

Contents

- Article*                    1        **Migration increases niche breadth in North American hummingbirds**  
Christopher J. Battey
- Article*                    11       **Fish community structure in a dynamic estuarine environmen**  
Brianna D. Pierce
- Article*                    22       **The spatial patterns and environmental constraints of antibiotic resistant bacteria in an urban river**  
Angela Klock
- Article*                    42       **Different words for the same thing: developing and testing composite vulnerability indices of fishing communities using multivariate techniques**  
J. Zachary Koehn

## Migration Increases Niche Breadth in North American Hummingbirds

Christopher J. Battey

*Department of Biology, University of Washington, Seattle, WA, USA*

*E-mail: [cjbattey@uw.edu](mailto:cjbattey@uw.edu)*

Received December 2015; accepted January 2016; published September 2016

### Abstract

Species capable of long-distance dispersal can deal with seasonal shifts in environment conditions with three basic strategies: individuals may stay in a single location and persist through annual variation in environmental conditions (“resident”); they can move to track a particular set of environmental conditions through the seasons (“niche tracking”); or they can move to take advantage of areas of niche space that are only available during certain times of year (“niche switching”). In order to assess the migratory strategies employed by North American hummingbirds, we used a principal components analysis and a kernel density estimator to quantify niche availability and occupancy across the breeding and nonbreeding ranges of 12 species. We found that all species of obligate migrant hummingbird would experience lower total niche breadth if they stayed on the wintering grounds rather than migrating to their observed breeding ranges. For most species analyzed, including all obligate migrants, breeding and nonbreeding niches are no more similar than expected by chance given the availability of environmental conditions in each season. Niche switching appears to be the dominant migratory strategy in Hummingbirds, suggesting that increased resource availability during the breeding season is the primary driver of migratory behaviors in the clade.

### Introduction

Seasonal shifts in climate and resource availability are an important feature of terrestrial ecosystems and play a major role in shaping species’ distributions. For species capable of long-distance dispersal, several strategies are available for dealing with changing environmental conditions: individuals can stay in a single location and persist through annual variation in environmental conditions; they can move to track a particular set of environmental conditions through the seasons; or they can move to take advantage of areas of niche space that are only available during certain times of year. These strategies can be summarized generally as “resident”, “niche tracking”, and “niche switching” (Nakazawa et al. 2004, Batalden et al. 2007).

In a biogeographic context, these strategies suggest different interpretations of the selective pressures leading to the evolution of migration. Most analyses in birds (Bell 2000, Zink 2011) suggest that migration evolved from natal dispersal and homing, with local movements gradually expanding into long-distance migration due to some combination of “push” (decreasing temperatures force individuals south during winter) and “pull” (lack of competition and abundance of resources in temperate latitudes during the spring and summer attracts southern species). If niche tracking is the dominant strategy driving migratory behavior in a species, migration can best be understood as evolving when conditions on the breeding ground become unsuitable for part of the year, forcing individuals to migrate south for the nonbreeding season (analogous to the “Northern Home” hypothesis, reviewed in Salewski & Burdener 2007, Winger et al. 2011).

Alternatively, a niche-switching strategy implies that species use migration to take advantage of ephemeral environmental conditions during the breeding season (e.g. “Southern Home”, reviewed in Salewski and Burderer 2007).

In addition, niche tracking and switching strategies require different analytical approaches and expectations when modeling species’ ranges under past or future climatic conditions. Recent efforts to forecast species’ ranges under future climate change scenarios have generally relied on correlative niche models built from occurrence data and environmental factors averaged across the whole year (Tingley et al. 2009), or seek to model only part of the annual cycle (National Audubon Society 2015). Biogeographic analyses seeking to describe the extent of niche conservatism in the diversification of species (McCormack et al. 2010) or model range shifts since the Last Glacial Maximum (Malpica & Ornelas 2014) also typically model a single niche for each species, based on average conditions experienced across the year in its range. Under a niche-tracking scenario, species ranges may be adequately represented by a single model, while niche-switching strategies indicate that separate models should be constructed for breeding and nonbreeding seasons. Niche switching species are also likely to have higher total annual niche breadth, and may be more able to adapt to changing conditions in the future.

A previous study of seasonal niches in North American songbirds (Nakazawa et al. 2004) found evidence for both migratory strategies, with some species appearing to occupy similar niches in breeding and nonbreeding seasons and others apparently switching climatic niche during different times of the year. Laube et al. (2015) recently expanded this framework by comparing breeding and nonbreeding niches occupied by Eurasian songbirds of the genus *Sylvia* with a hypothetical “resident” niche space that would be experienced by migratory species if individuals stayed on the breeding or wintering grounds year-round. This provides a natural null model for evaluating migratory strategies: if niche tracking is the primary strategy for migration we expect that the observed breeding and nonbreeding niches will overlap more than hypothetical “resident” and nonbreeding niches; if niche switching is the dominant strategy, we expect the opposite.

Here, we used the ordination approach proposed by Broennimann 2012 and extended by Laube et al. 2015 in migratory birds to evaluate seasonal changes in ecological niche occupancy in North American hummingbirds (Aves; Trochilidae), a family of small nectarivorous birds restricted to the New World, in which many species are reported to engage in annual movements tied to regional phenology (Rodewald et al. 2015). Our primary goal is to determine whether niche tracking or niche switching best explains hummingbird migration, and to assess the implications of our findings for the evolution of migration in the clade. Specifically, we ask: (1) are breeding and nonbreeding niches more similar than expected by chance, given the availability of environments in each season? And (2) would migratory species experience more or less niche overlap if they followed a resident strategy?

## Methods

### *Study species:*

We analyzed annual variation in the ecological niche space occupied by twelve species of hummingbird breeding in North and Central America (Table 1). We included all obligate long-distance migrant hummingbirds (*Selasphorus rufus*, *Selasphorus calliope*, *Archilochus alexandri*, and *Archilochus colubris*), five partially migratory species (*S. sasin*, *S. platycercus*, *C. anna*, *C. costae*, *C. Lucifer*), and three tropical montane species not reported to engage in large-scale annual movements (*A. heloisa*, *S. scintilla*, *S. flammula*).

All species included in the analysis are members of the “Bee” clade. This group includes all long-distance migratory hummingbirds breeding in temperate latitudes in North America. Biogeographic analyses strongly suggest that hummingbird diversification occurred primarily in South America (McGuire et al. 2014), with the Bee clade representing one of at least three colonizations of North America. Intriguingly, hummingbird fossils dated prior to the diversification of extant genera have also been found in northern Europe (Mayr 2004), suggesting that hummingbirds have a deep history of occupying high-latitude habitats. This biogeographic history, coupled with the

C.J. Battey, *Migration increases niche breadth in hummingbirds*

Table 1. Species included in the analysis. Breeding and nonbreeding months were identified as peak breeding and molting periods, respectively, in phenology charts given in Birds of North America (where available). \* Species reported as breeding in a very broad range throughout the year; Dec-Jan and Jun-Jul were used as proxy months to represent annual variation in environmental conditions.

Species	Migratory Status	Nonbreeding Months	Breeding Months
<i>Archilochus alexandri</i>	Obligate	Jan-Feb	May-Jun
<i>Archilochus colubris</i>	Obligate	Dec-Feb	Jun-Aug
<i>Atthis heloisa</i> *	Sedentary	Dec-Jan	Jun-Jul
<i>Calypte anna</i>	Partial	Aug-Oct	Dec-Feb
<i>Calypte costae</i>	Partial	Feb-Apr	Aug-Oct
<i>Calothorax lucifer</i>	Partial	Dec-Jan	Jun-Jul
<i>Selasphorus calliope</i>	Obligate	Dec-Jan	Jun-Jul
<i>Selasphorus flammula</i> *	Sedentary	Dec-Jan	Jun-Jul
<i>Selasphorus platycercus</i>	Partial	Nov-Jan	May-Jul
<i>Selasphorus sasin</i>	Partial	Oct-Dec	Mar-May
<i>Selasphorus rufus</i>	Obligate	Dec-Jan	May-Jun
<i>Selasphorus scintilla</i> *	Sedentary	Dec-Jan	Jun-Jul

abundance of migratory habits in the swifts (Apodidae), the sister family to hummingbirds (Jarvis et al. 2014), suggests that migration may be an ancestral trait for hummingbirds.

*Species occurrence and range data:*

Species occurrence reports were downloaded from the citizen-science database eBird (Sullivan et al. 2009; available at [www.ebird.org](http://www.ebird.org)). Digital range maps for all species of hummingbird were downloaded from Birdlife International (Birdlife International and NatureServe 2014). We trimmed each species' occurrence dataset to one point per 10arc-minute grid cell and subset occurrences by month to build separate datasets for breeding and nonbreeding seasons. Breeding and nonbreeding reports were cropped to include only occurrences within mapped ranges (to remove late- or early-migrating individuals), and 100 points per season were randomly selected for use in downstream analyses. When fewer than 100 reports were available, all available reports were used. All species had at least 15 occurrence points per season. For western species which are frequently reported as vagrants during the fall and winter in the eastern US (e.g. *S. rufus* and *S. sasin*), we also excluded all nonbreeding reports outside of Mexico or California, as these reports are outside the species' historic wintering ranges and are suspected to be driven by recent anthropogenic habitat alteration (Hill et al. 1998) rather than climate or phenology.

*Environmental data:*

We selected six environmental variables to represent the niche space occupied by each species. To represent annual variation in abiotic climatic conditions we extracted monthly mean values for temperature, total precipitation, daily temperature range, and number of frost days (days with low temperatures below 0°C) from the CRU CL 2.0 dataset (available at <http://www.cru.uea.ac.uk/cru/data/hrg/tmc/>).

Because hummingbirds are primarily nectivorous and are thought to time their movements to shifts in regional plant phenology (i.e. availability of blooming flowers), we calculated mean monthly Normalized Difference Vegetation Index (NDVI) from data produced by the Globalized Inventory Monitoring and Modeling Studies project (GIMMS; Pinzon & Tucker, 2014), using the “gimms” package in R (<https://cran.r-project.org/web/packages/gimms/gimms.pdf>).

Finally, to account for interspecific competition and biotic limitations on resource availability, we counted the number of co-occurring hummingbird species in each grid cell during each month of the year by overlaying species' range maps on a raster grid. All variables were processed and resampled to 10-arc-minute grid cells using the “raster” package in R (Hijmans & Etten, 2012).

*Breeding and nonbreeding months:*

To account for variation in the length and timing of breeding and nonbreeding seasons across species, we took the average of each of

environmental variable across peak breeding and wintering months (Table 1) for each species. Breeding months were identified through searches of published literature (Supp et al. 2015, Malpica et al. 2014, Rodewald et al. 2015). When data on molt was available, we classified the nonbreeding season as the peak months for flight and tail feather molt. Molt is an energetically intensive activity that occurs opposite the breeding season in most hummingbirds, suggesting that it is a suitable proxy for a stable nonbreeding season. For tropical resident species which are reported to breed throughout the year we used June-July as the breeding season and December-January as the nonbreeding season in order to represent a comparable temporal range of climatic conditions between migratory and sedentary species.

*Quantifying Niche Occupancy:*

We quantified species' occupancy of niche space by conducting a principal components analysis of environmental variables covering the full range of conditions in the study area during breeding and wintering seasons, following the methods described in Broennimann et al. 2012 and Laube et al. 2015. All variables were centered and scaled to equal weights prior to conducting the PCA. The first two PC axes were used to represent the niche space available to each species across the year. Values of environmental variables at species occurrence points during the relevant season (e.g. breeding vs. nonbreeding) were extracted and transformed to PC scores to visualize occupancy of niche space along the first two PC's. We also evaluated alternative migration strategies by extracting breeding-season environmental values for wintering occurrence points, and vice-versa. These data reflect a hypothetical "resident" strategy – the conditions individuals would experience if they stayed on either the breeding or nonbreeding grounds year-round.

To correct for spatial biases in reporting of occurrence data and allow for comparisons of niche space occupied in different seasons, we transformed point data for species occurrences to a smoothed density of environmental occupancy in niche space using a kernel density estimator. The niche space reflected by PC axes 1 and 2 was divided into a 100x100-cell grid with cell values  $v_{i,j}$ . A kernel density estimator was used to estimate the density of species

occurrences ( $\sigma_{i,j}$ ) and the density of available environments ( $\vartheta_{i,j}$ ), using the ecospat package (Broenniman et al. 2015, Broennimann et al. 2012) in R v. 3.1.3 (R Core Team, 2015). The occupancy of the environment ( $z_{i,j}$ ) was then calculated as follows (from Broennimann et al. 2012):

$$z_{i,j} = \frac{\sigma_{i,j}/\vartheta_{i,j}}{\max(\sigma/\vartheta)} \text{ if } \vartheta_{i,j} \neq 0$$

$z_{i,j}$  thus takes into account both the density of occurrence reports within a particular set of environmental conditions and the abundance of those conditions in the total available niche space. A  $z$  value of 0 indicates absence and 1 indicates presence.

*Comparing Niche Overlap and Similarity:*

To compare observed niche overlap with a hypothetical resident strategy, we calculated the  $D$  statistic (Schoener, 1970; Warren *et al.*, 2008) for (A) nonbreeding vs. breeding and (B) nonbreeding vs. "resident" occurrence densities.  $D$  is a measure of the overlap in niche space between two ordinations (or SDM predictions, in which case the test is typically applied in geographic rather than niche space). Specifically,  $D$  is equal to 1 minus the average difference in cell values between the gridded occurrence densities in different seasons:

$$D = 1 - \frac{1}{2} \left( \sum_{ij} |z_{1ij} - z_{2ij}| \right)$$

Values of  $D$  range from 0 (no overlap) to 1 (complete overlap).

We then compared mean values of (A) and (B) using paired  $t$ -tests applied separately to sedentary and migratory species. If niche switching is the primary migratory strategy in hummingbirds, we expect that  $D$  values for observed niches (A) will be greater than hypothetical resident strategies (B). If niche tracking is the dominant strategy, the opposite should be true.

To further test whether observed breeding and wintering season niches are more similar than expected by chance, we compared the observed niche overlap  $D$  between breeding and wintering range occurrences with a null distribution for  $D$  values between the wintering

range and a random set of points distributed across the niche space available in the breeding season (i.e. “Niche Similarity Test”, Warren et al 2008). We used 1000 randomizations to build our null model and repeated all analyses twice using different random subsets of occurrence points to avoid biases caused by randomly selecting a geographically constrained set of occurrences. If observed overlap values fell outside the 95% confidence intervals of the null distribution, we rejected the null hypothesis that niches were not more similar than expected given available environmental conditions.

In our analyses, we held the wintering range constant and simulated only breeding range points, rather than conducting the test in both directions. Extant hummingbirds have an unequivocally South American origin (McGuire 2014), and the recent glaciation across much of the breeding range of modern migratory species (i.e. during the LGM c. 21kya) suggests that the contemporary breeding ranges are unlikely to form the historic range core. More importantly, there is only one hummingbird species that occurs in temperate latitudes year-round (*C. anna*), and this species is a human associate currently undergoing dramatic northward range expansion, suggesting that persistence on the breeding grounds is, as a practical matter, not a viable strategy.

## Results

### *Niche quantification:*

PCA’s run on the full range of available environments in North America between breeding and nonbreeding seasons for each species resulted in similar variable weightings across species. The first two principal component axes represented an average of 76% of the total variance of the data across all species, though these axes did not explain more of the data than expected by chance under a “broken stick” model. Most PCA’s on monthly subsets placed NDVI, temperature, and frost days primarily along the first axis; and precipitation, number of co-occurring species, and daily temperature range primarily on the second axis (Figure 1).

Niche availability during breeding and nonbreeding seasons overlapped by an average of 84%. Total niche breadth was greater during

winter than summer (at least along the first two PC axes), reflecting the occurrence of low-temperature high-frost niche space only during the winter. Observed niches used by migrating hummingbirds during the breeding season was characterized primarily by variation along PC axis 1, reflecting shifts in temperature, NDVI, and frost days.

### *Niche overlap and similarity:*

Across all species included in the analysis, seasonal niche overlap was significantly greater for resident strategies than observed strategies ( $t = 3.1945$ ,  $df = 11$ ,  $p = 0.008537$ ; Figure 2). Migratory and partially migratory species also showed higher overlap for resident strategies ( $t = 4.0456$ ,  $df = 7$ ,  $p = 0.0049$ ). Sedentary species did not show significant differences in niche overlap between observed and resident strategies ( $t = 0.4491$ ,  $df = 3$ ,  $p = 0.6838$ ).

Breeding and nonbreeding niches were more similar than expected by chance in the two southernmost partial migrant species (*C. costae*, *C. lucifer*) and all sedentary species (Table 2,  $p < 0.02$ ). Niche space occupied during the breeding season was not significantly similar to the nonbreeding season in any obligate migratory species ( $p > 0.05$ ). All code for preparing data and conducting analyses is available on github (<https://github.com/cjbattey/nicheTrackR>).

## Discussion

### *Niche switching:*

Our analyses support a niche-switching model for migration in hummingbirds. All species of obligate migrant hummingbirds would experience lower total variation in environmental conditions if they stayed on the wintering grounds, as would four of five partial migrant species analyzed. Environmental conditions at breeding occurrence points are no more similar to those in the nonbreeding season than expected by chance in 7 of 12 species, including all obligate migrants and 3 of 5 partial migrants. The two partial migrant species which do appear to track their wintering niche during the breeding season (*C. costae* and *C. lucifer*) occur at the southern end of the distribution of migratory species, and occupy the majority of their total range year-round.

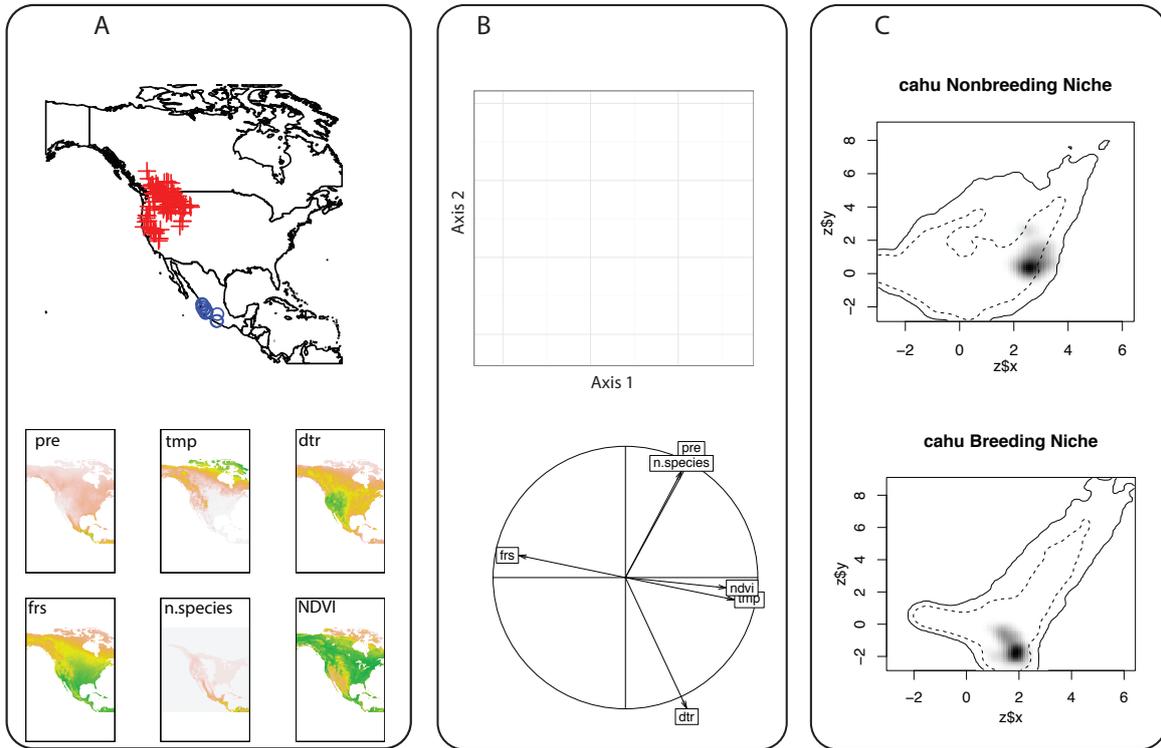


Figure 1. Niche quantification and comparison framework. (A) Species occurrence points for the breeding (red points) and nonbreeding (blue points) season are extracted from eBird reports and overlaid on a stack of raster grids. (B) A principal components analysis is used to decompose original variables into synthetic axes describing most of the variation in the data. Top: PCA results for *S. calliope* (red=breeding, blue=nonbreeding, green=resident), bottom: variable contributions to PC axes. (C) Point occupancy values for each season are transformed to a smoothed density of occurrences. Solid lines and dotted lines give 100% and 50% of the available environments in each season, respectively. Darker cells indicate regions of niche space with higher occupancy.

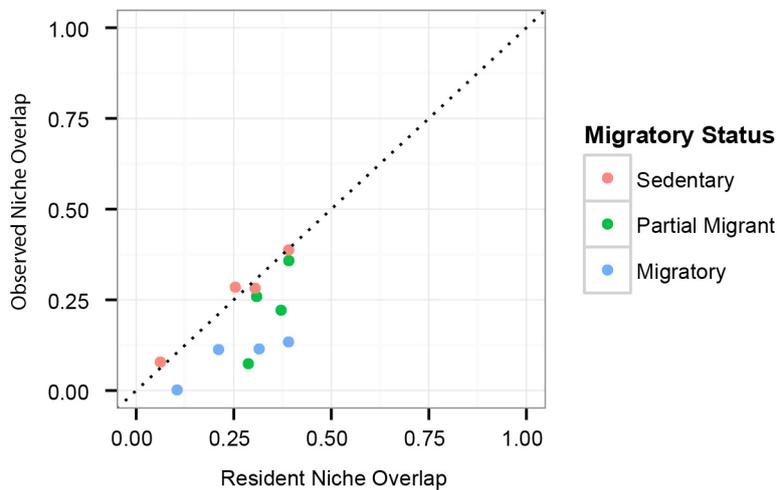


Figure 2. Observed overlap ( $D$ ) versus hypothetical “resident” niche overlap, by migratory status. Points below the dotted identity line indicate species which would experience greater niche overlap between breeding and nonbreeding seasons if they did not migrate.

Table 2. Niche overlap and similarity scores. D columns give degree of niche overlap between observed (nonbreed x breed) and hypothetical resident strategies (nonbreed x resident), on a 0 to 1 scale, with higher values indicating higher overlap. *P*(niche similarity) is the probability that the observed niche overlap is due to random chance – that is, it is no more similar to the nonbreeding niche than a random set of points in the breeding season. Values below 0.05 were accepted as significant.

Species	Migratory Status	<i>D</i> nonbreed x breed	<i>D</i> nonbreed x resident	<i>P</i> (niche similarity)
<i>Archilochus alexandri</i>	Obligate	0.115	0.315	0.871
<i>Archilochus colubris</i>	Obligate	0.113	0.211	0.100
<i>Atthis heloisa</i> *	Sedentary	0.285	0.254	0.020
<i>Calypte anna</i>	Partial	0.078	0.062	0.139
<i>Calypte costae</i>	Partial	0.391	0.358	0.100
<i>Calothorax lucifer</i>	Partial	0.308	0.259	0.019
<i>Selasphorus calliope</i>	Obligate	0.002	0.104	0.376
<i>Selasphorus flammula</i> *	Sedentary	0.391	0.387	0.020
<i>Selasphorus platycercus</i>	Partial	0.287	0.074	0.059
<i>Selasphorus sasin</i>	Partial	0.371	0.221	0.200
<i>Selasphorus rufus</i>	Obligate	0.390	0.134	0.415
<i>Selasphorus scintilla</i> *	Sedentary	0.304	0.282	0.002

Inspection of occurrence densities across seasons suggests that most migratory species occupy regions of niche space during the breeding season that are available during the nonbreeding season. This suggests that simple availability of conditions is not the primary factor driving our findings of niche switching. Breeding season niches are characterized primarily by higher frost days, lower competing species, and lower mean temperatures – all variables with predictable latitudinal gradients. Although these variables have much larger annual variance at temperate latitudes than in the subtropical wintering regions of most species analyzed here, our analysis suggests that the available annual variation is mostly within the range of conditions species occupy on the wintering grounds.

