Snake River Fish and Habitat Relationship Evaluation, Annual Project Report

Project 2019-006-00

Report covers work performed under BPA contracts #85403
Report was completed under BPA contract #85403

June 2020 – May 2021

Kevin See, Richie Carmichael, Sarah Hoffmann, Michael Ackerman, Steve Gorrono, Chris Beasley
Biomark, Inc.
Boise, Idaho

Report Created: May 2021

LIST OF TABLES

LIST OF FIGURES

This report should be cited as follows:

See, K., R.C. Carmichael, S.L. Hoffmann, M.W. Ackerman, S. Gorrono, and C. Beasley. 2021. Snake River Fish and Habitat, Annual Project Report. June 2020 – May 2021, BPA Project 2019-006-00, pp:1-30.

CHAPTER 1. Formalize Information Requirements for Drone Assisted Stream Habitat (DASH) Protocol

Coordination between federal, state, private, and tribal agencies has resulted in development of remote stream monitoring techniques on several related, but independently developed tracts. Over the last year (and several years prior) we have focused on sharing knowledge and collaborating on hardware and processing solutions for initial drone utilization. Through communication and collaboration, the community has expressed several needs necessary for development to better aide the community in employing drone derived products. The major barriers to entry expressed through multiple channels are; lack of standardized protocol for habitat monitoring and drone setup and operation including standardized metadata recordings, lack of centralized database/repository for viewing and sharing previously collected imagery, lack of information and tools specific to processing imagery for stream-habitat mapping and modelling, among others.

While the goals of drone supported monitoring may vary between agencies, the core principle of collaboration and coordination amongst groups should remain intact. The first step identified within our collaborations is to develop and then apply a universal protocol for habitat data collection. The second step is to support the data collection with a centralized repository and toolkit for imagery processing. The third is to build a space for collaboration amongst the developer community to capitalize on the ever-evolving machine learning, deep-learning, artificial intelligence, and computer science fields. While we have strived to ensure that we stay ahead of the curve, things like autonomous vehicle development are continuing to push imaging technology forward at a very rapid pace.

Plans for summer 2021, for multiple agencies, likely include finalizing protocols and survey data collection forms, finalize drone deployment protocols, continue to streamline the data pipeline and QA/QC tasks, and continue development on imagery processing and post-processing tools. The goal of winter/spring would be to finalize a joint protocol between multiple agencies to ensure data integrity and availability and then deploy the finalized protocol in the summer of 2022. Eventually, we hope that a single, drone-supported protocol can be adopted and deployed across the entire Columbia Basin watershed. Drone supported habitat monitoring could then be utilized across the basin for a wide variety of problems including status and trend monitoring, prioritization, effectiveness monitoring, etc.

Biomark's collaboration with the fisheries research and drone community has culminated in helping to reboot and participate in the Pacific Northwest Aquatic Monitoring Partnership's (PNAMP) remote sensing forum. The goals of the forum are to provide expert knowledge, guidance, and a space for collaboration. Biomark hopes to add to the forum with expert knowledge, protocol development, processing tool development, and data hosting services/centralize imagery repository. The outreach within the program will also allow for the acceptance of a standardize or semi-standardized methodology across the basin and hopefully ensure basin-wide buy-in of remote monitoring. The goal of all remote monitoring support by either drones or satellite derived products will be to collect more data, with higher confidence, across a much great spatial domain, while removing observer bias and crew to crew variability.

CHAPTER 2. PPK GPS Workflow and Automated Post-Processing

INTRODUCTION

The development and commercialization of small Unmanned Aerial Systems (sUAS, commonly drones) has fostered significant interest in their application for environmental research and monitoring. Recent and ongoing advancements in areas such as battery technology, structure from motion analyses, component miniaturization, and sensor refinement make UASs an increasingly popular tool. UASs are now commonplace among a variety of industries including forestry (Dainelli et al. 2021), precision agriculture (Tsuoros et al. 2019), archaeology (Adamopoulos and Rinaudo 2020), oil and gas (Yu et al., 2019), urban planning (Noor et al. 2018), and environmental research and monitoring (Manfreda et al., 2018) to name a few. Drone applications in riverine environments more specifically have accelerated in development and over the last decade. Measuring flow and velocity via drone is now a reality (Fujita et al. 1997, Tauro et al. 2015, Detert et al. 2015), in addition to determining water levels and water surface elevations (Ferreira et al. 2017, Bandini et al. 2017), and mapping vegetation both terrestrial and aquatic (Flynn et al. 2014). This review highlights the advancements of UAS technology in the environmental research field specifically, with recommendations for broadscale, standardized applications.

i. UAS applications in environmental research and monitoring

Remote sensing addresses a host of obstructive factors by facilitating safer (in comparison to piloted aerial surveys; Sasse et al., 2003), and less expensive data collection across expansive regions, remote and hard-to-access sites, cryptic and migratory species, and species sensitive to observer bias; at varying levels of resolution (Jones et al., 2006; Anderson and Gaston, 2013; Whitehead et al., 2014; Chabot and Bird, 2015; Linchant et al., 2015; Xiang et al., 2019). Advances in remote sensing technologies, particularly the refinement of sensors such as hig-resolution multispectral cameras, are continually progressing the ability to manage environmental resources effectively and efficiently.

Given the need for timely, high-resolution, and multispectral imagery, UAS have proliferated in the field of environmental monitoring and research (Manfred et al., 2018). Drones, in particular, are ideal for dynamic sites that require frequent re-sampling (Mullerova et al., 2017b). For example, drone-based invasive species monitoring is a common method for precise, rapid assessments of vegetation health and make-up, information necessary for appropriate management (Calviño-Cancela et al., 2014; Hill et al., 2017; Mullerova et al., 2017a). Additionally, the increased resolution of low-altitude imagery poses significant benefits to certain applications such as capturing variations in vegetative health (Assmann et al., 2020), improved imagery classification (Yang et al., 2019), and population monitoring (Wang et al., 2019),among others. UAS also tend to outperform satellites in variable weather conditions, such as being able to capture imagery below cloud cover and at various times throughout the day. Van der Wal et al., 2012, demonstrated that the probability of obtaining usable imagery from UAS was at least double, if not nearly triple, that of satellite imagery.

ii. Advantages, challenges, and future directions

Achieving broadscale adoption of UAS imagery in environmental monitoring requires advancement on a number of challenges. First, sensor calibration and error are common factors that may distort raw data, and thus implications from monitoring efforts (Gauci et al., 2018; Jones et al., 2010). Tools to account for minute errors in georeferencing, pixel deformation, and radiometric calibration are becoming more widespread and accessible, though increased education regarding their effect on data (and thus,

interpretation) is critical to ensuring that practitioners are accounting for this error. Data storage, processing requirements, and overall complexity of analysis is another common challenge faced in UAS adoption (Zimudzi et al., 2019). Imagery analysis and, in particular, the automation of data extraction from imagery have come a long way; however, making tools accessible and understandable to researchers remains a significant barrier to entry (Weinstein, 2017). Perhaps the most pervasive challenge in relevant, broadscale application of UAS to environmental research and monitoring is a lack of standardized operating procedures at all steps of the scientific process. Despite the many literature reviews heralding the utility of UAS, there is currently no centralized framework for research studies involving drones (Singh and Frazier, 2018; Manfreda et al., 2019; Tmusic et al., 2020; Dainelli et al., 2021). In fact, most of the challenges described above could be addressed with a centralized workflow to guide practitioners in the accurate, repeatable collection of imagery. This also has the potential to solve the issue of fragmentation, thereby fostering collaboration and data sharing among disparate organizations.

The Columbia Basin, more specifically, has seen many research groups from local, state, and private entities utilizing UAVs for Salmon and steelhead conservation; however, no centralized and/or coordinated effort has been developed. For example, UAVs lack in comparison to the many developed repositories and centralized efforts that are observed in satellite imagery applications, partly due to the infancy of the technology, but also due to lack of funding (Manfreda et al. 2018). The Columbia Basin provides an ample opportunity to apply a satellite-like approach to UAV data management, tool development, standardization, and coordination. Advantages to a basin-wide approach would include, but not limited to, data permanence in near real-time (Manfreda et al. 2018), improved field safety, efficiency, and accessibility (Sasse 2003, Xiang et al. 2019), and data flexibility (Xiang et al. 2019). Cost savings, when compared to other techniques must also be considered and UAVs have proven to be one of the most efficient data collection instruments available (Whitehead et al. 2014) and although UAVs have many promising attributes, severe limitation still exists.

Major challenges in UAV applications primarily arise after data collection. Processing and metric generation can be highly impacted by sensor calibration and error characterization, both spatially and within the reflectance of the object being imaged. Understanding how to correct for spatial uncertainty will lead to a more widely understood technology and a broader adoption. Other challenges after data collection include image registration, correction, and reflectance calibration (most notably radiometric). To further the challenges, thus far there has been no centralized study design or drone specifications, as well as no repository to store image metadata, resulting in a fragmented effort across the basin. This review seeks to synthesize current trends and recommendations regarding (i) platform and sensor choice, (ii) camera settings and UAS control software, (iii) flight configuration, (iv) georeferencing, and (v) post processing automation. Further, we provide recommendations specific to research and monitoring efforts for Pacific salmonids and their habitat.

METHODS

A comprehensive literature review was conducted on the use of UAS in environmental monitoring and research. Challenges in broadscale applications were synthesized from existing literature reviews (Singh and Frazier, 2018; Manfreda et al., 2019; Tmusic et al., 2020; Dainelli et al., 2021) and further investigated. Provided for each perceived challenge are background, suggested resolutions from the literature, use case examples, and recommendations for specific application in riparian habitat research and monitoring.

We also committed time to testing and developing several image classification algorithms and an automated centerline generation tool applicable to both UAV and satellite imagery. Algorithm development included testing of both a pixel-based (random forest) and a deep learning, object-based approach for imagery post-

processing and metric development. We applied the deep learning model to a very complex and downfall ridden site in Grouse Creek, a tributary to the Secesh River in central Idaho. In addition, we utilized both multispectral and red, green, blue (RGB) imagery to test an automated classification algorithm which leverages a random forest model with regional components.

RESULTS

(i) Platform and sensor choice

Fixed wing and rotary are the main options for UAS platforms. Fixed wing UAS tend to be more energy efficient, thus able to achieve greater spatial coverage. However, they are unable to operate at low altitudes, slower velocities, or in hovering, which can cause increased data uncertainty, especially for multispectral sensors (Singh and Frazier, 2018). In contrast, rotary wing (or vertical take-off and landing, VTOL) UASs are more commonly observed throughout the literature, likely due to ease of operation and flight path flexibility (Pádua et al., 2017; Singh and Frazier, 2018). Aside from user considerations, the major tradeoff appears to be between spatial coverage, in which fixed wing UAS present a primary advantage, and resolution, in which rotary wing are more favorable.

