PH.D. DISSERTATION PROPOSAL

Name: Rachel Fricke **Faculty Adviser: Julian Olden** **Date: 4/3/23**

This is to certify that this student's PhD Supervisory Committee had read and approved the Ph.D. Dissertation Proposal titled: ______

Approval of Supervisory Committee:

POSITION	PRINTED NAME	SIGNATURE	DATE
CHAIR	Julian Olden	Julian Olden	3/10/23
MEMBER	Andrew Berdahl		mar 12, 2023
MEMBER	Spencer Wood		3/13/23
MEMBER	Katie Arkema		3/27/2023
MEMBER			

Emerging technologies to assess human benefits from and risks to water resources

Rachel M. Fricke

A draft proposal for the degree of
Doctor of Philosophy

School of Aquatic and Fishery Sciences
University of Washington
Seattle, WA 98195

15 November 2022

Introduction

Freshwater resources provide a suite of cultural ecosystem services (CES) to humanity. For example, lakes and reservoirs alone provide drinking water, habitat and population maintenance, spiritual spaces, and numerous types of recreation (e.g., fishing, boating, swimming, camping) (Reynaud and Lanzanova 2017). However, the interaction between and linkage of these services to one another remains a critical gap in our understanding and valuation of them (Hanna *et al.* 2018). Previous work has found that the valuation of water-based services varies spatially within and between watersheds, and thus understanding the spatiotemporal distribution of freshwater CES is critical to managing water resources with the goal of meeting and preserving societal demands and values (Chen *et al.* 2018).

A robust understanding of the geography and dynamics of lake visitation is critical for anticipating and minimizing human impacts. While water-based activities contribute to human well-being, they simultaneously impose stressors on aquatic systems such as depleting aquatic and riparian habitat quality, altering species' behaviors, and changing the biogeochemical cycles of aquatic ecosystems (Venohr *et al.* 2018). Humans are also traversing the landscape to experience different types of CES, and this creation of overland pathways can pose further harm to water resources. For example, human movement is a significant and growing source of invasive species transmission (Drake and Mandrak 2014). Angler and boater activities that entangle invasive organisms on fishing gear, boat hulls, and outboard engines, or use non-native species as live bait, serve as modes of introduction.

Traditionally, human activity on waterbodies across time and space is inferred from sparsely conducted surveys that provide data with limited spatiotemporal scope (Davis and Darling 2017). Conventional approaches typically rely on in-person surveys (often at boat launches) or mail-in questionnaires and thus only provide a limited snapshot in time of activity at a particular location (Rothlisberger et al. 2010; Anderson et al. 2014). An additional drawback of such methods for holistically evaluating water-based CES is their narrow targeting of specific user groups and associated activities (e.g., anglers and boaters).

Previous studies in a terrestrial context have found strong correlation between mobile and ground-truthed visitation estimates for points of interest. At U.S. national parks, for example, the numbers of visitors that are estimated to visit each park by the National Park Service each month are highly correlated with the numbers of photographs that are shared on Flickr from the same parks (Sessions *et al.* 2016). Significant positive correlation between empirical visitation estimates and estimates from mobile applications has also been quantified at parks in New York City, Vermont state parks, and global recreational sites (Hamstead *et al.* 2018; Sonter *et al.* 2016; Wood *et al.* 2013). Despite these advances, such analyses have not been done for blue spaces (lakes), and data validation is lacking for aquatic resources over varying spatial scales (Keeler *et al.* 2015; Fisher *et al.* 2018). In recent years, computational tools have maximized the use of these public datasets which can provide a close to real-time understanding of human behavior and cost-effective monitoring (Deville *et al.* 2014; Venturelli *et al.* 2017; Xia *et al.* 2020).

Numerous potential uses exist for mobile application data within freshwater management. Foremost, the simple quantification of the magnitude and frequency of visitation over space and time offers an opportunity to largely automate visitor use surveys over previously unfeasible scales. Furthermore, an evaluation of water-based CES over a large spatiotemporal scale can aid in prioritizing additions and updates to waterbody infrastructure supporting recreation, evaluating the impact of environmental perturbations (e.g., algae blooms, wildfires, droughts) on the spatiotemporal distribution of waterbody use, estimating the likelihood of invasive species transmission based on propagule pressure, and assessing patterns in visitation before, during, and after the COVID-19 pandemic (e.g., Merrill et al. 2020, Ford et al. 2016, Fricke et al. 2020, Huang et al. 2020). In addition, climate change is projected to both increase and induce spatiotemporal shifts in demand for recreational ecosystem services on public lands (Fisichelli et al. 2015, Wilkins et al. 2021a). Recent studies have shown that daily weather conditions strongly influence patterns of recreation in parks, and social media data reveals that spatial trends in visitor behavior may change in response to increasingly frequent heatwaves (Wilkins et al. 2021b).

Mobile data is a highly promising resource, and yet its implementation presents many challenges. While most Americans own a cell phone (97%), ownership of smartphones in particular varies significantly by age, education, and household income. For example, only ~75% of individuals with a high school education or less, or making less than $30,000 per year, own smartphones (Pew Research Center 2021). Consistent use of mobile applications is often limited to a small proportion of total users who tend to fall at extremes in terms of their interest level and dedication to posting (Ruths and Pfeffer 2014; Gundelund *et al.* 2020). Significant variability in phone ownership, mobility, and social media use among demographic groups may also lead to biases in population-level estimates (Wesolowski *et al.* 2013; Malik *et al.* 2015; Feng et al. 2019).

