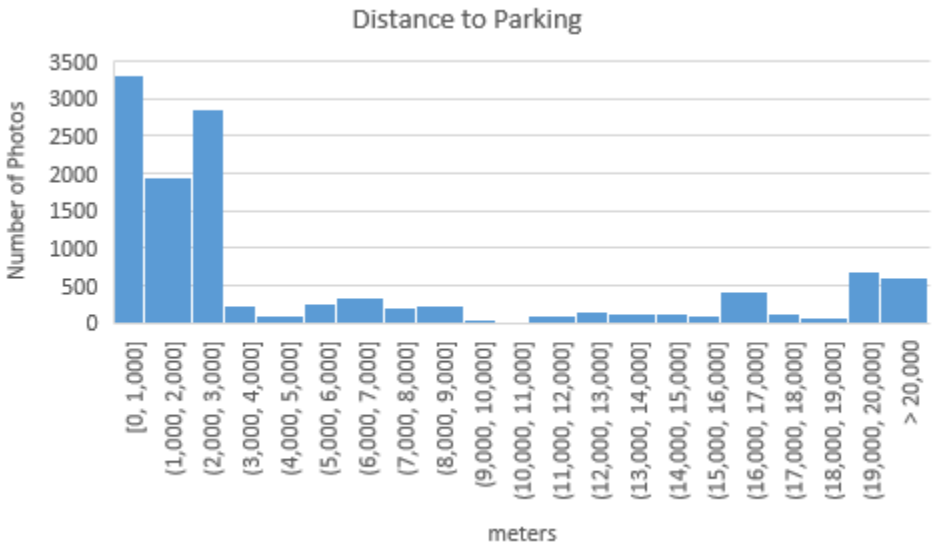
```
# Set location search parameters.
# Coordinates should be in decimal degrees, surrounded by single quotes
# Example: minLon = '-155.798371'
minLon =
minLat =
maxLon =
maxLat =

# Set date search parameters.
# Use date format 'YYYY-MM-DD' with single quotations.
minDate =
maxDate =

#First call for specified date and bbox
        ##This will return the URL that contains the information based on search variables above
        ##The remainder of the code will then read through the information and write it to a data frame
getURL(api_key, minDate, maxDate, minLon, minLat, maxLon, maxLat, pageNum=1)

#Read data returned from first call
target = getURL(api_key, minDate, maxDate, minLon, minLat, maxLon, maxLat, pageNum=1)
data = fromJSON(target)

# Get the total number of photo records returned using the current search parameters.
total = as.numeric(data$photos$total)

# Number of pages of records returned using search parameters.
numPages = data$photos$pages

# Create empty dataframe to populate with data.
df = NULL
# For each page of results, from the first to maximum page number extract photo information.
for (j in 1:numPages){
  pageNum = j
  target.loop = getURL(api_key, minDate, maxDate, minLon, minLat, maxLon, maxLat, pageNum)
  data1 = fromJSON(target.loop)
  numPhotos = length(data1$photos$photo)

  # Read photo information for each photo on the current page.
  for (i in 1:numPhotos){
    id = data1$photos$photo[[i]]$id
    title = data1$photos$photo[[i]]$title
    lat = data1$photos$photo[[i]]$latitude
    lon = data1$photos$photo[[i]]$longitude
    owner = data1$photos$photo[[i]]$owner
    taken = data1$photos$photo[[i]]$datetaken
    dateupload = data1$photos$photo[[i]]$dateupload

    # Convert UNIX epoch to date-time.
    upload = as.character.Date(as.POSIXct(as.numeric(dateupload), origin="1970-01-01"))
    views = data1$photos$photo[[i]]$views
```

