

# **River network structuring of climate and landscape effects in salmon watersheds**

by

**Kyle August Chezik**

M.Sc., University of Minnesota, 2013

B.A., St. Olaf College, 2009

Thesis Submitted in Partial Fulfillment of the  
Requirements for the Degree of  
Doctor of Philosophy

in the  
Department of Biological Sciences  
Faculty of Science

**© Kyle August Chezik 2019**  
**SIMON FRASER UNIVERSITY**  
**Spring 2019**

Copyright in this work rests with the author. Please ensure that any reproduction  
or re-use is done in accordance with the relevant national copyright legislation.

# Approval

Name: **Kyle August Chezik**

Degree: **Doctor of Philosophy (Biology)**

Title: **River network structuring of climate and landscape effects in salmon watersheds**

Examining Committee: Chair: Frank J. Lee  
Associate Professor

**Jonathan W. Moore**  
Senior Supervisor  
Associate Professor

**John D. Reynolds**  
Supervisor  
Professor

**David A. Patterson**  
Supervisor  
Adjunct Professor

**Leithen K. M'Gonigle**  
Internal Examiner  
Assistant Professor

**Aimee H. Fullerton**  
External Examiner  
Research Biologist  
Fish Ecology Division  
National Oceanic and Atmospheric Administration

Date Defended: **January 09, 2019**

# Abstract

Climate change is altering historic patterns of temperature and precipitation worldwide with significant implications for the abiotic and biotic dynamics of river ecosystems. Flowing downhill, precipitation aggregates into creeks, streams and eventually rivers, forming a branching network over the landscape architecturally similar to the branches, limbs and trunk of a tree. Linking disparate locations, these river networks integrate the varied expressions of climate within a watershed. Thus, the river network offers a framework for understanding how spatial patterns of climate are organized and become manifest in rivers. By considering the river network's structuring of climate and landscape interactions, we might better understand how climate and land-use change impact river ecosystems and more clearly identify particularly vulnerable biota. In chapter 2, I examine how river networks dampen signals of climate change in hydrologic flow by integrating varied flow trends from upstream. I demonstrate that by integrating a diverse climate portfolio, the network accumulates changing flow regimes of different volatility, direction and magnitude, such that on average downstream climate change trends are moderated. In chapter 3, I consider the match-mismatch potential of juvenile salmon migrating towards the springtime zooplankton resource pulse in the estuary. I show that populations further from, and whose climate is more dissimilar to, the estuary, are more likely to miss the peak zooplankton bloom. These findings suggest migratory distance influences phenological mismatch risk among populations. My fourth chapter develops an unsupervised machine learning method for cleaning stream temperature data to facilitate big data studies. In chapter 5, I gathered temperature data at over 100 locations throughout a watershed the size of Ireland, over 4 years at 2-hour intervals resulting in over 1 million data points. These data informed a spatial stream network model that quantified how landscape features and river connectivity control seasonal temperature dynamics. These temperature dynamics across space and time revealed that different adult salmon migrations have very different exposures to warm temperatures. Collectively, these findings illustrate that river networks: 1. integrate and dampen signals of climate change, 2. structure phenological match-mismatch patterns and 3. organize thermal exposure potential of biota.

**Keywords:** climate change; fluvial landscape; Fraser River; Thompson River; unsupervised learning

# Acknowledgements

Borrowing from the Finnish architect Alvar Aalto, "[Ideas]... are... like some big salmon or trout. They're not born fully grown; they're not even born in the sea or water where they normally live. They are born hundreds of miles away from their home grounds, where the rivers narrow to tiny streams. Just as it takes time for a spec of fish spawn to mature into a fully grown fish, so we need time, for everything that develops and crystallizes in our world of ideas." I would like to first and foremost thank my advisor Jonathan Moore for his unwavering support and guidance over the course of this most challenging intellectual endeavor. Jon has cultivated a truly wonderful environment for learning and exploration that has nurtured "specs" of ideas into high caliber research of which I am humbled to have been apart.

The culture of inclusiveness, respect, and encouragement within Jon's lab not only reflects its leadership but the quality of character of the individuals therein. Thank you to all the lab mates with whom I have shared the struggle of graduate school. Your friendship, enthusiasm and insightful discussions over the years are what make the pursuit of science fun. I would like to specifically thank, Michael Beakes, Justin Yeakel and Corey Phillis for their mentorship as senior lab members when I was only beginning my journey. Sam Wilson and Luke Andersson, who not only made my field work possible through impeccable organization but loads of fun despite the heat, cold, rain, devils club, bears and other logging-road excitement. Michael Arbeider, Ciara Sharpe, Will Atlas, Kara Pitman and Colin Bailey, for loads of fun at conferences, on the Keogh River, in the Nicola River, at lab retreats to Squamish and the Adams River.

I would also like to thank my committee members David Patterson and John Reynolds for many insightful discussions that challenged me to question my assumptions and fully develop my ideas. Beyond my official academic committee, I must thank Sean Anderson for his guidance and mentorship. Sean regularly went above and beyond to help me flesh out ideas, learn new tools and improve manuscripts. Sean's work ethic is unmatched and his dedication to the benefit of others is unassuming and inspiring. I am incredibly thankful for his leadership and friendship.

I must express my deepest gratitude to my family who have supported me through all these years of education. To my parents who have encouraged me my entire life and always believed in me. You've been there since I was just that "spec" and provided the loving

environment that gave me the courage and security to pursue my interests and succeed in all of my endeavors. To my sister Adrienne, who has been a good friend my entire life and has always been encouraging, thank you. To my grandmother Dolores Dageforde, your support and love is always felt.

Finally, I can't describe how grateful I am to have had such a loving and understanding partner in Louisa Harding. You have kept me grounded, assuaged the difficult moments and been there to celebrate my successes. Additionally, Louisa's family, the Hardings, Gilmores, Rues and Deena Burke, have been immensely supportive over the years. As a transplant from the mid-west I'm incredibly grateful for the sense of belonging and place they have provided.

# Table of Contents

<b>Approval</b>	ii
<b>Abstract</b>	iii
<b>Acknowledgements</b>	iv
<b>Table of Contents</b>	vi
<b>List of Tables</b>	ix
<b>List of Figures</b>	x
<b>1 Introduction</b>	1
1.1 River networks . . . . .	1
1.2 River network climate portfolios . . . . .	2
1.3 River network phenological match-mismatch . . . . .	3
1.4 Big data bottleneck . . . . .	4
1.5 River network cumulative effects . . . . .	4
1.6 Contributions . . . . .	5
<b>2 River networks dampen long-term hydrological signals of climate change</b>	7
2.1 Abstract . . . . .	7
2.2 Introduction . . . . .	7
2.3 Materials and methods . . . . .	10
2.3.1 Discharge data and flow metrics . . . . .	10
2.3.2 River discharge trend analysis . . . . .	11
2.3.3 Quantifying climate variability . . . . .	11
2.3.4 Quantifying the river network portfolio effect . . . . .	12
2.3.5 River attenuation null model simulations . . . . .	13
2.3.6 Dampening summary statistics . . . . .	13
2.4 Results . . . . .	14
2.4.1 Hydrologic trend diversity . . . . .	14
2.4.2 Climate trend diversity . . . . .	14

2.4.3	Climate portfolios and flow trend dampening . . . . .	14
2.4.4	Seasonally shifting flow trend dampening . . . . .	16
2.5	Discussion . . . . .	16
2.6	Conclusion . . . . .	18
2.7	Acknowledgements . . . . .	19
2.8	Supplementary material . . . . .	19
<b>3</b>	<b>Spatial structuring of phenological match-mismatch in a highly migratory fish</b>	<b>30</b>
3.1	Abstract . . . . .	30
3.2	Introduction . . . . .	31
3.3	Materials and methods . . . . .	33
3.3.1	Study site . . . . .	34
3.3.2	Pink salmon . . . . .	34
3.3.3	Spawn timing . . . . .	36
3.3.4	Emergence timing model . . . . .	37
3.3.5	Emergence & arrival timing estimation . . . . .	38
3.3.6	Zooplankton bloom, match-mismatch and climate dissimilarity . . . . .	39
3.3.7	Spatial structure of climate dissimilarity and phenological match-mismatch . . . . .	40
3.4	Results . . . . .	41
3.4.1	Spawn timing spatial patterns . . . . .	41
3.4.2	Emergence timing model . . . . .	41
3.4.3	Estuary arrival, zooplankton bloom and climate dissimilarity . . . . .	43
3.4.4	Phenological match-mismatch . . . . .	45
3.5	Discussion . . . . .	47
3.6	Acknowledgements . . . . .	51
3.7	Supplementary material . . . . .	51
<b>4</b>	<b>Cleaning stream temperature data with hidden markov models</b>	<b>58</b>
4.1	Abstract . . . . .	58
4.2	Introduction . . . . .	58
4.3	Methods . . . . .	61
4.3.1	Hidden markov models . . . . .	61
4.3.2	States: water & air temperature models . . . . .	61
4.3.3	Temperature HMM . . . . .	63
4.3.4	Observed data . . . . .	63
4.3.5	Model priors & known quantities . . . . .	65
4.3.6	Model post-hoc calculations . . . . .	65
4.4	Results . . . . .	66

4.5	Discussion . . . . .	69
4.6	Acknowledgements . . . . .	71
<b>5</b>	<b>Impacts of network structure and landscape complexity on the cumulative thermal exposure of migrating salmon</b>	<b>73</b>
5.1	Abstract . . . . .	73
5.2	Introduction . . . . .	74
5.3	Methods . . . . .	76
5.3.1	Thompson River watershed . . . . .	76
5.3.2	Stream temperature data . . . . .	77
5.3.3	Statistical models . . . . .	77
5.4	Results . . . . .	85
5.4.1	Temporal stream temperature model . . . . .	85
5.4.2	Spatial stream network models . . . . .	85
5.4.3	Projected network thermal exposure . . . . .	89
5.4.4	Chinook salmon thermal exposure . . . . .	90
5.5	Discussion . . . . .	92
5.6	Acknowledgements . . . . .	95
5.7	Supplementary material . . . . .	96
<b>6</b>	<b>General discussion</b>	<b>104</b>
6.1	Integrating the climate mosaic . . . . .	105
6.2	Considering shifting interactions . . . . .	105
6.3	Shifting climate portfolios . . . . .	106
6.4	Towards linking local landscape features and biology . . . . .	106
6.5	River network structured life-history risk . . . . .	107
6.6	Monitoring at scale . . . . .	107
6.7	In conclusion . . . . .	108
	<b>Bibliography</b>	<b>109</b>

# List of Tables

Table 5.1 Mean and standard deviation (SD) values of predictor and response variables. Only the standard deviation value is reported for the response variable as these values were only standardized but not scaled. Moreover, these SD values are averages from the 250 samples drawn from equation 5.1 parameter posteriors. Predictor variables were both centered and scaled and the respective mean and SD values were calculated across all observed and predicted sites. Transformations of some variables by the natural log (ln) was done prior to calculating the mean and SD. $H_2O$ indicates predictors that constitute the average value among values averaging over the contributing area. For instance, the average contributing area elevation among predicted and observed sites was 1431 which is similar to, but not the same as, the Thompson's overall mean elevation of 1264 metres as sites are nested. . . . .	83
--	----