Nevertheless, previous work has confirmed that mobile application data enhance visitation estimates, particularly at unmonitored sites and when parameterized with ground estimates (Wood *et al.* 2020). Compared to ground survey data, in urban parks mobile application data provides a greater magnitude of observations across a given area, thus offering park managers visitation data for locations with little or no in-person survey presence (Donahue *et al.* 2018). Optimal approaches will incorporate multiple streams (applications) of mobile data, as data from different types of applications (e.g., social media, citizen science, activity sharing) represent specific user groups and vary in their temporal and spatial granularity (Heikinheimo *et al.* 2020).

The vast suite of available mobile app-based technologies offer a novel, yet rarely explored, opportunity to improve our understanding of human behavior and CES (Venturelli *et al.* 2017; Havinga *et al.* 2020). At the highest level, mobile data can be categorized by how actively the mobile phone user is involved in collecting the data. Data that are generated by a user intentionally engaging with their device (e.g., posting photos, sending a tweet, logging a species observation) are types of active data, while data that are collected in the background once a user has downloaded an app and given permission for data collection (e.g., GPS location tracking) are passive (Wenz *et al.* 2017). Active data from mobile applications can be further classified by an application's purpose, such as social media (e.g., Twitter, Flickr), citizen science portal (e.g., iNaturalist, eBird) and outdoor activity-sharing platform (e.g., Gaia GPS, Strava) (Havinga *et al.* 2020). These data can include geotagged user-generated records, text, and images.

Here, I propose the first research to leverage 'big data' from mobile applications to quantify human visitation to, movement between, and use of waterbodies. Lakes and reservoirs are widespread and support critical ecosystem functions, provide numerous goods and services, and contribute to sustainable local and regional communities (Klessig 2001, Schindler 2009, Reynaud and Lanzanova 2017). Furthermore, blue spaces (lakes) in urban and suburban settings and adjacent green spaces (parks) serve as hotspots of connections to nature and offer heat stress relief in the midst of urban heat island effects (Gunawardena *et al.* 2017). Given lakes' multidimensional role in supporting human well-being, understanding human activity on them over vast spatiotemporal scales and assessing the benefits and risks associated with human water-based activities is critical for lake valuation and conservation.

My dissertation aims to assess the CES humans derive from and quantify the risk (via potential transmission of invasive species) we pose to lakes and reservoirs (hereafter, waterbodies) across western Washington by leveraging user-generated, mobile data sources and automating content analyses with machine learning approaches. Specifically, my objectives are to (i) summarize existing literature on use of emerging technologies to manage invasive species (ii) quantify spatiotemporal patterns in waterbody visitation from social media data, (iii) understand sociodemographic differences in waterbody use and accessibility, (iv) assess the magnitude and location of angler and boater activity across lakes as it relates to the potential for invasive species transmission under shifting temperature and precipitation regimes, and (v) classify the CES affiliated with water visitation by analyzing post and image content. These goals will be achieved in five chapters I describe further in the following sections.

<u>Chapter 1:</u> Technological innovations enhance invasive species management in the Anthropocene

Objectives
I aimed to offer an overview of technology-based tools that aid in invasive species management, with the goal of capturing the pace and scope of emerging technologies applied to this discipline. This review is tailored toward resource managers, and presents examples of these approaches as they relate to specific management priorities ranging from pathway intervention to preventing spread to limiting impacts and increasing public engagement.

Methods
I collated existing knowledge on the use of emerging technologies – approaches which benefit from one or more forms of remote data collection, analytical automation, and crowdsourcing user-generated data – for invasive species management. I conducted a systematic literature review to identify scientific studies and gray literature describing the integration of these new technologies in the prevention, detection, and control of invasive species and summarized my findings in a review article.

Results
Dr. Olden and I resubmitted a revised manuscript titled "Technological innovations enhance invasive species management in the Anthropocene" to *BioScience* on September 19, 2022 (first submitted January 29, 2022), and am currently awaiting a decision from the journal editor. As described by reviewers, our paper is "timely and relevant for a variety of audiences," and "should be really well received by a broad audience interested in thinking about and implementing solutions for vexing global change problems."

<u>Chapter 2:</u> Inferring waterbody visitation from a suite of social media and cellular data sources

Objectives

In this chapter, I will establish the utility of social media data for estimating lake visitation at waterbodies in western Washington and assess the drivers of lake visitation. These goals will be achieved by (i) validating social media estimates with field-based counts of visitation and (ii) modeling lake visitation as a function of lake and public park attractiveness.

Methods

Data acquisition and filtering

I will acquire a suite of user-generated data from social media (Twitter, Flickr), citizen science portals (iNaturalist, eBird), and outdoor activity-sharing platform (Gaia GPS) applications via data sharing agreements and application programming interface (API) queries (Fox *et al.* 2020). I am collating data from multiple platforms because previous studies have demonstrated significant population biases within each independent mobile data source, contributing to varying abilities to reflect actual visitation (Tenkanen *et al.* 2017; Hausmann *et al.* 2018; Heikinheimo *et al.* 2020). I will extract post metadata (containing the geographic coordinates of the device, a place name, and a distinct user ID) and images using the Twitter Developer API. Flickr metadata will be acquired with UW eScience's existing API. All other data sources are obtained through data sharing agreements or direct data downloads. The spatiotemporal extent and types of available data vary by mobile data source, and each individual application has its own strengths and weaknesses with regards to my study purposes and methods (Table 1). I will identify posts on waterbodies and in adjacent public parks based on geographic overlay with the USGS' National Hydrography Dataset and extract visitation frequency for each site (expressed as visitor user-days).