In contrast to Laube et al. 2015 and Nakazawa et al. 2004, both of which identified a mix of niche-switching and niche-tracking strategies in their study species, our results suggest that the few species of migratory hummingbirds which have evolved obligate migration all pursue a niche switching strategy. These differences could stem from real divergence in migratory strategies employed by different taxonomic groups, or could be the result of different analytic strategies and data types. Our study explicitly tested alternative migratory strategies and used randomization to build null models for niche space under a hypothesis of no relation between breeding and

winter niches, allowing us to reach stronger conclusions than Nakazawa et al. 2004, which relied on ad-hoc comparisons of niche model cross-predictions between observed wintering and breeding ranges. In addition, although the background environmental data was similar between our study and previous analyses, we used point occurrences rather than range map polygons to assess occupied niches in each season as in Laube et al. 2015. The improved spatial accuracy of point localities may allow finer resolution on niche quantifications, for instance reflecting altitudinal range restrictions within a broader breeding range made up of a matrix of suitable and unsuitable habitats. Finally, this is the first study the authors are aware of in which biotic interactions (in the form of number of co-occurring species) was included in a niche quantification to study seasonal changes in niche space.

#### *Utility of ordination approaches in niche quantification:*

Hutchinson's 1957 definition of the niche placed it in an explicitly multivariate context as the "N-dimensional hypervolume, enclosing the complete range of conditions under which an organism can reproduce itself". This concept moved the niche from a character spread in geographic space as Grinnell had generally envisioned – that is, the niche as a property of a *place* – to a trait intrinsic to each species. The conflict between these two conceptions of the

niche, and the associated biogeographic impact of splits between geographic and intrinsic niche space, is known as “Hutchinson’s Duality” (reviewed in Colwell & Rangel 2009). In modern analyses, this divide can be seen in the approaches of Species Distribution Modeling (SDM) versus multivariate ordination, with comparisons of niche space in SDM’s typically focused on cross predictions of occupancy probabilities in cell values spread across geographic space, and ordinations focused on comparing overlap between occupancy densities in niche space.

Multivariate methods which attempt to describe occupancy in a space bounded by multiple environmental variables are a natural fit for analyses of niche occupation and overlap that are consistent with Hutchinson’s original formulation (*if*, crucially, the underlying data can indeed reflect a species’ niche). Ordination methods have been shown to display less bias in tests of niche overlap and equivalency than SDM’s based on linear (GLM) or machine-learning (Maxent) approaches (Broennimann et al. 2012), suggesting that they should be employed over SDM approaches for most comparative studies. These methods are also substantially less computationally intensive, and to some degree can mitigate the “black box” problem of some existing SDM approaches, which build complex statistical models with relatively little user input and tuning. In favor of SDM methods, ordination methods do not allow simple ranking of variable contributions as is available in programs like MAXENT, and moving between niche and geographic space is more complicated in ordination approaches because at least two (rather than just a single) axis of predictions are produced.

#### *Conclusions and future directions*

Our results suggest several directions of further inquiry. First, the methods applied in this study should be extended to other species and clades to test the generality of our conclusions across taxonomic groups. In theory our methodological framework is applicable to all species engaging in annual movements for which sufficient point occurrences are available to adequately describe the niche space occupied in distinct seasons. In the case of hummingbirds our analysis was conducted across a small subset of species because only one clade of hummingbirds has evolved obligate long-

distance migration, and because occurrence data for most neotropical species is scarce. Sedentary species included in this analysis could be analyzed because they occur in countries with large numbers of birders who report occurrences to citizen-science databases (e.g. *S. flammula* and *S. scintilla* in Costa Rica and Panama), or are highly sought after species by birders (*A. heloisa*, the Bumblebee Hummingbird). Future analyses could utilize occurrence data from natural history museums to improve sampling across the nonbreeding range, or employ the range map approach as in Laube et al. 2015.

One major caveat for all efforts to quantify or model species’ niches is that the environmental variables selected may not accurately reflect elements of the niche driving selective pressures. Hutchinson’s conception of the niche theoretically includes all possible factors impacting population growth rates, with each factor independent of the others. This is an exceptionally high bar that has not been met by any existing analysis of the niche. In the case of hummingbirds, nectar availability likely has a strong impact on the seasonal distribution of species, but was not directly measured in our variables. The environmental variables we selected are expected to show strong correlations with nectar availability (i.e. nectar is more likely to be available in areas with high rather than low NDVI, or in seasons and areas with few frost days), but these measures are at best correlates of the factors impacting rates of population increase or decrease for the species. Including background data with clear biological relevance to the study species should improve our quantification of species’ use of niche space. For example, future analyses of hummingbird niche occupancy could include background layers describing the percent of plant species blooming during each month, or other site- and species-specific phenology data.

Hummingbird migration is a surprisingly poorly known phenomenon. Despite their status as among the most charismatic organisms occurring in temperate North America, we still know little about the routes or strategies used by migrating hummingbirds. Banding records provide little resolution of seasonal movements and circannual survival, as very few banded individuals are recaptured. No studies have yet examined the genetic bases of migration in hummingbirds, which has been conclusively

demonstrated in several species of songbird (Helbig 1991, Pulido 2014). Seasonal movements in neotropical species are widely reported (Rodewald et al. 2015), but no published study has quantified the frequency, season, or extent of cyclical movements in tropical hummingbirds. Given the paucity of available data, citizen-science programs and natural history museum records remain the best source of information on the routes and strategies employed by migrating hummingbirds. Here we utilized global environmental data and species occurrence points submitted by volunteer birders to demonstrate that migratory hummingbirds occupy distinct niches during breeding and wintering seasons, and do not appear to use migration to track a particular set of environmental conditions. Efforts to forecast species' ranges either under future climate change scenarios or in the geologic past should build separate models for breeding and wintering seasons whenever possible, and theoretical models of the evolution of migration in hummingbirds should account for the apparent use of a niche-switching strategy in extant migratory species.

#### Acknowledgements

The author would like to thank members of the Klicka Lab for advice on technical aspects of collecting and filtering data environmental and occurrence data. Analyses used in this paper are based on methods developed in Broennimann 2012 and extended to the case of migratory birds in Laube et al. 2015. Nakazawa et al. 2004 also served as direct inspiration for investigating niche tracking strategies in migratory species. In addition, the authors thank Dr. Xingli Giam and classmates in the fall 2015 FISH 560A class at the University of Washington.

#### References

- Batalden, R.V., Oberhauser K., and Peterson A.T. 2007. Ecological Niches in Sequential Generations of Eastern North American Monarch Butterflies (Lepidoptera: Danaidae): The Ecology of Migration and Likely Climate Change Implications. *Environmental Entomology* 36(6): 1365–73.
- Bell, C.P. 2000. Process in the Evolution of Bird Migration and Pattern in Avian Ecogeography. *Journal of Avian Biology* 31(2): 258–65.
- BirdLife International and NatureServe (2014) Bird Species Distribution Maps of the World.
- Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier L., Yoccoz, N.G., Thuiller, W., Fortin, M-J., Randin, C., Zimmerman, N.E., Graham, C.H., and Guisan, A. 2012. Measuring Ecological Niche Overlap from Occurrence and Spatial Environmental Data. *Global Ecology and Biogeography* 21(4): 481–97.
- Broennimann, O., Petitpierre, B., Randin, C., Engler, R.; Di Cola, V., Breiner, F., D'Amen, M., Pellissier, L., Pottier, J., Pio, D., Mateo, R.G., Hordijk, W., Dubuis, A., Scherrer, D., Salamin, N., and Guisan, A. 2015. ecospat: Spatial Ecology Miscellaneous Methods. R package version 1.1. <http://CRAN.R-project.org/package=ecospat>
- Colwell, R.K., and Rangel, T.F. 2009. Hutchinson's Duality: The Once and Future Niche. *Proceedings of the National Academy of Sciences* 106 (Supplement 2): 19651–58.
- Martínez-Meyer E., Peterson A.T., Navarro-Sigüenza A.G. 2004. Evolution of Seasonal Ecological Niches in the Passerina Buntings (Aves: Cardinalidae). *Proceedings of the Royal Society of London B: Biological Sciences* 271 (1544): 1151–57.
- Helbig, A.J. 1991. Inheritance of Migratory Direction in a Bird Species: A Cross-Breeding Experiment with SE- and SW-Migrating Blackcaps (*Sylvia atricapilla*). *Behavioral Ecology and Sociobiology* 28 (1): 9–12.
- Hijmans, R.J., and van Etten, J. 2012. raster: Geographic analysis and modeling with raster data. R package version 2.0-12. <http://CRAN.R-project.org/package=raster>
- Hill, G.E., Sargent, R.R., and Sargent, M.B. 1998. Recent Change in the Winter Distribution of Rufous Hummingbirds. *The Auk* 115 (1): 240–45.
- Hutchinson, G.E. 1957. Concluding remarks. *Cold Spring Harbor Symp Quant Biol* 22:415–427.
- Illán, J.G., Thomas, C.D., Jones, J.A., Wong, W-K., Shirley, S.M., and Betts, M.G. 2014. Precipitation and Winter Temperature

- Predict Long-Term Range-Scale Abundance Changes in Western North American Birds. *Global Change Biology* 20(11): 3351–64.
- Jarvis, E.D., Mirarab, S., Aberer, A.J., Li, B., Houde, P., Li, C., Ho, S.Y.W. et al. 2014. Whole-Genome Analyses Resolve Early Branches in the Tree of Life of Modern Birds. *Science* 346(6215): 1320–31.
- Joseph, L., and Stockwell, D. 2000. Temperature-Based Models of the Migration of Swainson's Flycatcher (*Myiarchus Swainsoni*) across South America: A New Use for Museum Specimens of Migratory Birds. *Proceedings of the Academy of Natural Sciences of Philadelphia* 150: 293–300.
- Laube, I., Graham, C.H., and Böhning-Gaese, K. 2015. Niche Availability in Space and Time: Migration in *Sylvia* Warblers. *Journal of Biogeography* 42(10): 1896–1906.
- Malpica, A., and Ornelas, J.F. 2014. "Postglacial Northward Expansion and Genetic Differentiation between Migratory and Sedentary Populations of the Broad-Tailed Hummingbird (*Selasphorus Platycercus*). *Molecular Ecology* 23(2): 435–52.
- Marini, M.A., Barbet-Massin, M., Martinez, J., Prestes, N.P., and Jiguet, F. 2010. Applying Ecological Niche Modelling to Plan Conservation Actions for the Red-Spectacled Amazon (*Amazona Pretrei*). *Biological Conservation* 143(1): 102–12.
- Mayr, G. 2004. Old World Fossil Record of Modern-Type Hummingbirds. *Science* 304(5672): 861–64.
- McCormack, J.E., Zellmer, A.J., and Knowles, L.L. 2010. Does Niche Divergence Accompany Allopatric Divergence In *Aphelocoma* Jays As Predicted Under Ecological Speciation?: Insights From Tests With Niche Models. *Evolution* 64(5): 1231–44.
- McGuire, J.A., Witt, C.C., Remsen Jr., J.V. , Corl, A., Rabosky, D.L., Altshuler, D.L., and Dudley, R.. 2014. Molecular Phylogenetics and the Diversification of Hummingbirds. *Current Biology* 24(8): 910–16.
- Nakazawa, Y., Peterson, A.T., Martínez-Meyer, E., and Navarro-Sigüenza, A.G.. 2004. Seasonal Niches of Nearctic-Neotropical Migratory Birds: Implications for the Evolution of Migration. *The Auk* 121(2): 610–18.
- Pinzon, J.E., and Tucker, C.J. 2014. A Non-Stationary 1981–2012 AVHRR NDVI3g Time Series. *Remote Sensing* 6(8): 6929.
- Pulido, F.. 2007. The Genetics and Evolution of Avian Migration. *BioScience* 57(2): 165–74.
- Salewski, V., and Bruderer, B. 2007. The Evolution of Bird Migration—a Synthesis. *Naturwissenschaften* 94(4): 268–79.
- R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Rodewald, P. (Editor). 2015. *The Birds of North America* Online: <http://bna.birds.cornell.edu/BNA/>. Cornell Laboratory of Ornithology, Ithaca, NY.
- Sullivan, B.L., Wood, C.L., Iliff, M.J., Bonney, R.E., Fink, D., and Kelling, S.. 2009. eBird: a citizen-based bird observation network in the biological sciences. *Biological Conservation* 142: 2282–2292.
- Supp, S.R., La Sorte, F.A., Cormier, T.A., Lim, M.C.W., Powers, D.R., Wethington, S.M., Goetz, S., and Graham, C.H.. 2015. Citizen-Science Data Provides New Insight into Annual and Seasonal Variation in Migration Patterns. *Ecosphere* 6(1): art15.
- Warren, D.L., Glor, R.E., and Turelli, M. 2008. Environmental Niche Equivalency Versus Conservatism: Quantitative Approaches To Niche Evolution. *Evolution* 62(11): 2868–83.
- Winger, B.M., Lovette, I.J., and Winkler, D.W. 2011. Ancestry and Evolution of Seasonal Migration in the Parulidae. *Proceedings of the Royal Society of London B: Biological Sciences* 279(1728): 610–18.
- Zink, R.M. 2011. The Evolution of Avian Migration. *Biological Journal of the Linnean Society* 104(2): 237–50.

## Fish community structure in a dynamic estuarine environment

Brianna D. Pierce

*School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA, USA*

*E-mail: [bdpierce@uw.edu](mailto:bdpierce@uw.edu)*

Received December 2015; accepted January 2016; published September 2016

### Abstract

Estuarine systems are highly productive and provide critical habitats for many culturally and economically valuable species. The estuarine environment is highly dynamic, creating variable habitats that are expected to influence the abundance and distribution of fish species, yet little is known about the factors that influence the estuarine fish community in Alaska. In the present study, I use a series of multivariate techniques to address the factors structuring the environment and fish community of the Anchor River estuary in southcentral Alaska. A principal components analysis indicates that tidal marsh channels have distinct yet considerably variable environmental conditions compared to mainstem estuarine channels. Nonmetric multidimensional scaling and principal coordinates analysis demonstrate that the fish communities also tend to be distinct across sampling sites. Pair-wise permutational multivariate analysis of variance confirms that these trends are significant ( $p < 0.05$ ) for all but the central site, which experiences considerable (but not significantly different) variation in relative species abundance. A Procrustes analysis and canonical analysis of principal coordinates reveal a significant relationship between environmental variables and relative species abundances ( $p < 0.01$ ). Understanding the factors that relate to fish community composition and distribution throughout the estuary is vital for the conservation and management of critical estuarine habitats.

### Introduction

The nonrandom way in which species are distributed across a landscape indicates that species preferentially occupy particular habitats (Fretwell and Lucas 1969). Preferable or optimal habitats provision more or superior resources and therefore provide organisms with a better chance of surviving and producing future generations (Rosenzweig 1981). However, humans are profoundly altering the natural environment at an unprecedented rate. In order to identify critical habitats and predict how organisms will respond to human-induced degradation or restoration, we must have a thorough understanding of the ways in which organisms select and move among differing habitats. Understanding patterns of habitat use and the resultant spatial and temporal distribution of species along environmental

gradients provides vital information for conservation and management decisions.

Habitat selection is “the process whereby individuals preferentially use, or occupy, a non-random set of available habitats” (Morris 2003). Animals select habitats that will directly or indirectly maximize their chances of surviving and producing offspring (Rosenzweig 1981). Individual fitness is influenced by the availability of resources, competition, predation, and the physiological state of the species. Therefore, we may expect to see distinct assemblage patterns emerge across a landscape as species seek necessary resources and their preferred environmental conditions while minimizing competition and predation risks. Furthermore, resource requirements often change over the course of an organism’s life, and patterns of habitat use correspondingly vary (Werner and Gilliam 1984). Habitat use is

therefore the result of multiple tradeoffs, and changes in habitat use patterns reveal how a species addresses these tradeoffs as the environment and needs of the species change.

Habitat selection has far reaching consequences on crucial processes such as species interactions and migrations, population regulation, community structuring, and the emergence and maintenance of biodiversity (Morris 2003). However, the complex and synergistic nature of ecological data makes it exceptionally difficult to analyze. Multivariate statistical methods have emerged as a means to explore the intricate ecological patterns and attempt to make causal inferences from this complex and highly correlated data (McGarigal et al. 2000). By identifying the underlying mechanisms governing species' distribution and environmental characteristics, we can develop conservation and management strategies that maximize the diversity of critical habitats such as estuaries, which provide essential habitat for many species.

Estuaries are highly productive systems that form where the river meets the sea. These dynamic systems support a robust food web and provide nursery habitat for many species including culturally and economically valuable animals such as salmon (Sibert 1979, Beck et al. 2001). Estuaries are physically structured by river flow and tidal flood, resulting in spatially and temporally complex habitat mosaics along environmental gradients that strongly influence the fish community (Elliot and Whitfield 2011). For juvenile salmon, estuaries provide quality feeding grounds, refuge from predators, and transition from fresh to salt water (Healey 1982, Thorpe 1994). Juvenile salmon that rear in the estuarine environment show increased growth rates compared to those that rear predominantly in freshwater habitats (Simenstad et al. 1982, Thorpe 1994, Craig et al. 2014). Because larger individuals have a better chance of surviving in the marine environment (Beamish and Mahnken 2001, Quinn 2005), increased growth in the estuary may confer survival benefits in later life. This may be crucial in northern latitudes where the growing season is shorter and salmon therefore have less time to attain a size critical for marine survival (Beamish and Mahnken 2001).

The diversity and quality of estuarine habitats is highly important to juvenile salmon. Chinook salmon survival to maturity is positively related to the proportion of the

estuary in pristine natural condition (Magnusson and Hilborn 2003). The diversity and quality of estuarine habitats also increases the number of juvenile salmon life history strategies expressed (Bottom et al. 2005, Jones et al. 2014, Craig et al. 2014). Diversity in life history patterns spreads the risk of mortality for a given population across space and time, thus increasing the stability and persistence of salmon populations (Schindler et al. 2010). There has been considerable research evaluating the importance of estuaries for juvenile salmon in the Pacific Northwest (e.g. Simenstad et al. 1982, Bottom et al. 2005, Quinn et al. 2013, Jones et al. 2014, Craig et al. 2014). However, comparatively little is known about estuary use by juvenile salmon in Alaska, despite the fact that salmon are of great cultural and economic importance in the region yet must face substantially different conditions than those in the Pacific Northwest. To address this knowledge gap, I am examining the habitat use patterns of juvenile salmon and the rest of the fish community in relation to environmental characteristics in the Anchor River estuary in southcentral Alaska.

*Objectives:*

Evaluate trends in environmental variables across sample sites.

Hypothesis 1: Marsh channels are distinct from each other and from mainstem sites.

Hypothesis 2: Marsh sites have greater variation in environmental variables than mainstem sites.

2) Explore fish community composition across sample sites.

Hypothesis 3: The fish assemblages are similar between the two marsh channels and distinct from the freshwater-dominated upper mainstem site.

Hypothesis 4: The middle mainstem site has significantly greater variation in relative species abundance compared to the other sites.

3) Investigate if environmental variables explain trends in fish community composition.

Hypothesis 5: Sites that are environmentally similar also have similar fish communities.

Hypothesis 6: Environmental variables explain relative fish community abundance.

## Methods

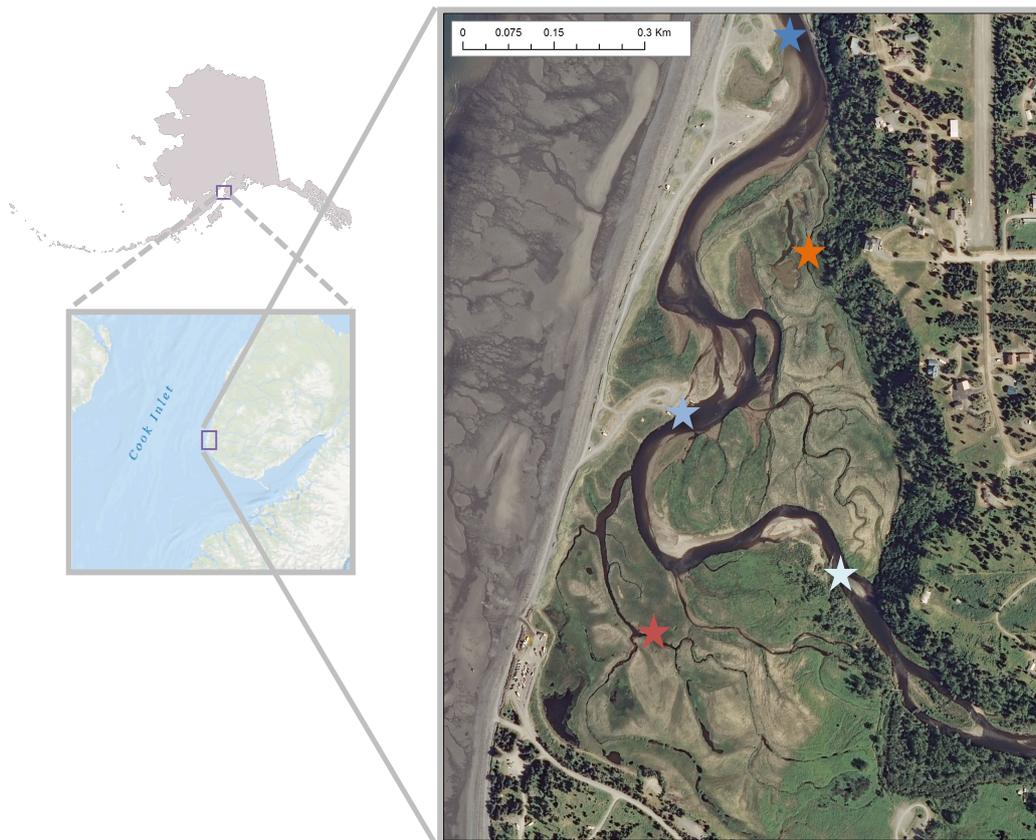
### *Study area:*

The Anchor River is a snowmelt and groundwater fed river located on the southwestern end of the Kenai Peninsula in southcentral Alaska (Figure 1). The Anchor River watershed encompasses over 580 km<sup>2</sup> and has 266 river km (rkm) accessible to anadromous fishes (Kerkvliet et al. 2013). At its mouth, the Anchor River forms a bar-built estuary approximately 8 rkm in length from the mean high tide mark to the confluence with Cook Inlet (Hoem Neher et al. 2013a). This region experiences a large (8 m) tidal range that creates diverse estuarine habitats including vegetated marshes and mudflats (Hoem Neher et al. 2013a,b).

### *Data collection:*

In the summer of 2015, we established five sites for repeat sampling in the Anchor River estuary.