# List of Figures

Figure 2.1	Climate trends (1970 - 2007) in mean annual temperature (MAT) and mean annual precipitation (MAP) within the Fraser River basin overlaid on a digital elevation model within British Columbia, Canada. Flow gauge sites (dots) are scaled by the size of the contributing area. The map is projected in Albers BC UTM to provide an equal area depiction of the region but labels are expressed in WGS84 latitude and longitude. . . . .	9
Figure 2.2	Climate variability index as a function of catchment area. As watershed area increases the variety of climatic trends within the watershed increases at a diminishing rate. . . . .	14
Figure 2.3	Annual flow trend attenuation within the Fraser River basin. ( <b>Left</b> ) Trend estimates ( $\hat{b}_s$ ) $\pm$ one standard error (SE, grey) plotted against watershed area ( $\text{km}^2$ ), colored by climate portfolio strength (green = small, blue = large). Blue lines represent observed attenuation; orange and red represent simulated attenuation that is weaker and stronger than observed. ( <b>Right</b> ) Density plots show null model simulated variance exponents ( $\hat{\delta}$ ) and the proportion on either side of the observed variance exponent (blue). Flow metrics include long-term flow-timing shifts (change per decade in day-of-year (DOY) to half annual flow), where decreasing trends suggest more annual flow is occurring earlier in the year (a), and the percent change per decade in minimum (b), maximum (c) and median (d) annual flow. Simulated lines ignore variance in the intercept and slope to focus visually on attenuation. . . . .	15

Figure 2.4 Monthly maximum flow trend attenuation within the Fraser River basin. (**Left**) Fraser River's basin-wide maximum-flow trend estimates (i.e., intercept = vertical grey lines) by month with density distributions of null-model simulations. Observed values falling further from the center of the density distribution suggest greater evidence for changes in maximum flow and a greater shift in magnitude. (**Center**) Observed monthly Fraser River maximum-flow variance exponent ( $\hat{\delta}$ , blue) and associated density distribution of simulated  $\hat{\delta}$  estimates. Decimal values represent the percent of simulated data exhibiting weaker attenuation (yellow) than observed. (**Right**) Trend estimates ( $\hat{b}_s$ )  $\pm$  one standard error (SE, grey) plotted against watershed area ( $\text{km}^2$ ), colored by climate portfolio strength (green = small, blue = large), for four seasonally representative months. These reflect months in the prior columns and describe the variation in percent change per decade of maximum flow among sites. Simulated lines ignore variance in the intercept and slope to focus visually on attenuation. . . . .

17

Figure 2.5 Median annual flow plotted at an annual time step between 1970 and 2007 for 55 flow gauge stations within the Fraser River basin. Points are colored by the percent of the total watershed area harvested in the previous five years. Grey (blue: sites with no harvest) lines indicate the trend in flow over time using a linear model. Light grey around each trend line represents the 95% confidence interval. . . . .

21

Figure 2.6 The response of monthly median flow trends (1970-2007) to logging in 55 sub-basins (colors) of the Fraser River watershed. Timber harvest is quantified as the percent of watershed area cut in the previous five years. Flow trends represent the mean change in median flow over time after being scaled by site and log transformed to facilitate cross site comparsion. Lines, estimated using a linear model, depict the change in flow trends with increasing timber harvest for each flow gauge station. . . . .

22

Figure 2.7	The global response of median annual flow trends in the Fraser River basin to timber harvest between 1970 and 2007. Points are colored by flow gauge station. Flow trends represent the mean change in the median flow over time after being scaled by site and log transformed for cross site comparison. The black line represents the fixed effect estimate of a generalized linear mixed model with a Gamma distribution and log link function, where random intercepts were allowed to vary by flow gauge station with a single shared slope. This model describes the typical impact of harvest on median flow trends across the Fraser River basin. . . . .	23
Figure 2.8	Example site specific flow simulations. Annual median-flow simulations (yellow) and observations (blue) for all 55 sites considered in this study found within the Fraser River basin. Simulations at each site were parameterized using the observed data but no trend was imposed. Fifty-five site simulations equate to one basin-wide simulation. Each flow metric was simulated 1000 times basin-wide per response variable. See supplemental figure 2.9 for resultant basin-wide simulation. . . . .	24
Figure 2.9	Example flow simulation results. Trend estimates (dots) $\pm$ SE (vertical lines) of simulated null-model data (yellow) and the observed data upon which simulations were based (blue), plotted against watershed area. Curved lines represent the estimated variance function given the simulated data. This is a single example result of the 1000 null-model simulations produced from the data in supplemental figure 2.8. . . . .	25

Figure 2.10 Monthly minimum flow trend attenuation within the Fraser River basin. (**Left**) Fraser River's basin-wide minimum-flow trend estimates (i.e., intercept = vertical grey lines) by month with density distributions of null-model simulations. Observed values falling further from the center of the density distribution suggest greater evidence for changes in minimum flow and a greater shift in magnitude. (**Center**) Observed monthly Fraser River minimum-flow variance exponent ( $\hat{\delta}$ , blue) and associated density distribution of simulated  $\hat{\delta}$  estimates. Decimal values represent the percent of simulated data exhibiting weaker attenuation (yellow) than observed. (**Right**) Trend estimates ( $\hat{b}_s$ )  $\pm$  one standard error (SE, grey) plotted against watershed area ( $\text{km}^2$ ), colored by climate portfolio strength (green = small, blue = large), for four seasonally representative months. These reflect months in the prior columns and describe the variation in percent change per decade of minimum flow among sites. Simulated lines ignore variance in the intercept and slope to focus visually on attenuation. . . . .

26

Figure 2.11 Monthly median flow trend attenuation within the Fraser River basin. (**Left**) Fraser River's basin-wide median-flow trend estimates (i.e., intercept = vertical grey lines) by month with density distributions of null-model simulations. Observed values falling further from the center of the density distribution suggest greater evidence for changes in median flow and a greater shift in magnitude. (**Center**) Observed monthly Fraser River median-flow variance exponent ( $\hat{\delta}$ , blue) and associated density distribution of simulated  $\hat{\delta}$  estimates. Decimal values represent the percent of simulated data exhibiting weaker attenuation (yellow) than observed. (**Right**) Trend estimates ( $\hat{b}_s$ )  $\pm$  one standard error (SE, grey) plotted against watershed area ( $\text{km}^2$ ), colored by climate portfolio strength (green = small, blue = large), for four seasonally representative months. These reflect months in the prior columns and describe the variation in percent change per decade of median flow among sites. Simulated lines ignore variance in the intercept and slope to focus visually on attenuation. . . . .

27

Figure 2.12 Timber harvest between 1970 and 2007 within 55 sub-basins of the Fraser River watershed in British Columbia, CA. Timber harvest is quantified as the percent of each sub-basin's area harvested in the previous five years. Lines are loess smoothers indicating general trends over time for each sub-basin and are colored by sub-basin drainage area. . . . .

28

Figure 2.13	Smoothed, five-day rolling average flow rates for each day in each year between 1970 and 2007 at 55 flow gauge stations in the Fraser River watershed. Years are colored oldest (red) to most recent (blue). The distributions demonstrate that flow in the Fraser River basin is defined by a spring melt and typically low summer flows. Flow timing, duration and magnitude vary between sites but are largely grouped by coastal and interior locations. For instance, sites 08MH006, 08MH029, 08MH076 and 08MH090 experience an earlier freshet and higher fall and winter flows due to their coastal proximity. . . . .	29
Figure 3.1	Fraser River watershed in British Columbia Canada, colored by the Climate WNA (Wang et al., 2016) estimated mean annual air temperature in 1970 ( <b>left</b> ) and 2010 ( <b>right</b> ). . . . . .	35
Figure 3.2	Fraser River basin pink salmon spawn locations, timing and projected shifts in timing. Points indicate the location of observed pink salmon spawning and are colored by the median spawn date across observed years. The Fraser River network is colored by the mean predicted shift in spawn date between 1970 and 2010 given our SSNM described in equation 3.1. Shifts in spawn date are expressed in days with positive values representing later seasonal spawning. . . . .	42
Figure 3.3	Days to emergence given the cumulative degree-days since fertilization. Cumulative degree-days are calculated as the summed mean daily temperatures above freezing ( $^{\circ}\text{C}_{T_0}$ ) from spawn (S) to emergence (day $n$ ). Points represent observation estimates derived from published studies. The data were fit to equation 3.2 and the dotted line describes the average decline in days to emergence with increasing cumulative degree-days. The solid lines indicate the variance around the mean given all posterior estimates of $\eta$ . . . . .	43
Figure 3.4	Estimated distributions of pink salmon estuary arrival timing overlayed on zooplankton bloom date estimates between 1968 and 2010, stratified by each populations natal stream climate dissimilarity with the estuary in a given year. Distributions are colored by the number of days the most common estuary arrival date (i.e., mode) missed the zooplankton bloom in the given year. . . . .	44

Figure 3.5 Percent phenological match between juvenile pink salmon arrival and the zooplankton bloom, given how dissimilar natal stream climates are from the estuary. Points describe the phenological match and climate dissimilarity estimates for each population in each year data were available and are colored by river distance to the estuary. The mean effect of climate dissimilarity on phenological match given a fitted estimate of equation 3.4 is described by the solid black line. Density distributions describe the spatial relationship of climate dissimilarity in the Fraser River basin (i.e., Supplementary Figure 3.8). . . . .

Figure 3.6 (left) Predicted phenological match (%) in estuary arrival with the 2010 zooplankton bloom throughout the Fraser River basin. Open circles indicate locations of temperature monitoring and are colored by the estimated median percent overlap across available years. (right) The difference between predicted 1970 and 2010 match-mismatch percentage across the Fraser River basin given equation 3.4. . . . .

Figure 3.7 The presence (1) and absence (0) of pink salmon populations in the Fraser River basin given the phenological match (%) between estuary arrival of juvenile pink salmon and the zooplankton bloom. The points are colored by their natal stream climate dissimilarity region, (i.e., Figure 3.5 & Supplementary Figure 3.8) and are jittered both vertically and horizontally for visibility. The dotted grey line describes the mean effect of a binomial generalized linear model surrounded by a light grey standard error estimate. . . . .

Figure 3.8 Climate regions by year among emergence timing prediction sites in the Fraser River basin in British Columbia Canada. Points (green) indicate temperature monitoring locations where emergence timing can be estimated. The ring around monitoring locations are made up of climate divergence estimates (points) for each year between 1970 and 2010 colored by climate divergence group. Four distinct climate divergence groups are estimated, located in the lower Fraser basin (blue), the Nicola and Thompson region (green), the central and northeast region (red) and the west and northwest region (yellow). Climate group boundaries vary from year to year within edge sites. The river network is colored by mean climate group and the Fraser River watershed is colored by mean annual temperature across years. . . . .

Figure 3.9 (Left) Leave one out cross validation (LOOCV) given the observed spawn timing estimates (solid line 1:1). (Right) LOOCV prediction standard error estimates in days given the observed data. . . . .