Table 1. Mobile app and cell data sources summarized by types of data available for extraction, spatiotemporal extent, strengths and weaknesses.

Data Source	Available Content	Spatial Extent	Temporal Extent	Strengths	Weaknesses
Flickr	Records Text Photos	Continental U.S.	2004 - present	Multiple data types for inferring activity type and derived CES.	Use is on the decline, and photo storage for free accounts is capped.
Twitter	Records Text Photos	Continental U.S.	2006 - present	Multiple data types for inferring activity type and derived CES.	Limited number of tweets are geotagged (< 1%), though previous work suggests ample data still.
eBird	Records	Continental U.S.	2002 - present	Minimum effort to acquire.	Single type of use.

iNaturalist	Records	Continental U.S.	2008 - present	Minimum effort to acquire.	Single type of use.
Gaia GPS	Records	Western WA	2009 - present	Better coverage of waterbodies in remote regions.	Summarized on a lake level, unable to quantify activity on individual basis.
Airsage	Records Home Locale Demographics	Western WA	2019 - present	Provides additional user locale and demographic info unavailable from any other source.	Expensive, and for limited spatiotemporal extent.

Visitation Model

I will validate my estimates of human visitation by comparing mobile post- and field-based estimates of visitor user-days on lakes (Wood *et al.* 2013; Donahue *et al.* 2018). Empirical estimates of user visitation available for validation come from a suite of datasets with varying sampling frequencies and total durations (Table 2). I will first model visitation at 70 lakes in King and Snohomish counties with both on-site and social media data using a GLMM which . Next, I will use this model to predict visitation at ~40 additional lakes for which I only have social media data.

Table 2. Preliminary list of ground-collected datasets collated for validation of mobile data, summarized by spatiotemporal extent and additional notes. Additional data sources will be investigated.

Data Source	Spatial Extent	Temporal Extent	Description
UW Data Collection	18 lakes in King and Snohomish Counties, WA	June-Sept 2021	Monthly single-day weekday and weekend counts of users by activity type.
King County	38 lakes in King County, WA	May-Oct 1994-2019	Bimonthly single-day counts of users by activity type.
Snohomish County	39 lakes in Snohomish County, WA	May-Oct 2014-2021	Monthly single-day counts of users by activity type.
Whatcom County	Samish and Whatcom (multiple launch sites) Lakes	2015-2020	Daily boat inspection records with last waterbodies visited and home state of vessel registration.
Wisconsin DNR	Hundreds of boat launches across Wisconsin	2003-2020	Single-day counts of boats and number of people at boat launches.

Revealed Preference Model
Lastly, I will model blue space (lake) visitation derived from social media as a function of lake and park attractiveness (e.g., boat launch, water quality, fish stocking, park amenities) and relative accessibility (public transportation) for 50 popular lakes in western Washington. This follows previous studies where social-media visitation rates were examined (Wood *et al.* 2020).

Results
Outputs from this analysis will include projections of lake visitation and key amenities driving lake use across 50 lakes throughout the Puget Sound region. These findings will be shared with partners via a visually dynamic web application and published as a peer-reviewed manuscript.

<u>Chapter 3:</u> Sociodemographic trends in western Washington urban lake use

Objectives
Understanding sociodemographic patterns in lake use is critical for ensuring equitable access to public lake resources. Here, I will use lake user surveys and in-person counts of lake use to (i) quantify spatiotemporal trends in public lake access use and (ii) understand how sociodemographic identities influence lake use and perceived barriers.

Methods
Survey distribution
In Summer 2021, my collaborators and I generated our own lake visitation and use dataset by surveying activity at 20 popular lakes in western Washington state via in-person hourly counts and distribution of digital and paper surveys asking respondents to identify the CES they associate with lakes and lakeside parks. We asked respondents to categorically rank their participation in different lake and park-associated activities and likelihood of park amenity use. In addition, we requested respondents' sociodemographic information (anonymized), home zip code, communication preferences, and barriers they associated with lake or park use. Lastly, we installed traffic counters at a subset of sites to relate visitation to vehicle traffic.

Survey analyses
First, I will calculate summary statistics of lake use and perceived barriers by sociodemographic group. Second, I will bin lake attributes into clusters by natural aesthetics (e.g., shade, open green space), infrastructure (e.g., parking lots boat launches), and amenities (e.g., swimming beaches, playgrounds). I will then use redundancy analysis and variation partitioning to identify significant lake attributes driving lake visitation. Third, I will assemble matrices of lake use types and perceived barriers by sociodemographic identities and use principal component analysis to assess trends in lake use and perceived barriers by sociodemographic group. Fourth, I will calculate average distance traveled per lake from home zip code responses. Lastly, I will categorically code open-ended survey responses to the prompt "List any challenges or barriers that make it difficult for you to access this lake" (Vaughn and Turner 2016).