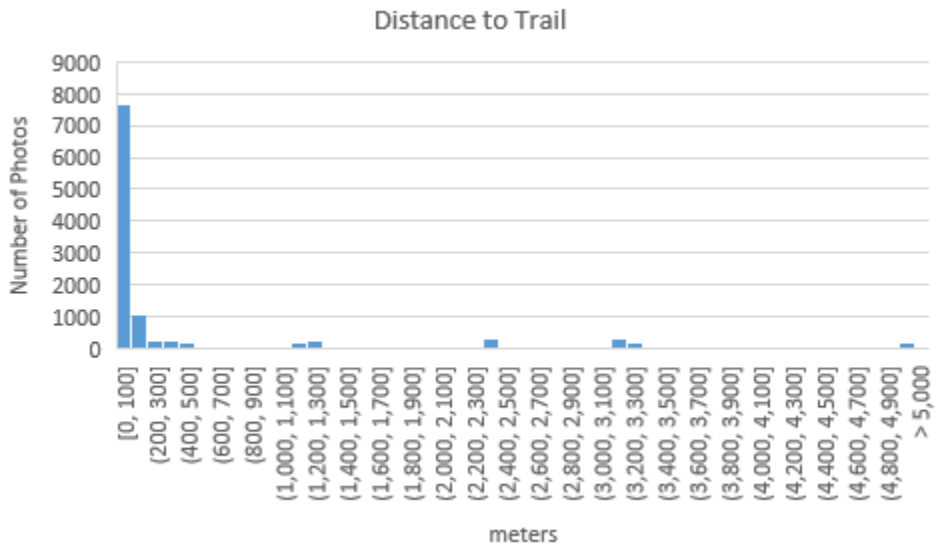
```
    tags = data1$photos$photo[[i]]$tags
    secret = data1$photos$photo[[i]]$secret
    server = data1$photos$photo[[i]]$server
    farm = data1$photos$photo[[i]]$farm
    imageURL = paste("https://farm", farm, ".staticflickr.com/", server, "/", id, "_", secret, ".jpg", sep="")
    getURL2(api_key, id, secret)
    target2 = getURL2(api_key, id, secret)
    exifData = fromJSON(target2)

if (exifData$stat!="fail") {
      device = exifData$photo$camera
    } else{
       # If a value is not provided, then skip.
      device='NA'
    }
    row = cbind(lon, lat, id, owner, taken, upload, views, tags, title, imageURL, device)
    rbind(df, row)-> df

    if (j + 1 < numPages) {
      pageNum = j + 1
    }
  }
}

# Write data in dataframe to a csv file.
write.csv(x=df, file="flickrPhotos.csv", row.names = TRUE)
```

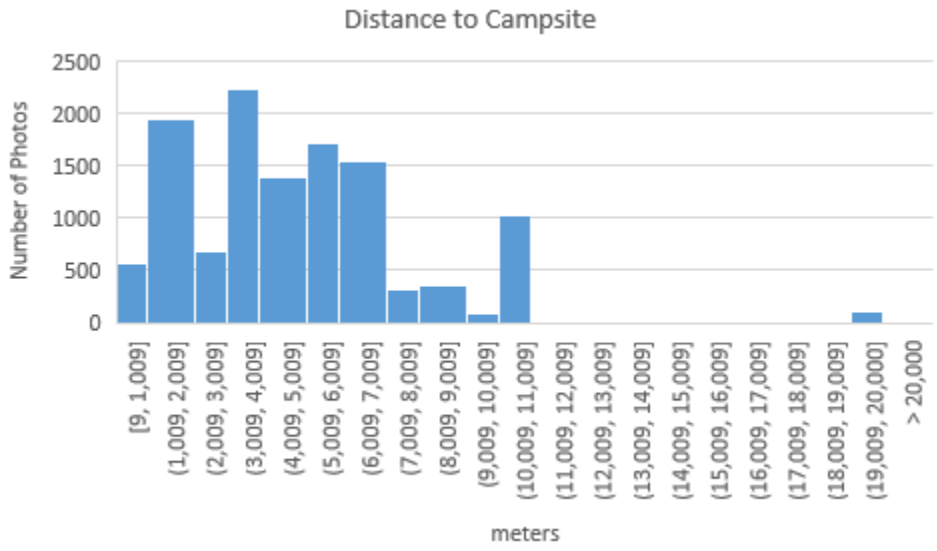
Appendix D: Histogram of photo records compared to model variables