Figure 3.10	Torgogram describing autocorrelation in spawn timing among flow connected and unconnected locations in the Fraser River basin where lower semivariance indicates higher autocorrelation. The size of the circles are proportional to the number of pairs binned in the given stream distance (Hoef et al., 2014). . . . .	54
Figure 3.11	Overlap between pink salmon estuary arrival and zooplankton peak bloom window in the Strait of Georgia from 1968 to 2010. Pink salmon arrival distributions represent populations residing in the lower Fraser River below Hells Gate (blue) and above Hells Gate in the Thompson and Nicola region (green). Zooplankton bloom date range (grey ribbon) is derived from model output in Allen & Wolfe (2013). . . . .	54
Figure 3.12	(Left) Leave one out cross validation (LOOCV) given the observed phenological match (%) between estuary arrival and zooplankton bloom among sites and years in the Fraser River watershed in British Columbia Canada (solid line 1:1). (Right) LOOCV prediction standard error estimate (%) given the observed data. . . . .	55
Figure 3.13	Phenological match (%) between juvenile pink salmon arrival and the zooplankton bloom, given the standardized degree of climate dissimilarity between each natal stream and the estuary in the given year. Colored points describe the overlap and climate divergence estimates for each population in each year data were available and are colored by river distance to the estuary. Black points represent prediction estimates given leave one out cross validation per a fitted equation 4 in the main text. . . . .	56
Figure 3.14	Changing estuary arrival timing inter-quartile range given the strength of the deviation from the mean zooplankton bloom date as described in Allen & Wolfe (2013). The points are colored by climate region above (green) and below (blue) Hells Gate. Lines describe the mean trend line given a simple linear model. . . . .	57
Figure 4.1	Observed water and air temperature locations within the Thompson River watershed, nested within the Fraser River watershed (dark grey) in central British Columbia Canada (light grey). Water temperature locations (blue) were monitored between July 2014 and August 2017 concurrently with air temperature stations (yellow) maintained by Environment Canada. Water temperature locations are labeled by their site number indicating the location of model results shown in figure 4.3. . . . .	64

Figure 4.2	Shifting disagreement between manual and automated cleaning of stream temperature data as the water certainty break-point decreases. The 'type' of disagreement assumes manual cleaning to be the 'true' state, where a type I error indicates when the HMM labels data water while manual methods indicate air to be the 'true' source. Type II 'errors' occur when the HMM labels data air while the data were manually labeled water. The total percentage of data in disagreement between these data cleaning methods is indicated by the long-dashed line. . . . .	67
Figure 4.3	(Left) Source estimates for temperature observations at 10 locations within the Thompson River watershed in central British Columbia, Canada . Air (yellow) and water (blue) source estimates were calculated using break-points of greater than 1% probability for water at the observation level and greater than $1 \times 10^{-7}\%$ when averaged to the daily level. HMM state estimate probabilities are described by the color bar above each time series. State model 95% confidence intervals describe the probability density window for air (green) and water (purple). Panels are labeled by site number, associated with locations in figure 4.1 and total disagreement (%). For visual consideration, mean daily air temperature observations at the nearest monitoring station are displayed in red. (Right) Temperature observations for site 49 in 2015 and associated source probability estimates. A rug plot in the lower panel identifies where manual and HMM models disagree and solid lines describe the mean estimate of the annual air and water profiles. . . . .	68
Figure 5.1	Stream temperature monitoring locations (blue) and Chinook salmon populations (yellow) within the Thompson River watershed. The inset indicates the North (green) and South Thompson's (purple) watershed locations within the Fraser River watershed (dark grey) in British Columbia Canada (light grey). The shape of the Chinook salmon population locations indicate the timing of migration through the Albion Test Fishery in early (June), mid (July) and late (August) summer, with the exception of Louis which is considered to migrate in the spring (March-May) but was grouped into the early summer period for this study (see Parken et al., 2008). . . . .	78

Figure 5.2	Examples of model fits to equation 5.1 at 4 sites in Thompson River watershed. The red ribbon describes the 95% credibility interval of the mean temperature estimate, while the orange dashed-line describes the average 95% variance window above the mean. The Barrière River and Bridge Creek are tributaries to the North Thompson River (NT), while the Shuswap River drains Mabel Lake and flows into the South Thompson River (ST) . . . . .	86
Figure 5.3	Evaluation of SSNM model predictive power using leave-one-out cross validation for each temporal model parameter. A 1:1 line describes a perfect fit between observed and SSNM predicted values. Parameter estimates were randomly drawn from equation 5.1 parameter posteriors and predicted by SSNM models (see Figure 5.4 for covariates) . . . . .	87
Figure 5.4	Coefficient estimates for mean annual water temperature ( $\alpha$ ), water temperature amplitude ( $A$ ) and seasonal hysteresis ( $\phi$ ) SSNMs by column respectively. Top row coefficients have been scaled while the bottom row have been logit transformed making parameter estimate effect sizes comparable by row within models. . . . .	88
Figure 5.5	Thermal exposure above 19 and 22°C in the North and South Thompson River watersheds as exemplified by the mid-summer migration in 2015. <b>(left)</b> The median probability of experiencing above threshold temperatures at each point in the network. <b>(center)</b> The cumulative migratory thermal exposure from the outlet to each point on the network. The color bar ranges from 0 (low) to <i>max n</i> (high) rather than depicting a probability in this column. <b>(right)</b> The median value of the average migratory thermal exposure probability. . . . .	89
Figure 5.6	Thermal exposure above 19 and 22°C for 16 Chinook salmon populations in the South and North Thompson River watersheds from 2014 to 2017. <b>(left)</b> The cumulative migratory thermal exposure along the populations migratory route. <b>(right)</b> The thermal exposure probability upon arriving on the spawning grounds. All values are calculated given the populations run-timing and populations are ordered from furthest to shortest migration within their respective watershed. Migration refers to run-timing through the Albion Test Fishery near the outlet of the Fraser River in the early (June:Circle), mid (July:Triangle) and late (August:Square) summer. . . . .	91

Figure 5.7	Probability (0 = blue, 1 = red) of thermal exposure above 19°C in the North and South Thompson River watersheds. Median migratory period estimates projected during early (June), mid (July) and late (August) summer between 2014 and 2017. Points indicate spawning locations of Chinook populations in the watersheds and their shape indicates their migration timing through the Albion Test Fishery near the mouth of the Fraser River. . . . .	97
Figure 5.8	Probability (0 = blue, 1 = red) of thermal exposure above 22°C in the North and South Thompson River watersheds. Median migratory period estimates projected during early (June), mid (July) and late (August) summer between 2014 and 2017. Points indicate spawning locations of Chinook populations in the watersheds and their shape indicates their migration timing through the Albion Test Fishery near the mouth of the Fraser River. . . . .	98
Figure 5.9	The median cumulative number of likely (>50% probability) thermal exposures above 19°C during migration up the North and South Thompson River watersheds. Values are scaled between 0 (blue) and 1 (red) by the largest value across migratory periods and years ( <i>max n</i> = 26,235). Migratory period estimates projected during early (June), mid (July) and late (August) summer between 2014 and 2017. Points indicate spawning locations of Chinook populations in the watersheds and their shape indicates their migration timing through the Albion Test Fishery near the mouth of the Fraser River.	99
Figure 5.10	The median cumulative number of likely (>50% probability) thermal exposures above 22°C during migration up the North and South Thompson River watersheds. Values are scaled between 0 (blue) and 1 (red) by the largest value across migratory periods and years ( <i>max n</i> = 1,271). Migratory period estimates projected during early (June), mid (July) and late (August) summer between 2014 and 2017. Points indicate spawning locations of Chinook populations in the watersheds and their shape indicates their migration timing through the Albion Test Fishery near the mouth of the Fraser River.	100
Figure 5.11	The average migratory thermal exposure probability (0 = blue, 1 = red) above 19°C during upstream migration in the North and South Thompson River watersheds. Run-timing estimates projected during early (June), mid (July) and late (August) summer between 2014 and 2017. Points indicate spawning locations of Chinook populations in the watersheds and their shape indicates their migration timing through the Albion Test Fishery near the mouth of the Fraser River.	101

Figure 5.12 The average migratory thermal exposure probability (0 = blue, 1 = red) above 22°C during upstream migration in the North and South Thompson River watersheds. Run-timing estimates projected during early (June), mid (July) and late (August) summer between 2014 and 2017. Points indicate spawning locations of Chinook populations in the watersheds and their shape indicates their migration timing through the Albion Test Fishery near the mouth of the Fraser River. 102

Figure 5.13 The average migratory thermal exposure probability of experiencing above threshold temperatures for each Chinook population within each run-timing period in the North (N) and South (S) Thompson River watersheds between 2014 and 2017. Migration period represents the run-timing of Chinook populations through the Albion Test Fishery in the lower Fraser River in early (June), mid (July) and late (August) summer. Points are horizontally jittered to improve visibility with 2(N), 3(N), 3(N), 2(S), 1(S) and 5(S) populations in the early, mid and late migration respectively. . . . . 103

# Chapter 1

## Introduction

### 1.1 River networks

Rivers support an immense amount of biological diversity (Strayer & Dudgeon, 2010) and provide the services foundational to culture and economies (Palmer & Richardson, 2009) that have founded civilization (Macklin & Lewin, 2015). Rivers' essential role in human advancement has resulted in their widespread development, overexploitation, pollution and degradation (Strayer & Dudgeon, 2010). Because rivers integrate their catchments, they are exposed to human activities well beyond the rivers' banks, such as logging, mining, agriculture and urbanization throughout their watershed (Palmer et al., 2009). Over the past century, anthropogenic activities from beyond the watershed have also begun to change the very climate that defines river flow regimes (Bindoff et al., 2013; Ficklin et al., 2018) further impacting riverine habitat (Palmer et al., 2009). Despite these numerous disturbances at various scales, rivers exhibit patterns of stability (Moore et al., 2015) that seem to emerge from their network shape and structure (McCluney et al., 2014; Yeakel et al., 2014). Improving our understanding of how river networks organize biotic and abiotic processes is important to predicting and mitigating the impacts of continued perturbation on this essential ecosystem (Lowe et al., 2006).

Grown from the interplay of climate, hydrologic forces and underlying topology (Seybold et al., 2017), river networks exhibit a dendritic or arborescent branching architecture (Campbell Grant et al., 2007; Stralher, 1964) that organize the physical and biological patterns of the lotic environment (Benda et al., 2004; Rodriguez-Iturbe et al., 2009). Water is the elemental feature that acts as the medium by which energy and matter pass from the atmosphere and landscape to the river network. As branches merge and water volume grows, the chemical (Hoef & Peterson, 2010), thermal (Peterson et al., 2013) and physical (Isaak et al., 2014) signature of the upslope area is integrated downstream, increasingly reflecting a greater diversity of climates, topologies, landscape features and anthropogenic impacts (Palmer et al., 2009).

The shifting riverscape, from the headwaters to the mouth, place spatial and temporal controls on the biological processes within the network (Vannote et al., 1980; Rodriguez-Iturbe et al., 2009). Just as oceans, mountains, deserts or canyons limit and direct the dispersion of terrestrial plants and animals (MacArthur, 1984), river networks constrain the movement of aquatic life bound to the network (Lowe et al., 2006). Specifically, the network's dendritic structure does not allow for loops such that the route and distance between any two points is strictly defined (Rodriguez-Iturbe et al., 2009). As a result, populations relatively 'close' in space may be separated by significant distance along the network. The limited connectivity derived from the networks branching complexity increases isolation among network populations and encourages adaptation to local sub-basin conditions (Terui et al., 2018). The varied population dynamics that derive from local adaptation produce asynchrony among populations (e.g., Crozier & Zabel, 2006) such that local temporal or spatial disturbance is not reflected uniformly across the meta-population (Schindler et al., 2010). Inversely, processes that enhance synchrony, such as network fragmentation (e.g., culverts, dykes, dams) or basin wide disturbance (e.g., clear cutting, climate change) undermine the network processes that stabilize meta-populations (Fagan, 2002; Moore et al., 2010).