Results
Previous studies on a national scale have revealed ethnic minorities and marginalized groups identify personal safety, language, and transport as barriers constraining their ability to access the outdoors (Ghimire et al. 2014). I expect my surveys to identify similar barriers explicitly

related to lake environments (e.g., water quality concerns related to swimming safety). Findings from this project will be communicated directly to our partners at King and Snohomish Counties, and we will share all data and analyses with them via a public GitHub repository. In addition, I will write and submit a manuscript of our work for publication in *Northwest Science.*

<u>Chapter 4:</u> Revealing human-mediated vectors for aquatic invasive species from mobile technology

Objectives

This study will couple user-generated location records from mobile applications and cellular phones with fine-scale weather data and records of invasive species distributions to (i) assess how the magnitude and location of boater and angler activity across lakes in the Puget Sound lowlands may change due to shifting temperature and precipitation regimes, and (ii) visualize and quantify connections between waterbodies in terms of the magnitude, direction and timing of angler and boater movements, with the goal of identifying potential invasion hubs. This chapter is funded by the Northwest Climate Adaptation Science Center, and an additional goal of this work is to co-produce actionable science with state agency partners.

Methods

Temperature and precipitation trends

First, I will calculate fine-scale spatiotemporal trends in maximum daily temperature and precipitation accumulation across waterbodies in western Washington with the gridMET dataset from University of California Merced's Climatology Lab (Abatzoglou 2013). Preliminary analyses of on-site visitation counts for Summer 2021 show a slight correlation of lake visitation with maximum temperatures from a much coarser NOAA dataset (Figure 1).

Mobile application, cellular, and on-site datasets

During my graduate studies, I have acquired geotagged user-generated records, text, and images from various cellular applications (e.g., Twitter, iNaturalist, Gaia GPS) (Chapter 2). In addition, archived cellular phone records were purchased from Airsage – a private telecommunications company – which collates anonymous location data from approximately 30% of the cell phone owning population. The spatial extent of my mobile application and cell data is 50 lakes and adjacent public parks in western Washington along an urban to rural gradient, dating back to the inception of each mobile app (ranging from 2002 to 2008) or cellular dataset (2019). Mobile application and cell record metadata includes unique

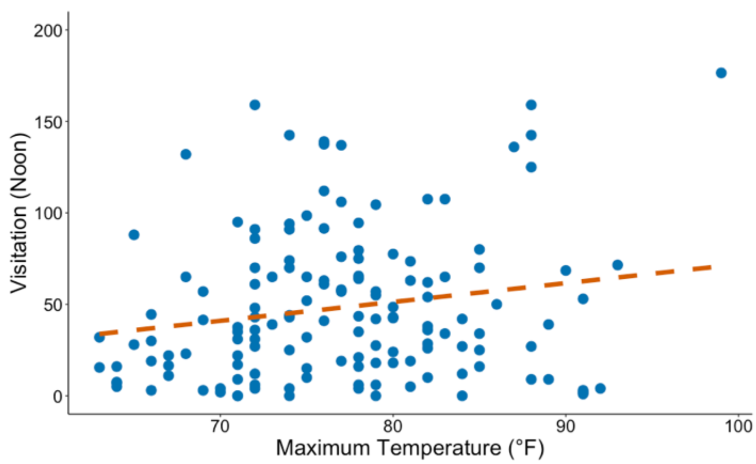


Figure 1. Noon visitation correlates with maximum daily temperature (°F) at 18 lakes in King and Snohomish counties surveyed bi-weekly from June-September 2021.

user IDs allowing me to construct trip matrices of movement between lakes over time (Weir et al. 2022). These data have been contrasted to field-based empirical estimates of actual lake use according to citizen science surveys provided by collaborators at King and Snohomish counties (Chapter 2).

Assessing climate influences on human behavior and predicting risk of invasion
I will deploy machine-learning approaches to model lake visitation (according to mobile application and phone user-days) as a function of temperature and precipitation trends and day of the week. This follows previous studies where social-media visitation rates were examined (Wood et al. 2020). Then, I will use a large database of invasive species distributions from USGS' Nonindigenous Aquatic Species database to identify angler and boater movements from invaded to uninvaded waterbodies, which will further be prioritized by time estimates of desiccation tolerance obtained from the literature. Graph-theoretical methods will be used to identify potential invasion hubs from a network model quantifying the magnitude, direction, and timing of overland connections between waterbodies from angler and boater movements (Stewart-Koster et al. 2015). Model outputs will be shared directly with stakeholder partners via a web application allowing for dynamic visualization of results and an educational webinar summarizing key findings.

Results
Final products will include models forecasting climate influence on future visitation across fifty lakes in the Puget Sound lowlands and identifying potential invasion hubs, a web application allowing stakeholder partners to visualize and download summary model projections of human visitation and movement (building on work from Chapter 2), and a peer-reviewed publication in Nature Communications. In addition, I will outreach with my stakeholder partners and other relevant parties via an educational webinar summarizing key findings from our study and offering instructions for using the web application user interface.

<u>Chapter 5:</u> Classifying water-based cultural ecosystem services from images and text posted to social media

Objectives
The goal of this chapter is to develop a classification of freshwater CES which merges all streams and types (images, text, point locations) of mobile application data. Specifically, I aim to (i) identify and classify key words in post text and images, and (ii) link these classifications to broader CES categories which can be derived from social media data (Figure 2).

Methods
Data acquisition
Using previously acquired geotagged posts from a suite of social media data sources (Chapter 2), I will extract associated post image and text data by referencing post metadata. Twitter image URLs, text, and tags are already available within post metadata , and Flickr images, text, and tags will be extracted using the photosearcher R package (Fox et al. 2020). Ecosystem services for all other data sources will be assigned by the type of activity the phone user is engaging in, which is either available in the metadata or can be inferred from the app itself.