Most commercially available, off the shelf UAS currently include a red-green-blue (RGB) complementary metal-oxide-semiconductor (CMOS) sensor, which has demonstrated applications for a variety of environmental analyses (see Manfreda et al., 2019 for a comprehensive list of UAS borne sensors and applications). The incorporation of multispectral (e.g., near-infrared, short wave infrared, thermal infrared) and hyperspectral (the collection of many narrower bands, typically between visible light and near infrared) sensors are becoming more common, particularly in environmental monitoring applications (Laliberte et al., 2011; Lu et al., 2017). For example, data from multispectral sensors have proven success in vegetation monitoring given the ability to evaluate normalized difference vegetation index (NDVI) and track variables such as growth, infection, and stress (Dash et al., 2018; Cardil et al., 2019; Stow et al., 2019). Hyperspectral imaging can be more challenging, given the necessary calibrations and corrections; however, hyperspectral sensors have been shown to perform with increased accuracy in estimating leaf carotenoid content in vineyards (Zarco-Tejada et al., 2013), bark beetle infestation (Näsi et al., 2015), and ground cover and vegetation discrimination (Mitchell et al., 2012). Advances in light detection and ranging (LiDAR) technology also make it a candidate for UAS application. LiDAR tends to produce more accurate three-dimensional data and has the added benefit of being able to penetrate water and the canopy (Tonina et al. 2019, Hyyppä et al., 2020). Advances in multispectral and hyperspectral imaging, along with Structure from Motion (SfM) techniques, allow for the generation of end products similar to LiDAR, where the main tradeoff is between cost and (particularly 3D) accuracy (Yin et al., 2019; D'Oliveira et al., 2020; Hyyppä et al., 2020).

Significant study design efforts should take place prior to platform selection. There are tradeoffs between each sensor and planning for error should be in place; multispectral, hyperspectral, and LiDAR data are particularly susceptible to errors in calibration. However, much work has been invested into multispectral imagery processing and calibration (see Micasense Altum Processing). In general, using multiple sensors in combination increases the accuracy of the end products (Dinlus et al., 2012; Komarek et al., 2018; Sankey et al., 2017); though, added steps will be necessary to spatiotemporally rectify the disparate bands. Platform and sensor choice fundamentally come down to the objectives of the study. Tmusic et al., 2020, collated a comparison of available options to consider during the study design phase (Table 1, Table 2).

Table 1: (Dis)advantage of different platforms. From Tmusic et al., 2020.

Platform	Advantages (+) and Disadvantages (-)	Flight Time/Coverage

Rotary-wing	+ flexibility and ease of use + stability + possibility for low flight heights and low speed + possibility to hover - lower area coverage - wind may affect the vehicle stability	Flight time typically 20–40 min Coverage 5–30 x 10^3 m^2 depending on flight altitude
Fixed-wing	+ capacity to cover larger areas + higher speed and reduced time of flight execution - take-off and landing require an experienced pilot - faster vehicle may have difficulties in mapping small objects or establish enough overlaps	Flight time up to hours Coverage e.g., >20 km^2 depending on flight altitude
Hybrid VTOL (Vertical Take O and Landing)	+ ability to hover, vertical take-off and landing + ability to cover larger areas - complex systems mechanically (i.e., tilting rotors or wings, mixed lifting and pushing motors)	Flight time up to hours, but usually less than fixed Wings Coverage x 10^6 m^2

Table 2: Type of sensors mounted on UAS and their possible applications. Adapted from Tmusic et al., 2020.

Sensor Type	Specifics	Main Applications
RGB	Optical	aerial photogrammetry, SfM-based 3D modeling, change detection, fluid flow tracking
Multispectral	Multiple wavelengths	vegetation mapping, water quality, classification studies
Hyperspectral	Overlapping, contiguous bands; analyzing the shape of the spectrum	vegetation mapping, plant physiology, plant phenotyping studies, water quality, minerals mapping, pest-detection
Thermal	Brightness surface temperature	thermography, plant stress, thermal inertia, soil water content, urban heat island mapping, water temperature, animal detection.
LiDAR	Surface structure	3D reconstruction, digital terrain mapping, canopy height models, plant structure, erosion studies

(ii) Camera settings

Regardless of sensor choice, a number of steps must be taken ensure that the data collected are usable and produce meaningful (i.e., accurate) results. Camera settings are particularly important for Structure from Motion (SfM) applications, such that pixel matching is suggested to be the primary factor in photogrammetry and post processing (Gruen, 2012). Pixel matching directly affects the photo alignment and, subsequently, tie point success and point cloud density. Matching pixels is primarily a function of resolution, illumination, and environmental complexity (Mosbrucker et al., 2017). In fact, increasing the accuracy of pixel matching has been demonstrated to increase point density by two orders of magnitude (Smith et al., 2016).

Camera technology has advanced significantly and many consumer grade, off the shelf cameras can achieve favorable SfM results. The two intrinsic factors thought to have the greatest effect on maximized data density are sensor resolution and dynamic range (Cao et al., 2010). Resolution is defined by the number of pixels (or more likely, megapixels, MP) that a sensor is capturing- effectively. The minimum suggested sensor resolution for SfM is 16MP, ensuring that the there is no anti-aliasing filter, which can cause unwanted artifacts (Mosbrucker et al., 2017). While resolution determines the quantity of data being

captured, dynamic range describes the quality of data in regards to luminance. High dynamic range sensors are able to sense a broader spectrum of pixel luminance, helping to prevent data loss in very bright or very dark areas; a minimum 14-bit non-linear analogue-to-digital convertor is recommended for SfM (Mosbrucker at al., 2017).

In addition to sensor settings, physical properties of the camera, including lens shape and shutter type, can have significant impacts on SfM accuracy. Lens dimensions directly affect the focal length and, thus, the optical properties in conjunction with the sensor size. Small focal lengths (wide-angle, fish-eye), increase field of view but are subject to increased radial distortion at the edges of the image (Mosbrucker et al., 2017). However, wide angle lenses can still be used in SfM and may be an ideal option for rapid monitoring that do not require as high-resolution (Zhang et al., 2019).

(iii) Flight configuration

In addition to pre-flight sensor choice and calibration, the mechanics of the flight will have significant effects on the imagery collected. Factors such as altitude, velocity, camera angle, image overlap, and time of day will all influence the type, quality, and end-use of the imagery. These factors are especially important when planning for Structure from Motion (SfM) analyses to produce products such as orthomosaics, dense clouds, digital elevation models (DEMs), and 3D models (Whitehead and Hugenholtz, 2014).

Flight altitude is a major determining factor of pixel resolution, where higher altitude flights will generate lower pixel ground sample distance (GSD), but cover a larger spatial extent. Altitude can also affect parallax variation, thereby influencing how objects are observed and measured (Johansen et al., 2018, Tu et al., 2018, Remondino et al., 2014). Altitude will also change the functional amount of overlap between images, which has been shown to decrease point cloud densities (i.e., resolution; Tu et al., 2020). Forward overlap is demonstrated to have the greatest effect on photo alignment and point cloud density, with a recommended threshold of 80% (Dandois et al., 2015; Tu et al., 2020). Similarly, side overlap has a suggested range of 70-80% to minimize negative effects on SfM products (Dandois et al., 2015; Tu et al., 2020).

In addition to ensuring sufficient overlap for adequate tie point and point cloud densities, flight pattern and camera angle can also significantly impact the resolution and accuracy of SfM products. A number of studies suggest that overlapping flight patterns may increase accuracy of SfM products (Figure 1; Gerke et al., 2016; Assmann et al., 2018). Similarly, capturing multiple camera angles (e.g., nadir, oblique, high angle, low angle) can improve imagery products, particularly in complex environments such as riverbanks and forests (Gerke et al., 2016; Rusnák et al., 2018; Manfreda et al., 2019; Martinez et al., 2020). However, care must be taken as the radiometric qualities may change as the drone changes heading, leading to varying illumination levels or "hot spots" (Stow et al., 2014; Tu et al., 2018).
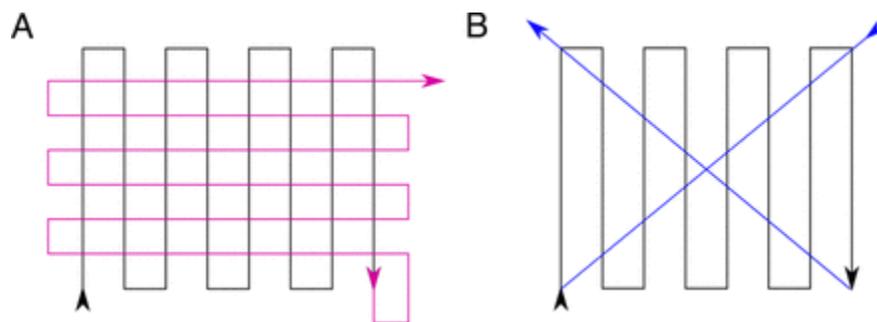
Figure 1: (a) Lawn-mower flight pattern (black) with perpendicular flight lines (pink) to achieve higher overlap and reduce BRDF effects when overlap is limited by aircraft or sensor triggering speed, and (b) Lawn-mover pattern flight path (black) with additional diagonal flight lines (blue) that may aid reconstruction. From Assmann et al., 2018.

Flight speed can also affect illumination and image quality, where faster speeds may increase motion blur (Roth et al., 2018), which is intrinsically linked to camera shutter type and speed. Increased flight velocities may also increase camera pitch angle for rotary wing aircraft lacking gimbal control (Tu et al., 2020). Though there is a tradeoff in resolution at very fast flight speeds, increased velocity has the benefit of increasing spatial coverage while maintaining similar environmental conditions. For example, solar elevation affects illumination and, thus, reflectance, which is critical for a number of analyses such as vegetative health and soil composition (Lu et al., 2017). Spectral information can also affect the success of machine learning processing, and extreme variations may pose a significant challenge for analyses and interpretation.

(iv) Georeferencing

Understanding the tradeoffs between efficiency and accuracy is essential to identifying the appropriate solution for your given application. Utilizing the internal GPS of the drone to geolocate initial image center points and processing via standard photogrammetry/stitching software such as *Agisoft Metashape* or *Drone Deploy* results in the most time efficient data collection, but the lowest spatial accuracy (errors ranging from 0.2-0.5m mean absolute error [Unger et al. 2018]). Studies have also shown that leveraging the internal GPS results in decreasing accuracy as you move from the center of the image outward to the margins.