The dendritic architecture and flow directed auto-correlation of the network offers a framework for quantifying basin-wide contributions of landscape and climate heterogeneity to the riverscape (Hoef & Peterson, 2010; Peterson et al., 2013). The work herein aims to use these relationships to uncover some of the structural properties of river networks and their implications for stream biota. Culturally and economically important, salmon are given particular attention as they inhabit and migrate through these systems.

## 1.2 River network climate portfolios

Ecological networks are known to buffer against perturbation, often through source-sink dynamics or functional redundancies (Fortuna et al., 2006). For instance, mycorrhizal networks in forests share nutrients and information during periods of stress, using "hub" like networks that facilitate the inoculation of yet-affected individuals, leading to greater forest-wide stability (Simard, 2015). The hierarchical structure of river networks is increasingly being recognized to facilitate similar stabilizing properties but through the propagation and integration of asynchrony also known as portfolio effects. Just as a well diversified financial portfolio can buffer against market instability (Markowitz, 1952; Doak et al., 1998), a river network buffers against volatility by integrating across a climatically-diverse watershed. For instance, flow and temperature dynamics have been shown to become more stable as watershed area increases (Wood et al., 1990; Moore et al., 2015). Moore et al. (2015) introduced the "watershed stability hypothesis" as a potential general framework for understanding patterns of riverine processes as a function of upstream riverine diversity or catchment size.

Under this framework, stability of key river processes lower in the watershed is reflective of diversity throughout the watershed.

Applying the watershed stability hypothesis concept at longer time scales, chapter 2 investigates whether river networks buffer against climate change. Impacting every aspect of precipitation (timing, frequency, state, etc.), climate change has and will continue to become manifest in stream discharge ( $\text{m}^3 \cdot \text{sec}^{-1}$ ) (Bindoff et al., 2013). By estimating trends in discharge throughout a large and topographically complex watershed, I demonstrate the networks' structural propensity to integrate a diverse climate portfolio and naturally stabilize flow regimes despite increasingly variable and diverse climate patterns throughout the basin. I argue that these findings suggest naturally free-flowing, connected and structurally complex river networks may have an intrinsic capacity to mitigate the local effects of climate change on rivers.

### 1.3 River network phenological match-mismatch

Phenology refers to the cyclical pattern of natural phenomena such as the annual foliation and defoliation of deciduous trees in temperate regions. In this example, pollinating insects may "time" their spring development to coincide with the bloom of specific plants during the growing season, thereby synchronizing or "matching" their phenologies. Missing important resource windows due to phenological mismatch could be devastating to some species (Thackeray et al., 2016). As climate change progresses, previously reliable environmental cues may no longer act as accurate predictors of resource phenology (e.g., Visser et al., 2006). The problem of phenological match-mismatch may be particularly pronounced for species migrating long distances as the climate at the origin may be decoupled from that of the destination and trending differently (Both et al., 2006).

Combining the idea of climate heterogeneity in chapter 2 with the concept of network driven spatial segregation and adaptation, my third chapter considers the spatial challenges of maintaining phenological synchrony with a distant resource in the face of climate change. Specifically, this chapter considers Pink salmon (*Oncorhynchus gorbuscha*) populations spread throughout a large and climatically-diverse watershed and quantifies their juvenile, outmigration match-mismatch potential with shifting estuarine phytoplankton blooms (Allen & Wolfe, 2013; Malick et al., 2015). In so doing, I demonstrate the constraints of distance and climate on juvenile outmigration as climate change impacts are expressed heterogeneously throughout the watershed. More distant populations had a greater degree of mismatch with estuary resources, suggesting strong spatial patterning to phenological mismatch driven by the spatial dynamics of river networks.

## 1.4 Big data bottleneck

Temperature controls the biology of the aquatic environment (Brett, 1969; Kingsolver, 2009; Angilletta, 2009), commonly referred to as a "master" variable. Thus, monitoring stream temperatures is an essential and routine component of any aquatic ecological study. Miniaturization and affordability of long-lived and environmentally robust sensors have made temperature monitoring in recent decades relatively trivial. The democratization of this technology has led to increased volumes of spatially- and temporally-resolved temperature databases containing millions of individual records (Isaak et al., 2017). These data volumes are incredibly important for monitoring temperatures within and across river networks in the face of climate change and other broad anthropogenic impacts. Unfortunately, quality assurance of these data becomes increasingly difficult as data volumes grow.

In chapter 4, I develop an unsupervised machine learning approach to data cleaning with the aim of improving and standardizing data quality while reducing manual labor. Using a Hidden Markov model framework, I compete models describing the expectation of air and water temperatures. Leveraging information in previous and subsequent time steps and incorporating known climatic constraints under a Bayesian framework, this fully data-driven analytical approach to error detection is flexible and consistent. New methodologies such as these offer the potential for objective error detection that improve and standardize data quality which reduces barriers to data sharing and centralized database development. Moreover, automated tools liberate researchers from the tedium of these essential but time consuming tasks that distract from the primary research.

## 1.5 River network cumulative effects

Stream temperatures are ultimately driven by climate but are heavily mediated by landscapes (Isaak et al., 2015) and strongly correlated along flow connected pathways (Hoef & Peterson, 2010). As a result, responses to global warming and shifting land use practices are expected to be extremely variable and reflect cumulative affects from the entire contributing area. Monitoring and quantifying the thermalscape (e.g., Tonolla et al., 2012) stands to greatly enhance diagnosis of troubled ecosystems and facilitate improved conservation strategies (Isaak et al., 2017). For fish species such as salmon that migrate through and inhabit entire watersheds, basin-wide monitoring paired with state-of-the-art stream network statistical methods (Hoef & Peterson, 2010; Peterson et al., 2013), will improve meta-population assessment and portfolio management (e.g., Anderson et al., 2015).

In salmon, adults return from the ocean between spring and fall and migrate upriver to spawn in their natal stream (Quinn, 2011). This migration constitutes some of the most thermally strenuous conditions of their lives. Acute thermal extremes during migration can result in localized mass mortality while non-lethal but stressful temperatures increase the risk of mortality through varied pathways (e.g., disease, energy depletion) (Richter &

Kolmes, 2005). As climate change progresses, stressful migratory stream temperatures are likely to increase in frequency and magnitude and reduce the likelihood of salmon survival (Palmer et al., 2009).

In chapter 5, I quantify the varied cumulative thermal exposure potentials along different migration routes that constitute varied landscape features and topological complexity. To do this I placed temperature sensors throughout the Thompson River watershed in central British Columbia, Canada. Using these data, I uniquely married a descriptive Bayesian temporal model and spatial stream network model to predict temperatures at any place within the network at any time during the observed period. This allowed me to probabilistically quantify the potential for thermal exposure of different populations migrating to different locations in the basin. In so doing, this chapter provides a measure of sub-lethal temperatures that will help identify particularly at-risk populations and inform future conservation management.

## 1.6 Contributions

My thesis contributes fundamental research on the dynamics of large river systems and their impacts on riverine biota with implications for conservation and management. Considering three of the four dimensions of the lotic ecosystem (Ward, 1989), I quantify the nature of integrating over space (latitudinal), along the network (longitudinal) and over time (temporal). The patterns that arise offer insights that can inform conservation management such as population or meta-population threat assessment and recovery planning. Specifically, chapters 3 and 5 offer new spatial considerations of phenological match-mismatch and thermal exposure potential respectively that could help direct conservation planning and portfolio assessment in the face of changing climate and landuse practices. Additionally, my work aims to facilitate big data studies in riverine science by offering tools that improve data quality and reliability. Together, I am hopeful that this thesis stands as a meaningful contribution to fluvial landscape ecology that will inform future work and improve our capacity to manage, protect and improve these essential ecosystems.

Although this work is largely my own, I developed three of my chapters in intellectual collaboration. As such I have written this introduction, Chapter 4 and the general discussion (Chapter 6) in the first-person singular but will use the first-person plural in chapters 2, 3 and 5. In all chapters I gathered the data, coded the analysis, interpreted results and wrote the first drafts of the text. Chapter 2 was developed, analyzed and written in collaboration with Jonathan (Jon) Moore and Sean Anderson. The idea was conceived during discussions with Jon and the analysis brought to maturity under Sean's guidance. Chapter 3 developed out of discussions with Jon and Samantha Wilson and both were instrumental in writing the manuscript. Chapter 5 was first conceived in discussions with Jon, Sean and David

Patterson and brought to final completion by me with manuscript contributions from Jon and David.

## Chapter 2

# River networks dampen long-term hydrological signals of climate change<sup>1</sup>

### 2.1 Abstract

River networks may dampen local hydrologic signals of climate change through the aggregation of upstream climate portfolio assets. Here we examine this hypothesis using flow and climate trend estimates (1970–2007) at 55 hydrometric gauge stations and across their contributing watersheds' within the Fraser River basin in British Columbia, Canada. Using a null hypothesis framework, we compared our observed attenuation of river flow trends as a function of increasing area and climate trend diversity, with null-simulated estimates to gauge the likelihood and strength of our observations. We found the Fraser River reduced variability in downstream long-term discharge by >91%, with >3.1 times the attenuation than would be expected under null-simulation. Although the strength of dampening varied seasonally, our findings indicate that large free-flowing rivers offer a powerful and largely unappreciated process of climate change mitigation. River networks that integrate a diverse climate portfolio can dampen local extremes and offer climate change relief to riverine biota.

### 2.2 Introduction

As billions are spent adapting to climate change by such means as building dikes or improving stormwater drainage (Narain et al., 2011), certain habitats and processes provide a natural defense system that mitigates climate change impacts (Jones et al., 2012). For example, coastal habitats such as oyster reefs, mangrove forests, and eelgrass beds shield 67%

<sup>1</sup>A version of this chapter appears as Chezik, K.A., S.C. Anderson, J.W. Moore. 2017. River networks dampen long-term hydrological signals of climate change. *Geophysical Research Letters*. 44 <http://doi.org/10.1002/2017GL074376>.

of the United States coastline and 1.4 million people from sea level rise and storm impacts (Arkema et al., 2013). Protection and enhancement of these natural systems has become a key component of many climate mitigation strategies (Guerry et al., 2015), as they return multiple cost-effective services in contrast to engineered alternatives (Jones et al., 2012). For instance, revitalization of the Mississippi delta not only improves storm protection, but also increases production and sustainability of fisheries and wildlife resources, protects the water supply, and reduces the impact of wastewater effluent (Louisiana Department of Natural Resources, 1998). The value of climate-buffering habitats stands only to increase as average global temperatures climb. The Paris Agreement of 2016 aims to limit global temperature rise to 1.5°C (Hulme, 2016), a goal that concedes a minimum doubling of observed warming (Hartmann et al., 2013), therefore, an urgent challenge and opportunity is discovering, maintaining and restoring systems that naturally buffer against the oncoming impacts of climate change.

In the face of changing global precipitation patterns (Donat et al., 2016), river networks may offer an unappreciated defense against shifting flow regimes under climate change (Hartmann et al., 2013; Palmer et al., 2009). Earlier snowmelt (Rauscher et al., 2008), reduced snow pack (McCabe & Wolock, 2014), shifts from snow to rain, and changes in the annual distribution of precipitation are all impacting stream flow (Hartmann et al., 2013), resulting in flashier flows with increased frequency of flooding (Hirabayashi et al., 2013), longer periods of drought and critically low flows for wildlife and agriculture (Melillo et al., 2014). However, we hypothesize that large free-flowing river networks may naturally dampen these signals of climate change by integrating across varied landscapes and different manifestations of climate. Processes that aggregate across asynchronous components tend to dampen the aggregate, a process known as the portfolio effect (Doak et al., 1998). Climate is expressed differently and asynchronously over the landscape (e.g., Wang et al., 2012) and on account of their branching architecture and directional flow, river networks integrate these varied climate manifestations (Peterson et al., 2013), potentially dampening the local climatic response of sub-catchments. Large free-flowing rivers may have diverse climate portfolios, with sub-catchments acting as assets in their portfolio, thereby dampening local climate impacts. Portfolio effects in rivers are already known to dampen short-term variation (Moore et al., 2015; Yeakel et al., 2014) but attenuation of long-term climate driven shifts have yet to be examined.