Identifying ecosystem services

For the purposes of defining types of ecosystem services, I have adopted a categorical system previously developed by Havinga et al. (2020) (Figure 2). I plan to constrain my CES analysis to a regional scale (western Washington), as completing an analysis over a greater spatial scale is unfeasible given my time constraints. I will classify ecosystem services associated with each record of mobile application activity by first assigning those which can be directly inferred from activity type (e.g., Naturalist from eBird and iNaturalist records). For applications with text and tags associated with posts (Twitter and Flickr), I will automate keyword detection in extracted post text with artificial intelligence built on natural language processing. This will allow me to build a dictionary of terms associated with waterbody use from which I can match keywords in posts to identify activity type and distinguish associated CES (Monkman *et al.* 2018; Schirpke et al. 2021). I will use a convolutional neural network (CNN) to analyze and classify image content (shore vs. water, boats, fishing gear), assign activity type (e.g., boating, photography), and categorize the CES derived from waterbody visitation. My training dataset will consist of positive (waterbody use) and negative (non-water activity) image sets, and the test image set will contain images from social media. If necessary, I will augment training data for underrepresented activities by geometrically transforming images (Winder et al. 2022). After labeling training images, image processing and recognition will be automated to identify waterbody user types in Microsoft Azure's CNN. This image recognition algorithm returns keywords associated with the content of each photograph, and I will hierarchically cluster photographs by their keywords after processing to expedite manual human review (Richards and Tunçer 2018). Text and image analysis outputs will be analyzed in tandem to synthesize contextual information (Väisänen et al. 2021).

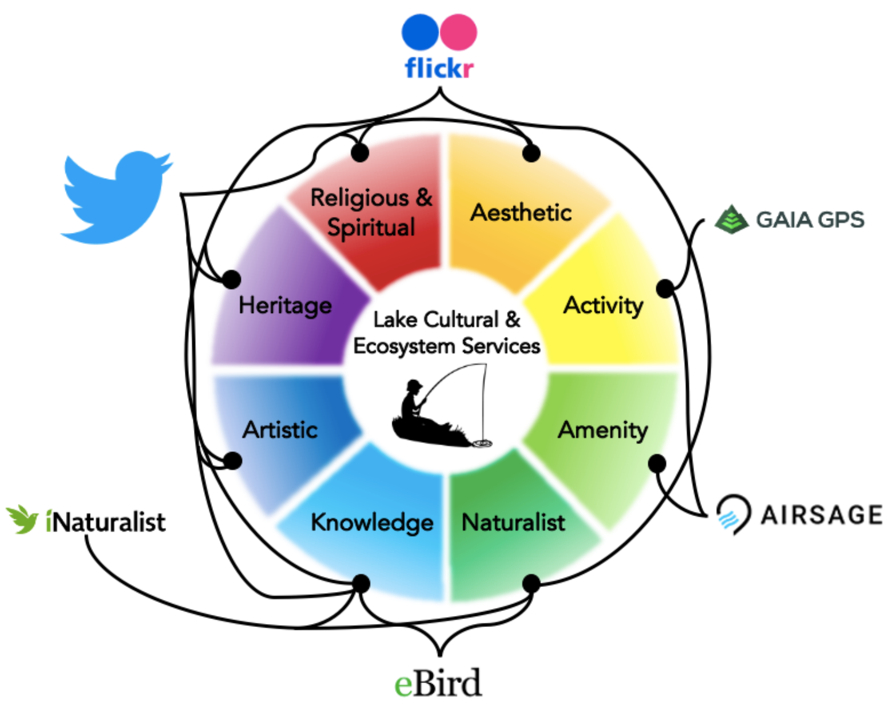


Figure 2. Streams of data from different mobile applications and cellular data sources inform water-based cultural and ecosystem services based on types of data (images, text, point locations) available for extraction from each platform (*see Table 1*).

Results
This analysis will be the first to synthesize automated text and image content from social media posted on lakes. The various streams of data will be integrated in a public GitHub repository with standardized data and code, and I will visualize CES across lakes in a web-mapping application (initiated in Chapter 2), thus ensuring accessibility for resource managers. Findings from this study will also be published as a peer-reviewed manuscript.

Interpretation

This research will be among the first to leverage 'big data' from mobile applications to quantify human movement over vast spatiotemporal scales and assess the benefits and risks associated with recreational water activities. My review paper will collate existing studies from disparate sources across the peer-reviewed and gray literature to summarize emerging applications of technological advances for detecting and managing invasive species – thus synthesizing new technologies with the goal of encouraging their uptake by managers. Chapter two will be one of the first analyses within the field of social media and outdoor recreation to validate social media with on-site instantaneous counts rather than continuous time series from counter devices, and this is a valuable step toward utilizing pre-existing datasets from resource managers in similar studies.

Furthermore, my third chapter will serve as one of only a handful of existing lake user surveys for CES. Understanding the sociodemographic breakdown of these services and their accessibility is critical for informing lake and park management. In addition, my network models of human activity connecting lakes will aid resource managers in identifying target locations for preventative measures. Lastly, my fifth chapter will be the first comprehensive effort of CES valuation based on mobile application data for freshwaters, an important step for informing conservation strategies and developing large-scale, cost-efficient monitoring (Boulton *et al.* 2016).