Two sites were located in tidal marshes that differed in their groundwater input and connectivity to the mainstem and three sites were located along a salinity gradient in the mainstem from the upper to middle estuary (Figure 1). Sites were sampled approximately once per week from late-July to early-September resulting in 30 sampling events. Environmental data included maximum depth and temperature, salinity, and dissolved oxygen taken at three points in the water column (just below the surface, mid-water column, and just above the substrate). In marsh channels, we placed block nets (0.3 cm mesh) at both ends of the 25 m reach and sampled fish with a pole seine using three depletion passes. At mainstem sites, a pole seine (2.2 x 6 m, 0.3 cm mesh) was pulled 25 m parallel to the bank in the upstream direction. All fish were counted, identified to species, and return to the channel.



*Figure 1.* Map showing the location of the Anchor River estuary study area in southcentral Alaska and the location of the sampling sites: two tidal marsh channels identified as the Bluff site (orange) and Tractor2 site (red) and three mainstem sites reaching from the upper estuary (MainUp, light blue) through the middle of the estuary (MainMid, medium blue) to the downstream end of the estuary (MainDown, dark blue).

*Analyses:*

For each sampling event, environmental data was summarized by taking the mean and standard deviation of the three water column point measurements (taken at the surface, mid-column, and bottom). The standard deviation therefore represents the degree of stratification and is hereafter simply referred to as stratification. Because environmental data were measured on multiple scales, the data were column standardized (z-score transformation) prior to analysis. Species that were observed in less than 5% of samples were excluded from analyses as they could have undue influence on results. Young of the year fish (<20 mm fork length) were sometimes present in considerable abundance, yet have different ecological roles from conspecific adults and are difficult to identify to species. Therefore, I chose to analyze them as separate entities instead of combining them with the adult abundances. Because sampling effort was slightly different between mainstem and marsh sites, species abundance data were row-standardized to compute relative species abundance and then log-transformed because the data were highly right-skewed. Throughout this paper, the terms “fish community” and “fish assemblage” refer to relative species abundances (not presence/absence). All analyses were performed in R version 3.2.2 (R Core Development Team) using the vegan package (Oksanen et al. 2015).

I used a principal components analysis (PCA) to evaluate trends in environmental characteristics across the study area (Gower 1966). PCA is an unconstrained multivariate ordination technique, meaning that it assesses relationships among interdependent variables in a single data set. Environmental data is well suited for PCA because the variables are often highly correlated. PCA reduces redundancy in the data by forming new uncorrelated synthetic dimensions (principal components) that are weighted linear combinations of the original variables. PCA maximizes the amount of variation explained in each successive axis so that the majority of the information is condensed in the first few axes while the “noise” is pushed to the later axes. Thus, complex multidimensional datasets can be expressed by a few (often two or three) principal components with minimal loss of information. Ecological meaning can be interpreted by the weight of each original

variable on the principal component axes. Variables with larger weights are of higher importance in that dimension. The relative position of each sampling event on the synthetic gradients allows us to explore the relationships among samples. If the environment is relatively stable, we would expect all samples from a given site to lie in close proximity in ordination space. The relative position of samples from each of the five different sites will reveal the similarities (or dissimilarities) of the different habitats based on the environmental metrics we collected. A randomized permutation test with 100 iterations was used to evaluate the significance of the principal components.

I used nonmetric multidimensional scaling (NMDS) to investigate how the fish community was structured across sites (Kruskal 1964). Like PCA, NMDS is an unconstrained multivariate ordination technique. However, while PCA compares the Euclidean distance from the correlation matrix, NMDS is more flexible allowing for the use of any appropriate distance metric. Furthermore, NMDS preserves the rank ordering relationship among objects as opposed to the distance values themselves. This relaxes the implicit assumption of linear relationships among variables in favor of the more gentle assumption of monotonicity. Thus, NMDS is well suited for species abundance data where it is more desirable to preserve the relative differences in species abundances across sampling sites where heterogeneous habitats can result in variable sampling efficiency and thus tenuous absolute abundance estimates. Like PCA, NMDS condenses the information in the dataset into synthetic dimensions. However, in NMDS, the number of dimensions is specified *a priori*. The NMDS algorithm places samples in this reduced dimensional space so that their distance rank order is as similar as possible to their distance rank order in original multivariate space. The degree to which the original rankings align with the ordination rankings is referred to as “stress”. Lower stress indicates a better match between samples in NMDS ordination. I used a scree plot to determine the number of dimensions necessary to maximize the information contained in the axes while minimizing stress. A random permutation test with 100 iterations was used to evaluate the significance of the stress value. Like PCA, the NMDS ordination plot can be used to assess

trends in species composition across samples and sites.

When analyzing species abundance data, the percentage difference distance coefficient (Motyka 1950; frequently referred to as the Bray-Curtis distance coefficient) is most appropriate. It is specifically designed for count data and ignores shared species absences between sites. In other words, two sites that both lack a given species are not considered more similar. This is particularly important because niche theory (Hutchinson 1957) suggests that species exhibit a unimodal distribution along an environmental gradient, meaning that the species will be absent at opposing ends of the environmental gradient where sites are quite dissimilar. Therefore, I used the percentage difference distance metric prior to conducting the NMDS for species abundance.

Based on the pattern observed in NMDS (see results), I used permutational multivariate analysis of variance (perMANOVA) to test if samples were significantly different among sites (Anderson 2001). PerMANOVA partitions and analyses the sample sums of squares using any metric or semi-metric distance measure, such as the percentage difference (“Bray-Curtis”) distance. A permutational procedure is used to evaluate the significance using an F-test. Permutations of the raw data are used to create a null distribution against which the observed F-value is compared. I used 1000 iterations for all perMANOVA tests. To evaluate trends in fish community composition, I first computed a global perMANOVA, and then conducted pairwise perMANOVAs to explore site-by-site comparisons based on the relationships observed in the NMDS plot. PerMANOVA results may indicate significant differences based on either the location or variation of samples in multivariate space (or a combination of both). Therefore, I conducted a test of multivariate homogeneity of group dispersions to see if the variation in fish abundance was significantly different among sites (Anderson 2006). This method first uses principal coordinate analysis (PCoA), which is similar to PCA but can use any distance measurement. PCoA orients objects in multidimensional space according to the specified distance measure and then embeds them in Euclidean space. I used the percentage difference distance (“Bray-Curtis”) coefficient to compute the PCoA for relative

species abundance. The principal coordinate axes are then used to calculate the Euclidean distance of group members to the group centroid. Groups with higher dispersion will have a larger average distance to the centroid. The significance of these results is evaluated using analysis of variance.

I used a Procrustes analysis to assess whether samples that had similar environmental measurements also had similar fish community compositions (Gower 1971). This method attempts to find the best superimposition of two data matrices of the same object (or their ordination results) by rotating, scaling, and/or mirror reflecting one of the matrices about the other. The best superimposition minimizes the sum of squared distances between the corresponding objects. A permutation procedure known as PROTEST is then used to assess the statistical significance of the correlation between the two datasets resulting from Procrustean superimposition (Jackson 1995). The Procrustes analysis is flexible, allowing for the use of any distance measure. I used the percentage difference (“Bray-Curtis”) distance for species abundance and Euclidean distance for environmental variables prior to conducting a two-dimensional NMDS for each data set for use in the Procrustes analysis.

Based on the results of the Procrustes analysis (see results), I wished to see if the environmental variables could be used to explain the fish community composition within the estuary. To accomplish this, I used a constrained analysis of principal components (CAP; also known as distance-based redundancy analysis, Legendre and Anderson 1999). CAP belongs to a family of constrained (or canonical) ordination methods that are conceptually similar to linear regression. Linear combinations of the environmental variable matrix are formed in such a way as to minimize the variance in the species abundance matrix. Like perMANOVA, the first step of CAP is to conduct a PCoA on the percentage difference (“Bray-Curtis”) distance matrix of species abundance to embed the samples in Euclidean space in the composite dimensions. The next step in CAP is to perform a redundancy analysis, which is essentially a PCA where the ordination of samples in species space is constrained to linear combinations of the environmental variables. CAP results can be visually displayed in a triplot, which

simultaneously displays the samples, species, and environmental metrics and enables the interpretation of relationships between the two datasets. The relative position of plot elements indicates the degree of similarity. Sites near each other in CAP ordination space have similar fish assemblages, which tend to be dominated by the nearby species. The magnitude and direction of the arrows indicate the relative importance of each environmental variable and their contribution to each axis, respectively. Angles between arrows and other arrows (or sites or species) show the correlation between variables; a 90° angle means no correlation while angles near 0° or 180° indicate strong positive or negative correlation, respectively.

## Results

We sampled a total of 5,953 fish belonging to 15 species in the Anchor River estuary between July 22 and September 10, 2015. Two species (pink salmon and an unidentified sculpin species) were only observed once and were thus excluded from analysis. Coho salmon (*Oncorhynchus kisutch*), staghorn sculpin (*Leptocottus armatus*), threespine stickleback (*Gasterosteus aculeatus*), and starry flounder (*Platichthys stellatus*) were the dominant species accounting for 98% of all fish sampled.

The principal components analysis demonstrates that environmental variables vary across sample sites. Stratification in temperature, stratification in salinity, average salinity, and average dissolved oxygen are the driving variables of the primary principal component. Pearson product-moment correlations (structure coefficients) between each variable and the first two principal components are given in Table 1 along with the observed mean and range of each variable. The randomized permutation test indicates that the first principal component is significant ( $p < 0.001$ ), but the second axis (and all subsequent axes) did not significantly explain more variation than expected under the null distribution ( $p > 0.5$ ). Thus, when viewing the ordination plot in Figure 2, it is important to remember that the significant patterns reside on the x-axis and the y-axis merely aids visualization. As can be seen, stratification is a strong driver and is generally associated with the tidal marsh channel identified as Tractor2. The mainstem sites (indicated by varying shades

of blue) tend to be warmer, with higher dissolved oxygen and low salinity. The mainstem channels and marsh Bluff channel have fairly consistent environmental conditions while the Tractor2 marsh channel show considerable variability.

Nonmetric multidimensional scaling results indicate that fish community abundances show distinct trends across sites. The scree plot indicated that the majority of stress could be minimized in the first two axes. Thus, I performed a two dimensional NMDS (Figure 3). The resulting stress value of 0.15 indicates that the NMDS was fairly good at maintaining the rank order distances, and the permutation test indicates that the NMDS model fits the observed data better than expected by chance ( $p = 0.01$ ). The relationship between the original dissimilarities and the resultant Euclidean distances from the NMDS result in a nonmetric  $R^2$  of 0.979 and linear  $R^2$  of 0.887, which further supports the NMDS ordination. As can be seen in Figure 3, samples tend to group by site. Steelhead and coho drive the fish community in mainstem sites while young of year sculpin and sticklebacks dominate the fish community in the Tractor2 marsh site and starry flounder are most influential in the lower estuary mainstem site. The middle mainstem site has a highly variable fish community as observed by the spread of the sample points.

Permutational multivariate analysis of variance was used to test if the sites have significantly different fish community composition. The results of the global and pairwise perMANOVAs are presented in Table 2. All perMANOVA comparisons were significant except when comparing the fish community abundance of the middle mainstem site to the upstream and downstream mainstem sites. Because perMANOVA results can arise from either location or variation of samples in multivariate space, the test of multivariate homogeneity of group dispersions was used to compare within-site variation among sites. All comparisons were nonsignificant ( $p > 0.05$ ), indicating no difference in within-site variation in species abundance across sites.

The Procrustes analysis and PROTEST indicate there is a significant relationship between the environmental and species abundance datasets ( $p = 0.001$ ). Using canonical analysis of principal components (CAP), the environmental variables explain 51% of the

variation in the fish abundance data. The permutation test indicates the relationship is significant ( $p=0.001$ ), as are the first two CAP

axes ( $p=0.001$ ). The nature of the relationship between environmental variables and the fish

*Table 1.* Structure coefficients (Pearson product-moment correlations) between environmental variables and the first two axes of principal coordinate analysis as well as the observed mean and range of each variable.

Variable	Mean (range)	Structure coefficients	
		PC1	PC2
Temperature stratification	0.20 (0-2.3) °C	-0.485	-0.223
Salinity stratification	0.6 (0-7.5) ppt	-0.466	-0.124
Salinity	2.3 (0-17.7) ppt	-0.415	0.283
Dissolved oxygen	8.1 (1.5-12.57) mg/L	0.410	-0.284
Dissolved oxygen stratification	0.3 (0-1.6) mg/L	-0.379	-0.317
Temperature	12.9 (6.3-18.8) °C	0.249	-0.100
Depth	0.7 (0.3-1.55) m	0.034	-0.815

*Table 2.* Results from global and certain pairwise perMANOVAs comparing fish community abundance across sites.

Comparison	F-value	R <sup>2</sup>	p-value
Global	9.71	0.61	0.001*
Mainstem Upper vs Down	11.48	0.56	0.003*
Mainstem Upper vs Middle	1.55	0.13	0.167
Mainstem Down vs Middle	3.04	0.25	0.066
Bluff vs Tractor 2	14.33	0.57	0.002*
Bluff vs Mainstem Middle	6.92	0.38	0.002*
Bluff vs Mainstem Down	7.36	0.42	0.004*
Tractor2 vs Mainstem Middle	8.24	0.45	0.003*

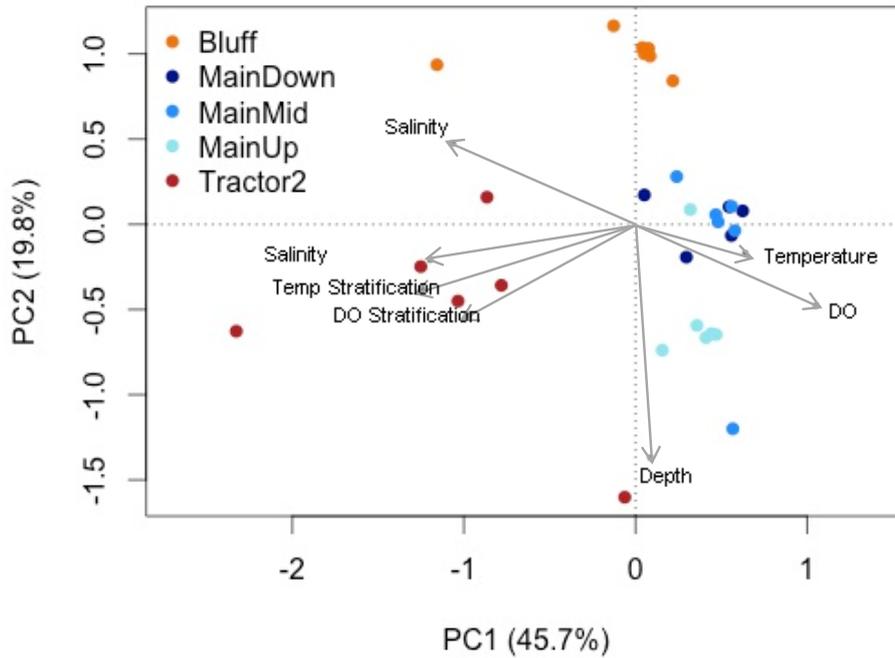


Figure 2. Principal component analysis of environmental variables. Points indicate sampling events and are colored by site. Orange and red indicate tidal marsh sites and blues represent mainstem sites. Arrow magnitude and direction indicate the importance and variable loadings on the principal component axes. PC1 is significant ( $p < 0.001$ ), but PC2 is not ( $p > 0.5$ ). DO = dissolved oxygen.

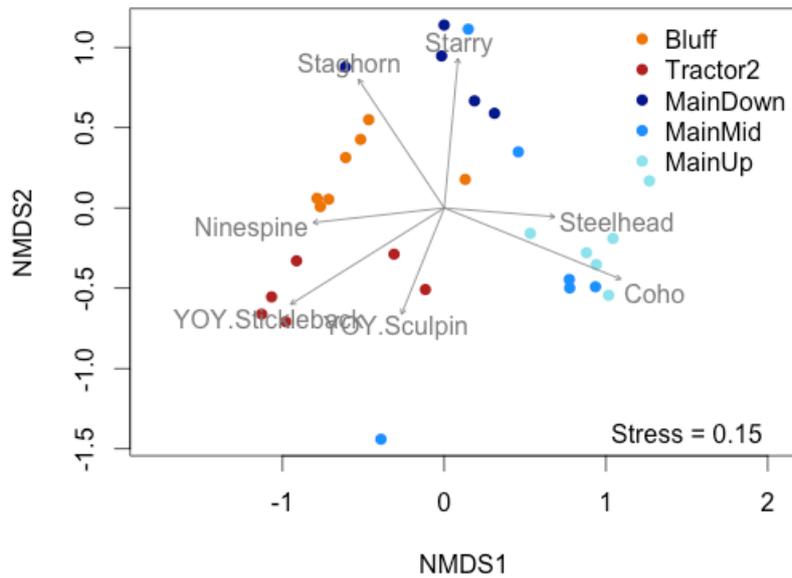


Figure 3. NMDS plot of species abundance. Points represent samples and are colored according to site. Arrows indicate the magnitude and direction of species loadings (variable weights) on the composite axes. Only species that significantly ( $p < 0.01$ ) contributed to the axes are displayed. YOY = young of year.

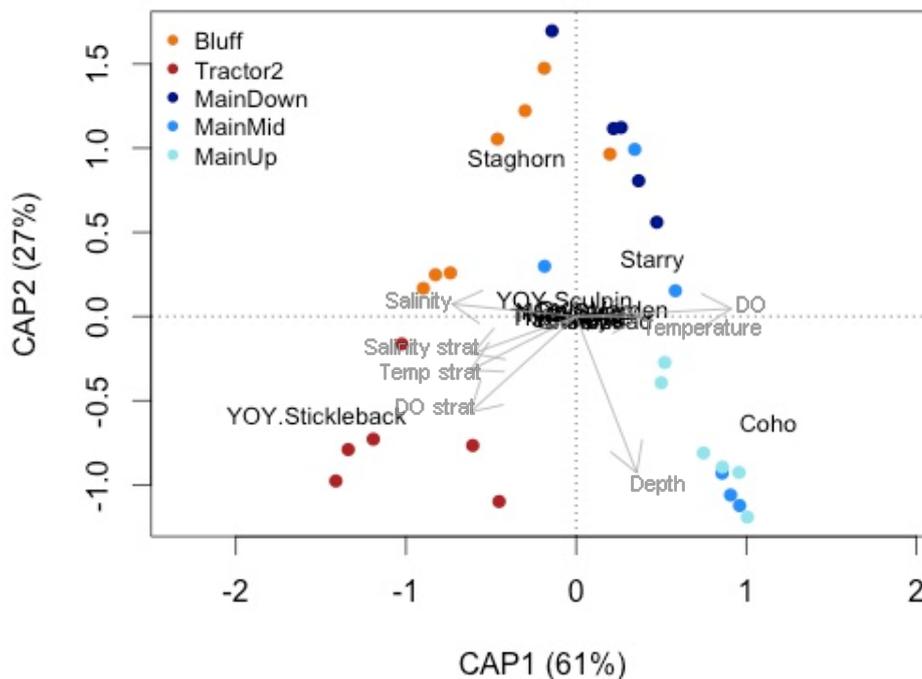


Figure 4. Triplot of the canonical analysis of principal coordinates (CAP), which ordines the samples in species space as a function of the environmental variables, scaled by species. Samples are colored by site. Magnitude and direction of arrows indicate the importance of the variable and its relative contribution to each axis, respectively. YOY = young of year, DO = dissolved oxygen, strat = stratification.

community is observed in Figure 4. Coho tend to be present in the greatest relative abundance in the upper mainstem site where the channel is deeper with higher dissolved oxygen, while showing no correlation with stratified channel conditions. Staghorn sculpin dominate the shallower waters of the Bluff marsh channel, which has no correlation to dissolved oxygen. Young of year stickleback typify the fish community in the Tractor2 marsh channel, which is environmentally characterized by low dissolved oxygen and stratified conditions. The middle mainstem site shows the greatest variation in fish community composition and channel depth. Channel stratification is not correlated to channel depth.

### Discussion

Environmental variables showed distinct trends across sampling sites in the Anchor River estuary. As hypothesized, PCA revealed that the two marsh channels were environmentally distinct from each other and from the mainstem sites. However, I hypothesized that marsh channels would show greater environmental

variability than mainstem channels, as observed by dispersion of the sample points in PCA ordination. This hypothesis did not entirely hold. The Tractor2 marsh channel showed substantial variability, yet the Bluff marsh channel was generally quite stable. This is likely due to each channel's connectivity to the mainstem: Tractor2 is always connected but a silt sill at the mouth of the Bluff channel requires the tide to reach approximately 4.5 m before the channel is inundated. This physical feature enables the Bluff channel to maintain environmental stability during low and moderate tides. It is important to account for such environmental heterogeneity when considering conservation and management of estuarine habitats. Estuarine systems are generally discussed as a single entity. However, the results presented here demonstrate that distinct environmental conditions exist even within rather small estuaries such as the Anchor. The dynamic nature and strong gradients of the estuarine environment result in a mosaic of habitat types facing varying environmental conditions.

The present study supports both hypotheses regarding the relationship between the fish assemblage and environmental characteristics: sites that were environmentally similar also had similar fish communities and environmental variables explained a significant portion of the variation in relative species abundance. This indicates that the dynamic estuarine environment plays an important role in structuring the fish community. Furthermore, these multivariate analyses reveal trends in species co-occurrence and associations with the environment, which can lend insight into possible interactions. For example, the CAP indicates that coho salmon dominate the fish community in the deeper waters of the upper mainstem site. However, coho were present in 90% of all samples and are thus the most pervasive species in the estuary. By evaluating the CAP ordination triplot (Figure 4), we see that coho relative abundance is negatively correlated with channel depth. Coho have their lowest relative abundance in the shallow Bluff channel where staghorn sculpin are dominant. Staghorn sculpin are benthic sit-and-wait predators that will predate on juvenile salmon. Therefore, it is possible that coho avoid areas where water depth brings them into closer proximity to this potential predator. An important next step would be to evaluate the size distribution of fish species across sampling sites to examine the potential for predation and competitive interactions.

Like environmental characteristics, fish assemblages based on relative species abundances showed distinct trends across sample sites in the Anchor River estuary. Indeed, fish assemblages were more distinct than originally hypothesized with all sites hosting significantly different fish communities. The only exception was the middle mainstem site, which had a highly variable fish community and was not distinct from the upper or lower mainstem sites. Although this middle site had a changing fish community, the variance was not significantly different from the other sampling sites, contrary to my initial hypothesis. Similar to these results, Nyitrai et al. (2012) documented that fish diversity is highest and most variable in the oligohaline reach of the estuary and suggested this is due to the environmental variability. This corresponds with the environmental patterns we observed in the PCA: the middle mainstem site had the

greatest sample dispersion. Variation in the fish community at this site could also result from the movement of some fish species as they seek preferred conditions within the dynamic environment while more tolerant species may remain in the same location. To answer this question, we have initiated a study using passive integrated transponder (PIT) tags to document the movement of individual salmonids, dolly varden char, and staghorn sculpin. This information, combined with the multivariate analyses presented here, will offer powerful insight into fish habitat selection within a highly variable environment. By identifying the underlying mechanisms governing species' distribution and environmental characteristics, we can develop conservation and management strategies that maximize the diversity of critical habitats such as estuaries, which provide essential habitat for many species.

#### **Acknowledgements**

The study presented here was conducted in collaboration with the research staff at the Kachemak Bay National Estuarine Research Reserve as part of an ongoing project under the direction of Coowe Walker and Steven Baird. Funding for fieldwork was provided by the Alaska Department of Fish and Game and the State Wildlife Grant Program. This work fulfilled a portion of the original research component of my Master's degree under the co-direction of Charles Simenstad and Timothy Essington. I am supported by a fellowship through the H. Mason Keeler Endowment for Excellence.