Following utilization of the internal GPS solely, the next most efficient approach is leveraging a post-processing or PPK correction with spatial accuracies typically resulting in the sub-decimeter range (Hill 2018). Although PPK can be achieved via CORS base-stations, it is recommended that the base be within 100 miles of the survey site. To achieve the highest ppk accuracy, it is suggested to purchase a stand-alone base station to collect static measurements used in the post-processing of the imagery location information. This type of approach can be quite efficient, and only requires additional time post-processing as opposed to additional field time.

The most time consuming, but most accurate and precise method of drone surveying is to utilize ground control targets manually laid out across the survey site. Those ground control points/targets are then surveyed with a high resolution rtk gnss rover and base station for sub centimeter accuracy. The surveyed targets can then be leveraged to georeference the drone survey through common post-processing softwares, resulting in the most accurate drone survey possible, with mean absolute errors for elevation roughly ~0.04m (Cao et al. 2017). Cao et al. (2017) also illustrates how accuracy can be improved through a variety or combination of approaches including higher resolution GNSS receivers that only require satellite connectivity for sub ~0.3m accuracy.

(v) Post Processing Automation

A machine learning pipeline is under development, which uses PPK GPS imagery as an input data stream, to extract meaningful features for the automated generation of ecohydrological and aquatic habitat modeling metrics. Outcomes and elements of this process include: 1) automated classification of relevant topographical features from high resolution multispectral aerial imagery, 2) Computing feature scope (areas and proportions) in true dimensions (as opposed to pixels), 3) generation of the flow-axis (centerline profile) and waterway length, 4) computing wetted channel width, sinuosity, floodplain characteristics and cover at variable scales, 5) channel-unit distribution, and 6) identify large woody debris (LWD).

Several variants of machine learning algorithms were developed and tested. Among these were a super-pixel classifier, Mask Regional Convolutional Neural Networks (Mask-rCNN), random forests, as well as more traditional morphological methods and parametric transforms. Additionally, drone-based geo-referenced aerial imagery was projected and transformed to geographic coordinate reference systems in order to maintain measurements at real scales. All tools were developed with open source compatible freeware and/or open license (e.g. MIT, Apache) packages.

Due to our previously demonstrated successes and expert knowledge in object detection and classification, we first explored a Mask-rCNN for feature identification and extraction. The Mask-rCNN was found to produce successful results with high accuracy in classifying and delineating water channels. Water channel segments were consistently classified above 90% (from softmax output), but the Mask-rCNN was plagued with issues of overfitting the data in the presence of LWD (high type II errors) without an adequate abundance of data. Despite this shortcoming, Mask-rCNN remains a viable and promising alternative that requires more training time, both in terms of number of epochs (training time was 3 hours on an RTX 3500 12MB GPU) as well as time in labeling of truth data. This type of approach must also be supported not only by large volumes of data, but also data from a large variety of regions and geographies to avoid overfitting when trying to apply to a new spatial domain. For this reason, we began exploration of other classification types found within the literature.

Traditional morphological methods (intensity gradients, etc.) and parametric transforms also yielded adequate results. Segmentation based on intensity gradients, followed by parametric line transforms produced 94% accuracy in identifying LWD count in preliminary trials. It is anticipated that further processing (non-linear filtering, etc.) will be necessary in the presence of any dirt roads, to avoid misclassification of tire tacks as LWD (type I errors). Though not an overly complicated task, this (along with the ever menacing thresholds that are associated with gradients) was cause enough to turn to the exploration of further methodologies.

Finally, a super-pixel based classifier was explored which showed promising results, yielding preliminary accuracies for the water, roads and trees classes at above 96%, while that of low vegetation and dirt roads was 89% and 82% respectively.

Once progress toward aerial image classification reached favorable levels of accuracy and confidence, feature extraction for deriving aquatic habit metrics was explored. One primary metric from which many others can be fundamentally derived was approached first; this is the flow-axis (or center line) profile. Employing techniques used in signal and image processing (specifically, vector quantization), an optimization algorithm, constrained to the inequality constraint defined by centroidal boundaries, offers a method to identify the bisecting dividing line which between two irregularly shaped edges. This approach readily affords an accurate delineation of the water channel flow-axis. Once the flow axis is determined, it is a simple matter to determine the Full-bank channel width, sinuosity and length of various reaches.

Our initial application of various classification and detection algorithms have yielded promising results in challenging environments. Imagery utilized in the testing of post-processing techniques was collected during previous projects; spatial domain covering the upper Secesh River and its tributaries and the Lemhi River and its tributaries. In order to further development and training of models, it is necessary to foster data sharing and collaboration across the entire Columbia Watershed. This would require regional coordination and cooperation as well as a centralized repository and collaboration space for invested parties. Future development should focus not only on drone imagery utilization, but also maximization of satellite imagery and satellite derived products for fisheries conservation and watershed management.

DISCUSSION

Understanding the products, resolution, and precision/accuracy necessary to meet your project needs is critical when choosing a UAV + sensor package. For example, hydrodynamic modelling of wadeable streams requires centimeter level accuracy (Tonina et al. 2019), but canopy and riparian density mapping may only require sub meter accuracy (Pal et al. 2018), illustrating how spatial accuracy and resolution could determine the necessary payload to fulfill the requirements. Furthermore, the tradeoff between effort and product resolution is also a key consideration when planning future UAV sampling.

To further and more efficiently utilize UAVs across the Columbia Basin, several steps are recommended. The interconnectedness of the variables described in this review all point to the necessity of better metadata sharing and storage. A central repository/database and collaboration space may be necessary to capture the rapidly growing and changing drone and UAV technologies. This type of initiative could be modelled after several already developed satellite LiDAR consortiums. In addition, protocols and recommendations must continue to adapt to the changing technologies, but also maintain a level of consistency in data products to ensure compatibility across years. Through formal and informal collaboration with the Columbia River Iinter-Tribal Fish Commision (CRTFC) and the Washington Department of Fish and Wildlife (WDFW), the two protocols have been developed on separate tracks, with the hopes to work towards merging our initiatives in 2022.

LITERATURE CITED

Adamopoulos, E. and Rinaudo, F., 2020. UAS-based archaeological remote sensing: Review, meta-analysis and state-of-the-art. *Drones*, *4*(3), p.46.

Anderson, K. and Gaston, K.J., 2013. Lightweight unmanned aerial vehicles will revolutionize spatial ecology. *Frontiers in Ecology and the Environment*, *11*(3), pp.138-146.

Assmann, J.J., Kerby, J.T., Cunliffe, A.M. and Myers-Smith, I.H., 2018. Vegetation monitoring using multispectral sensors—Best practices and lessons learned from high latitudes. *Journal of Unmanned Vehicle Systems*, *7*(1), pp.54-75.

Assmann, J.J., Myers-Smith, I.H., Kerby, J.T., Cunliffe, A.M. and Daskalova, G.N., 2020. Drone data reveal heterogeneity in tundra greenness and phenology not captured by satellites. *Environmental Research Letters*, *15*(12), p.125002.

Avino, F.S. and Prata, G.A. 2020. Aboveground biomass estimation in Amazonian tropical forests: A comparison of aircraft-and gatoreye UAV-borne LIDAR data in the Chico mendes extractive reserve in Acre, Brazil. Remote Sens. 12, 1754.

Bandini, F., Jakobsen, J., Olesen, D., Reyna-Gutierrez, J.A. and Bauer-Gottwein, P., 2017. Measuring water level in rivers and lakes from lightweight Unmanned Aerial Vehicles. *Journal of Hydrology*, *548*, pp.237-250.

Calviño-Cancela, M., Mendez-Rial, R., Reguera-Salgado, J. and Martín-Herrero, J., 2014. Alien plant monitoring with ultralight airborne imaging spectroscopy. *PloS one*, *9*(7), p.e102381.

Cao, J., 2017, April. Block adjustment of zy-3 multi-temporal images without ground control points. In *2017 2nd International Conference on Frontiers of Sensors Technologies (ICFST)* (pp. 251-255). IEEE.

Cardil, A., Otsu, K., Pla, M., Silva, C.A. and Brotons, L., 2019. Quantifying pine processionary moth defoliation in a pine-oak mixed forest using unmanned aerial systems and multispectral imagery. *Plos one*, *14*(3), p.e0213027.

Chabot, D. and Bird, D.M., 2013. Small unmanned aircraft: precise and convenient new tools for surveying wetlands. *Journal of Unmanned Vehicle Systems*, *1*(01), pp.15-24.

D'Oliveira, M.V.N. Broadbent, E.N. Oliveira, L.C. Almeida, D.R.A. Papa, D.A. Ferreira, M.E. Zambrano, A.M.A., and Silva, C.A. 2017.

Dainelli, R., Toscano, P., Di Gennaro, S.F. and Matese, A. 2021. Recent Advances in Unmanned Aerial Vehicle Forest Remote Sensing—A Systematic Review. Part I: A General Framework. *Forests*, *12*(3), p.327.

Dandois, J.P., Olano, M. and Ellis, E.C., 2015. Optimal altitude, overlap, and weather conditions for computer vision UAV estimates of forest structure. *Remote Sensing*, *7*(10), pp.13895-13920.

Dash, J.P. Pearse, G.D. and Watt, M.S. 2018. UAV multispectral imagery can complement satellite data for monitoring forest health. Remote. Sens. 10, 1216.

Detert, M. and Weitbrecht, V., 2015. A low-cost airborne velocimetry system: proof of concept. *Journal of Hydraulic Research*, *53*(4), pp.532-539.

Ferreira, E., Chandler, J., Wackrow, R. and Shiono, K., 2017. Automated extraction of free surface topography using SfM-MVS photogrammetry. *Flow Measurement and Instrumentation*, *54*, pp.243-249.

Flynn, K.F. and Chapra, S.C., 2014. Remote sensing of submerged aquatic vegetation in a shallow non-turbid river using an unmanned aerial vehicle. *Remote Sensing*, *6*(12), pp.12815-12836.

Fujita, I., Muste, M. and Kruger, A., 1998. Large-scale particle image velocimetry for flow analysis in hydraulic engineering applications. Journal of hydraulic Research, 36(3), pp.397-414.

Gauci, A., Brodbeck, C., Poncet, A., and Knappenberger, T. 2018. Assessing the Geospatial Accuracy of Aerial Imagery Collected with Various UAS Platforms. Trans. ASABE 61, 1823–1829.

Gerke, M. and Przybilla, H.J., 2016. Accuracy analysis of photogrammetric UAV image blocks: Influence of onboard RTK-GNSS and cross flight patterns. *Photogrammetrie, Fernerkundung, Geoinformation (PFG)*, (1), pp.17-30.