Significance

This work is an interdisciplinary collaboration between data scientists, ecologists, social scientists, and resource managers. My study will be one of the first large-scale efforts to automate monitoring of water-based activity and assessment of CES by applying artificial intelligence-based approaches to social media imagery, thus pushing the envelope of current approaches for analyzing mobile application content (Toivonen *et al.* 2019). Ultimately, I will create a pipeline delivering summary data from private applications to public resource managers for the purpose of enhancing and protecting our water resources. Implementation of infrastructure supporting data and analysis sharing encourages prolonged coordination between technologists and conservation practitioners (Berger-Tal and Lahoz-Monfort 2018). Final products will include a public GitHub repository containing code and summary-level data, web applications for visualizing risk of invasive species transmission and CES across lakes in western Washington state, and a peer-reviewed publication for each dissertation chapter outlined above.

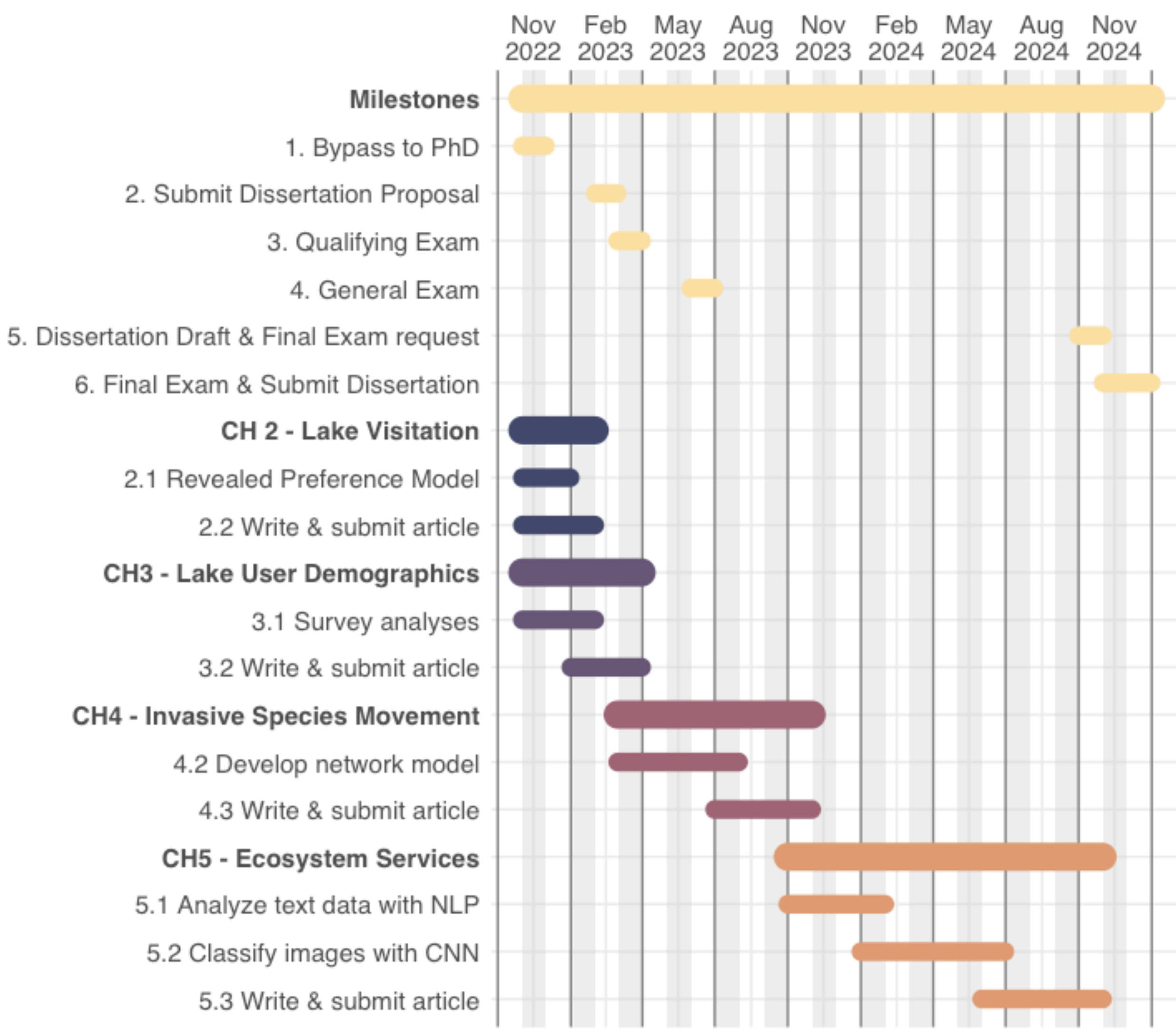
Through my work I will engage waterbody users and regulatory entities. Stakeholder collaborators in my region-scale analyses and field studies include King County Water Resources Division and Snohomish County Public Works, and summer field surveys were conducted in partnership with researchers from UW Bothell and UW Tacoma. Results from our lake user surveys will help King and Snohomish counties equitably allocate amenity upgrades to parks, plan future transportation access to public lakes (e.g., parking lots, transit stops), and improve signage around lake visitor concerns (e.g., water quality).

For my fourth chapter, I will partner directly with resource management partners, including the Washington Department of Fish and Wildlife, Washington State Parks, and the Washington Invasive Species Council – an assembly of representatives from federal, state, local, and tribal entities that shapes state policy to reduce invasive species' impacts. Collaborators employed by these state agencies and councils will provide input on data analyses and the design of the web application, and serve as co-authors on the culminating publication. Identification of specific waterbodies at highest risk of invasive species introductions will allow these stakeholders to prioritize waterbody locations for preventative measures such as educational signage, boat inspection stations, and gear cleaning services.