#### **References**

- Anderson, M.J. 2001. A new method for non-parametric multivariate analysis of variance. *Austral Ecology* 26: 32-46.
- Anderson, M.J. 2006. Distance-based tests for homogeneity of multivariate dispersions. *Biometrics* 62:245-253.
- Beamish, R. J., and C. Mahnken. 2001. A critical size and period hypothesis to explain natural regulation of salmon abundance and the linkage to climate and climate change. *Pacific climate variability and marine ecosystem impacts* 49(1-4):423-437.

- Bottom, D. L., K. K. Jones, T. J. Cornwell, A. Gray, and C. A. Simenstad. 2005. Patterns of Chinook salmon migration and residency in the Salmon River estuary (Oregon). *Estuarine, Coastal and Shelf Science* 64(1):79–93.
- Craig, B. E., C. A. Simenstad, and D. L. Bottom. 2014. Rearing in natural and recovering tidal wetlands enhances growth and life-history diversity of Columbia Estuary tributary coho salmon *Oncorhynchus kisutch* population. *Journal of Fish Biology* 85(1):31–51.
- Elliott, M., and A. K. Whitfield. 2011. Challenging paradigms in estuarine ecology and management. *Estuarine, Coastal and Shelf Science* 94(4):306–314.
- Fretwell, S. D., and H. L. Lucas Jr. 1969. On territorial behavior and other factors influencing habitat distribution in birds. *Acta Biotheoretica* 19(1):16–36.
- Gower, J.C. 1971. Statistical methods of comparing different multivariate analyses of the same data. *In*: Hodson F.R., D.G. Kendall, and P. Tautu (eds). *Mathematics in the archaeological and historical sciences*. Edinburgh University Press, Edinburgh, pp 138–149.
- Healey, M.C., 1982. Juvenile Pacific salmon in estuaries: the life support system. Pages 315–341 in V. S. Kennedy, editor. *Estuarine Comparisons*. Academic Press, New York.
- Hoem Neher, T. D., A. E. Rosenberger, C. E. Zimmerman, C. M. Walker, and S. J. Baird. 2013a. Use of glacier river-fed estuary channels by juvenile Coho Salmon: transitional or rearing habitats? *Environmental Biology of Fishes* 97(7):839–850.
- Hoem Neher, T. D., A. E. Rosenberger, C. E. Zimmerman, C. M. Walker, and S. J. Baird. 2013b. Estuarine Environments as Rearing Habitats for Juvenile Coho Salmon in Contrasting South-Central Alaska Watersheds. *Transactions of the American Fisheries Society* 142(6):1481–1494.
- Hutchinson, G.E. 1957. Concluding remarks. Cold Spring Harbor Symposium on Quantitative Biology 22:415–427.
- Jackson, D.A. 1995. PROTEST: a Procrustean randomization test of community environment concordance. *Ecoscience* 2:297–303.
- Jones, K. K., T. J. Cornwell, D. L. Bottom, L. A. Campbell, and S. Stein. 2014. The contribution of estuary-resident life histories to the return of adult *Oncorhynchus kisutch*. *Journal of Fish Biology* 85(1):52–80.
- Kerkvliet, C. M., M. D. Booz, and B. J. Failor. 2013. *Recreational fisheries in the Lower Cook Inlet Management Area, 2011–2013, with updates for 2010*. Alaska Department of Fish and Game, Fishery Management Report No. 13-42, Anchorage.
- Kruskal, J.B. 1964. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* 29:1–27.
- Legendre, P. and M.J. Anderson. 1999. Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. *Ecological Monographs* 69: 1–24.
- Magnusson, A., and R. Hilborn. 2003. Estuarine Influence on Survival Rates of Coho (*Oncorhynchus kisutch*) and Chinook Salmon (*Oncorhynchus tshawytscha*) Released from Hatcheries on the U. S. Pacific Coast. *Estuaries* 26(4):1094–1103.
- McGarigal, K., S.A. Cushman, and S.G. Stafford. 2000. *Multivariate Statistics for Wildlife and Ecology Research*. Springer, New York, New York, USA.
- Morris, D. 2003. Toward an ecological synthesis: a case for habitat selection. *Oecologia* 136(1):1–13.
- Motyka, J. B. Dobrzanski, and S. Zawadski. 1950. Preliminary studies on meadows in the south-east of the province Lublin. *Annales Universitatis Mariae Curie-Sklodowska. Lublin, Polonia. Sectio E*, 13:367–447.
- Nyitrai, D., F. Martinho, M. Dolbeth, J. Baptista, and M. A. Pardal. 2012. Trends in estuarine fish assemblages facing different environmental conditions: combining diversity with functional attributes. *Aquatic Ecology* 46(2):201–214.
- Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, and H. Wagner. 2015. *vegan: Community Ecology Package*. R package version 2.3-1. <http://CRAN.R-project.org/package=vegan>.
- Quinn, T. P. 2005. *The behavior and ecology of Pacific salmon and trout*. American Fisheries Society and Washington University Press, Bethesda, MD and Seattle, WA.

- Quinn, T. P., N. Harris, J. Anne Shaffer, C. Byrnes, and P. Crain. 2013. Juvenile Coho Salmon in the Elwha River Estuary Prior to Dam Removal: Seasonal Occupancy, Size Distribution, and Comparison to Nearby Salt Creek. *Transactions of the American Fisheries Society* 142(4):1058–1066.
- R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Rosenzweig, M. L. 1981. A Theory of Habitat Selection. *Ecology* 62(2):327–335.
- Schindler, D. E., R. Hilborn, B. Chasco, C. P. Boatright, T. P. Quinn, L. A. Rogers, and M. S. Webster. 2010. Population diversity and the portfolio effect in an exploited species. *Nature* 465(7298):609–612.
- Sibert, J. R. 1979. Detritus and juvenile salmon production in the Nanaimo estuary: II. Meiofauna available as food to juvenile chum salmon (*Oncorhynchus keta*). *Journal of the Fish Research Board of Canada* 36:497-503.
- Simenstad, C. A., K. L. Fresh, and E. O. Salo. 1982. The role of Puget Sound estuaries in the life history of Pacific salmon: An unappreciated function. Pages 343-364 in V.S. Kennedy, editor. *Estuarine comparisons*. Academic Press, New York.
- Thorpe, J. E. 1994. Salmonid Fishes and the Estuarine Environment. *Estuaries* 17(1):76–93.
- Werner, E. E., and J. F. Gilliam. 1984. The Ontogenetic Niche and Species Interactions in Size-Structured Populations. *Annual review of ecology and systematics* 15:393–425.

## The spatial patterns and environmental constraints of antibiotic resistant bacteria in an urban river

Angela Klock

*School of Environmental and Forest Sciences, University of Washington, Seattle, WA, USA*

*E-mail: [angelk9@uw.edu](mailto:angelk9@uw.edu)*

Received December 2015; accepted January 2016; published September 2016

### Abstract

Antibiotic resistant infections are projected to become the number one cause of human death by 2050 primarily as a result of the systematic overuse and misuse of antibiotics in medicine and most egregiously in industrial animal agriculture. Most surveys of resistant bacteria are conducted in surface waters with known point sources of antibiotic resistant determinants such as waste water treatment plants, but little is known of the spatial patterns of antibiotic resistance in urban rivers impacted solely by runoff contaminates. The environment is a source of novel resistance mechanisms and urban rivers may act as reservoirs of resistance genes and pathways for pathogens within the urban water cycle. Heavy metal pollution in urban rivers compound the impacts of antibiotic resistance due to genetic co-selection. The objectives of this analysis were to determine if antibiotic resistance is clustered along the Green-Duwamish River, and to identify the significant environmental factors that explain the distribution pattern. To accomplish these objectives hierarchical clustering and canonical correspondence analysis were used. Results indicate that antibiotic resistance is spatially variable due in part to stream power, and the presence of natural bacteria and wildlife.

### Introduction

Clean fresh water is requisite for human wellbeing and a pressing global issue. As the population grows the demand for fresh-water resources will increase but availability will decline due to pollution and contamination. By 2025 over 70% of the accessible renewable fresh water on the planet will be co-opted for human use (Postel et al. 1996). In 2008-2009 (the most recent survey available) the Environmental Protection Agency determined 44% of the streams and rivers surveyed in the US are impaired and the top two causes are high levels of bacteria and habitat alteration. The source of the bacteria is primarily industrial animal agriculture and most of the habitat alteration is due to hydromodification (USEPA 2002). What these facts indicate and goes largely unstated is that the availability of

unpolluted water to non-human biota is declining on a global scale.

The character of an urban watershed influences stream health in many direct and indirect ways. Changes are commonly made to natural river morphometry including channelization (straightening), confinement (narrowing), hardening (impermeable), bank stabilization (steepening) dredging (deepening), and the impounding of water to alter flow (safety/convenience). Simultaneously, urban rivers receive a wide variety of anthropogenic-based constituents from point and non-point sources. Stormwater runoff delivers a relentless input of heavy metals, toxic chemicals, pharmaceuticals, and fecal bacteria from non-point sources which mix in the urban water cycle (Bedan & Clausen 2002, Liu et al. 2011). To confuse the problem the arrival of these pollutants are extremely variable as runoff

events occur as pulses with precipitation. Further, the total volumes and types of pollutants again vary by land use (Hipp et al. 2006) and even by storm event at the same site (Johnson et al. 2011). Urban rivers represent some of the most severely impacted ecosystems which are routinely subjected to multiple abiotic and biotic stressors which overlap in time and space. The synergies of these stressors determine the community composition and distribution of macro-invertebrates and micro-organisms (Mazaris et al. 2010) and most desperately complicate our ability to quantify their impacts.

#### *The role of riverine bacteria*

Bacteria are the fundamental drivers of all biogeochemical cycles (Falkowski et al. 2008) and their exquisite genetic and metabolic diversity optimize ecological function (Mcgrady-steed et al. 1997). Bacteria also demonstrate very high functional redundancy which has a stabilizing effect on ecosystems and maintains ecological resilience. The counter to resilience are the physical impacts of urbanization which simplify riverine habitats and disturb the ecological equilibria of bacterial community composition (Cabello 2006), niche architecture (Drudge et al. 2012), and the processes of community assembly and succession (Lowell et al. 2009, Lear & Lewis 2009). Urban river segments that are highly impacted are extreme environments with high temperatures and/or nutrients, and wide ranges of pH which constrains the microbial community and result in a loss of biodiversity (Palmer et al. 2010) and a decline of ecological function (Singer et al. 2010).

#### *Riverine bacteria are impacted by development*

To appreciate the impacts of urban development on bacterial populations we should understand bacterial evolutionary processes (Mcarthur et al. 2011). The mechanisms by which bacteria confront environmental stressors are vertical gene transfer or cell division, and by horizontal gene transfer (HGT) between species or the acquisition of mobile genetic elements (MGE) (Rankin et al. 2011). MGEs such as plasmids, transposons, and gene cassettes are important components of bacterial diversity and adaptability in challenged environments (Smalla & Sobczyk 2002) but these elements are also self-replicating and capable of living free in the environment. MGEs represent a stable source of

resistant determinates (Franco et al. 2009) in polluted rivers when they carry genes for phenotypic traits for resistance to heavy metals and antibiotics (Baker-Austin et al. 2006). In environments where there are strong selective pressures the survival advantage conferred by MGEs may be acquired quickly by bacteria to meet the physiological demands of the environment (van Elsas & Bailey 2002) and this includes pathogenic and environmental strains.

#### *Fecal contamination and indicator species*

In the US one of the most important regulatory indicators of fecal-contaminated water is *Escherichia coli* (*E. coli*). *E. coli* is a gram-negative enteric coliform that is a normal member of the gut of warm blooded vertebrates (Gordon & Cowling 2003) and is abundant in human sewage and food animal manure, particularly feedlot and dairy cattle (Islam et al. 2014). *E. coli* is a well suited indicator for contamination by these sources as it common in runoff, is adept at HGT with other species including pathogens (Maal-Bared et al. 2013), and exhibits very high rates of antibiotic resistance. Unfortunately there are several drawbacks to the efficacy of *E. coli* as an indicator. For example, it is commonly found in urban streams (Rompré et al. 2002) where it is capable of persisting (Anderson et al. 2005) which makes the determination of recent contamination problematic. There is also little correlation between the presence of *E. coli* and many zoonotic pathogenic forms including *E. coli* 0157 (Drozd et al. 2013).

#### *Antibiotic resistance in urban rivers*

Antibiotic resistant infections are a critical global health problem (WHO 2014) and is predicted to be the leading cause of death, killing 10 million people a year, by 2050 (O'Neill 2014). The overuse and misuse of antibiotics in animal agriculture is the leading source of antibiotic resistant determinates in the environment (Pruden et al. 2013), and most antibiotics taken by humans or given to animals are excreted in a bioactive form (W. Zhang et al. 2009). Contaminated waterways are an efficient mode of dissemination of ARGs (Collignon 2015), borne out by the identification of ARGs in all types of waters such as lakes, reservoirs, streams, rivers, ground water, and drinking water (Baquero et al. 2008).

Antibiotic resistant determinants (ARD) include human- and animal- associated fecal

bacteria, antibiotic resistant bacteria (ARB), antibiotic resistant genes (ARG), antibiotics, antibiotic metabolites, heavy metals, and other pollutants, all of which have biological implications for natural environmental bacteria (Singer et al. 2006). ARDs are environmental pollutants (Pei et al. 2006) which have ecological influence on environmental bacteria community structure and function (Rizzo et al. 2013).

Gram-negative bacteria are increasingly responsible for antibiotic resistant infections (Ash et al. 2002) and efflux pumps which pump antibiotics out of the cell make them highly resistant to many antibiotics. This mechanism is used by multi- and pan-drug resistant *E. coli* strains (Martinez, Sánchez, et al. 2009) which have recently acquired the New Delhi Metallo- $\beta$ -lactamase-1 (NDM-1) enzyme from *K. pneumonia* which is a serious public health concern (Fair & Tor 2014). The NDM-1 gene provides resistance to all beta-lactam antibiotics including the newer carbapenems (Keen & Patrick 2013). Although some antibiotics such as aminoglycosides and fluoroquinolones do not contain beta-lactam rings, most bacteria that have acquired NDM-1 have previously acquired other resistance factors and most are already resistant to aminoglycosides and fluoroquinolones (Lupo et al. 2012). If NDM-1 persists in urban rivers it could transform common bacteria like *E. coli* which cause most of the common nosocomial and community-acquired infections into superbugs resistant to virtually all commonly used antibiotics (WHO 2014).

Environmental risk assessments should include understanding how bacteria and plasmids move, mix, and interact between the medical, community, and environmental ecologies and the role the water cycle plays in these interactions. This is important as water environments host ARGs displaying a wide range of resistance mechanisms (X.-X. Zhang et al. 2009) and diverse population of organisms including many opportunistic pathogens which possess an inherently low level of susceptibility to antibiotics (Martinez & Sánchez, et al. 2009). Opportunistic pathogens are important organisms as they cause more infections in immune-compromised individuals which comprise a growing number of people in our community. This is a converging environmental-medical health issue as

increasing resistance is tracking with an increasingly susceptible population.

The natural function of ARGs is probably different in polluted waters than in unpolluted waters (Dantas & Sommer 2012) and in light of the significant roles natural bacterial communities play in maintaining ecological function, could be important distinctions (Falkowski et al. 2008).

The occurrence of fecal bacteria in biofilms is relevant since it indicates potential reservoirs of pathogens in this compartment which can be mobilized during river-bed scouring and sediment transport (Mcarthur et al. 2011). Once mobilized, they are more available for transmission to humans, pets or wildlife during exposure to contaminated water (Donnadieu et al. 2013).

Large inputs of fecal bacteria are also competitors with natural bacterial communities, and if antibiotic residues are present may exert sufficient selective pressure on natural populations to shift them to more resistant communities over time (Kümmerer 2009). It is therefore important to consider antibiotic resistance as a social-ecological problem that accounts for the interaction of bacterial populations originating from very different sources which mix, shuffle, and recycle through each of these communities via the water cycle.

Fecal-based bacteria and environmental strains mix at hot spots such as tributary confluences, runoff pour points, and sites of effluent discharge where gene exchange between the microbial ecologies readily occurs. The mixing is substantially compounded by heavy metal pollution due to co-selection of genes which tend to stabilize phenotypes resistant to antibiotics and metals (Mcarthur & Tuckfield 2000), making either determinant a driver for resistance to both. While the focus of the problem of antibiotic resistance has been primarily on medical and veterinary practices it is clear the environment acts as an antibiotic resistance source, reservoir (Drudge et al. 2012) and pathway that needs exploring (Pruden 2014).

Antibiotic compounds are a natural product of bacterial physiology but the increase in antibiotic contamination of surface waters make the ability to detect changes in the levels of environmental antibiotics important (Rizzo et al. 2013). Unfortunately surveys for background levels of antibiotics are impossible due to a lack of historical baselines before the use of

antibiotics became widespread. The natural production of antibiotics also complicates our ability to determine the origin of antibiotics as either anthropogenic or environmental.

The distribution of antibiotic resistance is not “everywhere” but is spatially variable (Eisenberg et al. 2012). For example, in many undisturbed environments resistance is very low (Singer et al. 2006) but it is sometimes found in pristine environments without a legacy of antibiotic contamination (Leff et al. 1993). Conversely, in areas with substantial antibiotic pollution sensitive bacterial strains are sometimes found (Garcia-Armisen et al. 2011). There are many studies that have evaluated environmental resistance in polluted sites (Leff et al. 1993, McArthur & Tuckfield 2000, McArthur 2006), and along environmental gradients of pollution (Akiyama & Savin 2010, Viau et al. 2011) but there is a dearth of studies that compare pristine sites.

The evolutionary mechanism where functional changes of a gene occurs purely due to environmental context is exaptation (Martínez & Baquero 2014). In urban rivers the exaptation of resistance is confounded by anthropogenic activity (Garcia-Armisen et al. 2011) but it appears that constraints limit adaptation even when there are strong selective pressures and niche opportunities (Kassen et al. 2004). The evolutionary and ecological mechanisms of resistance are probably contextually, temporally and spatially dependent. These are intriguing observations that highlight our lack of understanding the role resistance plays in natural populations (Martinez, Fajardo, et al. 2009) and how resistance responds to land use along the longitudinal distance of a river.

Determining the spatial patterns of resistance across a watershed and aligning bacterial traits with common factors will identify how resistance is a survival advantage in the built environment. The identification of how anthropogenic inputs impose selective pressure on riverine bacteria is critical for assessing the health risks of urban waters. These are fundamental questions in urban river ecology but most surveys for resistance only consider point sources of ARDs such as waste water treatment plants (Hirsch et al. 1999, Czekalski et al. 2012), animal agricultural (W. Zhang et al. 2009), industrial effluents (McArthur & Tuckfield 2000), and heavy metals (McArthur et al. 2011). Therefore, a survey of a watershed

without point sources of ARDs is needed to identify the resistance of resident bacteria where there are no tractable sources of ARDs. Water quality assessments are critical to evaluating stream health, but characterizing the bacterial community traits that are actively being expressed in the environment will shed light on the immediate risks of pathogens. Urban rivers are highly disturbed and simplified ecosystems and efforts to protect human and environmental health must include how off-stream and in-stream processes work together. The goals of this survey were to identify the distribution and environmental constraints of antibiotic resistance among riverine bacteria across an urban watershed.

## Methods

### *Study area:*

The Green-Duwamish River is located in southern King County Washington and is the main river corridor for water resource area nine (WRIA 9). The upper Green River is the municipal water supply for the city of Tacoma. The watershed has a forested to industrial land use gradient with no wastewater treatment plants or combined sewage overflow effluents above the lower Duwamish waterway. The forested headwaters provide some of the cleanest water in the nation but the last 8.5 kilometers of the Duwamish River is extremely polluted and an EPA Superfund site. This makes for a fascinating contrast from which to evaluate the levels of antibiotic resistance in relation to land use where there are no direct point sources of ARB or ARG (Figure 1).

### *Water and biofilm sampling:*

Water and biofilm samples were collected at twenty nine sites along the main stem and tributaries of the Green-Duwamish River (King County, Washington) over several weeks during dry weather conditions. Water samples were collected by near the river bottom in approximately 1m depth. Epilithic biofilms were collected by placing a square grid (100 square mm) on a submerged rock and scraping the delineated surface with a sterile Specisponge at a water depth of approximately one meter and replacing the Specisponge to its sterile Whirl-pac. All samples were held on ice in the field and analyzed within 12 hours of collection.

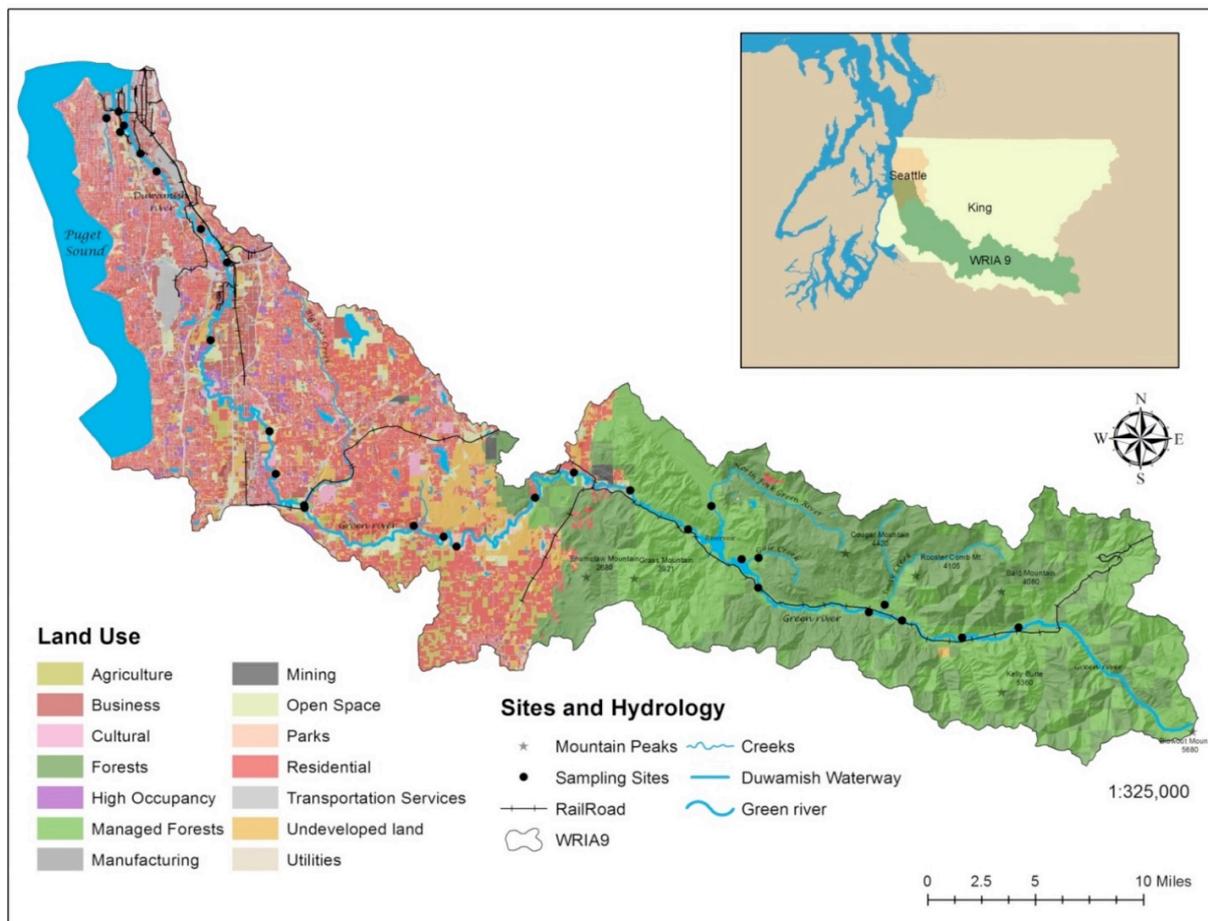


Figure 1. Study area including the land use and hydrology features. Sites are black points along the main corridor and tributaries.