Gruen A. 2012. Development and status of image matching in photogrammetry. The Photogrammetric Record 27(137): 36–57. DOI:10.1111/j.1477-9730.2011.00671.x.

Hill, A.C., 2019. Economical drone mapping for archaeology: Comparisons of efficiency and accuracy. Journal of Archaeological Science: Reports, 24, pp.80-91.

Hill, R. A., Wilson, A. K., George, M., and Hinsley, S. A. 2010. Mapping tree species in temperate deciduous woodland using time-series multi-spectral data. Applied Vegetation Science, 13, 86–99. doi:10.1111/j.1654-109X.2009.01053.x.

Hyyppä, E., Hyyppä, J., Hakala, T., Kukko, A., Wulder, M.A., White, J.C., Pyörälä, J., Yu, X., Wang, Y., Virtanen, J.P. and Pohjavirta, O., 2020. Under-canopy UAV laser scanning for accurate forest field measurements. *ISPRS Journal of Photogrammetry and Remote Sensing*, *164*, pp.41-60.

Johansen, K., Raharjo, T. and McCabe, M.F., 2018. Using multi-spectral UAV imagery to extract tree crop structural properties and assess pruning effects. *Remote Sensing*, *10*(6), p.854.

Jones, G.P, Pearlstine, L.G. and Percival, H.F., 2006. An assessment of small unmanned aerial vehicles for wildlife research. *Wildlife society bulletin*, *34*(3), pp.750-758.

Jones, H.; Vaughan, R. Chapter 621, Preparation and Manipulation of Optical Data, Image Correction, Geometric Correction. In Remote Sensing of Vegetation: Principles, Techniques, and Applications; Oxford University Press: Oxford, UK, 2010.

Komárek, J., Klouček, T., Prošek, J., 2018. The potential of Unmanned Aerial Systems: A tool towards precision classification of hard-to-distinguish vegetation types? Int J Appl Earth Obs. 71, 9–19. https://doi.org/10.1016/j.jag.2018.05.003.

Laliberte, A.S. Goforth, M.A. Steele, C.M., and Rango, A. 2011. Multispectral Remote Sensing from Unmanned Aircraft: Image ProcessingWorkflows and Applications for Rangeland Environments. Remote Sens. 3, 2529–2551.

Linchant, J., Lisein, J., Semeki, J., Lejeune, P. and Vermeulen, C., 2015. Are unmanned aircraft systems (UAS s) the future of wildlife monitoring? A review of accomplishments and challenges. *Mammal Review*, *45*(4), pp.239-252.

Lu, B. and He, Y. 2017. Species classification using Unmanned Aerial Vehicle (UAV)-acquired high spatial resolution imagery in a heterogeneous grassland. ISPRS J. Photogramm. 128, 73–85

Manfreda, S., Dvorak, P., Mullerova, J., Herban, S., Vuono, P., Arranz Justel, J.J. and Perks, M., 2019. Assessing the accuracy of digital surface models derived from optical imagery acquired with unmanned aerial systems. *Drones*, *3*(1), p.15.

Manfreda, S., McCabe, M.F., Miller, P.E., Lucas, R., Pajuelo Madrigal, V., Mallinis, G., Ben Dor, E., Helman, D., Estes, L., Ciraolo, G. and Müllerová, J., 2018. On the use of unmanned aerial systems for environmental monitoring. *Remote sensing*, *10*(4), p.641.

Martinez, J.G., Albeaino, G., Gheisari, M., Volkmann, W. and Alarcón, L.F., 2021. UAS Point Cloud Accuracy Assessment Using Structure from Motion–Based Photogrammetry and PPK Georeferencing Technique for Building Surveying Applications. *Journal of Computing in Civil Engineering*, *35*(1), p.05020004.

Mitchell, J.J., Glenn, N.F., Anderson, M.O., Hruska, R.C., Halford, A., Baun, C., and Nydegger, N. 2012. Unmanned aerial vehicle (UAV) hyperspectral remote sensing for dryland vegetation monitoring. In Proceedings of the 2012 4th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Shanghai, China, pp. 1–10.

Mosbrucker AR, Spicer KR, Major JJ, Saunders, DR, Christianson, TS, Kingsbury, CG. 2015. Digital Database of Channel Cross-section Surveys, Mount St. Helens, Washington, US Geological Survey Data Series 951. US Geological Survey: Reston, VA; 9 pp. DOI. 10.3133/ds951

Müllerová, J., Bartaloš, T., Brůna, J., Dvořák, P. and Vítková, M., 2017a. Unmanned aircraft in nature conservation: an example from plant invasions. *International Journal of Remote Sensing*, *38*(8-10), pp.2177-2198.

Müllerová, J., Brůna, J., Bartaloš, T., Dvořák, P., Vítková, M. and Pyšek, P., 2017b. Timing is important: Unmanned aircraft vs. satellite imagery in plant invasion monitoring. *Frontiers in plant science*, *8*, p.887.

Näsi, R., Honkavaara, E., Lyytikäinen-Saarenmaa, P., Blomqvist, M., Litkey, P., Hakala, T., Viljanen, N., Kantola, T., Tanhuanpää, T. and Holopainen, M., 2015. Using UAV-based photogrammetry and hyperspectral imaging for mapping bark beetle damage at tree-level. *Remote Sensing*, *7*(11), pp.15467-15493.

Nex, F. and Remondino, F., 2014. UAV for 3D mapping applications: a review. *Applied geomatics*, *6*(1), pp.1-15.

Noor, N.M., Abdullah, A. and Hashim, M., 2018. Remote sensing UAV/drones and its applications for urban areas: a review. In *IOP conference series: Earth and environmental science* (Vol. 169, No. 1, p. 012003). IOP Publishing.

Pádua, L., Vanko, J., Hruška, J., Adão, T., Sousa, J.J., Peres, E., and Morais, R. UAS, sensors, and data processing in agroforestry: A review towards practical applications. Int. J. Remote Sens. 38, 2349–2391.

Pal, S.C., Chakrabortty, R., Malik, S. and Das, B., 2018. Application of forest canopy density model for forest cover mapping using LISS-IV satellite data: a case study of Sali watershed, West Bengal. Modeling Earth Systems and Environment, 4(2), pp.853-865.

Roth, L., Hund, A., and Aasen, H. 2018. PhenoFly Planning Tool: Flight planning for high-resolution optical remote sensing with unmanned aerial systems. Plant Methods, 14, 116.

Rusnák, M., Sládek, J., Kidová, A. and Lehotský, M., 2018. Template for high-resolution river landscape mapping using UAV technology. *Measurement*, *115*, pp.139-151.

Sankey, T., Donager, J., McVay, J., and Sankey, J.B. 2018. UAV lidar and hyperspectral fusion for forest monitoring in the southwestern USA. Remote Sens. Environ. 195, 30–43.

Sasse, D.B., 2003. Job-related mortality of wildlife workers in the United States, 1937-2000. *Wildlife society bulletin*, pp.1015-1020.

Singh, K.K. and Frazier, A.E., 2018. A meta-analysis and review of unmanned aircraft system (UAS) imagery for terrestrial applications. *International Journal of Remote Sensing*, *39*(15-16), pp.5078-5098.

Smith, M.W., Carrivick, J.L. and Quincey, D.J., 2016. Structure from motion photogrammetry in physical geography. *Progress in Physical Geography*, *40*(2), pp.247-275.

Stow, D., Nichol, C.J., Wade, T., Assmann, J.J., Simpson, G., and Helfter, C. 2019. Illumination Geometry and Flying Height Influence Surface Reflectance and NDVI Derived from Multispectral UAS Imagery. Drones 3, 55.

Tauro, F., Pagano, C., Phamduy, P., Grimaldi, S. and Porfiri, M., 2015. Large-scale particle image velocimetry from an unmanned aerial vehicle. *IEEE/ASME Transactions on Mechatronics*, *20*(6), pp.3269-3275.

Tmušić, G., Manfreda, S., Aasen, H., James, M.R., Gonçalves, G., Ben-Dor, E., Brook, A., Polinova, M., Arranz, J.J., Mészáros, J. and Zhuang, R., 2020. Current practices in UAS-based environmental monitoring. *Remote Sensing*, *12*(6), p.1001.

Tonina, D., McKean, J.A., Benjankar, R.M., Wright, C.W., Goode, J.R., Chen, Q., Reeder, W.J., Carmichael, R.A. and Edmondson, M.R., 2019. Mapping river bathymetries: Evaluating topobathymetric LiDAR survey. Earth Surface Processes and Landforms, 44(2), pp.507-520.

Tsouros, D.C., Bibi, S. and Sarigiannidis, P.G., 2019. A review on UAV-based applications for precision agriculture. *Information*, *10*(11), p.349.

Tu, Y.H., Johansen, K., Phinn, S. and Robson, A., 2019. Measuring canopy structure and condition using multi-spectral UAS imagery in a horticultural environment. *Remote Sensing*, *11*(3), p.269.

Tu, Y.H., Phinn, S., Johansen, K. and Robson, A., 2018, July. Assessing Radiometric Corrections for UAS Multi-Spectral Imagery in Horticultural Environments. In *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium* (pp. 5449-5452). IEEE.

Tu, Y.H., Phinn, S., Johansen, K., Robson, A., and Wu, D. 2020. Optimising drone flight planning for measuring horticultural tree crop structure. ISPRS J. Photogramm. Remote Sens. 160, 83–96.

Unger, D., Hung, I., Kulhavy, D., Zhang, Y. and Busch-Peterson, K., 2018. Accuracy of unmanned aerial system (drone) height measurements.

Van der Wal, T., Abma, B., Viguria, A., Previnaire, E., Zarco-Tejada, P.J., Serruys, P., Van Valkengoed, E., Van der Voet, and P. Fieldcopter. 2013. Unmanned aerial systems for crop monitoring services. Precis. Agric. 13, 169–175.

Wang, D., Shao, Q. and Yue, H., 2019. Surveying wild animals from satellites, manned aircraft and unmanned aerial systems (UASs): A review. *Remote Sensing*, *11*(11), p.1308.

Weinstein, B.G., 2018. A computer vision for animal ecology. *Journal of Animal Ecology*, *87*(3), pp.533-545.

Whitehead, K. and Hugenholtz, C.H., 2014. Remote sensing of the environment with small unmanned aircraft systems (UASs), part 1: A review of progress and challenges. *Journal of Unmanned Vehicle Systems*, *2*(3), pp.69-85.

Xiang, T.Z., Xia, G.S. and Zhang, L., 2019. Mini-unmanned aerial vehicle-based remote sensing: techniques, applications, and prospects. *IEEE Geoscience and Remote Sensing Magazine*, *7*(3), pp.29-63.