Here we consider whether river networks attenuate local long-term climate change through analysis of hydrometric data from one of the largest free-flowing rivers in North America. Located in British Columbia, Canada, the Fraser River drains an area approximately the size of the United Kingdom ( $\sim 217,000 \text{ km}^2$ ) and in a rare combination is fairly well-monitored while also having no dams on its main stem (Vörösmarty et al., 2010) (Figure 2.1). This watershed drains interior high plateau, coastal mountains, and the Canadian Rockies, and thus integrates a diverse mosaic of landscapes, weather, and climate. As with

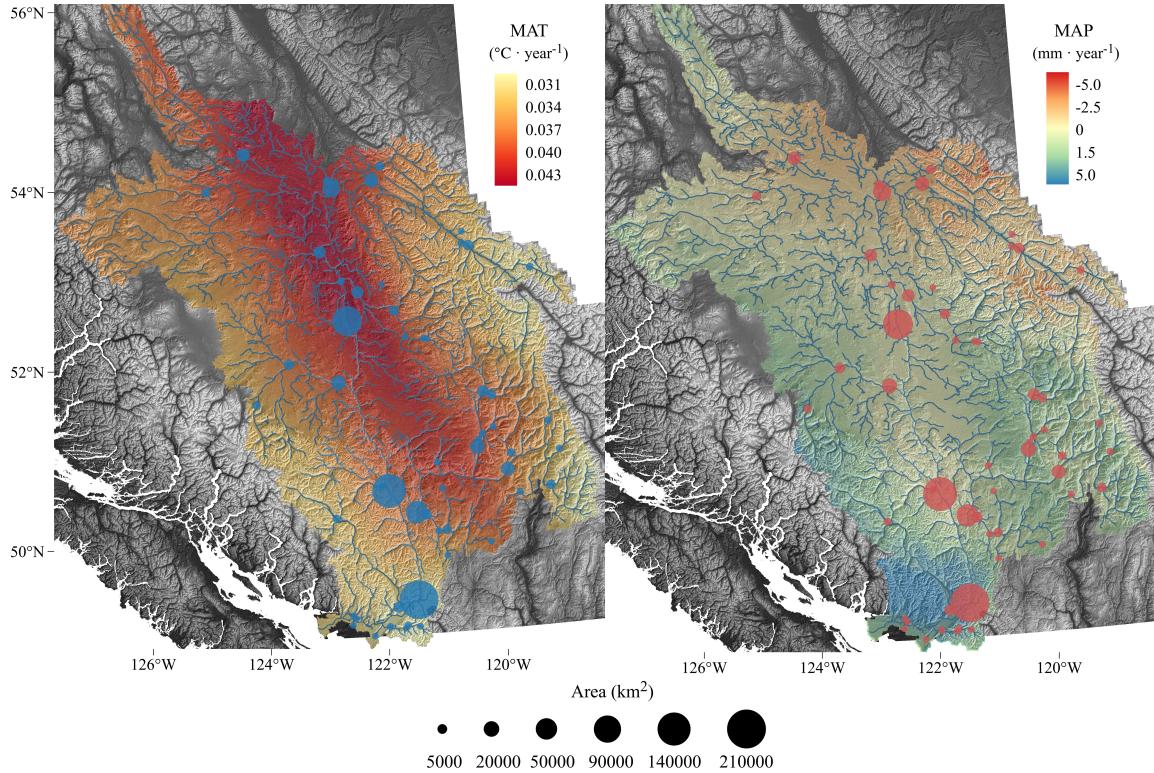


Figure 2.1: Climate trends (1970 - 2007) in mean annual temperature (MAT) and mean annual precipitation (MAP) within the Fraser River basin overlaid on a digital elevation model within British Columbia, Canada. Flow gauge sites (dots) are scaled by the size of the contributing area. The map is projected in Albers BC UTM to provide an equal area depiction of the region but labels are expressed in WGS84 latitude and longitude.

other mid-latitude rivers (Bindoff et al., 2013), the Fraser River’s discharge volume and variability have increased over the last several decades (Déry et al., 2012; Morrison et al., 2002). Following projected patterns of climate change, the Fraser River basin is increasingly shifting from a snow dominated to rain-snow hybrid system (Kang et al., 2016). Using 38 years of flow data collected concurrently at 55 hydrometric gauge stations throughout the watershed, we use a novel analytical approach to examine whether this large river exhibits signals of long-term climate dampening.

## 2.3 Materials and methods

### 2.3.1 Discharge data and flow metrics

To assess signals of climate change in hydrology within the Fraser River basin we downloaded daily flow data from Environment Canada’s HYDAT database. Within HYDAT we selected Fraser River basin hydro-metric gauge stations that were not observed to be dam influenced and which collected data in each month between 1970 and 2007, where no month was missing more than 5 days data and at least 35 complete years were present. We selected this 38-year range because this period maximized the number of gauge stations in the Fraser River basin ( $n=55$ ) operating concurrently over a long enough time period in which climate trends may be observable. We linearly interpolated missing data in log-space, but of the 746,544 days of data only 23 days were interpolated. These daily data were smoothed using a five-day rolling average before summarizing into annual and monthly response variables. This smoothing technique was used to reduce the influence of erroneous and unusually extreme values (e.g., Déry et al., 2009). Because dams control flow and stabilize flow trends, we removed sites where daily flow exhibited unusually high autocorrelation and low variability relative to other sites of similar size, and sites that were observed downstream of, and atypically close to, a dam ( $n=3$ ). The downstream distance of gauge stations from dams was determined using GIS spatial layers in the *BC Data Catalog*, maintained by the BC provincial government.

As climate change shifts precipitation patterns from snow to rain in temperate regions, rivers are generally predicted to have lower low-flows and higher high-flows, as well as earlier snowmelt (Nijssen et al., 2001). These manifestations of rising temperatures are impacting species (Xenopoulos & Lodge, 2006) and people (Hirabayashi et al., 2013). We summarized our daily rolling average discharge data into annual and monthly flow metrics that capture these transitions. We used maximum- and minimum-flow estimates to capture trends in extreme discharge and median-flow estimates to capture general annual and seasonal trends. Because timing of flow is particularly important to the phenology of many organisms, we also calculated the day-of-year in which half the annual flow was reached.

### 2.3.2 River discharge trend analysis

To test for a river network portfolio effect, we need to detect varied and asynchronous, long-term changes in flow throughout the watershed. To do this we estimated annual and monthly flow trends for each response variable at each gauge station using a generalized least squares (GLS) model with an AR1 autocorrelation function:

$$Q_{s,t} = a_s + b_s Y_{s,t} + \epsilon_{s,t}, \quad \epsilon_{s,t} \sim \mathcal{N}(\phi \epsilon_{s,t-1}, \sigma_Q^2), \quad (2.1)$$

where  $Q_{s,t}$  represents a flow metric [log(maximum), log(minimum), log(median), logit(day-of-year to half-annual-flow)] at each site ( $s$ ) and time point ( $t$ ), and  $Y_{s,t}$  represents the time in years (e.g., 1970, 1971, ..., 2007) with 1988 subtracted to approximately center the predictor. Flow metrics were scaled by site to facilitate cross-site comparisons. Parameters  $b_s$  (slope) and  $a_s$  (intercept) represent the estimated mean effect of time on  $Q_{s,t}$  and the estimate of  $Q_{s,t}$  at time zero (i.e., 1988), respectively. Error ( $\epsilon_{s,t}$ ) of the current time step for a given site was allowed to be correlated with that of the previous time step by  $\phi$  and was assumed to be normally distributed with a variance of  $\sigma_Q^2$ . To ensure our model estimates and simulations remained within the calendar year we logit transformed our flow-timing response variable (i.e., day-of-year to half-annual-flow) after scaling the data between 0 and 1 (i.e., dividing by 365).

### 2.3.3 Quantifying climate variability

For a river network to act as a portfolio of climate and dampen the effects of climate change, expressions of climate and their trends need to vary throughout the network. In order to quantify climate trend diversity we developed a climate index for each flow-gauge catchment. Using the Western North America Climate tool (ClimateWNA) (Wang et al., 2016), we estimated historic climate values on an evenly spaced  $1 \text{ km}^2$  grid across the Fraser River basin for each year and month in our study. Temperature and precipitation are likely to drive changing hydrology, therefore we included variables in our climate index that capture changes in extreme and mean temperature, as well as changes in precipitation volume and physical state (snow vs. rain). Climate variables included the mean annual temperature (MAT), and precipitation (MAP), extreme minimum (EMT) and maximum (EXT) temperature, and precipitation as snow (PAS). Climate variable trends were calculated at each grid point using a GLS model similar to eq. 2.1:

$$C_{s,t} = \alpha_s + \beta_s Y_{s,t} + \epsilon_{s,t}, \quad \epsilon_{s,t} \sim \mathcal{N}(\phi \epsilon_{s,t-1}, \sigma_C^2), \quad (2.2)$$

where  $C_{s,t}$  represents a climate variable and  $Y_{s,t}$  represents the time in years (e.g., 1970, 1971, ..., 2007). Parameters  $\beta_s$  (slope) and  $\alpha_s$  (intercept) represent the estimated mean effect of time on  $C_{s,t}$  and the estimate of  $C_{s,t}$  at time zero, respectively. Autocorrelated errors

were incorporated as in eq. 2.1. Spatial variation in climate change would suggest catchments of different size integrate different amounts of climate variability. In order to calculate a climate index for each gauge station we delineated flow-gauge catchments and summarized climate trends within those catchments. Using Whitebox Geospatial Analysis Tools (Lindsay, 2016), we delineated the Fraser River basin and the catchment areas of each flow-gauge station. Digital elevation models were provided by the provincial government of British Columbia, Canada and state government of Washington, USA and were pre-processed with a breaching algorithm and stream burned to facilitate proper flow path and accumulation models (Woodrow et al., 2016). We summarized the climate variability integrated at each flow gauge by calculating the standard deviation in climate trends within their catchments using the *rasterstats* module (<https://github.com/perrygeo/python-rasterstats.git>) in Python v2.7. The standard deviations of these climate variables were scaled between 0 and 1 and summed as a general climate index for each catchment.