*Detachment and dispersal of microorganisms:*

25 mL of deionized water were added to the Whirl pacs containing the sponge and attached biofilm and vigorously massaged for five minutes to dislodge the material from the sponge. These suspensions were then pipetted into sterile glass vials and vortexed for three minutes to break up the biofilm polymer matrix.

*Enumeration of microorganisms:*

Water samples were evaluated for total coliform and *E. coli* abundance, calculated as most probable number (MPN) using the Colilert Method with Quanti-Tray/2000 according to the manufacturer’s protocol (IDEXX). 100-mL water and biofilm suspensions were used directly or 10-fold diluted, poured into trays and incubated at 36.5C for 24 hours. Tray wells that were yellow (positive for coliforms), and yellow and fluorescent upon exposure to long-wavelength UV light (positive for *E. coli*) were

counted. The MPN counts were expressed as MPN/100mL water or MPN/100sqmm of biofilm. The MPN analyses were performed in duplicate and results were given as the arithmetic mean.

IDEXX wells with a positive reaction for coliform bacteria and *E. coli* were haphazardly selected and 10uL aliquots withdrawn and inoculated onto R2A agar plates (Reasoner & Geldreich 1985) with the spread plate method. Heterotrophic plate counts were enumerated after incubation at 36.5°C for 24 hours. Counts were recorded as colony-forming units (CFU) per 100mL of water or per 100 square millimeters of biofilm, respectively.

Antibiotic resistance was tested by plating as above but aliquots were inoculated onto three MacConkey agar plates produced with 25ug/ml of Ampicillin, Chloramphenicol, or Tetracycline, respectively, in triplicate. Plates

were incubated at 36.5C for 24 hours before enumeration.

*Data sets and assumptions:*

Data matrices were constructed for the response variables of antibiotic resistance and then partitioned by bacterial groups (environmental/enteric) and stream compartment (water/biofilm). The matrix objects (rows) were the sample sites and the descriptors (columns) were the colony counts of each response category (Table 1).

The explanatory matrix consisted of a mixed data set of field measurements of geographical coordinates in latitude and longitude, water physicochemistry, a characterization of the riparian habitat, wildlife, and land use. Additional environmental factors were derived within a geographic information system including stream sinuosity, migration, and stream power. It was assumed that the environmental variables would display a linear and gradual gradient across the watershed and the response variables would exhibit a unimodal distribution. The objects were the sites and the attributes were factors. All analyses was by R Studio Ver. 0.99-486 (©2008-2015 R Studio, Inc).

The entire data set was explored with descriptive statistics prior to any analysis. Ecological count data rarely conform to a normal distribution, equal variance or linear relationships so these assumptions were relaxed for the analysis. The response data exhibited significant outliers which would exert significant influence on the coordinate axes but these data points were considered important and retained for the analysis as their deletion would probably influence the signal of the factors. There were two objectives to determine, 1) if there was clustering in the distribution of antibiotic resistance by bacterial group, antibiotic class, or stream compartment, and ii) the primary environmental drivers of the observed spatial patterns.

*Statistical methods:*

Multivariate statistical techniques were used to analyze this complex and stratified data set. To accomplish this a combination of data exploration, visualization tools, hierarchical clustering, and canonical correspondence analysis (CCA) were used to quantify concordance between the explanatory and response variables. An Empirical Bayesian

kriging spatial statistic was run to map resistance probabilities across the watershed. These techniques provide information about the data structure, its relation to environmental gradients, and the risk to human health at multiple scales.

All data was inspected and visualized prior to any statistical analysis. A mosaic plot was generated with the response data. This method represents the probabilities of multiple variables in multi-dimensional space and allows for the rapid detection of spatial patterns, the relationship among the data and over spatial extent of the watershed (Figure 2). The explanatory variables were plotted in a correlation matrix or correlogram to observe the most strongly associated covariates (Appendix A). Those factors were tested with a detrended correspondence analysis (unconstrained method) using the function (decorana) followed by interpretation of the effects by fitting with the function (envfit). The results indicated that several physical, chemical, land use, and biological components were significant factors in the spatial variability of resistance and those that also made ecological sense for the analysis were retained.

*Hierarchical Cluster Analyses:*

A hierarchical cluster analysis was run on the response variables. Since there was little information known about the structure of antibiotic resistance in the environment the assumption was that no grouping structure was present. Therefore the clustering method was applied as an exploratory method and not for hypothesis testing.

There are many parameters for performing a cluster analysis which behave differently. To account for this several distance measures were explored and the results were all similar but some metrics improved the result which supported a natural clustering in the data. The metric of similarity was compared to dissimilarity. Similarity being a measure of the ratio of the number of factors shared by any two sites (objects) compared to all of the factors between them. Sites which shared all factors are therefore identical and have a similarity of 1.0, while sites which share nothing have a similarity of 0.0. Dissimilarity is the inverse of similarity, characterized by the number of attributes that are unique to two site compared to the total number of attributes between them.

A. Klock, *Antibiotic resistant bacteria in an urban river*

Table 1. Matrices of response variables: colony forming units (CFU) of antibiotic resistant bacteria by metabolic ability to ferment lactose, the stream compartment isolated from, and by antibiotic class resistance.

Matrix	Ampicillin	Tetracycline	Chloramphenicol	Biofilm	Water	Coliform	Enteric
A. Total Antibiotic Resistance	x	x	x	x	x	x	x
B. Water	x	x	x		x	x	x
C. Biofilm	x	x	x	x		x	x
D. Coliform Bacteria	x	X	x	x	x	x	
E. Enteric Bacteria	x	x	x	x	x		x
F. Ampicillin	x			x	x	x	x
G. Tetracycline		x		x	x	x	x

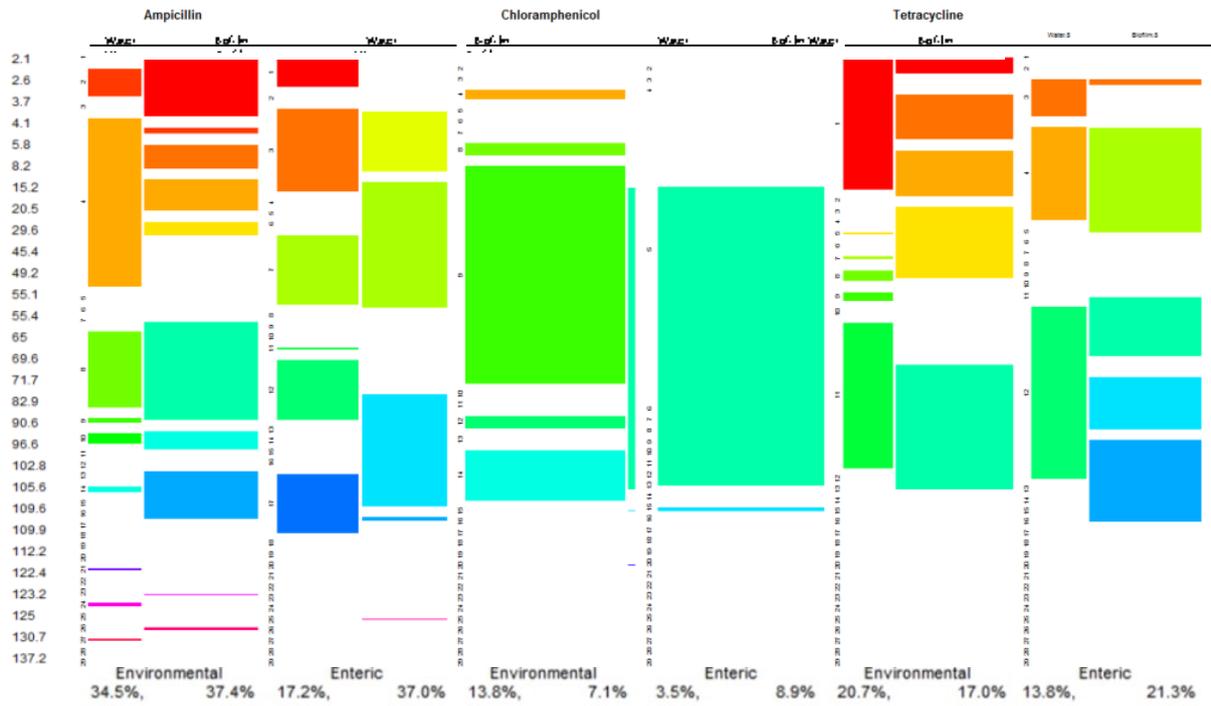


Figure 2. Mosaic plot representing the probability of resistance by antibiotic class and stream compartment (top), and among environmental (coliform) and enteric (*E. coli*) groups (bottom). Percentages across the bottom are the proportion of samples that were resistant in each stream compartment. The Y axis is the river kilometer beginning at Eliot Bay in Puget Sound extending upstream to the river headwaters at Stampede Pass in the Cascade mountain range. There is clear portioning of resistance by antibiotic class, bacterial groups, stream compartment and by river distance.

The clustering algorithm selected was an *exclusive sequential hierarchical agglomerative and polythetic* technique (i.e., each datum can participate in only one cluster through a recursive sequence of operations that arranges groups in a hierarchy of defined relationship after groups are built by considering all

variables. This technique is applicable when there are no known or anticipated clustering behaviors of the data. Since a normal distribution is not required, the data needed no transformation but due to the range of values the data was log transformed, and the

explanatory data was standardized as they were many measurement scales.

After transforming the response data a dissimilarity index was selected. This is a geometric measure of site proximity in semantic space defined by the values of the attributes. The function `vegdist()` within the `vegan` package was used with the dissimilarity indices `jaccard` ( $2B/(1+B)$ ), where  $B$  is the Bray–Curtis distance, and `gower` where  $M$  is the number of columns.

$$d[j][k] = (1/M) \sum (\text{abs}(x[i][j]-x[i][k]) / (\max(x[i][j]) - \min(x[i][j])))$$

binary:  $(A+B-2*J)/M$ ,  
 where  $M$  is the number of columns (excluding missing values) (1)

Dissimilarities are bounded [0,1] as once a pair of sites share no species they are completely dissimilar.

For the hierarchical cluster analysis the function `hclust()` from the `stats` package was used. The agglomeration method was `ward.D` or the *minimum variance* method as it tends to reduce the complexity (number of clusters) in the data which better explains the observations.

For the environmental data a separate cluster analysis was run by calculating a series of pairwise dissimilarities between sites. The R packages `vegan`, `cluster`, `pvclust`, `NbClust` and a custom package `biostats.R` provided by the UW School of Aquatic and Fishery Sciences were used.

The results of the cluster analysis were evaluated by examining the cluster table which summarizes the fusions and associated distances, and by visualizing the dendrograms. The cluster solution was then validated by the calculating the agglomerative coefficient [0,1] which is a measure of cluster strength, one being greater, and then by measuring the fidelity of the cluster representation to the original data set with the cophenetic correlation. The correlation is calculated by evaluating the intergroup dissimilarity (distance) to the original dissimilarity of the data in  $p$ -dimensional space. The results were reported and visualized with a plot. The number of clusters selected for each response variable was subsequently observed with a scree plot. Although this provides a guideline, bootstrapping was run to evaluate cluster stability by a random sampling with replacement of the data. This tells us how robust the structure of the clusters are robust to perturbation. The function `pvclust` calls the

bootstrap. 2000 permutations and an alpha of 0.95 was used.

A Mantel test was conducted to test for significance between the dependent and independent matrices by calculating the Mantel correlation statistic, statistical significance is determined by permuting the first dissimilarity matrix. The hierarchical clustering method validated by determining the agglomeration coefficient. This metric describes the clustering structure of the data. If the clusters agglomerate quickly into distinct clusters which form a single cluster at some distance the coefficient approaches 1.

The cluster analysis of the ARB data was conducted with the CFU counts were  $\log+1$  transformed and any rows that summed to zero (i.e., missing data) were removed. The cluster analysis minimizes within group differences while maximizing among-group differences as with stratification to represent the variability of the data structure. Previous studies have reported antibiotic resistance to exhibit distinct high and low levels and not a gradual gradient in the environment but known have examined the structure of clustering events.

#### *Canonical correspondence analysis:*

To investigate the relationships between resistant bacteria and the environmental variables a canonical correspondence analysis (CCA) was performed. CCA is a constrained ordination technique that was specifically developed to relate biological communities with variations in the environment. CCA can handle continuous and categorical data.

The ordination axes (based on the biotic community) are linear combinations of environmental variables that assume a unimodal species–environment relationship that best explain the relative position of the community in semantic space. The variance calculated by the analysis does not represent all of the variance, simply that which is explainable by the environmental variables that confront the analysis. Therefore the results are dependent on the set of constraints selected and their interactions in the model. CCA is appropriate for this data set as there is some *a priori* information about how the constraints influence the persistence of antibiotic resistance in streams. For example, highly diverse natural communities tend to mitigate the persistence of ARBs (Wellington et al. 2013). This allowed for

the improvement of the model over the “shotgun” ( $y \sim x$ ) approach of utilizing all of the environmental variables as constraints. To test for significance of the individual variables each was fitted onto an ordination and the factors that were significantly related to bacterial groups were determined by backward selection with 1000 Monte Carlo permutations.

#### *Empirical Bayesian kriging*

Since most bacterial ecological studies do not evaluate environmental occurrence in a spatially explicit manner and do not sample at the appropriate scale of investigation (Maphosa et al. 2010). Empirical Bayesian kriging (EBK) was used to map the risks of exposure to antibiotic resistance across the watershed.

## **Results**

#### *Hierarchical Cluster Analyses:*

The results of the cluster analysis were evaluated by a series of metrics. The agglomerative coefficient measures the clustering structure of the data by determining each observations dissimilarity to the first cluster it is merged with and divided by the dissimilarity of the final merge. The agglomerative coefficient is the average width of the plot (R documentation). The cophenetic correlations were also examined as a test of how reliable the dendrograms represent the pairwise relationships of the original data in distance. The cophenetic correlations were compared with a Mantel Test. All cluster groups scored greater than 0.71 and were assumed to a reliable indicator of the actual clustering in the environment (Appendix B).

Total antibiotic resistance (all response) was statistically significant but showed rather weak concordance with environmental variables (*ARB*:  $r=0.11$ ;  $p<0.05$ ; and geographic distance (*GEOG*:  $r=0.12$ ;  $p<0.04$ ); ampicillin resistance was significant for geographic distance (*GEOG*:  $r=0.12$ ;  $p<0.03$ ); resistance in biofilms was significant for distance (*GEOG*:  $r=0.37$ ;  $p<0.006$ ); resistance in environmental coliforms and fecal coliforms were significant for distance (*Geog*:  $r=0.22$ ;  $p<0.009$ ); and (*Geog*:  $r=0.20$ ;  $p<0.03$ ), respectively.

#### *Canonical Correspondence Analysis:*

The second objective of this survey was to identify the major environmental drivers of

antibiotic resistance and this was accomplished by using CCA. The choice of using CCA was verified by evaluating the axis length or standard deviation of each log transformed matrix with the call `decorana`. Following a common rule-of-thumb the DCA1 axis length for the response was 4.796 indicating a unimodal response, and the axis length of the environmental data was 0.712, indicating a most likely linear response. The environmental variables were then fit with (`envfit`) onto an ordination. The maximum correlation with corresponding environmental variables were stream power ( $Pr>0.008$ ) and the presence of animals (metazoans) ( $Pr>0.020$ ).

The purpose of ordination is to simplify the interpretation of complex data sets. With the large number of environmental variables in this analysis and in the interest of semiology the ordination plot was limited to 6-10 arrows so a reduction in the number of variables was required. To reduce the number of variables without a substantial loss of ecological meaning, variables were removed in two ways; one *a priori* and one *a posteriori*. The explanatory variables were examined for covariance (i.e., redundancy) with one another which would inflate the variance without improving the fit of the variables. For example, in WRIA 9, impervious surface, channel width, and development all increase and are correlated in the downstream direction. Therefore only the most ecologically meaningful variables were retained. Conversely, elevation, drainage density, and forest cover decrease in the same direction. This requires prior knowledge of the natural history of the organism or ecological system of interest, such as the factors important for bacteria growth and the environmental selective pressures on the bacterial community for antibiotic resistance. To reduce the dimensionality of the explanatory factors a Principal Components Analysis (PCA) was run on the environmental correlation matrix to choose the environmental variables most strongly associated with the first three principal axes. A random permutation and distribution of eigenvalues indicate that the first three principal components were significant. The list of variables above a cutoff of 0.3 are listed in Table 2. The total amount of variance explained by the first two PCs were (42.5%); PC1 (30.7%) PC2 (11.8%) (Figure 3). The variables with the

A. Klock, *Antibiotic resistant bacteria in an urban river*

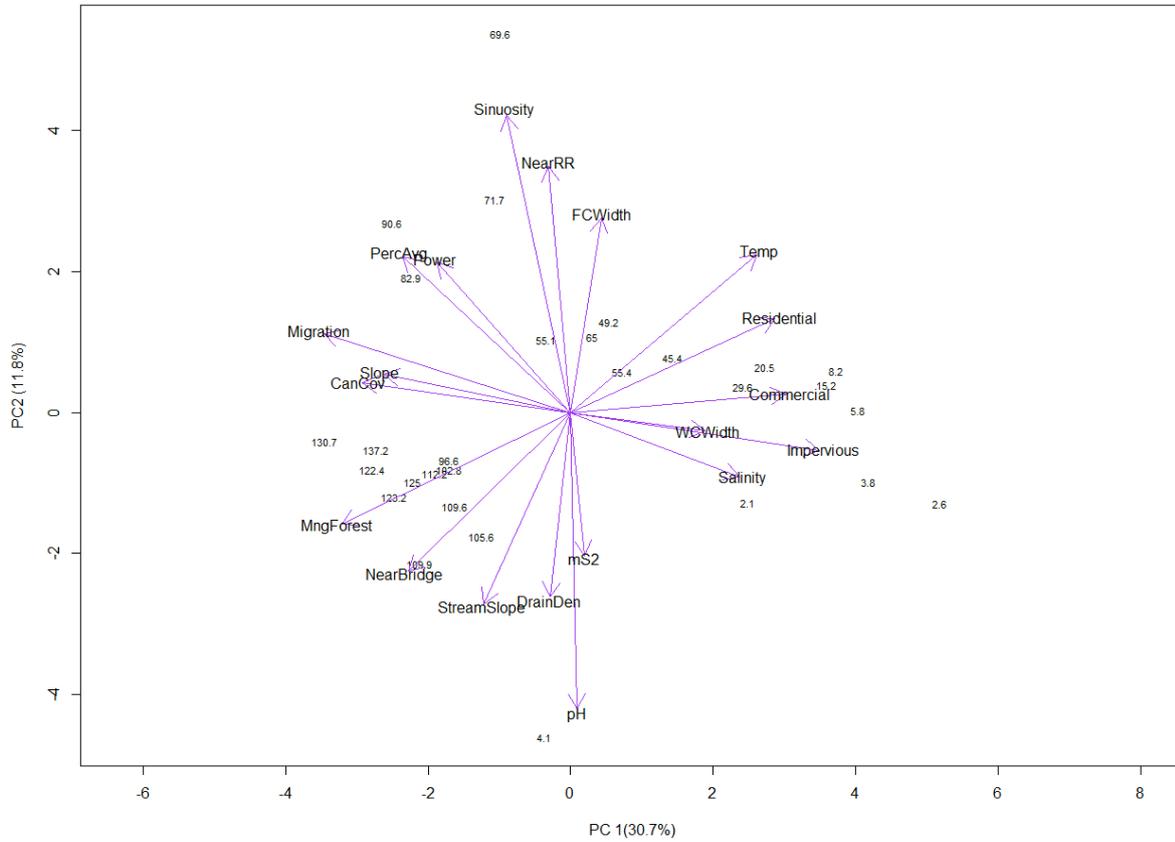


Figure 3. Biplot of the principal component analysis (PCA) for the environmental variables.

Table 2. The correlation and biplot scores from PCA. All data is log-transformed. The biplot scores are based on the correlations between the environmental variables and the LC ordination scores, weighted by the eigenvalue associated with each axis.

Variable	PC1	PC2	PC3
FC.MPNW		0.48	
FC.MPNB		0.508	0.489
EC.MPNW	-0.439		-0.432
EC.MPNB	-0.349		-0.318
MIGRATION	-0.876		
DRAINDEN		0.342	-0.674
WCWIDTH	0.484		0.652
FCWIDTH		-0.348	0.312
SINUOSITY		-0.716	
SLOPE	-0.659		
STREAMSLOPE		0.419	-0.568
POWER	0.496		
PERCAVG	0.645	-0.384	
NEARBRIDGE	0.531	0.503	
NEARRR		-0.592	

CANCOV	0.744		
MNGFOREST	0.779	0.428	
COMMERCIAL	-0.725		
RESIDENTIAL	-0.694	-0.437	-0.356
IMPERVIOUS	-0.895		
PH		0.811	
TEMP	-0.634	-0.445	
MS2			-0.82
SALINITY	-0.594		0.54

shortest arrows, most aligned in direction, and considered the least ecological relevant were removed prior to CCA in a backward (and sometime forward) stepwise selection until the model was improved and statistically significant.

The CCA analysis (Figure 4) showed that the resistance– environment correlation was good; i.e., there was a moderate relationship between the environmental variables and bacterial resistance (Table 3), and the improved model

A. Klock, *Antibiotic resistant bacteria in an urban river*

was statistically significant (ANOVA,  $F = 1.47$ ;  $P = 0.02$ ).

```
Model: cca(sp.tran~Power+FC.MPNB+
EC.MPNB+Metazoa+Sinuosity+SubAvg+
MngForest+Impervious+pH+Temp,
data = env.tran)
```

The total amount of variance explained by the first two CCAs in the model accounted for 3.57% of the variance of which the environmental variables accounted for 53.1%. The CCA also indicated that the sites scores were heavily skewed by sampling sites in tributaries. A further analysis should be undertaken to address this but additional samples would be required.

Table 3. Summary of the results of the CCA of the resistant bacterial community to the environmental variables.

Total inertia	3.517		
Sum of all constrained eigenvalues	1.868		
	CCA1	CCA2	CCA3
Eigenvalues	0.502	0.472	0.297
Cumulative	14.29	27.73	36.18

The environmental variables were tested for statistical significance by CCA axis and individually (Tables 4 & 5). The three strongest variables were highly correlated to the first two ordination axes: the significant variables in axis 1 were power ( $r=0.661$ ), MPN of environmental coliforms in biofilms ( $r=0.573$ ), and in axis two the presence of metazoans ( $r=0.651$ ) (Table 6). Since the eigenvalues were relatively high for both ordination axes it can be concluded that the distribution of antibiotic resistant bacteria is strongly related to the factors in these axis including many of the off-stream processes that were not measured directly and which drive the strongest variables. For example, stream power is a function of slope, geology, and flow which are all directly affected by hydromodification and land use, the natural environmental bacteria present in biofilms are competitors to resistant strains, and the presence of wildlife increases nutrients and non-resistant coliform counts. Impervious surface had a strong negative correlation in both axis but was not the primary driver as expected (Table 6).