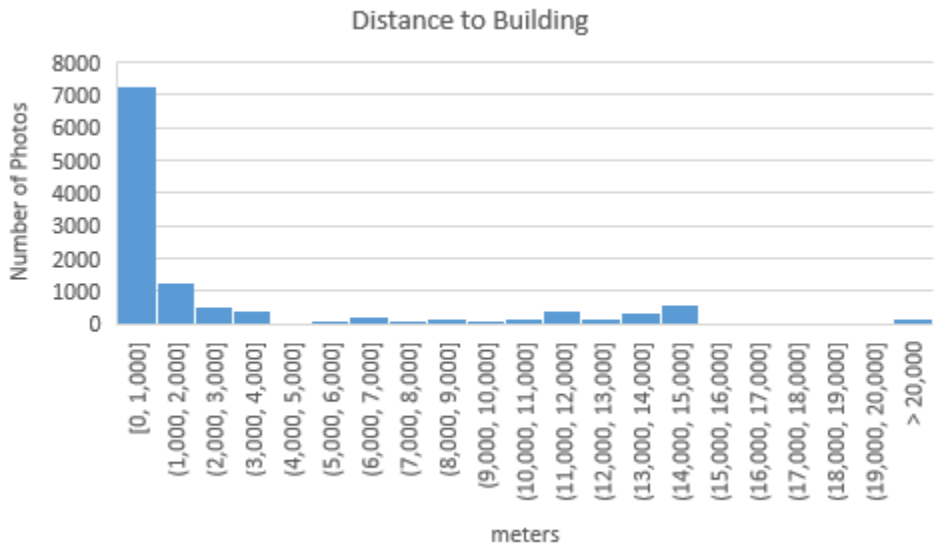
For each variable included in the MaxEnt model, the histograms below illustrate the number and distribution of photo records compared to the distance from visitor infrastructure or environmental features, or occurred within specific vegetation types or terrain values.