### 2.3.4 Quantifying the river network portfolio effect

We hypothesize that flow gauge sites with larger climate portfolios will exhibit greater long-term flow trend stability than sites with smaller climate portfolios. We used catchment area as a proxy for the climate portfolio diversity of a given site, given that larger Fraser River sub-catchments have more variable climate portfolios (Figure 2.2) as well as aggregate more water and carry more weight on downstream dampening. Therefore, to test whether networks dampen climate signals, we regressed site-specific trend estimates (e.g., % change·decade<sup>-1</sup>) onto each site's catchment area using a GLS model:

$$\hat{b}_s = c + d\sqrt{A_s} + \eta_s, \quad \eta_s \sim \mathcal{N}(0, f(A_s)), \quad (2.3)$$

where  $\hat{b}_s$  represents a flow trend at a given site ( $s$ ) and  $A_s$  represents site  $s$ 's watershed area. We adjusted  $\sqrt{A_s}$  by subtracting the mean  $\sqrt{A}$ , thereby centering our predictor. Fitted  $d$  and  $c$  parameters represent the mean effect of watershed area ( $\sqrt{A}$ ), on flow trends and the mean flow trend for an average sized watershed (i.e.,  $\sim 5,700$  km<sup>2</sup>), respectively. Captured in our climate portfolio proxy variable ( $A$ ), is the interaction between climate and landuse. Timber harvest, the dominant source of land cover change in the Fraser River basin, scales with sub-basin area and contributes to changes in flow. Although exploratory linear models at the site level do suggest an effect of timber harvest on long-term flow trends for some gauge stations (Supplementary Figures 2.5, 2.6), when information is shared across sites in a global model, climate effects dominate (Supplementary Figure 2.7). Thus we did not attempt to separate the contributions of land cover change, climate change and their interaction on long-term flow trends.

To quantify attenuation, we modeled the change in trend variability with the change in catchment area using an exponential variance function (Pinheiro & Bates, 2000, p. 211):

$$f(A_s) = \sigma_b^2 \exp(2\delta \sqrt{A_s}), \quad (2.4)$$

where the variance of the estimated error ( $\eta_s$ ) exponentially increases with  $\sqrt{A}$  as defined by the estimated  $\delta$  parameter. This variance function allowed us to predict the range of flow trend values that would be expected as watershed area increased. A decreasing range of flow trends with increasing watershed area indicates a dampening network effect.

### 2.3.5 River attenuation null model simulations

In the observed data, small watersheds exhibited greater variability around their trend estimates than large watersheds, likely because of greater short-term variation in small catchments (Moore et al., 2015). This relationship between watershed size and trend certainty may pull small watershed trend estimates away from zero, thereby creating the appearance of decreasing trend variability among sites as watershed size increases. Using a null-hypothesis framework we simulated time-series with no underlying trend from each sites' fitted model, using the estimated standard deviation and autocorrelation parameters ( $\hat{\sigma}_Q, \hat{\phi}$ ). Using the same GLS model and AR1 correlation structure as applied to the observed data (eq. 2.1), we estimated trends for 1000 simulated time series at each site for each response variable (e.g., Supplementary Figure 2.8). This simulation process, a form of parametric bootstrapping, created distributions of null-expectations with which we compared our observed results.

We then applied equation 2.3 to each of the 1000 basin-wide simulations (e.g., Supplementary Figure 2.9), resulting in 1 observed variance exponent parameter ( $\hat{\delta}$ ) and 1000 simulated  $\hat{\delta}$  for each flow metric. By comparing our observed attenuation with our basin-wide null-model simulations we addressed the potential that more variable flows in smaller catchments contribute to the observed pattern of flow trends as a function of watershed area (Figure 2.3). This null-model approach asks how likely our observation is due to a sampling effect versus a network portfolio effect.

### 2.3.6 Dampening summary statistics

We calculated attenuation certainty as the percentage of simulated  $\hat{\delta}$  that were less than the observed  $\hat{\delta}$ . We calculated attenuation strength (AS) as the ratio between the standard deviation at the smallest and largest watersheds as defined by the exponential variance function (i.e.,  $AS = \sqrt{f(A_{s=1})} / \sqrt{f(A_{s=2})}$ ). To compare the degree of attenuation observed with the null-simulated attenuation, we compared this ratio with the same ratio calculated from the null-simulated data (i.e.,  $AS/AS_{NULL}$ ). Finally, to estimate how the day-of-year to half-annual-flow was changing, we estimated the steepest point of each site's logistic curve and multiplied this slope value by 365 to revert from our 0–1 scale to our original annual

range of 0–365. To get a decadal rate we multiplied these results by 3.8 (number of decades in 38 years). We estimated the steepest part of the logistic curve using the *divide by four* rule (Gelman & Hill, 2008) (dividing the coefficient by four equals the first derivative of the logistic curve at its steepest point).

## 2.4 Results

### 2.4.1 Hydrologic trend diversity

Long-term changes in flow metrics varied substantially across the watershed. For instance, annual maximum flow decreased in 40 sites but increased in 15 over the last four decades (-26 to 10% change·decade<sup>-1</sup>, SD = 6%). Annual median flow showed similar variability but a greater tendency towards rising flow trends with 49 increasing and only six decreasing (-5 to 25% change·decade<sup>-1</sup> SD = 5%). Thus, within the Fraser River basin, there are a variety of long-term trends in flow, suggesting a diversity of climate responses and landscape effects.

### 2.4.2 Climate trend diversity

Our index of climate variability increased with area (Figure 2.2), asymptoting in catchments larger than 40,000 km<sup>2</sup>. Thus, rivers that drain larger catchments integrate a greater diversity of shifting climates. These findings suggest that larger areas contain greater landscape complexity, which produce a variety of climates locally and respond to global climate shifts differently.

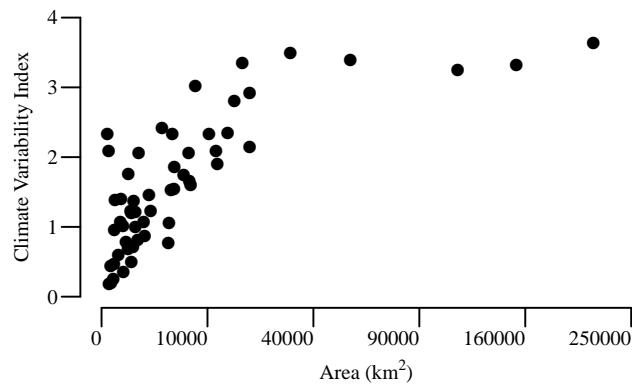


Figure 2.2: Climate variability index as a function of catchment area. As watershed area increases the variety of climatic trends within the watershed increases at a diminishing rate.

### 2.4.3 Climate portfolios and flow trend dampening

Long-term changes in hydrology were less variable in larger catchments, as predicted, demonstrating between 92 and 96% dampening of flow trends moving from headwater catchments to the network's outlet. The largest catchments were approximately 10 times more

stable in their climate response than the smallest catchments. For instance, trends ranged between a 6% reduction and 19% increase in median flow per decade among small watersheds, while large watersheds ranged between 0.8 and 1.6% change per decade. Similar attenuation was seen in flow timing with a 94% reduction in the day-of-year to half-annual-flow trends such that small watersheds were reaching half their annual flow between 8 days later and 25 days earlier while large watersheds have only shifted 2 to 4 days earlier between 1970 and 2007.

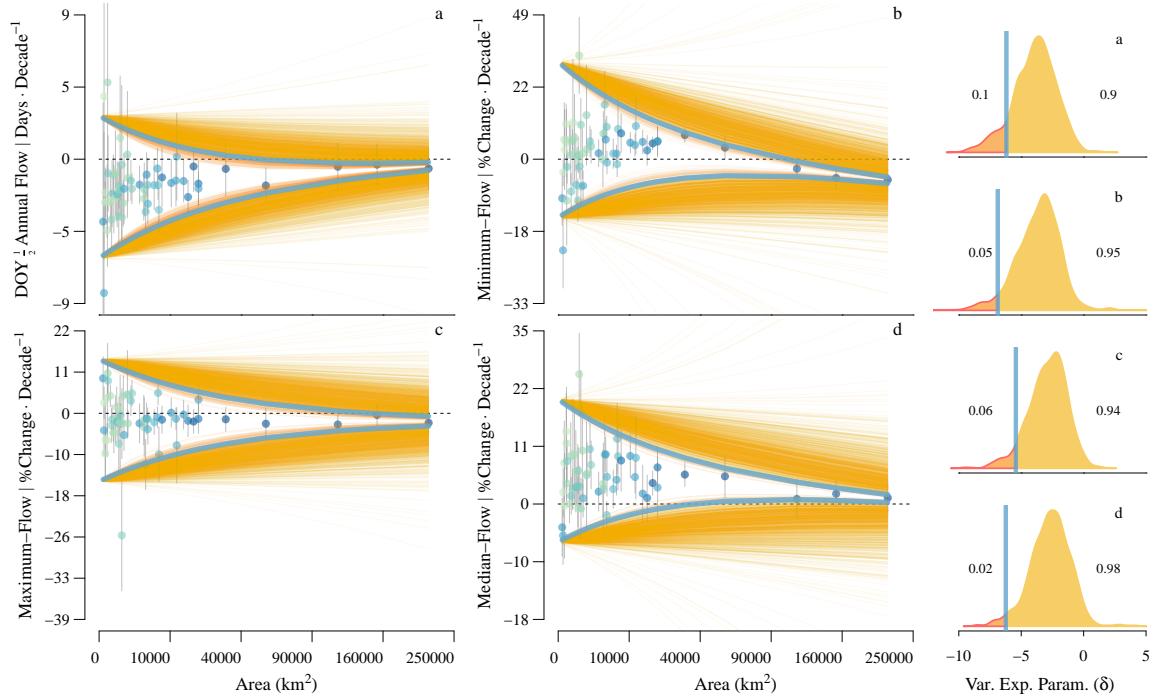


Figure 2.3: Annual flow trend attenuation within the Fraser River basin. **(Left)** Trend estimates ( $\hat{b}_s$ )  $\pm$  one standard error (SE, grey) plotted against watershed area ( $\text{km}^2$ ), colored by climate portfolio strength (green = small, blue = large). Blue lines represent observed attenuation; orange and red represent simulated attenuation that is weaker and stronger than observed. **(Right)** Density plots show null model simulated variance exponents ( $\hat{\delta}$ ) and the proportion on either side of the observed variance exponent (blue). Flow metrics include long-term flow-timing shifts (change per decade in day-of-year (DOY) to half annual flow), where decreasing trends suggest more annual flow is occurring earlier in the year (a), and the percent change per decade in minimum (b), maximum (c) and median (d) annual flow. Simulated lines ignore variance in the intercept and slope to focus visually on attenuation.

Our null model approach illustrates that support for a river network portfolio effect is greater than 90% for all four annual flow trend metrics and as high as 98% for median-annual-flow trends (Figure 2.3). Trend attenuation was 3.1 (0.6–9.8 this and hereafter are 90% intervals), 4.7 (1.1–14.7), 3.2 (0.9–9.0) and 5.2 (1.4–14.4) times greater than the null model for annual-flow timing, and minimum, maximum and median annual-flow, respec-

tively. These statistics provide strong support for the hypothesis that river network portfolio effects contribute to climate dampening in large rivers.

#### 2.4.4 Seasonally shifting flow trend dampening

Regional flow-trends and attenuation strength varied by season. Analysis of monthly trends revealed that winter flows are getting higher and summer flows are getting lower over time, but that these trends have been more variable in smaller catchments (Figure 2.4, Supplementary Figures 2.10, 2.11). For example, winter maximum flows in small catchments ranged from nearly no change to dramatic 47% increases per decade with decreases only as large as 9%. Despite these extreme locations, on average the basin has only experienced moderate winter maximum flow increases of 5-9%. Summer flows exhibited the opposite, with decreases in high flows over time. However, there was weak evidence of attenuation in the spring (Figure 2.4, Supplementary Figures. 2.10, 2.11). Overall, climate portfolios tended to have greater attenuation during stable periods of the year and decreased during seasonal transitions. In spring, the likelihood of network-driven attenuation of maximum flow trends was as low as 24% (Figure 2.4). Deterioration of the network's attenuation strength was coincident with the spring freshet, when snowmelt drives high flows across the basin. Synchronization events such as the freshet may subvert river network dampening mechanisms by homogenizing the region's response to seasonally driven climate shifts.