Timeline

During the first two years of my graduate studies, I have held my first committee meeting and had my thesis proposal approved (Spring 2021), conducted field surveys at lakes to inform my third chapter analyses (Summer 2021), completed preliminary analyses of survey data (Fall 2021), submitted my first chapter to *BioScience* (Winter 2022, resubmitted Fall 2022), acquired funding for and purchased Airsage cellular location data (Winter 2022), and made significant progress on my second chapter analyses. The timeline below outlines the remainder of my graduate work (Fall 2022 – Fall 2024).

Table 3. Planned PhD timeline by month (November 2022 – December 2024) including major milestones (yellow), and tasks to complete chapters three (blue), four (purple), and five (orange). Vertical lines delineate quarters.



References

Abatzoglou JT. 2013. Development of gridded surface meteorological data for ecological applications and modelling. *Int J Climatol* **33**: 121–131.

Anderson LG, White PCL, Stebbing PD, Stentiford GD, Dunn AM. 2014. Biosecurity and vector behaviour: Evaluating the potential threat posed by anglers and canoeists as pathways for the spread of invasive non-native species and pathogens. *PLoS ONE* **9**: e92788.

Berger-Tal O and Lahoz-Monfort JJ. 2018. Conservation technology: The next generation. *Conserv Lett* **11**: 1–6.

Boulton AJ, Ekebom J, and Gíslason G már. 2016. Integrating ecosystem services into conservation strategies for freshwater and marine habitats: a review. *Aquat Conserv Mar Freshw Ecosyst* **26**: 963–85.

Chen Y, Parkins JR, and Sherren K. 2018. Using geo-tagged Instagram posts to reveal landscape values around current and proposed hydroelectric dams and their reservoirs. *Landsc Urban Plan* **170**: 283–92.

Davis AJS and Darling JA. 2017. Recreational freshwater fishing drives non-native aquatic species richness patterns at a continental scale. *Divers Distrib* **23**: 692–702.

Deville P, Linard C, Martin S, *et al.* 2014. Dynamic population mapping using mobile phone data. *Proc Natl Acad Sci* **111**: 15888–93.

Donahue ML, Keeler BL, Wood SA, *et al.* 2018. Using social media to understand drivers of urban park visitation in the Twin Cities, MN. *Landsc Urban Plan* **175**: 1–10.

Drake DAR and Mandrak NE. 2014. Bycatch, bait, anglers, and roads: Quantifying vector activity and propagule introduction risk across lake ecosystems. *Ecol Appl* **24**: 877–94.

Feng GC, Zhang Y, and Lin Z. 2019. A meta-analysis of the effects of sociodemographic factors on social media adoption. *Int J Commun–US* **13**: 1996–2025.

Fischichelli NA, Schuurman GW, Monahan WB, and Ziesler PS. 2015. Protected area tourism in a changing climate: Will visitation at US National Parks warm up or overheat? *PLoS ONE* **10**: e0128226.

Fisher DM, Wood SA, White EM, *et al.* 2018. Recreational use in dispersed public lands measured using social media data and on-site counts. *J Environ Manage* **222**: 465–74.

Ford JD, Tilleard SE, Berrang-Ford L, *et al.* 2016. Opinion: Big data has big potential for applications to climate change adaptation. *P Natl Acad Sci* **113**: 10729-10732.

Fox N, August T, Mancini F, *et al.* 2020. "photosearcher" package in R: An accessible and reproducible method for harvesting large datasets from Flickr. *SoftwareX* **12**: 100624.

Fricke RM, Wood SA, Martin DR, and Olden JD. 2020. A bobber's perspective on angler-driven vectors of invasive species transmission. *NeoBiota* **60**: 97–115.

Ghimire R, Green GT, Poudyal NC, and Cordell HK. 2014. An analysis of perceived constraints to outdoor recreation. *J Park Rec Admin* **32**: 52–67.

Gunawardena KR, Wells MJ, Kershaw T. 2017. Utilising green and bluespace to mitigate urban heat island intensity. *Sci Total Environ* **584**: 1040–55.

Gundelund C, Arlinghaus R, Baktoft H, *et al.* 2020. Insights into the users of a citizen science platform for collecting recreational fisheries data. *Fish Res* **229**: 105597.

Hamstead ZA, Fisher D, Ilieva RT, *et al*. 2018. Geolocated social media as a rapid indicator of park visitation and equitable park access. *Comput Environ Urban* **72**: 38–50.

Hanna DEL, Tomscha SA, Ouellet Dallaire C, and Bennett EM. 2018. A review of riverine ecosystem service quantification: Research gaps and recommendations. *J Appl Ecol* **55**: 1299–311.

Hausmann A, Toivonen T, Slotow R, *et al.* 2018. Social Media Data Can Be Used to Understand Tourists' Preferences for Nature-Based Experiences in Protected Areas. *Conserv Lett* **11**: 1–10.

Havinga I, Bogaart PW, Hein L, and Tuia D. 2020. Defining and spatially modelling cultural ecosystem services using crowdsourced data. *Ecosyst Serv* **43**: 101091.

Heikinheimo V, Tenkanen H, Bergroth C, *et al.* 2020. Understanding the use of urban green spaces from user-generated geographic information. *Landsc Urban Plan* **201**: 103845.