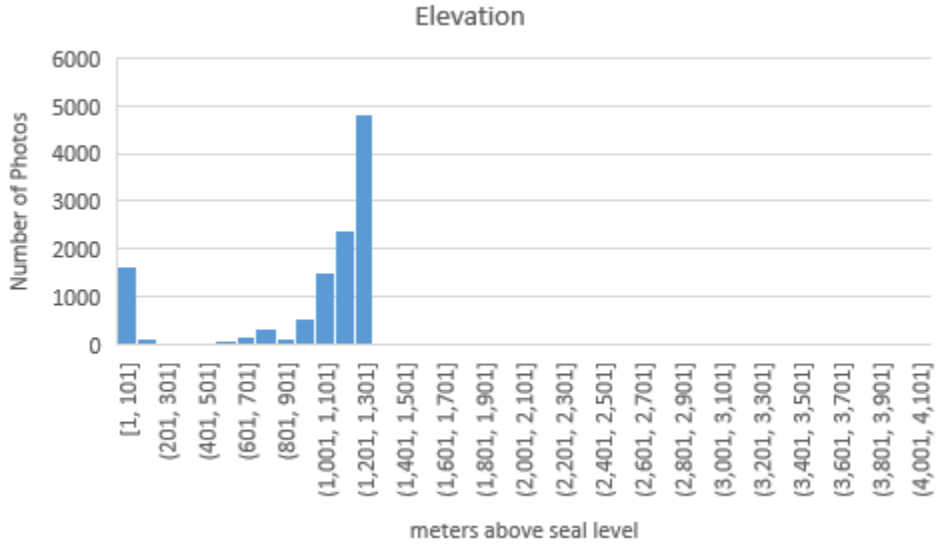


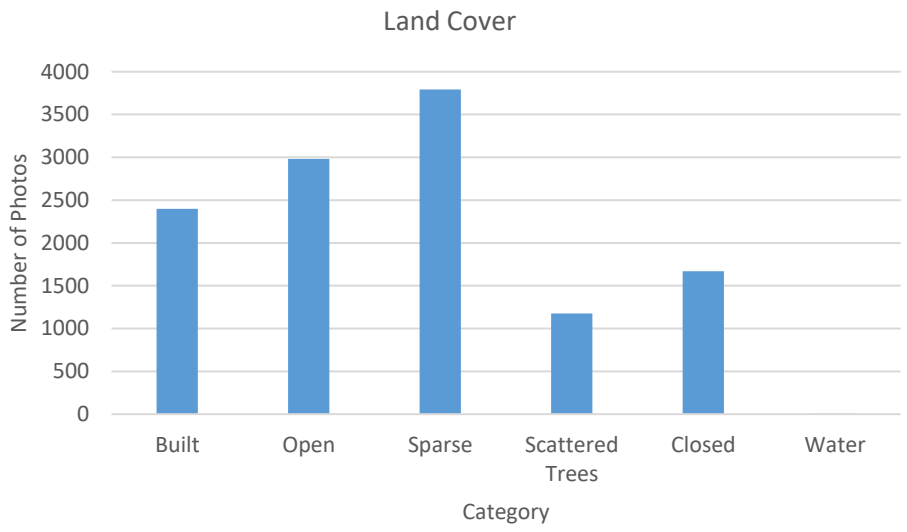


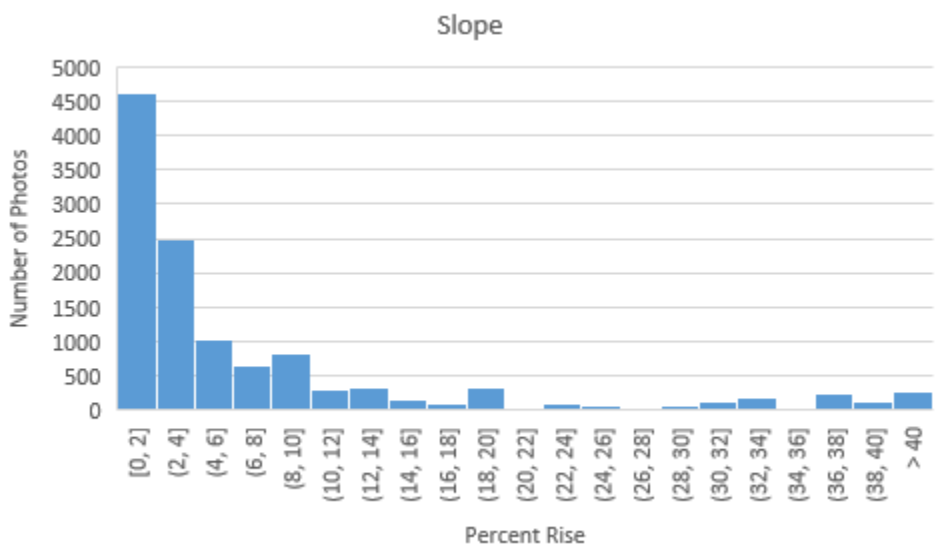

Distance to Trail
Number of Photos
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
[0, 100]
(200, 300]
(400, 500]
(600, 700]
(800, 900]
(1,000, 1,100]
(1,200, 1,300]
(1,400, 1,500]
(1,600, 1,700]
(1,800, 1,900]
(2,000, 2,100]
(2,200, 2,300]
(2,400, 2,500]
(2,600, 2,700]
(2,800, 2,900]
(3,000, 3,100]
(3,200, 3,300]
(3,400, 3,500]
(3,600, 3,700]
(3,800, 3,900]
(4,000, 4,100]
(4,200, 4,300]
(4,400, 4,500]
(4,600, 4,700]
(4,800, 4,900]
> 5,000
meters