### 2.5 Discussion

Our results suggest that river networks mitigate and dampen long-term hydrologic trends by integrating a diverse climate portfolio over a heterogeneous landscape. By quantifying the diversity of climate and flow trends within many sub-basins of the Fraser River watershed (Figures 2.1, 2.2) we demonstrate as much as a 10 fold decrease in flow trend variability with increasing watershed area (i.e., climate portfolio) (Figures 2.3, 2.4). Similar to other temperate rivers (Rauscher et al., 2008) and consistent with recent studies (e.g., Kang et al., 2016, 2014; Shrestha et al., 2012), we show the Fraser River basin is exhibiting climate driven shifts with increasing flows over winter transitioning to decreasing flows over the summer (Figure 2.4). These findings are consistent with a general body of work that suggests climate change is transitioning previously snow driven systems to rain, resulting in a redistribution of water towards winter flows (Bindoff et al., 2013). As climate change progresses, river hydrology will continue to shift, further stressing riverine ecosystems and subsequently demanding responsive management. However, our work suggests that river networks provide an underappreciated defense against these climate-change impacts as they are uniquely organized to leverage locally-filtered expressions of climate into stability. A simple product of network form and gravity, stability emerges as river networks integrate landscape heterogeneity, creating a downstream portfolio of climate that smooths

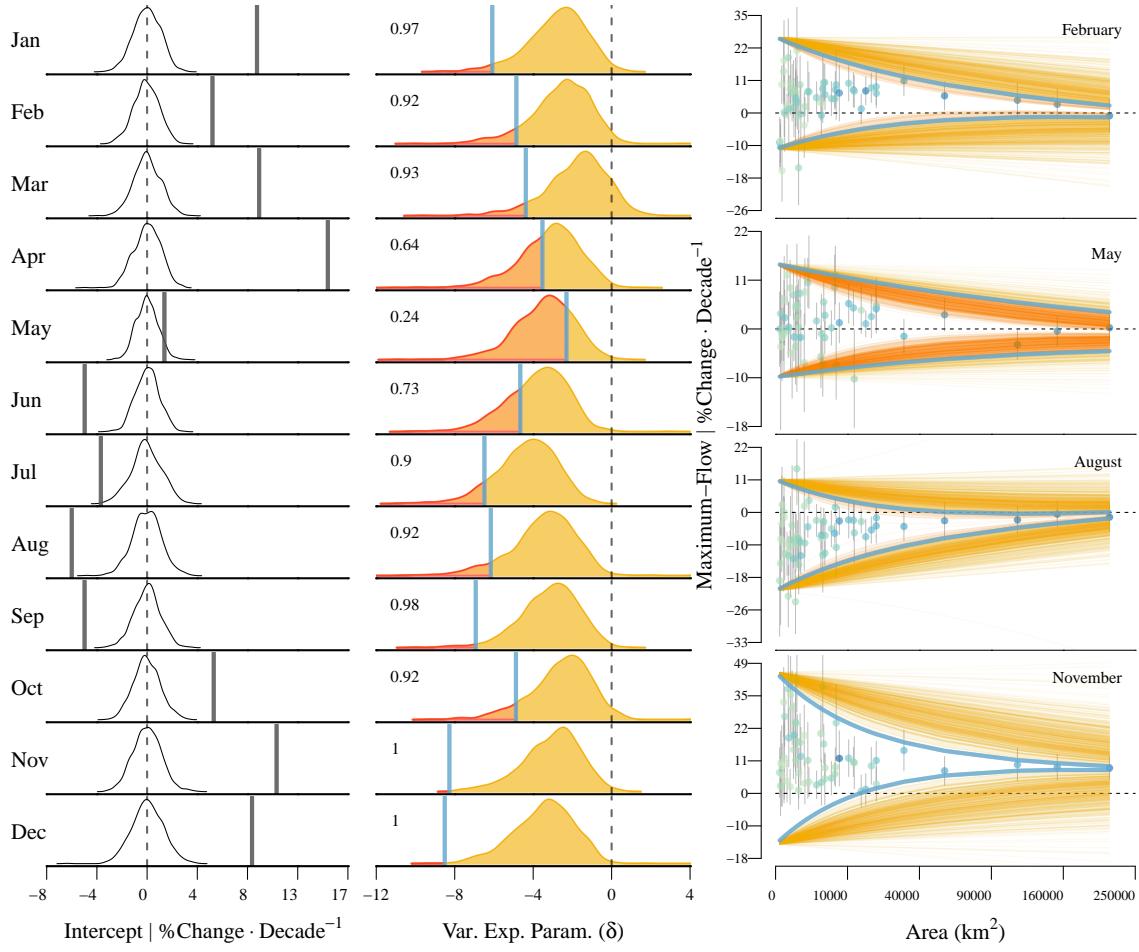


Figure 2.4: Monthly maximum flow trend attenuation within the Fraser River basin. **(Left)** Fraser River's basin-wide maximum-flow trend estimates (i.e., intercept = vertical grey lines) by month with density distributions of null-model simulations. Observed values falling further from the center of the density distribution suggest greater evidence for changes in maximum flow and a greater shift in magnitude. **(Center)** Observed monthly Fraser River maximum-flow variance exponent ( $\hat{\delta}$ , blue) and associated density distribution of simulated  $\hat{\delta}$  estimates. Decimal values represent the percent of simulated data exhibiting weaker attenuation (yellow) than observed. **(Right)** Trend estimates ( $\hat{b}_s$ )  $\pm$  one standard error (SE, grey) plotted against watershed area ( $\text{km}^2$ ), colored by climate portfolio strength (green = small, blue = large), for four seasonally representative months. These reflect months in the prior columns and describe the variation in percent change per decade of maximum flow among sites. Simulated lines ignore variance in the intercept and slope to focus visually on attenuation.

local extremes and tempers long-term flow trends. Therefore, larger rivers should have flow regimes that are less sensitive to local climate trends.

The Fraser River basin has seen changes in landcover between 1970 and 2007, which have impacted river discharge (e.g., Zhang & Wei, 2014). Over this period, timber harvest has been the dominant contributor to changes in forest cover (Supplementary Figure 2.12) but increasingly natural events are having a greater impact. For instance, abnormally hot, dry weather in 2003 led to fires that destroyed an unprecedented  $2,600 \text{ km}^2$  of forest largely within the Fraser River basin (Filmon, 2003). Beginning in the early 2000's, western mountain pine beetle outbreaks began consuming large swaths of forest, impacting 20% of British Columbia by 2013 (Schnorbus et al., 2010). These natural events are facilitated by climate change and are becoming increasingly common and extreme (Melillo et al., 2014; Maness et al., 2013). Pine beetle outbreaks are often subsequently heavily logged and timber removal can occur in anticipation of fires, resulting in harvest practices responding dynamically to climate change. Analyses indicated that while timber harvest rates did have some local impacts on hydrological change, harvest did not have substantial impacts on flows in a global model (Supplementary Figure 2.7); thus we focus on long-term hydrological change and make the simplifying assumption that different forest harvest rates are subsumed by long-term climate change. Indeed, timber harvest was relatively consistent over the last 38 years across the vast majority of our 55 study catchments, with the exception of four relatively small, heavily fire and pine beetle impacted sites (Supplementary Figure 2.12). Therefore, our study captures the interwoven direct and indirect impacts of climate change by considering changes in precipitation but also latent affects through climate filtering of altered landscapes due to climate change.

Our study provides insight into the spatial scaling of a river network's dampening capacity. Although we focused on a vast watershed with a remarkably diverse climate portfolio, we would still expect that smaller rivers or watersheds with less topographic complexity, will dampen climate variability. For instance, based on our data, we predict that rivers draining a catchment of  $60,000 \text{ km}^2$  would have 66% less variability in the rate of change in hydrology than a smaller river draining  $5000 \text{ km}^2$ . While this dampening is less than the 10-fold decrease observed in the entire  $217,000 \text{ km}^2$  Fraser River watershed, smaller watersheds can offer a defense against the impacts of climate change.

## 2.6 Conclusion

Climate change is causing economic and conservation challenges for river systems worldwide (Palmer et al., 2009; Pecl et al., 2017). In the Fraser River, sockeye salmon have seen as much as 80% pre-spawn mortality in years with later run-off and elevated temperatures, leading managers to reduce or close the fishery (e.g., 2013 and 2015). In the absence of climate dampening, the impacts of these climate extremes could have been considerably

worse. In contrast to previously studied types of climate mitigation where habitats modify climate drivers (e.g., carbon storage, (Jones et al., 2012)) or physically absorb climate change impacts (e.g., mangroves, (Arkema et al., 2013)), river network climate dampening smooths out extreme climatic trends thereby reducing the impact of local extremes. But just as the destruction of coastal habitats degrades their natural capacity to mitigate sea-level rise or storm events (Arkema et al., 2013), the management of river basins likely impacts the ability for river networks to attenuate climate trends. For example, dams synchronize the flow regimes of rivers (Poff et al., 2007). These anthropogenic activities may magnify the impacts of climatic shifts such as the transition from snow- to rain-dominated precipitation in some basins. The climate cost of these watershed activities should be considered in environmental decision making. Climate change is a global challenge; here we suggest large rivers dampen local change by leveraging landscape diversity into a portfolio of climate, an important tool in the climate-mitigation toolbox.

## 2.7 Acknowledgements

This paper relied on many open source tools including R v3.3.2 (<https://www.r-project.org/>), Python v2.7, QGIS v2.18 and WhiteboxGAT (Lindsay, 2016). All R code can be found at <http://github.com/kchezik/River-Network-Flow-Trends.git>. Data processing was facilitated greatly by GNU Parallel (Tange, 2011). This study relied on decades of flow data collected by those at Environment Canada. Raw HYDAT data can be obtained via Environment Canada or a simplified version can be found on GitHub. K.A. Chezik and J.W. Moore were supported by the Liber Ero Chair of Coastal Science and Management and Simon Fraser University. S.C. Anderson was supported by the David H. Smith Conservation Research Fellowship.

## 2.8 Supplementary material

The following supplementary material includes 9 additional figures. Figures 2.5, 2.6, 2.7 and 2.12 explore the impact of timber harvest on flow in the Fraser River basin in British Columbia, CA. These figures aim to demonstrate the complicated and unclear relationship between flow trends and forest cover change. Timber harvest data were obtained through British Columbia's Forest Practices Board, based on the Vegetation Resources Inventory (VRI) (<http://www.for.gov.bc.ca/hts/vri/>). Figures 2.10 and 2.11 are similar to figure 4 in the main text where monthly flow trends in the Fraser River basin are shown along with trend attenuation results. These supplementary figures highlight minimum and median seasonal flow trends rather than maximum seasonal flow trends. Figures 2.8 and 2.9 are example plots demonstrating a single null-model iteration for the median flow metric. Figure 2.13 shows annual discharge curves at each flow gauge station in the Fraser River basin. This figure demonstrates that our flow data are well approximated by normal and lognormal

distributions which suggests median, maximum, minimum flow metrics and our flow timing variable (i.e., day-of-year to half annual flow), are descriptive of the system and unlikely to be biased.