Yang, B., Hawthorne, T.L., Torres, H. and Feinman, M., 2019. Using object-oriented classification for coastal management in the east central coast of Florida: A quantitative comparison between UAV, satellite, and aerial data. *Drones*, *3*(3), p.60.

Yin, D. and Wang, L. 2019. Individual mangrove tree measurement using UAV-based LiDAR data: Possibilities and challenges. Remote Sens. Environ. 223, 34–49.

Yu, L., Yang, E., Ren, P., Luo, C., Dobie, G., Gu, D. and Yan, X., 2019, September. Inspection robots in oil and gas industry: a review of current solutions and future trends. In *2019 25th International Conference on Automation and Computing (ICAC)* (pp. 1-6). IEEE.

Zarco-Tejada, P.J., Guillén-Climent, M.L., Hernández-Clemente, R., Catalina, A., González, M.R., and Martín, P. 2013. Estimating leaf carotenoid content in vineyards using high resolution hyperspectral imagery acquired from an unmanned aerial vehicle (UAV). Agric. For. Meteorol. 171–172, 281–294.

Zhang, H., Aldana-Jague, E., Clapuyt, F., Wilken, F., Vanacker, V. and Van Oost, K., 2019. Evaluating the Potential of PPK Direct Georeferencing for UAV-SfM Photogrammetry and Precise Topographic Mapping. *Earth Surf. Dyn. Discuss*, pp.1-34.

Zimudzi, E., Sanders, I., Rollings, N., and Omlin, C.W. 2019. Remote sensing of mangroves using unmanned aerial vehicles: Current state and future directions. J. Spat. Sci., 1–18.

CHAPTER 3. Evaluate Alternatives to Transfer Capacity Estimates from Quantile Random Forest Models to a Linear Stream Network

INTRODUCTION

The quantile random forest (QRF) capacity model we have developed uses paired fish abundance and detailed habitat data from selected sites around the Columbia River Basin to estimate the carrying capacity at the 200-500 meter reach scale, where such detailed habitat data is available. Initially, and to date, the sites where such data was collected were monitored by CHaMP (Columbia Habitat Monitoring Program). This aspect of the QRF model is useful for examining empirical fish/habitat relationships, determining what habitat factors may be limiting capacity at a particular location, and examining the improvement to capacity after rehabilitation actions. However, there is also a need to generate capacity estimates on larger spatial scales (e.g. tributary, watershed, population, etc.).

To date, inference to areas without detailed habitat data and at larger spatial scales has relied on master sample points and attributes associated with them (Larsen et al. 2016). This method was developed because that dataset was available at the time, and covered the entire Columbia River Basin. However, each master sample point does not actually represent a stretch of river, rather they are a single location (latitude/longitude coordinates) that is meant to be representative of about a kilometer of stream. Some points are closer together than a kilometer though, due to tributary junctions or other issues. Because of how they're constructed and what they actually represent, the interpretation of capacity estimates at master sample points is slightly more complicated than we desire.

Recently, a group of NOAA researchers has developed a line layer that breaks down the Columbia River Basin into 200 meter reaches, with various attributes assigned to each reach. It covers the same area as the master sample point dataset, but provides better interpretation and visualization properties than the master sample point layer. In this document, we present details about how we have used both the master sample points and the line network as extrapolation tools, and include some comparisons of the results between them.

METHODS

We examined results from a total of six different QRF models, three each for spring/summer Chinook and steelhead. These consisted of a model for redds, and two versions of a QRF model for summer juveniles. The first of these used what we considered the best choice of metrics from the entire CHaMP dataset. The second focused on metrics that are collected by DASH (Drone Assisted Stream Habitat protocol) that can be calculated from CHaMP data as well. This version allows direct QRF estimates to be made for areas sampled by DASH, since the CHaMP protocol is no longer in use.

Master Sample Points

The master sample points were generated in the design phase of CHaMP. These 551,046 sites were selected from the NHD Plus 1:100,000 stream layer covering WA, OR and ID at an average density of one site per kilometer (Larsen et al. 2016). Each CHaMP site where direct QRF capacity estimates were made corresponds to one of these master sample points, identified and selected using a generalized randomized tessellation stratification (GRTS) design (Olsen et al. 2012, Stevens Jr and Olsen 2004). CHaMP generated a number of attributes for each master sample point, referred to here as globally available attributes (GAAs)

because they are associated with every master sample point across all watersheds. We chose 11 to include in the extrapolation model (Table 3).

Table 3: Attributes available at every master sample point, used as covariates in extrapolation model.

ShortName	Name	Description
TRange	Temperature Range	Mean Temperature Range from PRISM data
Elev_M	Elevation	Elevation of site as extracted from the 10 m Digital Elevation Model
CHaMPsheds	CHaMP Watershed	CHaMP Watershed site falls in if appropriate
NatPrin1	Natural Class PCA 1	Natural Classification PCA 1 Score
DistPrin1	Disturbance Class PCA 1	Disturbance Classification PCA 1 Score
SrtCumDrn	Drainage Area (sqrt)	Square root of the cumulative drainage
StrmPwr	Stream Power	Stream Power
Slp_NHD_v1	Slope	Slope of Flowline (m/m) from the NHD Plus file
Channel_Type	Channel Type	Geomorphic Channel Type from Beechie Layer
WIDE_BF	Bankfull Width - modeled	Modeled bankfull width of stream, (m)
S2_02_11	Average August Temperature	NorWeST 10 year average August mean stream temperatures for 2002-2011

The original extrapolation model used the log of capacity estimates at each CHaMP site (fish / m) as the response, and selected GAAs as covariates. The model was fit using the svyglm function in the survey (Lumley 2004) package with R software (R Core Team 2019), accounting for the various survey design weights within each CHaMP watershed. We then used that model to predict capacity at every master sample point that was not a CHaMP site. In other words, we fit a linear regression to establish associations between estimated habitat capacity, from QRF, at CHaMP sites and globally available attributes from those sites and then used those associations at locations where CHaMP habitat data was not available to predict capacity at those master sample points.

The design weights were based on the particular stratification used in each CHaMP watershed to select monitoring sites. The most common stratification used three categories of valley segment type (source, transport and depositional) and selected a fixed number of sites from each strata. Because the strata are not equally distributed across the watershed, the design weights account for that unequal distribution. There are potential consequences to ignoring those weights when analyzing data from these sites (Nahorniak et al. 2015).

To roll up capacity estimates to larger spatial scales, the average predicted capacity of master sample points along a stream was multiplied by the length of that stream, and then combinations of streams could be added together to generate overall capacity estimates for a watershed.

Line Network

We adapted this method to using a stream layer created by Morgan Bond and Tyler Nodine at the Northwest Fisheries Science Center. This layer consisted of a line file divided into 200m reaches with various attributes attached to each reach. The line file is based on the National Hydrography Dataset High Resolution (NHDPlus HR) dataset, which has a higher resolution, 1:24,000, compared to the older layer that the master sample points were chosen from.

Table 4: Attributes available at every 200m reach, used as covariates in extrapolation model.

ShortName	Description
slope	Stream gradient
rel_slope	Relative slope. Reach slope minus upstream slope
Sinuosity	Reach sinuosity. 1=Straight, 1< sinuous
regime	Flow regime. 1= mixed, 2=snow dominated, 3=rain dominated.
alp_accum	Number of upstream cells in alpine terrain
fines_accu	Number of upstream cells in fine grain lithologies
flow_accum	Number of upstream DEM cells flowing into reach
grav_accum	Number of upstream cells in gravel producing lithologies
p_accum	Number of upstream cells weighted by average annual precipitation.
fp_cur	Current unmodified floodplain width
S2_02_11	NorWeST 10 year average August mean stream temperatures for 2002-2011
DistPrin1	Disturbance Classification PCA 1 Score
NatPrin1	Natural Classification PCA 1 Score
NatPrin2	Natural Classification PCA 2 Score

We determined which reach was closest to each CHaMP site, and used the predicted QRF capacity density of those CHaMP reaches as the response with the attributes attached to each 200m reach as covariates (Table 4). We also took this opportunity to move to a random forest modeling framework. This accommodates possible non-linear or saturating effects of some of these covariates on capacity predictions, and prevents the extrapolation model from predicting capacity values well above or well below the range of predictions at CHaMP sites.

Range of Covariates

We examined the range of the covariates used in each method, for wadeable streams, and compared it to the range of values found at CHaMP sites or reaches. This exercise provides some context about how representative the suite of CHaMP sites are compared to the rest of the Columbia River Basin. These figures are found at the end of this document.

RESULTS

We computed the total capacity of each species in each population using both methods, for summer juveniles (using both CHaMP and DASH habitat metrics) and redds, and compared them. The correlations between the two estimates are shown in Table 5.

Table 5: Correlation coefficient between capacity estimates at the population scale using each method.

Species	Model	r
Chinook	CHaMP	0.934
Chinook	DASH	0.903
Chinook	Redds	0.908
Steelhead	CHaMP	0.866
Steelhead	DASH	0.990
Steelhead	Redds	0.986