Table 4. Summary table of test of significance of individual axis in CCA.

Axis	F	Pr(>F)
CCA1	3.961	0.006
CCA2	3.727	0.003
CCA3	2.343	0.030
CCA4	2.013	0.065

Table 5. Summary table of test of significance of individual variables in CCA.

Variable	F	Pr(>F)
Power	2.526	0.012
FC.MPNB	2.263	0.024
Metazoa	2.171	0.034

Table 5. Summary table of correlations of variables in CCA (\* significant).

Variable	CCA1	CCA2
Power*	0.661	-0.222
FC.MPNB*	0.573	0.359
Metazoa*	0.137	0.651
Impervious	-0.518	-0.268

## Discussion

The first goal was to determine if antibiotic resistance is spatially variable across a watershed that does have point sources of resistant determinants such as a waste water treatment plant. And yes the distribution of antibiotic resistance appears to cluster in most of the response matrices examined. The response of total antibiotic resistance in the cluster analysis demonstrated three significant splits that aligned with the land use gradient. The first split divided the headwater group from river kilometer (RK) 123 downstream to RK 105 just below the Howard Hansen dam. The second split divided the watershed at RK 45 which corresponds to where the Green River becomes the Duwamish River. This effectively represents a relatively pristine upper watershed, the middle watershed dominated by rural, residential and commercial development, and the industrial and severely polluted lower Duwamish. This is what is expected for the overall gradient in this watershed. Ampicillin resistance tracks perfectly with total resistance as ampicillin is the most prevalent resistance gene examined here. Tetracycline is less common and lacked significant clusters but demonstrated a different longitudinal profile

being primarily impacted by the agriculture in the middle to upper reaches. Tetracycline is a

common additive to animal feed and for animal

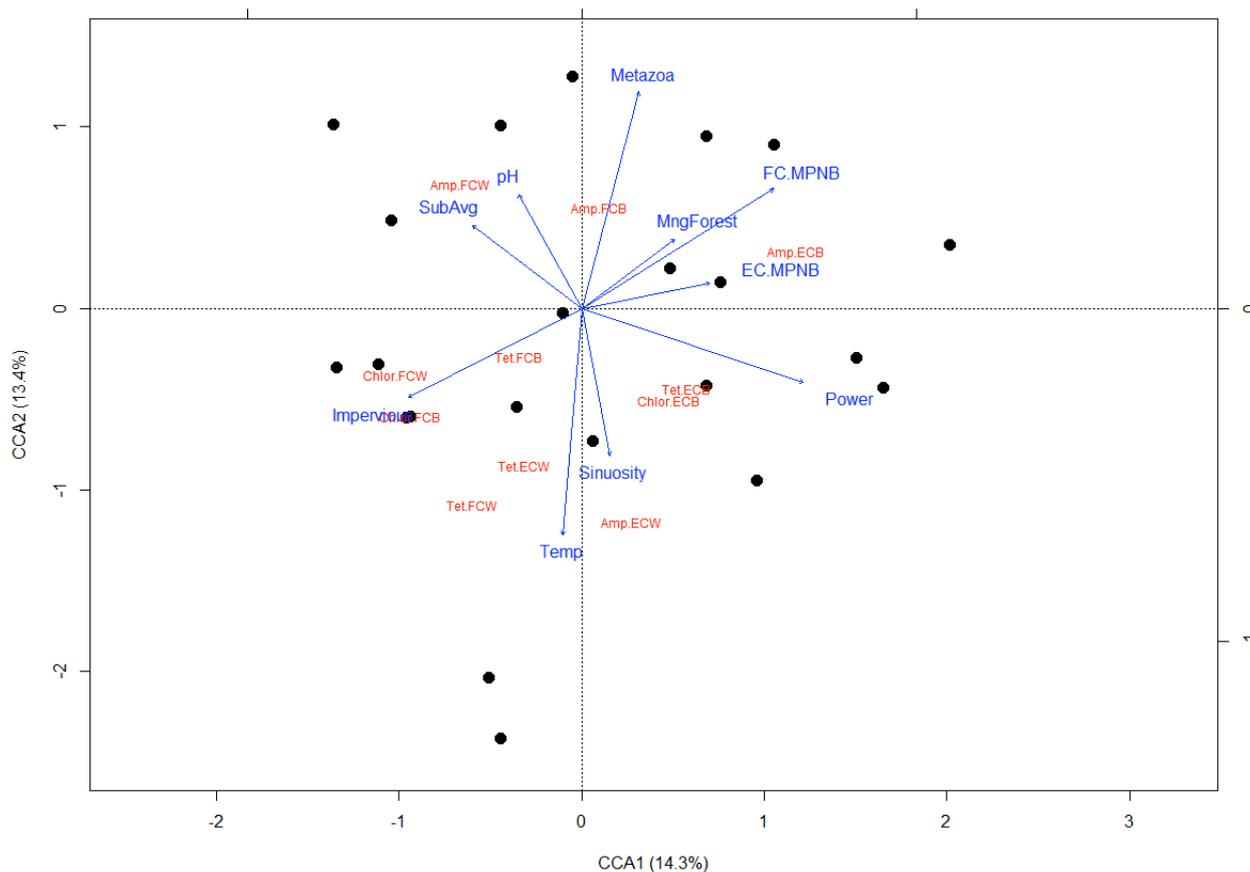


Figure 4. Triplot of the best canonical correspondence analysis (CCA) model. Results from CCA of the different antibiotic resistant components in relation to the strongest environmental variables among 65 samples.

disease treatment but is also quickly degraded by sunlight.

The total amount of resistant bacteria in the water column was less common but more interesting due to transport effects. Again there were three dominant clusters but they were more spatially heterogeneous showing mixed levels among the middle and lower watershed clusters. This could be explained by runoff hot spots such as overpasses, storm drains, or tributary confluences. The subsequent splits in the second cluster corresponded to the area where the Green River becomes the Duwamish. Interestingly the last cluster grouping includes a lower watershed tributary (Longfellow Creek) with sites much higher in the watershed indicating that the tributaries are similar to much less impacted waters although they traverse highly developed areas. This may due

to the retention of substantial amounts of riparian vegetation in these lower tributaries which usually drain residential areas.

Resistance in biofilms was dramatic and there were two dominant clusters. The split occurring at RK 15. The more cluster in the pristine upper parts of the watershed again included Longfellow Creek which enters the Duwamish River near Elliot Bay. The presence of resistance in a biofilm is well known to be advantageous for the long-term persistence of resistant infections and this is the primary mechanism of persistent *Pseudomonas* infections in people with cystic fibrosis (Mena et al. 2008). The physical structure of the biofilm creates a gradient for nutrients which limits the growth of the cells near the bottom of the biofilm. These bottom layers thus reside in a resting state. Antibiotics function by

interrupting processes involved with growth and replication in the targeted bacterial cells, and so these cells must be actively growing for antibiotics to work. Examples include building new cell walls or replicating nucleic acids. Therefore the persistence of resistance in biofilms is a confounding problem for rivers segments that have intense selective pressure due to ARD in the water column. For even if the outer layers of the biofilm are scoured away by a storm event the biofilm overall will remain resistant as the lost colonies are replaced by the resistant cells beneath when they spring to life as nutrients become available.

The difference of clustering between environmental and enteric bacteria was distinct. The first split for natural bacteria was the pristine upper watershed separating from the mid and lower regions. The opposite was true for enteric bacteria where the first split was in the lower watershed. This again is reasonable as enteric (fecal) bacteria that are resistant have gained that resistance from the environment and are more common in the downstream reaches. Environmental coliforms are common in the upper watershed and the Howard Hansen Reservoir had some of the highest levels of coliforms due to wildlife such as a large elk herd and waterfowl that are common in the area. Although these coliforms are common, they rarely exhibit antibiotic resistance due to a lack of selective pressure in the guts of wildlife and in the environment (Appendix B). A primary limitation of this cluster analysis was the small sample size and some important factors that were not collected, such as dissolved oxygen and flow velocity at the substrate. These may be fatal errors which obfuscates the true relationships and environmental drivers.

The second goal was to determine the primary environmental drivers of antibiotic resistance along a typical urban watershed land use gradient. There were three dominant environmental drivers; stream power, the amount of environmental coliforms present in biofilms, and the presence of animals such as stream invertebrates, birds, and large mammals. Less significantly was the percentage of impervious surface but the impact of runoff aligns with geographic distance as it is generally considered to be the primary source of antibiotic resistance determinants found in urban rivers. In this watershed it is the major source of ARD, along to a lesser extent septic

systems and hobby farms which are common in the middle watershed. Interestingly, impervious surface was not statistically significant but appears to be the primary driver at the watershed scale (maybe this is a bias). This is probably an indirect effect as increasing impervious surface increases runoff and the amount of pollutants entering urban rivers.

Stream power is a measure of energy dissipation on the streambed and can be a meaningful index of flow velocity, sediment transport, channel pattern, migration potential, and erosion. Stream power has not previously been associated with antibiotic resistance as there remains a divide between our understanding of macro- and micro-organismal ecology (Bent & Forney 2008). The distribution of any organism is a property of the ecological system and the other organisms.

The natural history of environmental bacteria is largely unknown but we do know that bacteria participate in all the same ecological relationships of more complex organisms including competition, mutualism, and symbiosis (Allison & Martiny 2008). But enteric bacteria are invasive species in the environment when they are delivered in the amounts that occur in urban waterways. These organisms are competitors with environmental bacteria but the presence of wildlife will increase nutrients and non-resistant coliform counts which may have direct impacts on the microbial community dynamics.

Most multivariate statistical techniques are descriptive and should be considered a tool for making inferences (Legendre & Gallagher 2001). Statistics provide a framework for determining how ecological data can be described generally such as determining groupings or clusters, or as an investigative tool to tease out how environmental factors, both individual and cumulatively, determine the structure of ecological communities.

My results support my two questions. There are unique spatial patterns of ARB and there are quantifiable environmental drivers for that distribution. Although WRIA 9 lacks clear point sources of ARD there is clearly a risk in the utilization of an urban waterway such as the Duwamish. There are a number of public parks, and people regularly utilize the river for recreation

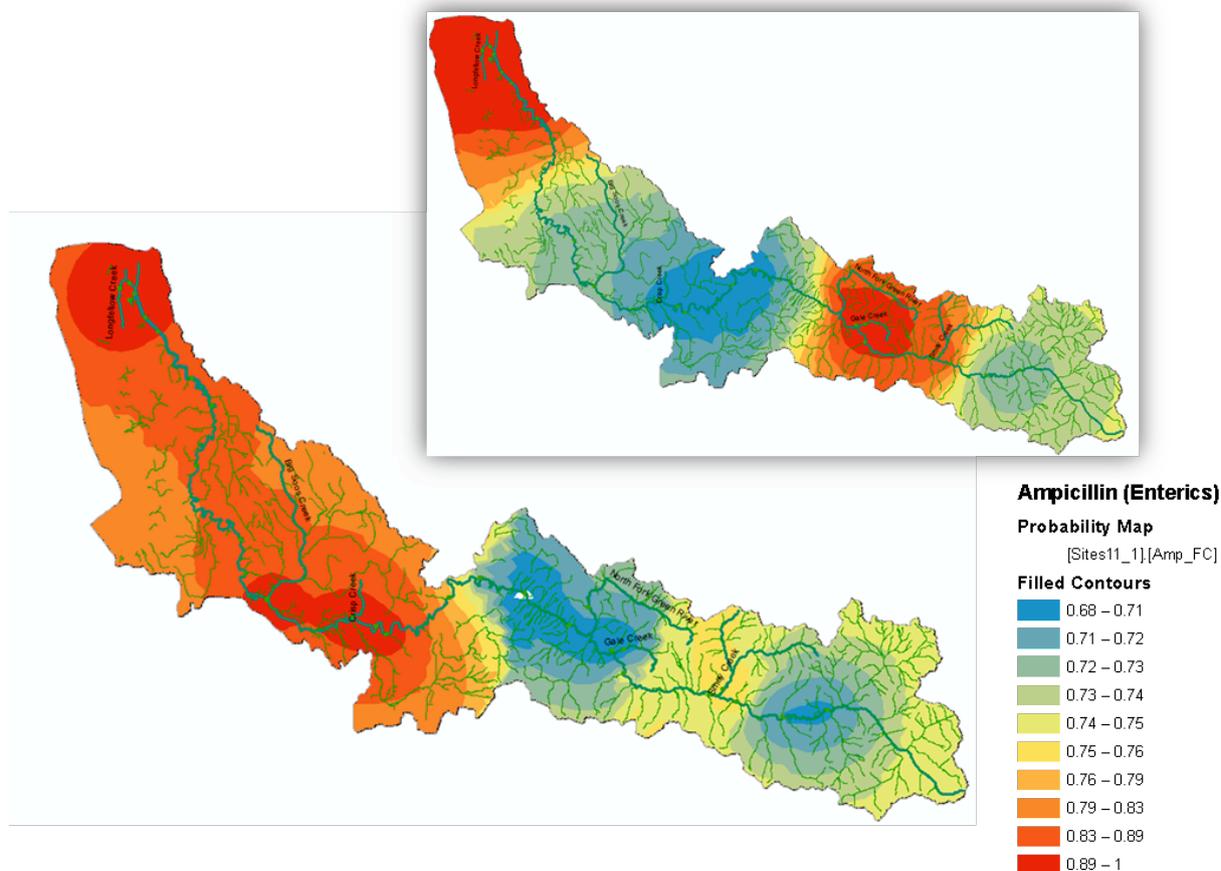


Figure 5. Probability distribution derived from empirical Bayesian kriging across the Green-Duwamish River watershed (WRIA 9) of enteric gram-negative bacteria within biofilms expressing resistance to ampicillin. Inset displays the results for the most probable number (MPN) of enterics from the same samples.

and subsistence fishing. The probability of exposure to the most common resistant bacteria in urban waterways was generated with EBK (Figure 5).

Microbes are the unseen majority and are important and potentially dangerous players in urban watersheds. Bacteria are key to our understanding of larger ecological questions such as nutrient cycling and productivity, and the health and function of the higher trophic levels that primary production support. Asking ecological questions through a microbiological lens may offer unique insights into understanding macro-organism conservation as well.

Ultimately the environmental “clean-up” of resistant genes will only occur if we understand how the environment naturally constrains, mitigates, or removes these agents from the

ecosystem, upon which the long-term health of humans, animals and the environment depend. In a time when the use and allocation of water is stressed this will have impacts on a global scale.

### Acknowledgements

I would like to thank Xingli for his input, availability and openness as an instructor. I would also like to thank my husband who has been exhaustively neglected this quarter.

### References

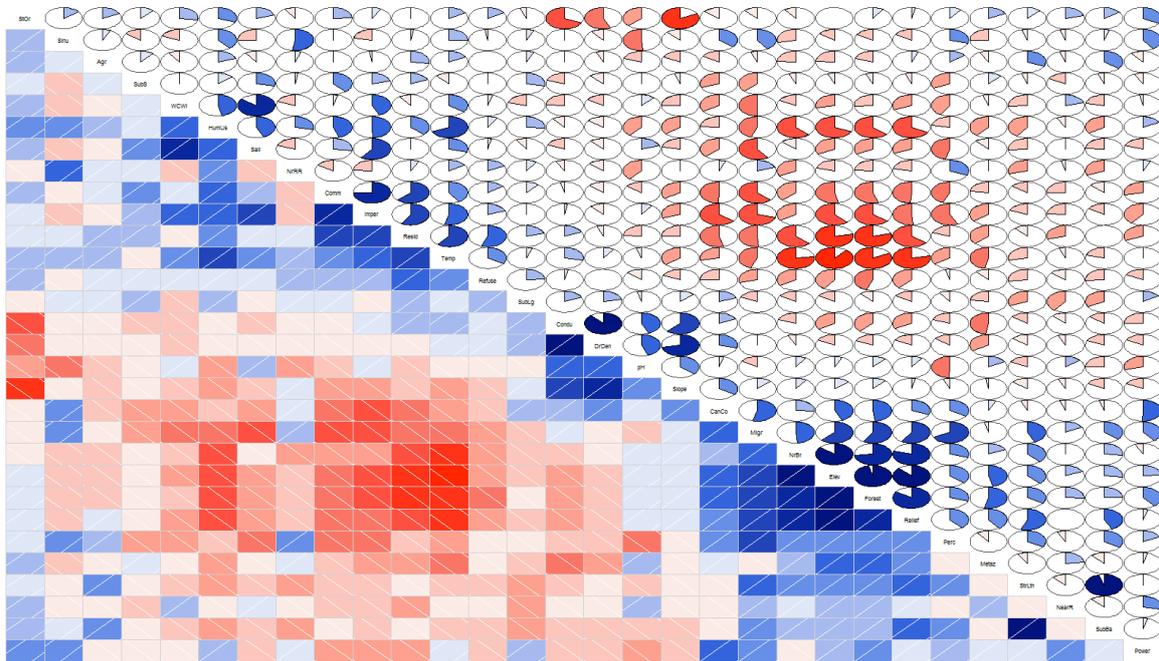
Akiyama, T., and Savin, M.C. 2010. Populations of antibiotic-resistant coliform bacteria change rapidly in a wastewater effluent dominated stream. *The Science of the total environment* 408(24):6192–6201.

- Allison, S.D., and Martiny, J.B.H. 2008. Resistance, resilience, and redundancy in microbial communities. *Proceedings of the National Academy of Sciences USA* 105 (Suppl 1): 11512–11519.
- Ash, R.J., Mauck, B., and Morgan, M. 2002. Antibiotic Resistance of Gram-Negative Bacteria in Rivers, United States. *Emerging Infectious Diseases* 8(7):713–716.
- Baker-Austin, C. et al. 2006. Co-selection of antibiotic and metal resistance. *Trends in Microbiolog*, 14(4): 176–182.
- Baquero, F., Martínez, J., and Cantón, R. 2008. Antibiotics and antibiotic resistance in water environments. *Current Opinion In Biotechnology* 19(3): 260–265.
- Bedan, E.S., and Clausen, J.C. 2002. Stormwater runoff quality and quantity from traditional and low impact development watersheds. *Journal Of The American Water Resources Association*, 45(4):998–1008.
- Bent, S.J., and Forney, L.J., 2008. The tragedy of the uncommon: understanding limitations in the analysis of microbial diversity. *The ISME Journal* 2(7): 689–95.
- Cabello, F.C. 2006. Heavy use of prophylactic antibiotics in aquaculture: a growing problem for human and animal health and for the environment. *Environmental Microbiology* 8(7): 1137–1144.
- Collignon, P. 2015. Antibiotic resistance: are we all doomed? *Internal Medicine Journal* 45(11): 1109–1115.
- Czekalski, N. et al., 2012. Increased levels of multiresistant bacteria and resistance genes after wastewater treatment and their dissemination into lake geneva, Switzerland. *Frontiers in Microbiology* <http://dx.doi.org/10.3389/fmicb.2012.00106>
- Dantas, G., and Sommer, M.O. 2012. Context matters - the complex interplay between resistome genotypes and resistance phenotypes. *Current Opinion in Microbiology* 15(5): 577–82.
- Donnadieu, F. et al. 2013. Characterization and evolution of natural aquatic biofilm communities exposed in vitro to herbicides. *Ecotoxicology and Environmental Safety* 88: 126–134.
- Drozd, M. et al. 2013. Evaluating the occurrence of host-specific, general fecal indicators, and bacterial pathogens in a mixed-use watershed. *Journal of Environmental Quality* 42(3): 713–725.
- Drudge, C.N. et al., 2012. Diversity of integron- and culture-associated antibiotic resistance genes in freshwater floc. *Applied and Environmental Microbiology* 78(12): 4367–72.
- Eisenberg, J.N.S. et al. 2012. In-roads to the spread of antibiotic resistance: regional patterns of microbial transmission in northern coastal Ecuador. *Journal of the Royal Society Interface* 9: 1029–1039.
- van Elsas, J.D., and Bailey, M.J. 2002. The ecology of transfer of mobile genetic elements. *Federation of European Microbiological Societies* 42(2): 187–97.
- Fair, R.J., and Tor, Y. 2014. Perspectives in Medicinal Chemistry Antibiotics and Bacterial Resistance in the 21st Century, *Perspectives in Medicinal Chemistry* 2014:6: 25–64.
- Falkowski, G. et al. 2008. The Microbial Engines That Drive Earth's Biogeochemical Cycles. *Science* 320(5879): 1034–1039.
- Franco, B.E. et al., 2009. The determinants of the antibiotic resistance process. *Infection and Drug Resistance*, 2(1), pp.1–11.
- Garcia-Armisen, T. et al. 2011. Antimicrobial resistance of heterotrophic bacteria in sewage-contaminated rivers. *Water Research* 45(2): 788–796.
- Gordon, D.M., and Cowling, A. 2003. The distribution and genetic structure of *Escherichia coli* in Australian vertebrates: Host and geographic effects. *Microbiology* 149(12): 3575–3586.
- Hipp, J.A. et al. 2006. Optimization of stormwater filtration at the urban/watershed interface. *Environmental Science and Technology* 40(15): 4794–4801.
- Hirsch, R. et al. 1999. Occurrence of antibiotics in the aquatic environment. *The Science of the Total Environment*. 225(1-2): 109–118.
- Islam, M.Z. et al. 2014. Regional variation in the prevalence of *E. coli* O157 in cattle: A meta-analysis and meta-regression. *PLoS ONE* 9(4): e93299.
- Johnson, K. a. et al. 2011. Biotic responses to low-concentration urban road runoff. *Journal of the North American Benthological Society* 30(3): 710–727.
- Kassen, R., Llewellyn, M., and Rainey, P.B. 2004. Ecological constraints on diversification in a model adaptive radiation. *Nature* 431: 984–989.