Huang X, Li Z, Jiang Y, Li X, Porter D. 2020. Twitter reveals human mobility dynamics during the COVID-19 pandemic. *PLoS ONE* **15**: e0241957.

Keeler BL, Wood SA, Polasky S, *et al.* 2015. Recreational demand for clean water: Evidence from geotagged photographs by visitors to lakes. *Front Ecol Environ* **13**: 76–81.

Klessig LL. 2001. The contribution of lakes to sustainable societies. *Lake Reserv Managet* **6**: 95-101.

Malik MM, Lamba H, Nakos C, and Pfeffer J. 2015. Population bias in geotagged tweets. *AAAI Work - Tech Rep* **WS-15-18**: 18–27.

Merrill NH, Atkinson SF, Mulvaney KK, *et al.* 2020. Using data derived from cellular phone locations to estimate visitation to natural areas: An application to water recreation in New England, USA. *PLoS One* **15**: 1–19.

Monkman GG, Kaiser MJ, and Hyder K. 2018. Text and data mining of social media to map wildlife recreation activity. *Biol Conserv* **228**: 89–99.

Pew Research Center. 2021. Mobile fact sheet.

Reynaud A and Lanzanova D. 2017. A global meta-analysis of the value of ecosystem services provided by lakes. *Ecol Econ* **137**: 184-194.

Richards DR and Tunçer B. 2018. Using image recognition to automate assessment of cultural ecosystem services from social media photographs. *Ecosyst Serv* **31**: 318–325.

Ruths D and Pfeffer J. 2014. Social media for large studies of behavior. *Science* **346**: 1063–4.

Rothlisberger JD, Chadderton WL, McNulty J, Lodge DM. 2010. Aquatic invasive species transport via trailered boats: What is being moved, who is moving it, and what can be done. *Fisheries* **35**: 121–132.

Schirpke U, Tasser E, Ebner M, and Tappeiner U. 2021. What can geotagged photographs tell us about cultural ecosystem services of lakes? *Ecosyst Serv* **51**: 101354.

Sessions C, Wood SA, Rabotyagov S, Fisher DM. 2016. Measuring recreational visitation at US National Parks with crowd-sourced photographs. *Environ Manage* **183**: 703-11.

Schindler DW. 2009. Lakes as sentinels and integrators for the effects of climate change on watersheds, airsheds, and landscape. *Limnol Oceanogr* **54**: 2349-2358.

Sonter LJ, Watson KB, Wood SA, Ricketts TH. 2016. Spatial and temporal dynamics and value of nature-based recreation, estimated via social media. *PLoS One* **11**: e0162372.

Stewart-Koster B, Olden JD, and Johnson PT. 2015. Integrating landscape connectivity and habitat suitability to guide offensive and defensive invasive species management. *J App Ecol* ***52***: 366–378.

Tenkanen H, Minin E Di, Heikinheimo V, *et al.* 2017. Instagram, Flickr, or Twitter: Assessing the usability of social media data for visitor monitoring in protected areas. *Sci Rep* **7**: 1–11.

Toivonen T, Heikinheimo V, Fink C, *et al.* 2019. Social media data for conservation science: A methodological overview. *Biol Conserv* **233**: 298–315.

Väisänen T, Heikinheimo V, Hiipala T, and Tuivonen T. 2021. Exploring human-nature interactions in national parks with social media photographs and computer vision. *Conserv Biol* **35**: 424–436.

Vaughn P and Turner C. 2016. Decoding via coding: Analyzing qualitative text data through thematic coding and survey methodologies. *J Libr Admin* **56**: 41-51.

Venohr M, Langhans SD, Peters O, *et al.* 2018. The underestimated dynamics and impacts of water-based recreational activities on freshwater ecosystems. *Environ Rev* **26**: 199–213.

Venturelli PA, Hyder K, and Skov C. 2017. Angler apps as a source of recreational fisheries data: opportunities, challenges and proposed standards. *Fish Fish* **18**: 578–95.

Weir JL, Vacura K, Bagga J, *et al.* 2022. Big data from a popular app reveals that fishing creates superhighways for aquatic invaders. *PNAS Nexus* **1**:1–9.

Wenz A, Jäckle A, and Couper MP. 2017. Willingness to use mobile technologies for data collection in a probability household panel. *ISER Working Paper* 2017-10.

Wesolowski A, Eagle N, Noor AM, *et al.* 2013. The impact of biases in mobile phone ownership on estimates of human mobility. *J R Soc Interface* **10**.

Wilkins EJ, Wood SA, and Smith JW. 2020. Uses and limitations of social media to inform visitor use management in parks and protected areas: A systematic review. *Environ Manage*.

Wilkins EJ, Howe PD, and Smith JW. 2021. Social media reveal ecoregional variation in how weather influences visitor behavior in U.S. National Park Service units. *Sci Rep* **11**: 2403.

Wilkins EJ, Chikamoto Y, Miller AB, and Smith JW. 2021. Climate change and the demand for recreational ecosystem services on public lands in the continental United States. *Glob Env Chang* **70**: 102365.

Wood SA, Guerry AD, Silver JM, and Lacayo M. 2013. Using social media to quantify nature-based tourism and recreation. *Sci Rep* **3**.

Wood SA, Winder SG, Lia EH, *et al.* 2020. Next-generation visitation models using social media to estimate recreation on public lands. *Sci Rep* **10**: 1–12.

Xia J, Wang J, and Niu S. 2020. Research challenges and opportunities for using big data in global change biology. *Glob Chang Biol* **26**: 6040–61.