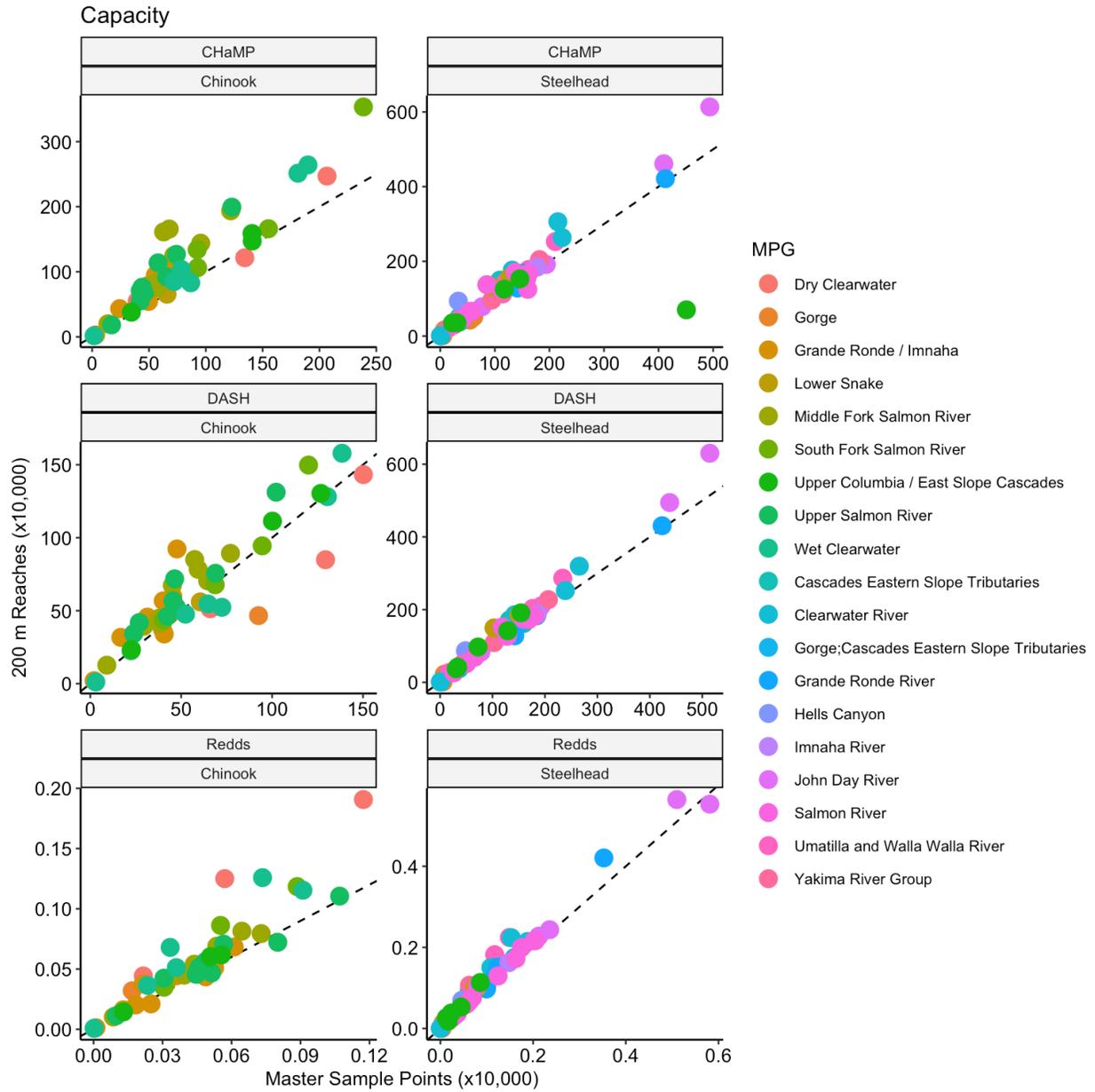We plotted one estimate against the other in Figure 2, and showed the relative difference in Figure 3.

Figure 2: Capacity estimates for each population, calculated with the master sample points method on the x-axis and the line network on the y-axis.
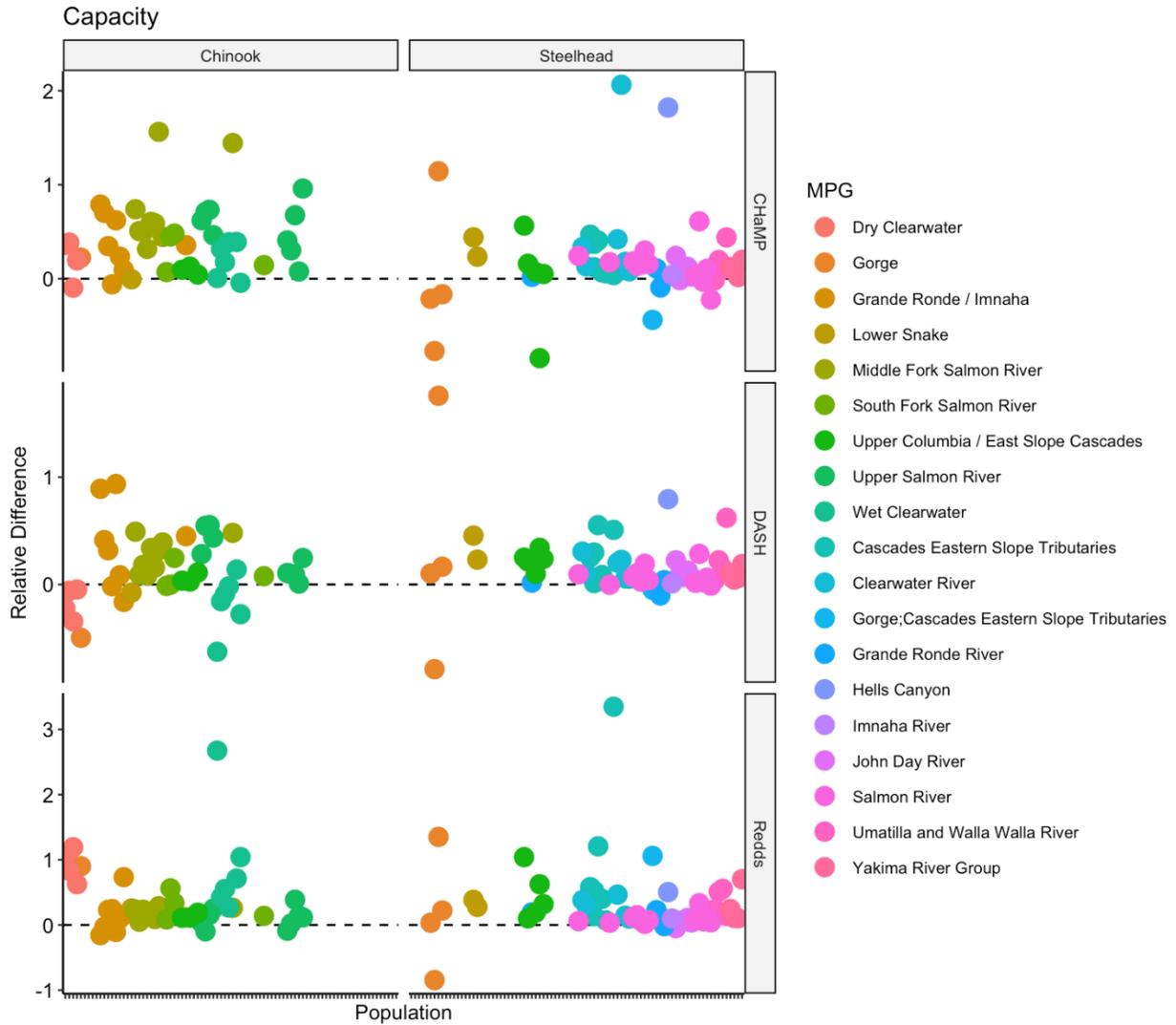
Figure 3: Relative difference between the capacity estimates for each population, using the master sample points method as the reference.

Maps

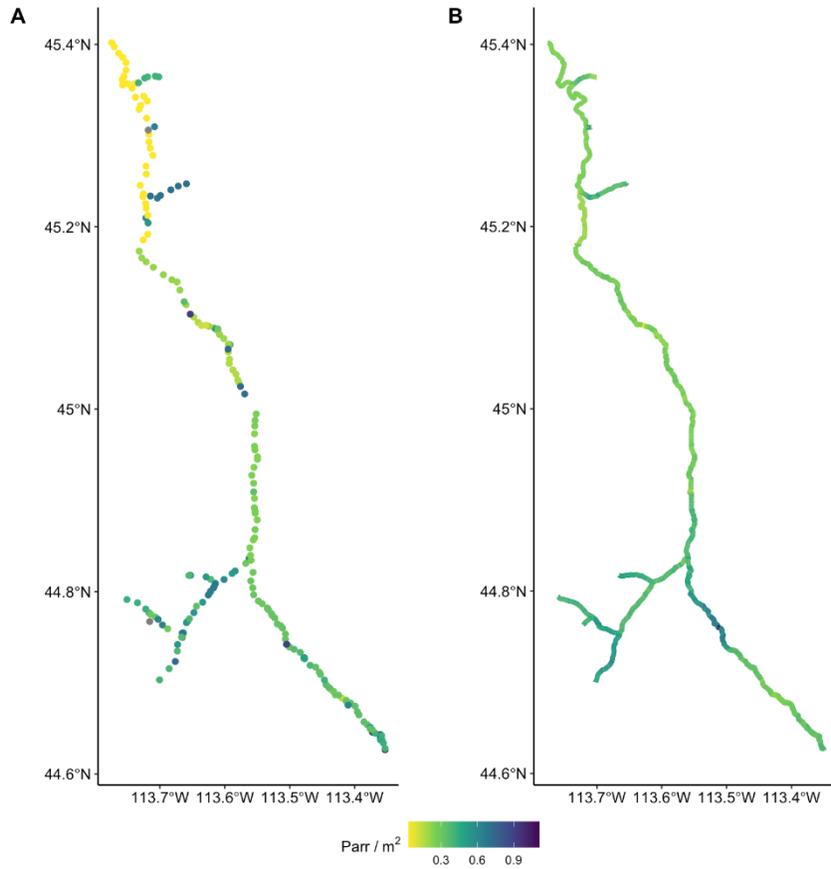This shows the difference in how the results can be visualized.

Figure 4: Plots of Chinook parr capacity in the Lemhi, using the master sample points method (A) and the 200 m reach method (B).
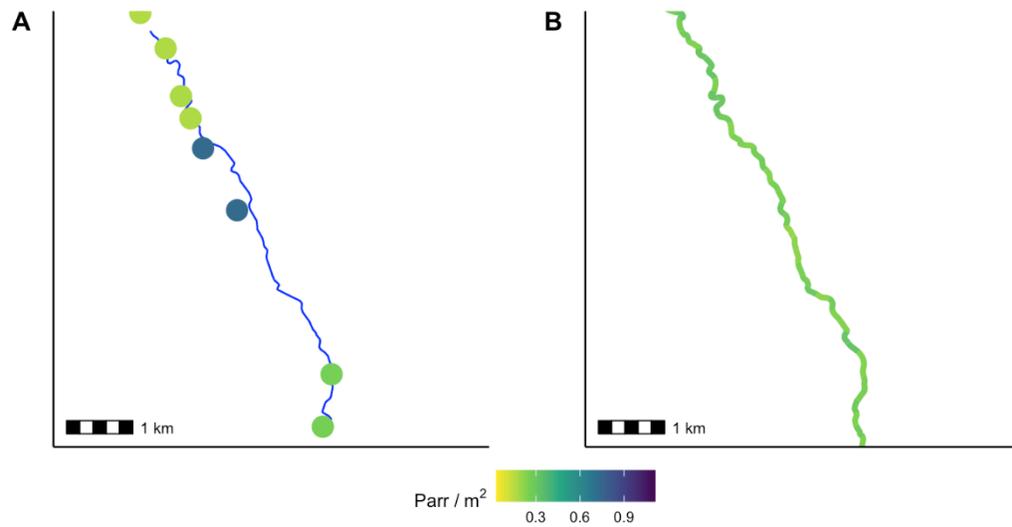

Figure 5: Plots of Chinook parr capacity in an approximately 8km stretch of the Lemhi, using the master sample points method (A) and the 200 m reach method (B). The NHDPlus layer has been added in (A).