- Keen, P.L., and Patrick, D.M. 2013. Tracking Change: A Look at the Ecological Footprint of Antibiotics and Antimicrobial Resistance. *Antibiotics* 2:191–205.
- Kümmerer, K. 2009. Antibiotics in the aquatic environment--a review--part II. *Chemosphere* 75(4): 435–41.
- Lear, G., and Lewis, G.D. 2009. Impact of catchment land use on bacterial communities within stream biofilms. *Ecological Indicators* 9(5): 848–855.
- Leff, L.G., Vaun McArthur, J., and Shimkets, L.J. 1993. Spatial and temporal variability of antibiotic resistance in freshwater bacterial assemblages. *FEMS Microbiology Ecology*, 13(2): 135–143.
- Legendre, P., and Gallagher, E. 2001. Ecologically meaningful transformations for ordination of species data. *Oecologia* 129(2): 271–280.
- Liu, Y. et al. 2011. Stormwater runoff characterized by GIS determined source areas and runoff volumes. *Environmental Management* 47(2): 201–217.
- Lowell, J.L. et al. 2009. Habitat heterogeneity and associated microbial community structure in a small-scale floodplain hyporheic flow path. *Microbial Ecology*, 58(3): 611–20.
- Lupo, A., Coyne, S., and Berendonk, T.U. 2012. Origin and evolution of antibiotic resistance: the common mechanisms of emergence and spread in water bodies. *Frontiers in Microbiology* 3: 18.
- Maal-Bared, R. et al. 2013. Phenotypic antibiotic resistance of *Escherichia coli* and *E. coli* O157 isolated from water, sediment and biofilms in an agricultural watershed in British Columbia. *The Science of the Total Environment* 443: 315–323.
- Martinez, J.L., Fajardo, A., et al. 2009. A global view of antibiotic resistance. *FEMS Microbiology Reviews* 33(1): 44–65.
- Martinez, J.L., Sánchez, M.B., et al. 2009. Functional role of bacterial multidrug efflux pumps in microbial natural ecosystems. *FEMS Microbiology Reviews* 33(2): 430–449.
- Martínez, J.L., and Baquero, F. 2014. Emergence and spread of antibiotic resistance: setting a parameter space. *Upsala Journal of Medical Sciences* 119(2): 68–77.
- Mazaris, A.D. et al. 2010. Biogeographical patterns of freshwater micro- and macroorganisms: a comparison between phytoplankton, zooplankton and fish in the eastern Mediterranean. *Journal of Biogeography* 37(7): 1341–1351.
- McArthur, J.V. et al. 2011. When rivers become reservoirs of antibiotic resistance: industrial effluents and gene nurseries. Proceedings of the 2011 Georgia Water Resources Conference, April 11-13, 2011, University of Georgia.
- McArthur, J.V., and Tuckfield, R.C. 2000. Spatial Patterns in Antibiotic Resistance among Stream Bacteria: Effects of Industrial Pollution Applied and Environmental Microbiology 66(9): 3722–3726.
- McArthur, T. and, 2006. Down-stream spatial distribution of antibiotic resistance along metal contaminated stream reaches. *Microbes and Environments*.
- Mcgrady-Steed, J., Harris, P.M., and Morin, P.J. 1997. Biodiversity regulates ecosystem predictability. *Nature* 390: 162–165.
- Mena, A. et al. 2008. Genetic adaptation of *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients is catalyzed by hypermutation. *Journal of Bacteriology*, 190(24): 7910–7917.
- O'Neill, J. 2014. Antimicrobial Resistance: Tackling a crisis for the health and wealth of nations. Review on Antimicrobial Resistance.
- Palmer, M. a., Menninger, H.L., and Bernhardt, E. 2010. River restoration, habitat heterogeneity and biodiversity: a failure of theory or practice? *Freshwater Biology* 55: 205–222.
- Pei, R., Storteboom, H., and Pruden, A., 2006. Antibiotic Resistance Genes (ARG) as Emerging Environmental Contaminants. World Environmental and Water Resource Congress, pp. 1-10.
- Postel, S.L., Daily, G.C., and Ehrlich, P.R. 1996. Human Appropriation of Renewable Fresh Water. *Science* 271(5250): 785–788.
- Pruden, A., 2014. Balancing water sustainability and public health goals in the face of growing concerns about antibiotic resistance. *Environmental Science, and Technology* 48(1): 5–14.
- Pruden, A. et al. 2013. Management options for reducing the release of antibiotics and antibiotic resistance genes to the environment. *Environmental Health Perspectives* 121(8): 878–885.

- Rankin, D.J., Rocha, E.P.C., and Brown, S.P. 2011. What traits are carried on mobile genetic elements, and why? *Heredity* 106: 1–10.
- Rizzo, L. et al. 2013. Urban wastewater treatment plants as hotspots for antibiotic resistant bacteria and genes spread into the environment: a review. *The Science of the Total Environment* 447: 345–60.
- Rompré, A. et al. 2002. Detection and enumeration of coliforms in drinking water: current methods and emerging approaches. *Journal of Microbiological Methods* 49(1): 31–54.
- Singer, G. et al. 2010. Physical heterogeneity increases biofilm resource use and its molecular diversity in stream mesocosms. *PloS ONE* 5(4): e9988.
- Singer, R.S., Ward, M.P., and Maldonado, G. 2006. Can landscape ecology untangle the complexity of antibiotic resistance? *Nature Reviews Microbiology* 4(12): 943–952.
- Smalla, K., and Sobczyk, P.A. 2002. The prevalence and diversity of mobile genetic elements in bacterial communities of different environmental habitats: insights gained from different methodological approaches. *FEMS Microbiology Ecology* 42: 165–175.
- USEPA. 2002. National Water Quality Inventory Report to Congress 2004 Reporting Cycle. (States Environmental Protection Agency EPA-841-R-02-001. Office of Water, Washington D.C.). Available at: [http://water.epa.gov/lawsregs/guidance/cwa/305b/upload/2009\\_01\\_22\\_305b\\_2004report\\_2004\\_305Breport.pdf](http://water.epa.gov/lawsregs/guidance/cwa/305b/upload/2009_01_22_305b_2004report_2004_305Breport.pdf).
- Viau, E.J. et al. 2011. Bacterial pathogens in Hawaiian coastal streams—associations with fecal indicators, land cover, and water quality. *Water Research* 45(11): 3279–3290.
- Wellington, E.M.H. et al. 2013. The role of the natural environment in the emergence of antibiotic resistance in gram-negative bacteria. *The Lancet Infectious Diseases*, 13(2): 155–65.
- WHO 2014. Antimicrobial Resistance Global Report on Surveillance.
- Zhang, W. et al. 2009. Accumulation of tetracycline resistance genes in aquatic biofilms due to periodic waste loadings from swine lagoons. *Environmental Science, and Technology* 43(20): 7643–7650.
- Zhang, X.-X., Zhang, T., and Fang, H.H.P. 2009. Antibiotic resistance genes in water environment. *Applied Microbiology and Biotechnology* 82(3): 397–414.

**Appendix A.** Correlation matrix of explanatory variables



A. Klock, *Antibiotic resistant bacteria in an urban river*

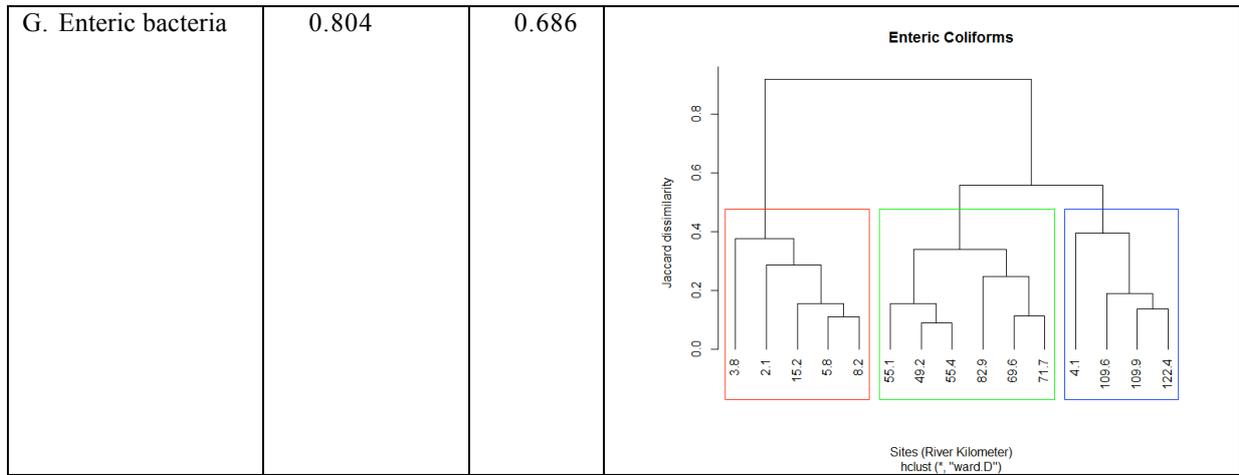
**Appendix B.** Hierarchical Cluster Analysis plot for total all response variables across WRIA 9.

Matrix	Agglomerative coefficient	Cor	Bootstrap AU ( $\alpha=0.95$ )
A. Total Antibiotic Resistance	0.888	0.585	<p style="text-align: center;"><b>Total Antibiotic Resistance</b></p> <p style="text-align: center;">Sites (River Kilometer) hclust("ward.D")</p>
B. Ampicillin	0.888	0.585	<p style="text-align: center;"><b>Ampicillin Resistance</b></p> <p style="text-align: center;">Sites (River Kilometer) hclust("ward.D")</p>
C. Tetracycline	0.716	0.794	<p style="text-align: center;"><b>Tetracycline Resistance</b></p> <p style="text-align: center;">Sites (River Kilometer) hclust("ward.D")</p>

A. Klock, *Antibiotic resistant bacteria in an urban river*

D. Water	0.781	0.571	<p style="text-align: center;"><b>Resistance in Water</b></p> <p style="text-align: center;">Sites (River Kilometer) hclust("ward.D")</p>
E. Biofilm	0.836	0.726	<p style="text-align: center;"><b>Resistance in Biofilms</b></p> <p style="text-align: center;">Sites (River Kilometer) hclust("ward.D")</p>
F. Environmental bacteria	0.83	0.58	<p style="text-align: center;"><b>Environmental Coliforms</b></p> <p style="text-align: center;">Sites (River Kilometer) hclust("ward.D")</p>

A. Klock, *Antibiotic resistant bacteria in an urban river*



## Different words for the same thing: developing and testing composite vulnerability indices of fishing communities using multivariate techniques

J. Zachary Koehn

*School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA, USA*

*E-mail: [zkoehn@uw.edu](mailto:zkoehn@uw.edu)*

Received December 2015; accepted January 2016; published September 2016

### Abstract

Successful management of fisheries depends on accurate assessments of fishing activity and its impacts, including an increasing focus on social aspects of the fishery. Recent amendments to federal law place increased value on social dimensions of the fishery. Responses from federal fishery managers have sought to create vulnerability indices to aid in measuring how fishery management changes affect dependent communities. Developing a cost-effective approach that can aid in the development of vulnerability indices is an important step forward. This research provides a novel approach to the development and testing of social vulnerability indices with a study of the 1995 transition of Alaska's pacific halibut long-line fishery to individual fishing quotas. Using nonmetric multidimensional scaling procedures, social and fishery indicators are composed into single-dimension index scores for all Alaskan communities. Communities were separated by their dependence on halibut into treatment or control groups. A differences-in-differences approach was used to test if differences in index scores associated with halibut dependent treatment communities before and after the 1995 change were significantly different from control communities. Using case studies to corroborate the index results, it was found that the analysis successfully generated composite indices that reflected known changes following the management change.

### Introduction

Indicator approaches have long been used to monitor the status of fisheries. While biological indices related to the fish stock are predominant, the increasing interest in ecosystem-based management, including impacts to surrounding biotic and abiotic components and human communities, has led to the development of a more holistic suite of social and ecological indices. Among federally managed fisheries in the United States, the 2006 Reauthorization of the Magnuson Stevens Fishery Act set 10 national standards to direct American fishery management plans and policy objectives. National Standard 8 (16 U.S.C. §1851(a)(8)) focused explicitly on the need for managers to consider human communities

related to fisheries, and “provide for their sustained provide for their sustained participation and to minimize adverse economic impacts.” The National Standard directed a research agenda for research focusing on the human dimensions of federal fisheries by the National Marine Fisheries Service (NMFS) including the development of vulnerability indices.

In response to this National Standard, social scientists in the NFMS set about developing a methodology to create cost-effective social indices of vulnerability and well-being using readily available census and fisheries data to inform future management and understand socio-economic impacts to fishing communities. Much of this work was based on the foundational Social Vulnerability Index

developed by Cutter and colleagues, which used socioeconomic and demographic data to create an index of social vulnerability to environmental hazards (2003). The Gulf of Mexico was the first region to tailor the social vulnerability indices to understand management change impacts to fisheries under National Standard 8 (Jepson and Jacob 2007). The NMFS Fishery Science Centers have since defined a working definition of vulnerability as the “states of susceptibility” of fishing communities creating the potential for harm, including pre-existing characteristics and patterns of differential access to resources (Jepson and Colburn 2013). To this date, social vulnerability indices have been developed for fishing communities on the East Coast, on the West Coast, in the Pacific Islands, and in Alaska (Jepson and Colburn 2013; Jacob et al 2013; Norman et al 2012). While the indices share many characteristics across regions; they are not identical and reflect unique traits of communities in the region. For example, social scientists at the Alaska Fishery Science Center modified the indices to reflect unique traits related to Alaskan fishing communities’ characteristics of rural communities and increased emphasis on subsistence harvest.

The purpose of this paper is to present an approach that tests the accuracy of the social vulnerability indices developed by NMFS with a retrospective analysis of case studies with known impacts. The analysis compares data before and after a selected disturbance to understand how sensitive a given suite of indices is to the change. In the present case, we will be testing “pulse disturbances” that occurred in a single event (i.e. management change), rather than a “press disturbance,” which occur to various intensities throughout the study period (e.g. ocean acidification). We have selected two specific fisheries management actions that has had known impacts on fishing communities but differed in the geographic distribution of the fishery as well as the diversity of stakeholders: the 1995 transition to ITQ management for the Alaskan commercial halibut fleet (hereafter Halibut Rationalization).

Alaskan fishing communities consistently rank among the highest landings and gross revenue among American harbors. In addition, most are characterized as small towns, and are largely dependent on their fisheries. Management changes that redistribute access to

fisheries such as Halibut rationalization may have disproportionately large adverse effects on smaller, less economically resilient communities (Carothers et al 2010). Alaskan communities should thus, in theory, be sensitive to large changes occurring in commercial fisheries they harvest. This study will employ AFSC’s social vulnerability indices to test how well the indices reflect changes that were felt in Alaskan fishing communities as a result of both the 1995 Halibut Rationalization. First, we provide a brief history of management in the halibut fisheries. Second, an approach is developed to compose indicators of vulnerability into individual index scores by community for snapshot years 1990 and 2000. Third, a methodology is developed to categorize communities for hypothesis testing. Finally, the indices are tested using the differences-in-differences econometric approach to measure if there were significant differences before and after the management change between halibut dependent fishing communities and fishing communities not focused on fishing for halibut. Finally, statistically significant indices are compared to see how well indices reflected known impacts in select communities. Ultimately, social indices could provide NMFS inexpensive, comparatively efficient way to assess potential impact of management or biophysical stressors on fishing communities using secondary data. Before using the indices as an assessment tool, they must first be tested for their ability to accurately reflect changes in the fishery system. This research presents a rigorous approach to the development and testing of vulnerability indices representing impacts to local fishing industry as well as dependent fishing communities and livelihoods.

#### *History of Alaskan halibut management*

Pacific halibut fishing in Alaska has enjoyed a long history, stretching back to the 1880s. The fishery occurs close to shore, and a vast majority of the landings occur in Southeast Alaska (McDowell Group Inc., 2007). Communities of varying size and demography participate in the halibut fishery, either through vessels residing in their harbors or processing facilities within community limits. Prior to rationalization, halibut was managed under a Total Allowable Catch (TAC) set for each region prior to the season opening. Due to profitability of the fishery, the number of vessels entering the fishery increased each year.

This resulted in ever-shortened seasons - in some areas the TAC was reached in 24 hours – that made it difficult to fish halibut full time (Hartley and Fina 2001). In the midst of increasing number of vessels and the desire to secure a profitable proportion of the landings, there was also growing concern that the fishery was becoming unsafe. Now considered a classic tragedy of the commons case, fishery managers in the North Pacific turned to a neoliberal, market-based approach to decrease overcapacity (Mansfield 2004).

In 1995, the Pacific Halibut fishery was transferred system using individual fishing quotas (IFQs) to address concerns about vessel overcapacity, economic efficiency, and safety on the water (U.S. Dep't of Commerce).<sup>1</sup> The IFQ program limited access to the fishery by allocating quotas to fishermen with historic participation in the fishery. A body of literature has developed around the economic and social impacts of the Halibut Rationalization. One direct effect was an immediate decrease in the number of vessels participating in the fishery. In the years following rationalization, the seasons elongated and allowed the industry shift from low-value frozen to high-value fresh product, and exvessel revenues of fishermen have also increased (Knapp 1997; Matulich and Clark 2003). Consolidation has been the focus of research on negative impacts regions where there was a net sale of quota away from the community. Knapp found that fishermen who received larger shares of quota were generally positive about the management change while those with smaller shares of quota were generally negative towards rationalization (1997). More recently, Carothers et al. found a disproportionate loss of IFQs in small remote communities, especially those with Alaska Native Heritage (2010), raising concern of inequitable impact on these communities that are less “efficient” harvesters of halibut but due to their remoteness are presumably more dependent on the resource.

## Methods

### *Data collection and aggregation:*

The method used to test the accuracy of the indices to reflect management change impacts on fisheries

engagement and reliance created a composite index of census and fisheries data and separated them into treatment and control groups. The present analysis included the variables that were used in the construction of social and fisheries indices during previous phases of this project (Himes-Cornell and Kasperski 2015, Kasperski and Himes-Cornell 2014). These variables were also utilized in the development of indices focused on fishing communities in the Gulf of Mexico (Jacob et al. 2013). The methodological approach used here leveraged secondary data sourced from multiple government-based databases, the periodicity with which the original data was collected varied by source and aggregating the data required considerable effort. Fisheries data largely came from the Alaska Fishery Information Network Database, which aggregates federal and state fisheries data annually and at the community level. Fisheries data was collected for each of the variables in each index for 1991. Halibut Rationalization analysis used indices of variables before and after (1991 data and 2000 data respectively) the 1995 management change. Note that 1991 data was used instead of 1990 because of concerns with data quality that were addressed during the database collection the following year. Also note that recreational engagement and reliance data was only available from 1993 onward, so the “before” data for Halibut Rationalization is actually 1993 data and not 1991 data. Caution should be used when interpreting the recreational indices for the 1990 index year, as many of these are based on a single variable. Recreational engagement and reliance for 2000 is based on better data.

Census data was not annually available for all communities. Prior to 2005, demographic and social census-based data among small communities was collected using the decennial census. Following 2005, the U.S. Census Bureau established the American Community Survey (ACS) in order to collect demographic and social information annually from a 7% national sample of the population. Fortunately, Halibut rationalization social variables both originated from the decennial census. Even between decennial censuses, the data collection was methodologies saw a slight change and variables were recorded differently. It should be noted that the slight methodological change could lead to misrepresentation in the results but that there was no better approach to handling this data.

---

<sup>1</sup> Note that halibut and sablefish both transitioned over into similar IFQ programs at the same time. Only a treatment of the halibut fishery will be conducted here.

Raw fisheries data was used to create values of engagement and reliance for each community, to better understand different dimensions of the fishing communities. Engagement is an unaltered value reflecting overall participation in fishing relative to other fishing communities. Reliance is a per capita value of the community's participation in fishing relative to its population as reported by the Department of Labor for year 1990 and then due to data limitation with Department of Labor population estimates switched to the United States Census Population estimate for 2000. Alaska-specific indices included both "Social" (Appendix 1) and "Fishery" variables (Appendix 2). The indices included between three and five variables each. Data were cleaned for each index, excluding communities that had missing variable values. Any community with one or more missing value was excluded. Furthermore, any community with a factor score of 0 for both variables was also excluded.

*Creating index scores using nonmetric multidimensional scaling:*

Each index was created by combining a set of conceptually relevant variables into a single score using nonmetric multi dimensional scaling (nMDS) procedures, incorporating the methodology developed by Jepson and Colburn (2012) and modified by AFSC to take into account Alaskan community characteristics (Himes-Cornell and Kasperski 2015 and Kasperski and Himes-Cornell 2014). This project diverged from previous NMFS Social Vulnerability research by electing to use the more flexible properties of nMDS and not to use factor analysis procedures. Multidimensional scaling is an established procedure across disciplines ranging from ecology to sociology to political science (Hout et al 2013).

Index scores were extracted from nMDS rather than factor analysis for a number of reasons.<sup>2</sup> While both nMDS and factor analysis are ordination techniques from which underlying structures (i.e. dimensions in nMDS and factors in factor analysis) can be derived when represented in "continuous coordinate space" (Jaworska and Chupetlovska-Anastasova 2009). At each iteration, nMDS focuses on

---

<sup>2</sup> Note that all indices were computed using both nMDS and factor analysis for comparison. However, where nMDS stress levels were low enough to suggest they successfully collapsed into a single dimension (i.e., index score), factor analysis results indicated that many of the social indices should have been further separated into multiple factors.

improving the relationship between variables until the stress, until the best fit is found (see stress in Section 3.1). One theoretical benefit useful to this research is that nMDS is better suited to modeling non-linear relationships between variables (Hout et al 2013; Olden 2015). While one could see where linear relationships exist among fishery indicator variables (e.g., increasing revenues increase linearly with the number of vessels in a community), it cannot be assumed that social variables have linear relationships. For example, personal disruption variables may exhibit nonlinear relationships where percentage of separated females the percentage of unemployed might not show correlation until particularly high percentages of both. If factor analysis were to be conducted, we would assume that the percentages of separated females and of unemployment would be linearly related. With the aim of the project to create single index score for each suite of indicators, nMDS proved the better ordination approach. The nMDS analysis was conducted using the *metaMDS* function in R package "vegan" (Oksanen et al 2015). Because of the large number of zeros and presence of data gaps, Gower's distance coefficient was used. Each index-year was computed using a maximum of 100 iterations to give the model more opportunity minimize stress and increase dimensional convergence in the model.

The variables were separated by year and single nMDS was conducted with a two-year panel, from 1990 and 2000 data for halibut rationalization. The limited number of variables in each index and the objective to create a single score for each community lead to the selection of single-dimension that parallels previous work at AFSC to create vulnerability indices (Himes-Cornell and Kasperski 2015 and Kasperski and Himes-Cornell 2014). Scores were then extracted for each index and each year period separately, which are the base data for the econometric approach utilized below.

*Hypothesis creation:*

Control and treatment groups must be selected for statistical tests on impact of rationalization on fishing communities. In this case, we used a simple metric of participation across a number of fishing and processing variables to indicate whether or not a community was dependent on halibut. Variables used to indicate halibut dependence included harvester exvessel

revenue, harvester landings, harvester count of vessels, processor revenue, processor pounds, and count of processors. Each halibut specific variable was divided by the total variable accounting all species of fish to create a ratio (i.e., halibut exvessel revenue / total exvessel revenue). A community was considered halibut dependent if that community had a ratio of halibut to total greater than or equal to 20%. The cutoff of 20% was chosen based on previous NMFS documentation. For example, the halibut dependent community of Mekoryuk had an exvessel value for halibut that was 88% of the total exvessel revenue generated in the community for a given year, whereas the community of Juneau had a harvester exvessel ratio of only 14% and was thus not halibut-dependent. In order to limit the influence of annual outliers, we conducted this analysis in the 4 years leading up to rationalization. Any community with a single variable ratio greater or equal to 20% was considered halibut dependent and included in the treatment group. All other communities were included in the control. Treatment and control groups were then merged with the nMDS index scores for each year.

*Statistical framework:*

A statistical framework was then developed to measure if there were significant differences between treatment and control groups over the two time periods using a difference-in-differences approach, used to analyze the impact on employment before and after minimum wage was increased in the treatment state of New Jersey versus the control state of Pennsylvania (Card and Krueger 1993). In the context of halibut rationalization, difference-in-differences approach was used to measure directional change for each social and fishery index before and after the 1995 rationalization and whether or not that change was significantly different between halibut dependent and fish reliant-engaged communities (Table 1). Because much index score data approximately followed normal distribution, with a mean of 0 and standard deviation near 1 for all indices, a Student's t-Test was then used to test whether there were significant differences between treatment and control populations.

Table 1. Difference-in-Differences Approach for fish (control -"c") and halibut (treatment- "t") communities between 1990 and 2000.

	1990	2000	Difference
Control	$V_{c,1990}$	$V_{c,2000}$	$V_{c,2000}-V_{c,1990}$
Treatment	$V_{t,1990}$	$V_{t,2000}$	$V_{t,2000}-V_{t,1990}$
Difference-in-differences (DiD) =			
$(V_{t,2000} - V_{t,1990}) - (V_{c,2000}-V_{c,1990})$			

**Results**

Results of the analysis are divided into two segments. First, the nMDS approach on how well the single-dimension index score represented the raw indicator data. Second, the differences-in-differences analysis of halibut-dependent treatment communities versus non-halibut dependent control communities showed promise as a flexible method to compare impacts of management on fishing dependent communities. In brief, the nMDS scores were found successful in representing the raw data and the differences-in-differences analysis picked up differences between treatment and control.