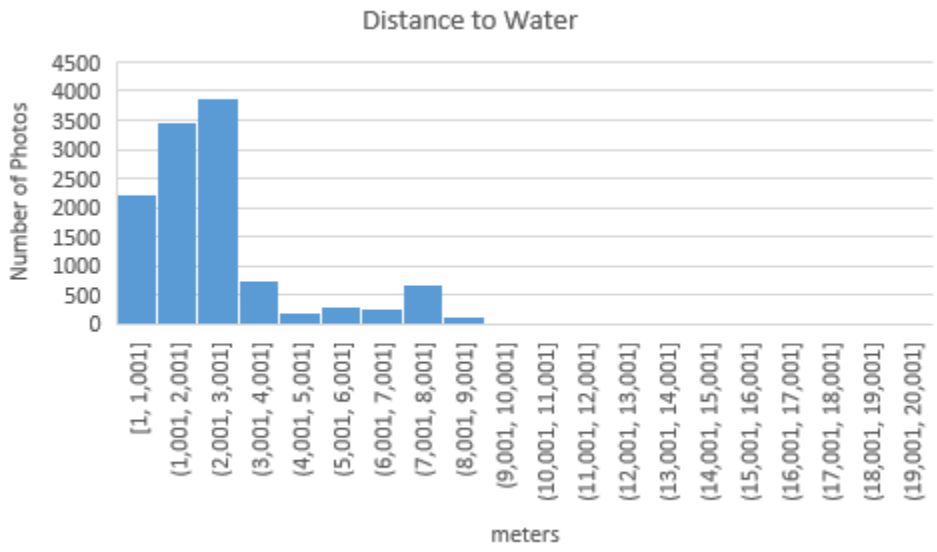

Distance to Campsite
Number of Photos
2500
2000
1500
1000
500
0
[9, 1,009]
(1,009, 2,009]
(2,009, 3,009]
(3,009, 4,009]
(4,009, 5,009]
(5,009, 6,009]
(6,009, 7,009]
(7,009, 8,009]
(8,009, 9,009]
(9,009, 10,009]
(10,009, 11,009]
(11,009, 12,009]
(12,009, 13,009]
(13,009, 14,009]
(14,009, 15,009]
(15,009, 16,009]
(16,009, 17,009]
(17,009, 18,009]
(18,009, 19,009]
(19,009, 20,000]
> 20,000
meters


Distance to Building
Number of Photos
8000
7000
6000
5000
4000
3000
2000
1000
0
[0, 1,000]
(1,000, 2,000]
(2,000, 3,000]
(3,000, 4,000]
(4,000, 5,000]
(5,000, 6,000]
(6,000, 7,000]
(7,000, 8,000]
(8,000, 9,000]
(9,000, 10,000]
(10,000, 11,000]
(11,000, 12,000]
(12,000, 13,000]
(13,000, 14,000]
(14,000, 15,000]
(15,000, 16,000]
(16,000, 17,000]
(17,000, 18,000]
(18,000, 19,000]
(19,000, 20,000]
> 20,000
meters


Elevation
Number of Photos
6000
5000
4000
3000
2000
1000
0
[1, 101]
(201, 301]
(401, 501]
(601, 701]
(801, 901]
(1,001, 1,101]
(1,201, 1,301]
(1,401, 1,501]
(1,601, 1,701]
(1,801, 1,901]
(2,001, 2,101]
(2,201, 2,301]
(2,401, 2,501]
(2,601, 2,701]
(2,801, 2,901]
(3,001, 3,101]
(3,201, 3,301]
(3,401, 3,501]
(3,601, 3,701]
(3,801, 3,901]
(4,001, 4,101]
meters above seal level


Land Cover
Number of Photos
4000
3500
3000
2500
2000
1500
1000
500
0
Built
Open
Sparse
Scattered Trees
Closed
Water
Category


Slope
Number of Photos
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
[0, 2]
(2, 4]
(4, 6]
(6, 8]
(8, 10]
(10, 12]
(12, 14]
(14, 16]
(16, 18]
(18, 20]
(20, 22]
(22, 24]
(24, 26]
(26, 28]
(28, 30]
(30, 32]
(32, 34]
(34, 36]
(36, 38]
(38, 40]
> 40
Percent Rise


Distance to Water
Number of Photos
4500
4000
3500
3000
2500
2000
1500
1000
500
0
[1, 1,001]
(1,001, 2,001]
(2,001, 3,001]
(3,001, 4,001]
(4,001, 5,001]
(5,001, 6,001]
(6,001, 7,001]
(7,001, 8,001]
(8,001, 9,001]
(9,001, 10,001]
(10,001, 11,001]
(11,001, 12,001]
(12,001, 13,001]
(13,001, 14,001]
(14,001, 15,001]
(15,001, 16,001]
(16,001, 17,001]
(17,001, 18,001]
(18,001, 19,001]
(19,001, 20,001]
meters

Appendix E: Nearest neighbor ratios for geotagged photo records, 2011-2015

Table A.4 Average nearest neighbor statistics by year

	Observed Mean Distance (meters)	**Expected Mean Distance** (meters)	**Nearest Neighbor Ratio**
2011	37.9721	389.1237	0.0976*
2012	42.9587	399.2325	0.1076*
2013	33.3781	360.1958	0.0927*
2014	31.0493	366.7038	0.0847*
2015	25.4932	480.9970	0.0530*

* Statistically significant ($p < 0.000$)