DISCUSSION

Extrapolations of QRF predictions are useful for higher-level spatial analyses or comparisons, such as at the watershed level. Examining predictions at individual master sample points or 200m reaches should be discouraged. For that scale, detailed habitat data should be collected, by using a protocol like DASH, and direct estimates of capacity can be made using a QRF model. On the other hand, extrapolation summaries of capacity at the watershed scale, for various species and life-stages, can be useful in broad prioritization discussions, to determine what life-stages and watersheds to target for rehabilitation.

For most of the GAAs, the range of values represented at CHaMP sites or reaches overlapped with the range of values in other places, with a few exceptions. The most notable is modeled precipitation (Precip) in the Clearwater basin (Figure 6). We did not use Precip as a covariate in the master sample points extrapolation model, but it does indicate that something about the conditions in the Clearwater may be different from other places with the interior Columbia River Basin, and therefore extrapolations to that area should be scrutinized carefully. The 2nd PCA of the natural classification (NatPrin2) also shows some deviation from the CHaMP dataset in the Willamette, Lower Columbia and Salmon watersheds. It could be worth investigating what part of that PCA (or combination of parts) are driving that deviation.

For both species, across all three QRF models, the two extrapolation models resulted in estimates of total capacity at the population scale that are very highly correlated (Table 5). The linear network estimates were often greater than the master sample point estimates, to a greater or lesser degree, but not always (Figure 2).

Changing the modeling framework from linear regression to a random forest has several benefits. Primarily, it provides a method to constrain extrapolation predictions naturally, even when the extrapolation covariates are beyond the range found at CHaMP sites. In addition, random forests accommodate potential non-linear associations between capacity predictions and GAAs while handling correlations among GAAs. The sample size of CHaMP sites with QRF predictions of capacity is sufficient to fit a random forest model, so we have no concerns about the "data-hungry" nature of this framework for this situation.

Although the master sample point method has been used for several years, there is no reason to believe estimates from that method are inherently superior to using a line network, so even in the cases when the two models result in different estimates of capacity, it is difficult to say which is "better". On the other hand, there are several reasons to support using the line network method, apart from the actual results, primarily based on the ease of interpretation. Extrapolation to a line network involves capacity predictions at actual 200 m reaches along a stream network, while the master sample point method provides estimates at instantaneous "points" on the landscape. The summation of capacity to larger spatial scales is more straightforward when using a line network, and the maps that can be created are easier to interpret (Figure 4).

Therefore, we conclude that the extrapolation to a linear network method presented here is superior to the master sample point method, and should be adopted moving forward for examining QRF outputs at large spatial scales.

LITERATURE CITED

Larsen, D. P., C. J. Volk, D. L. Stevens Jr, A. R. Olsen, and C. E. Jordan. 2016. An overview of the Columbia Habitat Monitoring Program's (CHaMP) spatial-temporal design framework. South Fork Research.

Lumley, T. 2004. Analysis of complex survey samples. Journal of Statistical Software 9(1):1–19.

Nahorniak, M., D. P. Larsen, C. Volk, and C. E. Jordan. 2015. Using inverse probability bootstrap sampling to eliminate sample induced bias in model based analysis of unequal probability samples. Plos one 10(6):e0131765.

Olsen, A., T. Kincaid, and Q. Payton. 2012. Design and analysis of long-term ecological monitoring studies. Pages 126–150 in R. Gitzen, J. Millspaugh, A. Cooper, and D. Licht, editors. Cambridge University Press England, UK.

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Stevens Jr, D. L., and A. R. Olsen. 2004. Spatially balanced sampling of natural resources. Journal of the American Statistical Association 99(465):262–278.
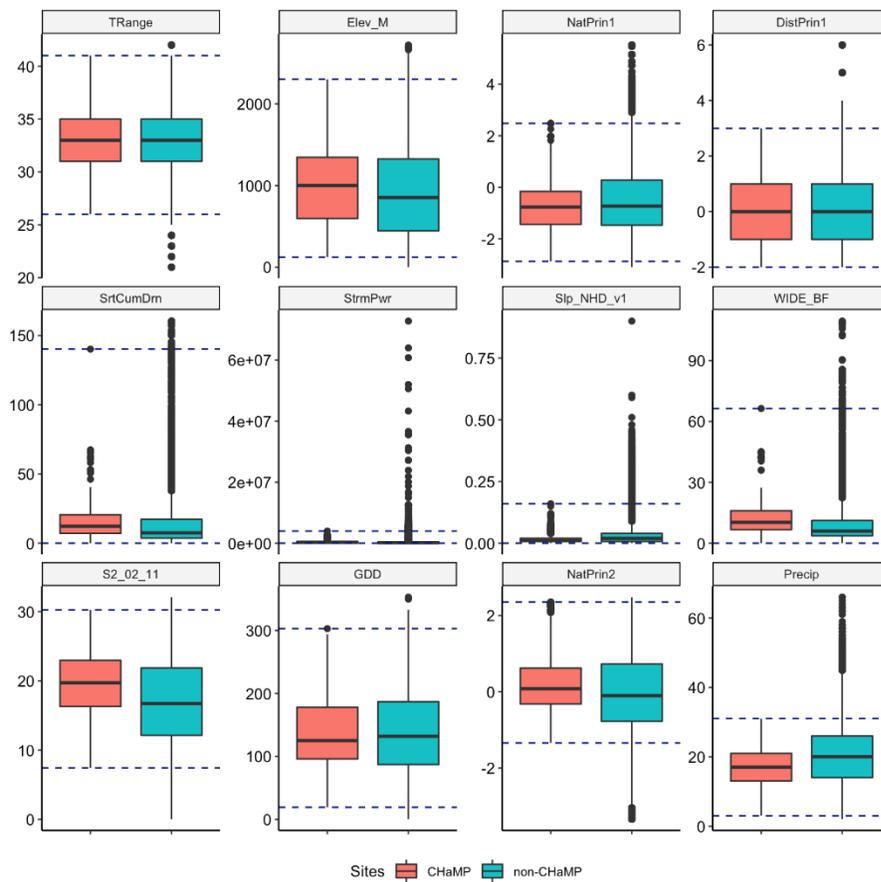
APPENDIX A: Covariate Range Figures