*Nonmetric multidimensional scaling:*

Composite index scores from multiple indicator variables were generated using nMDS by setting the desired number of dimensions to one. Because separate nMDS procedures were conducted on each of 12 indices, the complete summary of nMDS results for each index will not be reported here but is available upon request. Given the desire to reduce variables into a single dimension, the important summary statistic shown here is the stress value for each nMDS result. Stress can be viewed as a goodness-of-fit measure of the correspondence between ranked distances of the original raw data matrix and ranked multivariate distances in reduced-dimensional space. Having low-dimensionality increases stress, and as such is an important indicator of the accuracy of the nMDS to represent each indicator. In this case, it measures correspondence between the ranked distances in the original data matrix and the single-dimensional result. Typically, stress values over 0.3 would be rejected as not having a high enough correspondence, whereas extremely low (i.e,  $1*10^{-4}$ ) values would

suggest too few samples or variables for the required dimensionality. The general rule-of-thumb for acceptable stress varies would have stress values no more than 0.3, where good representations are under 0.1 and very good representations are under 0.05.<sup>3</sup>

Index stress levels for each year indicated that reduced, single-dimension result satisfactorily represented the original indicator variables: generally stress scores were between 0.1 and 0.25 (Table 2). Personal disruption saw the highest stress values in 1990 (0.35), suggesting that the index scores associated with personal disruption should be cautiously viewed. Furthermore, the lack of data on all but a single variable in recreational engagement and reliance index specifically for 1990 lead to a near-zero stress score, meaning that the single recreational variable was reducing to a single dimension almost perfectly, limiting the usefulness of the index score for 1990. Indices for processor engagement and reliance showed particularly strong correspondence across all years (stress<0.1), and thus should be considered the index scores most reflective of the raw indicator data. Other than personal disruption in 1990, and recreational engagement and reliance in 1990, it appears the nMDS acceptably reduced indicator variables to a single-dimension score.

Table 2. Summary of nMDS stress values for each index and each year.

Index	1990	2000
Personal disruption	0.346	0.048
Population composition	0.202	0.184
Poverty	0.234	0.267
Housing disruption	0.164	0.263
Housing characteristics	0.229	0.188
Labor	0.240	0.287
Fishing harvest engagement	0.267	0.198
Fishing harvest reliance	0.185	0.195
Fishing processor engagement	0.056	0.098
Fishing processor reliance	0.060	0.107
Recreational engagement	NA	0.249
Recreational reliance	NA	0.239

*Differences-in-differences:*

<sup>3</sup> A 1975 paper from Rabinowitz 1975 has slightly different, higher values for acceptance, in References section below.

Treatment communities dependent on halibut saw significant differences from non-halibut dependent control communities in many index scores (Table 3). Personal disruption, poverty, housing disruption, labor, fishing harvest engagement, fishing harvest reliance, recreational fishing and recreational reliance all had p-values lower than 0.1, indicating there was a significant difference between treatment and control groups. All index comparisons had large degrees of freedom. Note that personal disruption, recreational engagement and reliance also had very low p-values, but that the 1990 index score should be doubted due to either unacceptably high stress for personal disruption in 1990 or near-zero stress in the case of 1990 for recreational engagement and reliance. Poverty and fishing harvest engagement showed the strongest p-values (p=3.52E-7, 3.69E-10).

Further interpretation of these results is complicated by the complex nature of the indices (for more, Section 4.3). Remember that index values compared in treatment and control groups are the differences between index scores “after” and “before” halibut rationalization, presenting two important details for interpreting the results. 1) If a difference value is positive then that the index scores were closer to zero before treatment than after rationalization, indicating an index score after rationalization that is further from zero and more of a change. 2) If treatment values are greater than control values it suggests the magnitude of the change was more pronounced in treatment communities than in control communities.

Keeping those details in mind, trends emerged from the statistical approach when considering the differences in mean values of treatment and control for indices with significant differences. Fishing harvest engagement and reliance indices both showed very similar significant differences for halibut dependent treatment communities from non-halibut dependent control communities. Fishing harvest engagement and reliance indices both showed a much stronger, negative movement away from zero for treatment communities. This suggests that halibut dependent communities experienced more negative change after rationalization (for more, Section 4.1).

Table 3. Summary results of the differences-in-differences Student's t-Test.

Index	Treatment Deg. Free.	Control Deg. Free.	Treatment Mean	Control Mean	t- statistic	Lower Conf. Interval	Upper Conf. Interval	P- value
Personal disruption	40	160	0.122	-0.019	2.681	0.022	0.150	0.009
Population composition	41	160	-0.015	0.006	-0.872	-0.082	0.032	0.387
Poverty	31	103	-0.008	-0.102	5.496	0.259	0.553	4E-07
Housing disruption	21	39	0.020	0.070	-2.788	-0.476	-0.078	0.007
Housing characteristics	37	116	-0.001	-0.001	0.272	-0.055	0.073	0.786
Labor	41	160	-0.023	0.055	-3.220	-0.354	-0.083	0.002
Fishing harvest engagement	55	152	-0.503	-0.020	6.607	0.166	0.307	4E-10
Fishing harvest reliance	55	154	-0.509	-0.025	1.772	-0.006	0.107	0.078
Fishing processor engagement	31	15	0.087	-0.038	0.547	-0.310	0.534	0.589
Fishing processor reliance	30	15	0.120	-0.051	0.000	-0.428	0.428	1.000
Recreational engagement	56	245	-0.538	-0.044	5.679	0.190	0.396	2E-07
Recreational reliance	56	239	0.218	-0.022	4.750	0.104	0.253	8E-06

Interpreting the results of significant social indices reflect a similarly complicated interpretation. The poverty index showed a smaller magnitude of change away from zero among treatment groups, even though both groups showed decreases following management change. This suggests that control groups saw a more favorable decrease in mean poverty index for control communities (mean = -0.10) versus halibut-dependent treatment communities (mean = -0.01). The housing disruption index saw a smaller magnitude of change among treatment than control communities, but both treatment and control groups showed an increase in housing disruption following rationalization. Finally, the labor index saw a smaller magnitude but negative change in treatment communities and a larger magnitude but positive change among control communities. These significant results suggest that the labor index dropped among halibut-dependent communities after rationalization and increased among control communities, but that the t drop among

treatment communities was not as large a change as that experienced by non-halibut dependent communities.

### Discussion

*Engagement and reliance, harvesting and processing: Different words for the same thing?:*

Engagement and reliance were separated in order to reflect potential differences between smaller and larger communities. It was anticipated that engagement would better reflect of total participation in the fishery (i.e., the biggest communities and players on the halibut industry) whereas the per-capita measure of reliance would better reflect smaller communities that might have lower output but where fishing was still important (i.e., communities dependent on fishing). It was further anticipated that communities involved in fishing would differ from those involved in processing. The results of this analysis indicated almost identical levels of significance between

engagement and reliance for fishing harvest and for processing. While fish harvest showed significant, negative trends for halibut dependent communities following rationalization, processing did not show a significant difference. When comparing respective engagement and reliance indices, both fishing indices showed nearly identical results. A similar observation could be made when comparing engagement and reliance indices in the processing sector, though there was no significant difference between treatment and control in processing engagement ( $p=0.58$ ) and in processing reliance ( $p=0.99$ ). First, this suggests that for the case of halibut rationalization, impacts on total engagement and per-capita reliance could be considered equivalent. It also suggests that halibut dependent communities participating in halibut harvest sector were more sensitive to rationalization than were halibut dependent communities participating in the halibut processing sector. While engagement and reliance were largely reporting the same results of Halibut Rationalization, it should be noted that in regions with larger populations, it is possible that per capita reliance indices would more accurately measure unique aspects of fishing communities.

*Limitations of census-based data with low-population American communities:*

Census-based social data at the same scale as fisheries data was difficult to find. As discussed in Section 2.1, Alaskan fishing communities are, with only three exceptions across the state, in the smallest census category, where data only collected once every five to ten years. While this did not affect the difference-in-difference analysis, which only requires one set of data before and one set of data after the intervention, the lack of annual data limits the utility of census-based social vulnerability indices. Prior to 2000, much of the census data is decadal, further limiting the applicability of the census-based suite of indices to measure historic change in fisheries. The housing disruption index, in particular, was a composite percent-change variables that looked. This requirement severely limited sample size for two reasons. First, because a large number of Alaskan communities surveyed in the more recent decennial census were not surveyed in the previous census. Second, census variables are

organized and described differently between decennial censuses, numeric values between the two were compared to see if the variable in question conveyed identical or “similar enough” information to the previous decennial census. The lack of certainty that identical information was being conveyed between census groups decreases the confidence of the percent-change variables derived from early censuses.

*Interpreting composite indices, the fishing harvest narrative, and management applications:*

Using a composite index to unify the multiple facets a community into a single score might oversimplify the complexity of the system, especially when comparing indices over time. Given the complexity of social-ecological systems and global fishery resource markets, it is difficult to know if the nMDS scores alone accurately portray all conditions that may be influencing social changes occurring in fishing communities. Despite these concerns, a well-designed index that can be used by the social impact assessments now mandated by the Magnuson Stevens Act’s National Standard 8 would have immediate policy implications.

Interpreting the results of composite scores made up of variables changing over time was difficult. Where an increase in some indices would suggest a social benefit to the community such as the labor index, increases in other indices might lead to social detriment, such as the poverty index. Most indices are either explicitly linked to positive or negative social responses,<sup>4</sup> or are a mix of indicators whose increase might confer social benefit and other indicators conferring negative social impact.<sup>5</sup> As a result, a deeper dive into interpreting the results of an index is useful for future analysis.

Creating a narrative for the significant result seen in between treatment and control communities for fishing harvest engagement and fishing harvest reliance indices a unique opportunity for a more general explanation on

---

<sup>4</sup> Population composition is an example of a composite index with no clear impact on society: indicators include percent of self-identifying as white, percent of female head of household, percent of population speaking English less than well, and percent of population under 5 years of age.

<sup>5</sup> Housing disruption includes indicators whose increase could be interpreted to lead to negative community impacts including housing in median mortgage costs or the number of mortgages over 35% of income and others that might lead to positive impacts like home values.

how composite indices are flexible enough to represent a nuanced story of fishery impacts. Their interpretation should be validated by case studies. Both harvest engagement and reliance indices are based on a composite of vessel counts, exvessel revenue, landings, and permits. A short review of community-level impact case studies following halibut rationalization corroborates this significant result. Given that the rationalization was meant to reduce capacity, studies in the years following the change reported the expected decrease in vessel but also that revenues would increase likely due to increased economic efficiency of the market-based approach. Both indices contained variables for vessel counts and revenue. But how should the significance of the results be interpreted when the composite index is made up of variables that case studies showed were related to both positive and negative economic impacts?

Remember that both indices saw a larger, negative impact on halibut dependent treatment communities than on non-halibut dependent control communities. With stress scores between the acceptable range 0.1 and 0.2 for both fishing harvest engagement indices<sup>6</sup> we can assume the index scores should generally reflect decreases in vessel counts, exvessel revenue, landings, and permits. However, it must also be noted that halibut dependent communities might see different responses from other non-halibut dependent communities that could explain this significant result. Turning to case studies helps illustrate the strongly negative response of halibut-dependent communities. Surveys reported consolidation away from individuals in smaller communities (Carothers et al. 2010). Furthermore, while thousands of individuals received quota share initially, much of this quota was consolidated to a few hundred (Knapp 1997). Because halibut-dependent treatment communities were designated using data prior to rationalization, it is likely that many of the communities identified here were those communities that lost out during the consolidation.

Existing literature can be used to interpret the results of historic management changes, as shown by the narrative in this section. Caution should be used when attempting to use these indicators to predict potential impact of future

management change where impact studies have not been conducted. It is possible that constituent indicators in each index are context specific, and might vary under different sociopolitical and demographic contexts. For example, AFSC was forced to exclude an index of gentrification because it did not make sense within the context of remote Alaskan communities. Other, latent variables might better explain the results than those presented in the pre-existing indicators developed by each regional NMFS science center. Nevertheless, it does appear that the approach accurately categorized some social vulnerability in fishing communities before and after halibut rationalization, but most likely did not reflect all changes known to occur. For this reason, when tasked with understanding the impacts of management on a community's social vulnerability indices might be best used as a cost-effective to temper a deeper, more intensive research agenda.

#### *Conclusion:*

Management change can have significant impacts on the communities dependent on fishery resources. Index scores based on secondary data can provide a cost-effective measurement of social vulnerability and are important aspect of fishery management. This research outlined a novel, hypothesis-driven approach test the impact on of management change on fish dependent communities using the example of the 1995 management change of Alaskan pacific halibut fishery. First, nonmetric multidimensional scaling was shown to be an efficacious method to create single-value composite indices from nonlinear, real-world data. Second, the approach categorized and assigned halibut-dependent treatment communities and non-halibut-dependent control communities to create an experimental design that could be integrated with basic hypothesis testing. The results of the hypothesis test showed a narrative of indices with significant differences between treatment and control that are supported by historic case studies. Finally, the complex reality of interpreting composite index scores was discussed, but it was found that the approach could be used to generate meaningful, albeit nuanced, results useful to management when supplemented with case studies.

---

<sup>6</sup> Fish harvest engagement index stress score was 0.267, barely acceptable under the rule of thumb

## Acknowledgements

I would like to thank Xingli for his input, availability and openness as an instructor. I would also like to thank my husband who has been exhaustively neglected this quarter.

## References

- Alaska Fisheries, Community Development Quota (CDQ) Program, <http://alaskafisheries.noaa.gov/cdq/cdqprogs/ummary.pdf>
- Branch, T.A., Hively, D.J. and Hilborn, R. 2013. Is the ocean food provision index biased?. *Nature* 495: E5-E6.
- Card, D., and Krueger, A.B. 1993. Minimum wages and employment: A case study of the fast food industry in New Jersey and Pennsylvania (No. w4509). National Bureau of Economic Research.
- Carothers, C., Lew, D.K., and Sepez, J. 2010. Fishing rights and small communities: Alaska halibut IFQ transfer patterns. *Ocean & Coastal Management* 53(9): 518-523.
- Cutter, S.L., Boruff, B.J., and Shirley, W.L. 2003. Social vulnerability to environmental hazards\*. *Social science quarterly*, 84(2), 242-261.
- Hartley, M., & Fina, M. 2001. Changes in fleet capacity following the introduction of individual vessel quotas in the Alaskan Pacific halibut and sablefish fishery. *FAO Fisheries Technical Paper*, 186-207.
- Himes-Cornell, A. and Kasperski, S. 2015. Using indices to assess the vulnerability and resiliency of Alaskan fishing communities to climate change. *Fisheries Research* 162:1-11.
- Hughes, S.E., and Woodley, C. 2007. Transition from open access to quota based fishery management regimes in Alaska increased the safety of operations. *International maritime health*, 58(1-4): 33-45.
- Hout, M.C., Papesh, M.H., and Goldinger, S. D. 2013. Multidimensional scaling. *Wiley Interdisciplinary Reviews: Cognitive Science* 4(1): 93-103.
- Jacob, S., Weeks, P., Blount, B., and Jepson, M. 2013. Development and evaluation of social indices of vulnerability and resiliency for fishing communities in the Gulf of Mexico. *Marine Policy* 37: 86-95.
- Jaworska, N., & Chupetlovska-Anastasova, A. 2009. A review of multidimensional scaling (MDS) and its utility in various psychological domains. *Tutorials in Quantitative Methods for Psychology* 5(1): 1-10.
- Jepson, M., & Jacob, S. 2007. Social indices and measurements of vulnerability for Gulf Coast fishing communities. *Napa Bulletin*, 28(1), 57-68.
- Jepson, M., and Colburn, L.L. 2013. Development of Social Indicators of Fishing Community Vulnerability and Resilience in the US Southeast and Northeast Regions. NOAA Technical Memorandum NMFS-F/SPO-129. 64pp.
- Kasperski, S., and Himes-Cornell, A.. 2014. Indicators of Fishing Engagement and Reliance of Alaskan Fishing Communities. Alaska Fisheries Science Center Quarterly Report, January to March 2014. Accessible at: [http://www.afsc.noaa.gov/Quarterly/jfm2014/JFM14\\_Feature.pdf](http://www.afsc.noaa.gov/Quarterly/jfm2014/JFM14_Feature.pdf).
- Knapp, G. 1997. Initial effects of the Alaska halibut IFQ program: survey comments of Alaska fishermen. *Marine Resource Economics* 12:239-248.
- Knapp, G. 1999. Effects of IFQ management on fishing safety: survey responses of Alaska halibut fishermen. Institute of Social and Economic Research, University of Alaska.
- Mansfield, B. 2004. Neoliberalism in the oceans: "rationalization," property rights, and the commons question. *Geoforum* 35(3): 313-326.
- Matulich, S.C., and Clark, M.L. 2003. North Pacific halibut and sablefish IFQ policy design: quantifying the impacts on processors. *Marine Resource Economics* 18: 149-166.
- McDowell Group, Inc. 2007. Economic Impact of the Commercial Halibut Fisheries in Areas 2C and 3A. Report Prepared for The Halibut Coalition.
- Oksanen, J., F., Blanchet G., Kindt R., Legendre P., Minchin P.R., O'Hara R.B., Simpson G.L., Solymos P., Stevens M.H.M., and Wagner H. 2015. *vegan: Community Ecology Package*. R package version 2.3-1. <http://CRAN.R-project.org/package=vegan>
- Olden, J. Tutorial 8-0 Nonmetric multidimensional scaling. *Fish 560: Applied Multivariate Statistics for Ecologists*.

- Pascoe, S., Kirkley, J.E., Greboval, D.F., and Morrison, C.J. 2003. Measuring and assessing capacity in fisheries: Issues and methods. Food & Agriculture Org.
- Pollnac, R.B., Abbott-Jamieson, S., Smith, C., Miller, M.L., Clay, P.M., and Oles, B. 2006. Toward a model for fisheries social impact assessment. *Marine Fisheries Review* 68(1-4): 1-18.
- Rabinowitz, G.B. 1975. An introduction to nonmetric multidimensional scaling. *American Journal of Political Science* 19: 343-390.
- Sepez, J., Norman, K., and Felthoven, R. 2007. A quantitative model for ranking and selecting communities most involved in commercial fisheries. *NAPA Bulletin* 28(1): 43-56.
- US Department of Commerce, National Oceanic and Atmospheric Administration. Community purchase, final rule. *Federal Register* 69(84):23,681e94.

**Appendix. List of all social (Table A1) and fishery (Table A2) variables in each index along with description and source**

<b>Social Indices</b>		
<b>Variable</b>	<b>Description</b>	<b>Source</b>
<b>Personal Disruption</b>		
pd_XX_unemp	% of residents that are unemployed	U.S. Census Bureau
pd_XX_nodipl	% of residents without a high school diploma	U.S. Census Bureau
pd_XX_pov	% residents in poverty	U.S. Census Bureau
pd_XX_femsep	% of females aged 15 and over that are separated	U.S. Census Bureau
<b>Population Composition</b>		
pc_XX_white	% of population that self identifies as white	U.S. Census Bureau
pc_XX_femhh	% of households with a female head of household	U.S. Census Bureau
pc_XX_u5	% of population that is aged between 0 and 5 years old	U.S. Census Bureau
pc_XX_spkeng	% of residents that speak English less than well	U.S. Census Bureau
<b>Poverty</b>		
pov_XX_assist	% of residents receiving cash public assistance	U.S. Census Bureau
pov_XX_fam	% Families in poverty	U.S. Census Bureau
pov_XX_o64	% of residents over age 65 in poverty	U.S. Census Bureau
pov_XX_u18	% of residents under age 18 in poverty	U.S. Census Bureau
<b>Labor Force Structure</b>		
lab_XX_labforce	% of residents in the labor force	U.S. Census Bureau
lab_XX_fem	% of female residents in the labor force	U.S. Census Bureau
lab_XX_selfemp	% of residents that are self employed	U.S. Census Bureau
lab_XX_ssincome	% of residents on social security (1990 data is % of households on social security / total households)	U.S. Census Bureau
<b>Housing Disruption</b>		
hd_XX_medmort_change	% change in median mortgage cost	U.S. Census Bureau
hd_XX_homeval_change	% change in median home values	U.S. Census Bureau

hd_XX_moc35	% of households with mortgage costs that exceed 35% of their household income	U.S. Census Bureau
<b>Housing Characteristics</b>		
he_XX_fuel	% of households heating with fuel oil, kerosene, etc	U.S. Census Bureau
he_XX_plumbing	% of households lacking complete plumbing facilities	U.S. Census Bureau
he_XX_rooms	Median number of rooms per residence	U.S. Census Bureau
he_XX_grossrent	Median monthly gross rent	U.S. Census Bureau
he_XX_mortgage	Median monthly mortgage costs	U.S. Census Bureau
<b>Status of Schools *</b>		
sch_XX_schools	Number of schools in a community	Alaska Dep't of Education, 2012
sch_XX_students	Number of students in a community	Alaska Dep't of Education, 2012

**Table A1:** Constituent variables for each social index, where *\_XX\_* is a variable for an individual census year of 1980, 1990, 2000, or 2010. \*Schools was excluded from this analysis because of a lack of data on the number of students and the Alaskan context of virtual schools.

<b>Fishery Indices</b>		
<b>Variable</b>	<b>Description</b>	<b>Source</b>
<b>Commercial Processing Engagement</b>		
port_val	Ex-vessel value of commercial catch landed in a community	ADF&G and CFEC
port_lbs	Pounds of commercial catch landed in a community	ADF&G and CFEC
processors	Number of processors located in a community	ADF&G
<b>Commercial Harvesting Engagement</b>		
res_val	Ex-vessel value of commercial catch from vessels owned by residents	ADF&G and CFEC
res_lbs	Pounds of commercial catch from vessels owned by residents	ADF&G and CFEC
permits	Number of CFEC permits held by residents	CFEC
vessels	Number of vessels owned by residents	CFEC

crew	Number of crew licenses held by residents	ADF&G
<b>Commercial Processing Reliance</b>		
port_val_pc	Ex-vessel value of commercial catch landed in a community per capita	ADF&G and CFEC
port_lbs_pc	Pounds of commercial catch landed in a community per capita	ADF&G and CFEC
processors_pc	Number of processors located in a community per capita	ADF&G
<b>Commercial Harvesting Reliance</b>		
res_val_pc	Ex-vessel value of commercial catch from vessels owned by residents per capita	ADF&G and CFEC
res_lbs_pc	Pounds of commercial catch from vessels owned by residents per capita	ADF&G and CFEC
permits_pc	Number of CFEC permits held by residents per capita	CFEC
vessels_pc	Number of vessels owned by residents per capita	CFEC
crew_pc	Number of crew licenses held by residents per capita	ADF&G
<b>Recreational Engagement</b>		
clog_ttl_business_count	Total CLOG business count	
sportfish_licenses_sold	Sportfish licenses sold	
sportfish_guide_business	Total number of sportfish guide businesses	
sportfish_guide_licenses	Total number of sportfish guide licenses	
<b>Recreational Reliance</b>		
pc_clog_ttl_business_count	Total CLOG business count per capita	
pc_sportfish_licenses_sol	Sportfish licenses sold per capita	
pc_sportfish_guide_business	Total number of sportfish guide businesses per capita	
pc_sportfish_guide_licenses	Total number of sportfish guide licenses per capita	

**Table A2: Constituent variables for each fishery index.**