Figure 6: Boxplots of GAA values at CHaMP sites and non-CHaMP master sample points. Horizontal lines represent range of values at CHaMP sites.
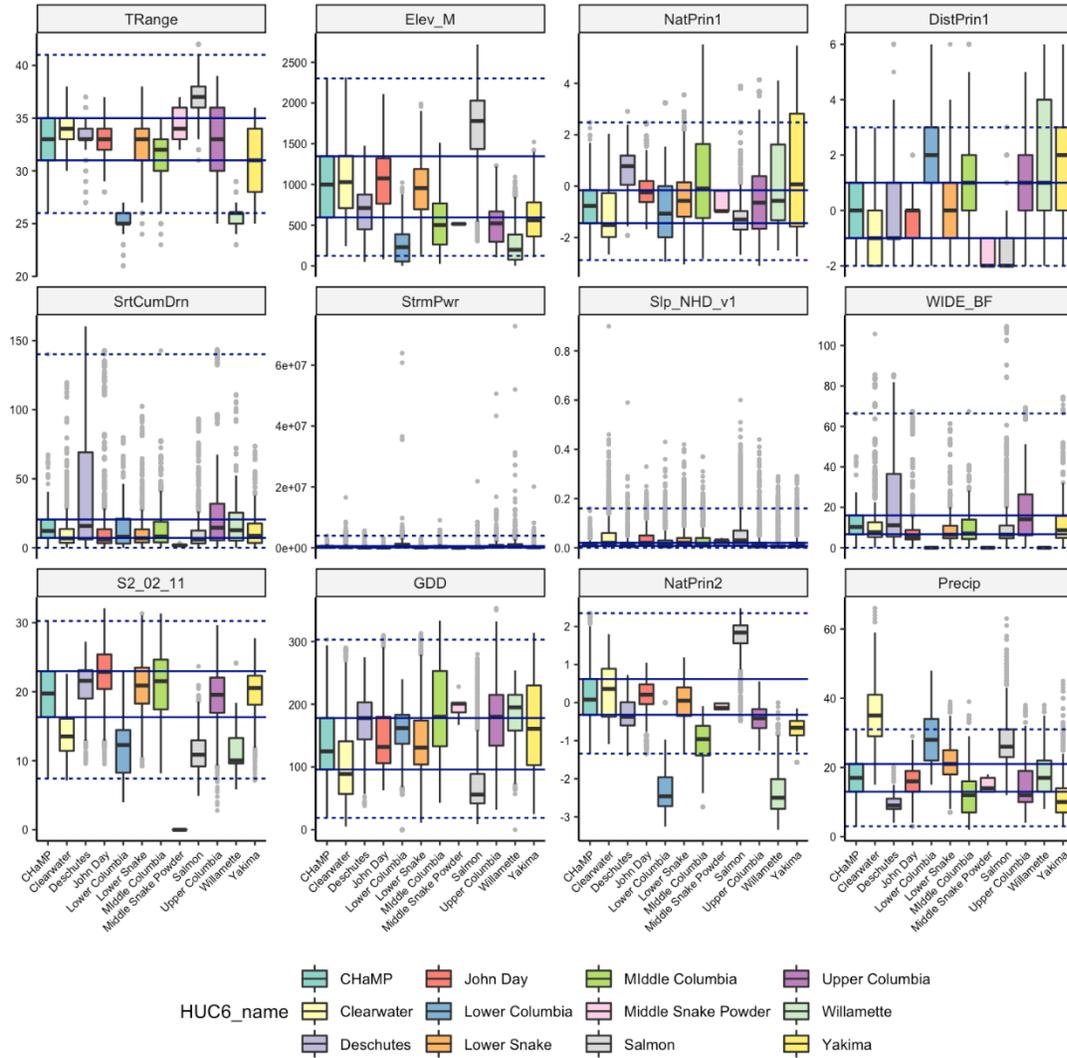
Figure 7: Boxplots of GAA values at CHaMP sites and non-CHaMP master sample points, colored by HUC6. Horizontal lines represent range of values at CHaMP sites (dashed) and the 25th and 75th quantiles of the CHamP sites (solid).
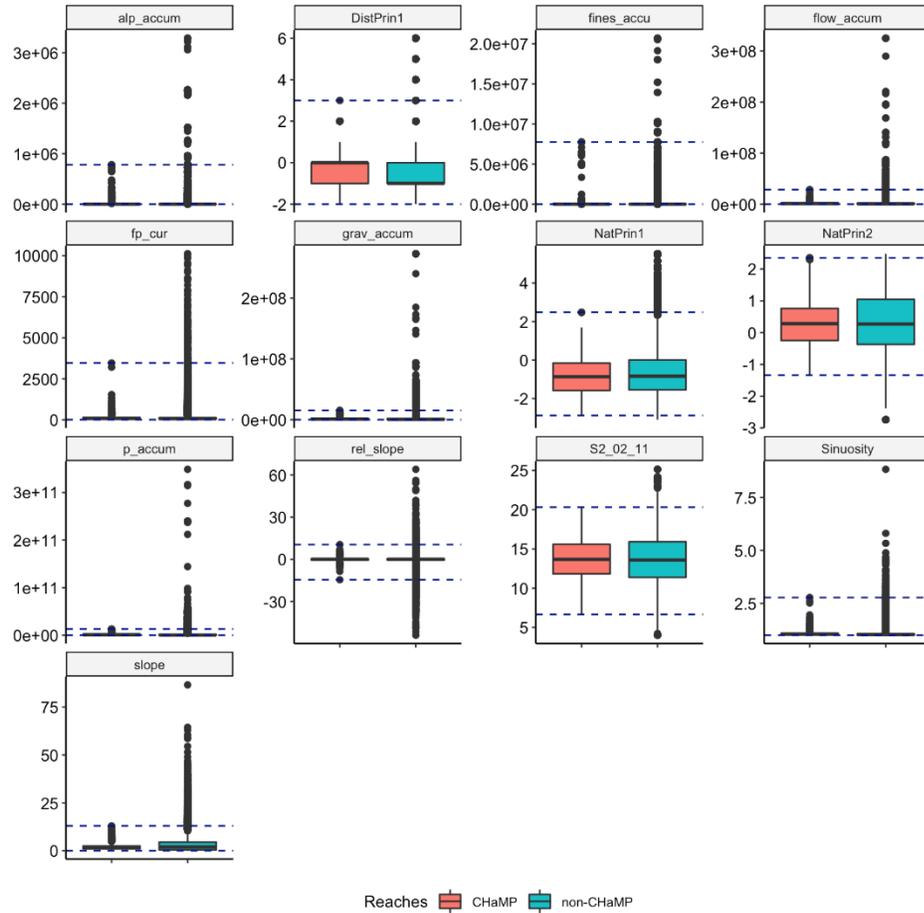
Figure 8: Boxplots of GAA values at CHaMP reaches and non-CHaMP 200 m reaches. Horizontal lines represent range of values at CHaMP sites.
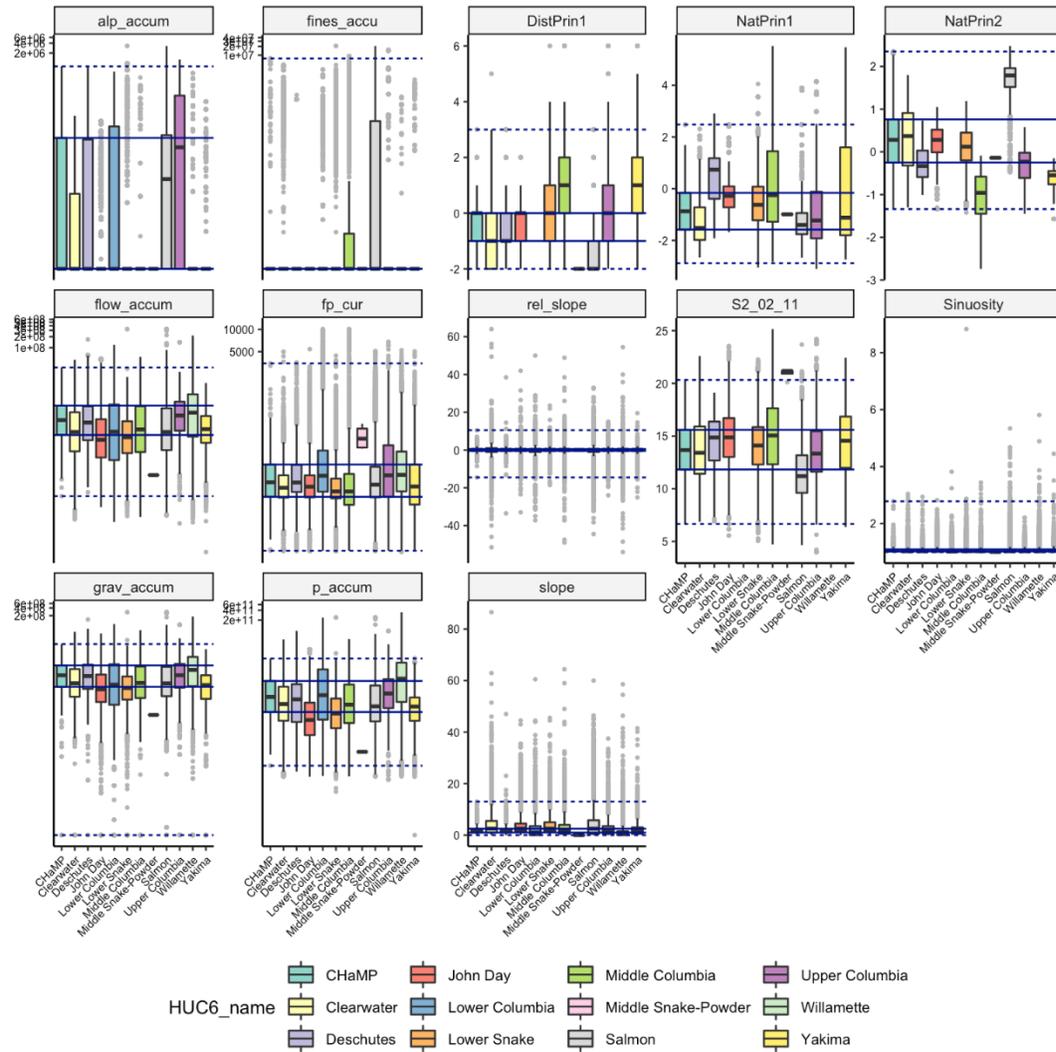
Figure 9: Boxplots of GAA values at CHaMP reaches and non-CHaMP 200 m reaches, colored by HUC6. Horizontal lines represent range of values at CHaMP sites (dashed) and the 25th and 75th quantiles of the CHamP sites (solid).

CHAPTER 4. Improve Flexibility and End-User Support for the Dam Adult Branch Occupancy Model (DABOM)

INTRODUCTION

The **D**am **A**dult **B**ranch **O**ccupancy **M**odel (DABOM) was initially developed as a proof-of-concept model to generate age and sex structured escapement of wild spring/summer Chinook and steelhead above Lower Granite Dam. Briefly, the methods consist of PIT tagging a representative sample of the returning fish, and using the subsequent detections of those tags at various sites upstream (or possibly downstream) of the starting location to estimate escapement or abundance past each detection site. The DABOM model estimates transition (or movement) probabilities past various detection sites while accounting for imperfect detection at those sites, essentially a multi-state variation of a spatial Cormack-Jolly-Seber model. The DABOM package implements this kind of model in a Bayesian framework. Further mathematical details of the model can be found in Waterhouse et al. 2020.

In recent years, the use of DABOM has expanded, and there are now several versions implemented in various locations around the Columbia River Basin, including:

- Wild spring/summer Chinook past Lower Granite Dam
- Wild steelhead past Lower Granite Dam
- Wild & hatchery steelhead past Priest Rapids Dam
- Wild & hatchery spring Chinook past Tumwater Dam
- Wild steelhead past Prosser Dam
- Wild & hatchery spring Chinook past Priest Rapids Dam
- Hatchery Coho past Priest Rapids Dam

As each of these versions has been developed, the initial code to make the model run has often been modified or added to as necessary in order to accommodate new users' needs. This has resulted in some redundant coding, multiple functions to accomplish the same purpose and unwieldy code that is difficult to update. On the other hand, new PIT tag detection sites are being added each year, sometimes with the express purpose of being incorporated into DABOM.

With multiple versions (meaning different release points and detection sites in each) exisiting, and the potential need to update each of them each year, our goal in this work element was to streamline the underlying code as much as possible and provide better instructions for running and updating DABOM models. To accomplish this, we have updated two different but related R software packages, named PITcleanr and DABOM. PITcleanr was developed to take observations from PTAGIS and transform them into inputs for DABOM, while the DABOM package contained the functions to run the DABOM model.

METHODS

PITcleanr

PITcleanr was originally written to take data from the Lower Granite trap database, and detections from PTAGIS and merge them into the format needed by the original DABOM model. In our efforts to update PITcleanr, we strove to make it useful for a broad range of potential analyses that utilize PIT tag data, not just DABOM.

DABOM

We rewrote many of the functions in the DABOM package to make them more generic, so that one function would take the place of 4 related but slightly different functions (each focused on the DABOM version for a particular release site). The types of functions we consolidated include:

- Write the JAGS model
- Setting initial values
- Creating inputs for JAGS
- Setting the number of branches at each branching node
- Determining which parameters to track

We aimed to make these work generically based on a user-defined configuration file that maps detections from specific antennas to "nodes" in the DABOM model, a parent-child table that defines which sites are upstream of other sites and a processed detection history from PITcleanr.

RESULTS

PITcleanr

We released PITcleanr 2.0.0 on GitHub. Instructions for installation, as well as three helpful vignettes to help users navigate the functionality of PITcleanr, can be found on this website: https://biomarkabs.github.io/PITcleanr/.

DABOM

We released DABOM 2.0.0 on GitHub. Instructions for installation, as well as a helpful vignette to help users navigate the functionality of DABOM, can be found on this website: https://biomarkabs.github.io/DABOM/.

DISCUSSION

These updates to the PITcleanr and DABOM software packages have improved them in many ways. PITcleanr has become more useful for all kinds of PIT-tag based analyses, including juvenile survival through the hydro-system, smolt-to-adult-return estimates, many varieties of mark-recapture type analyses for both anadromous and resident species, as well as preparing adult salmonid data for DABOM. Meanwhile, the DABOM software has become much more streamlined, able to be deployed in new watersheds, as well as making it easier to incorporate new sites into existing models in the future.

We developed extensive documentation for both packages to make them more accessible to more users and promote their utility within the Columbia River Basin. Our hope was by making many of the help pages and software vignettes available on a website for each package, even people who are not necessarily well-versed in R software could read more about their functionality and learn about their capabilities.

LITERATURE CITED

Waterhouse, Lynn, Jody White, Kevin See, and Andrew Murdoch. 2020. "A Bayesian Nested Patch Occupancy Model to Estimate Steelhead Movement and Abundance." *Ecological Applications* 30 (8). https://doi.org/10.1002/eap.2202.