Figure 2.5: Median annual flow plotted at an annual time step between 1970 and 2007 for 55 flow gauge stations within the Fraser River basin. Points are colored by the percent of the total watershed area harvested in the previous five years. Grey (blue: sites with no harvest) lines indicate the trend in flow over time using a linear model. Light grey around each trend line represents the 95% confidence interval.

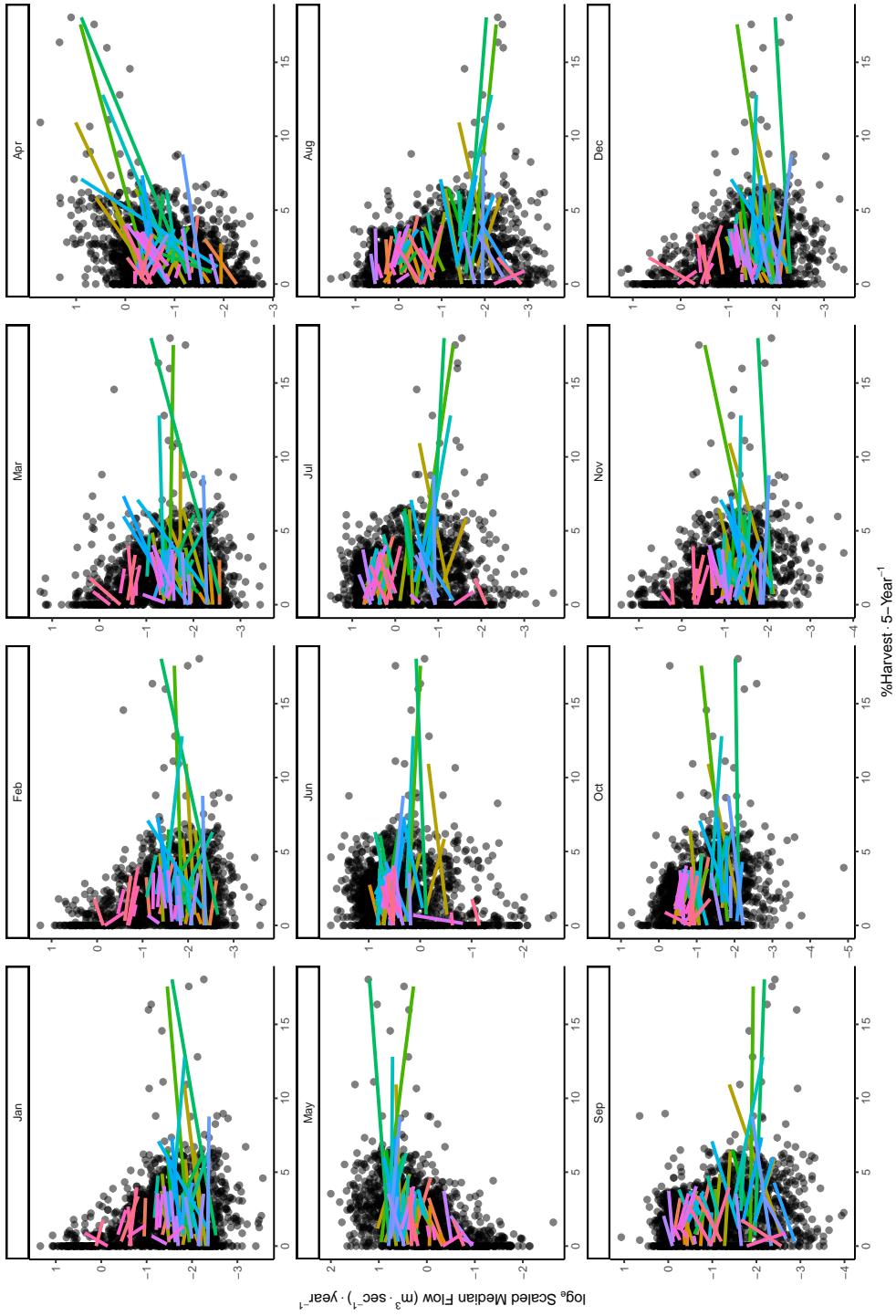


Figure 2.6: The response of monthly median flow trends (1970-2007) to logging in 55 sub-basins (colors) of the Fraser River watershed. Timber harvest is quantified as the percent of watershed area cut in the previous five years. Flow trends represent the mean change in median flow over time after being scaled by site and log transformed to facilitate cross site comparison. Lines, estimated using a linear model, depict the change in flow trends with increasing timber harvest for each flow gauge station.

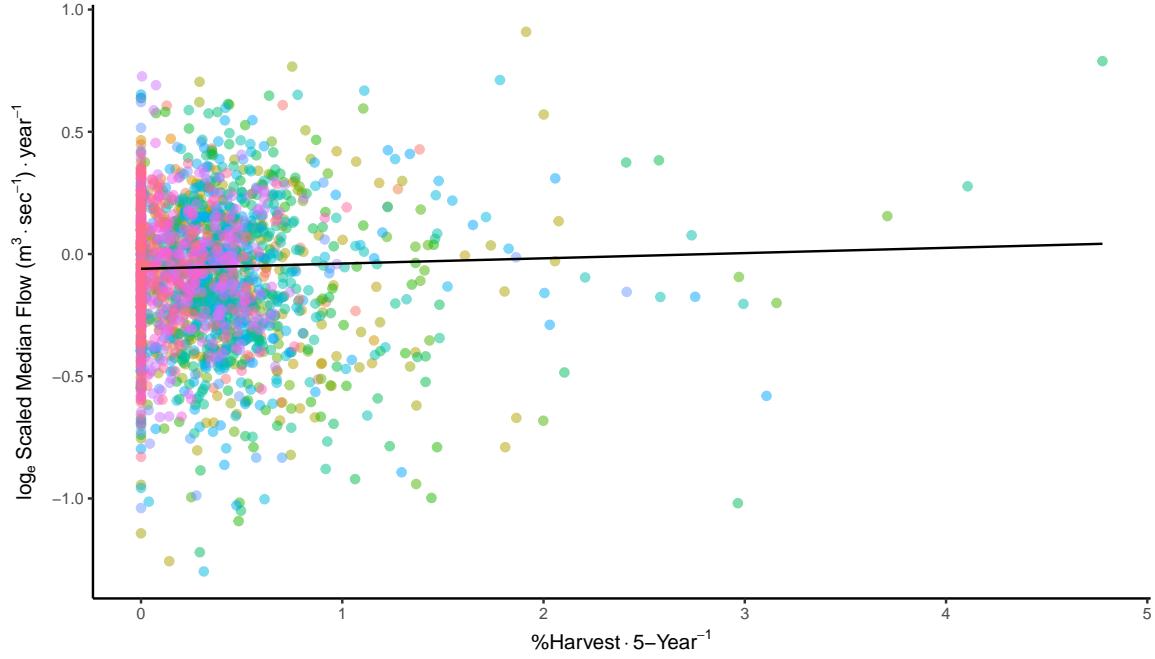


Figure 2.7: The global response of median annual flow trends in the Fraser River basin to timber harvest between 1970 and 2007. Points are colored by flow gauge station. Flow trends represent the mean change in the median flow over time after being scaled by site and log transformed for cross site comparison. The black line represents the fixed effect estimate of a generalized linear mixed model with a Gamma distribution and log link function, where random intercepts were allowed to vary by flow gauge station with a single shared slope. This model describes the typical impact of harvest on median flow trends across the Fraser River basin.

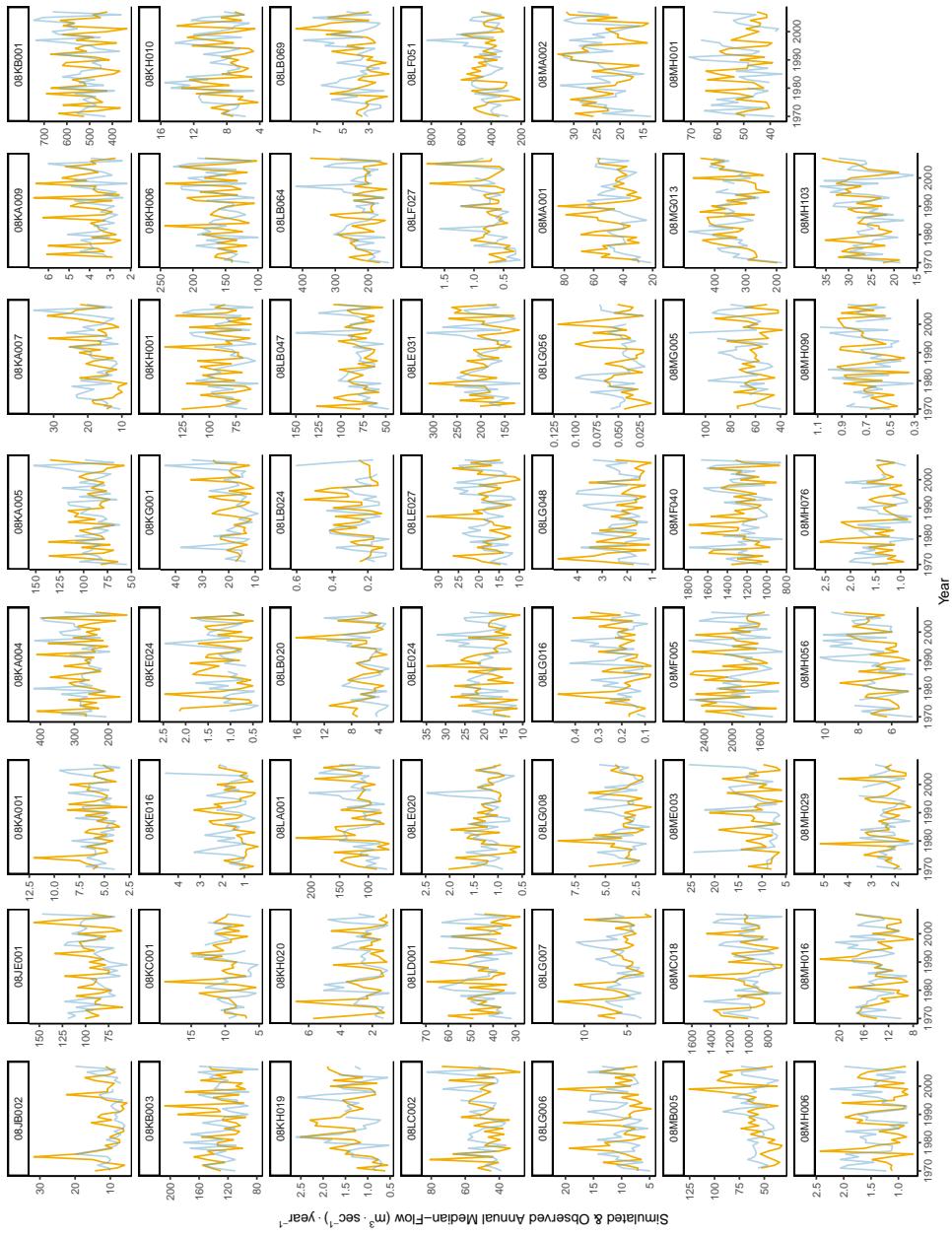


Figure 2.8: Example site specific flow simulations. Annual median-flow simulations (yellow) and observations (blue) for all 55 sites considered in this study found within the Fraser River basin. Simulations at each site were parameterized using the observed data but no trend was imposed. Fifty-five site simulations equate to one basin-wide simulation. Each flow metric was simulated 1000 times basin-wide per response variable. See supplemental figure 2.9 for resultant basin-wide simulation.

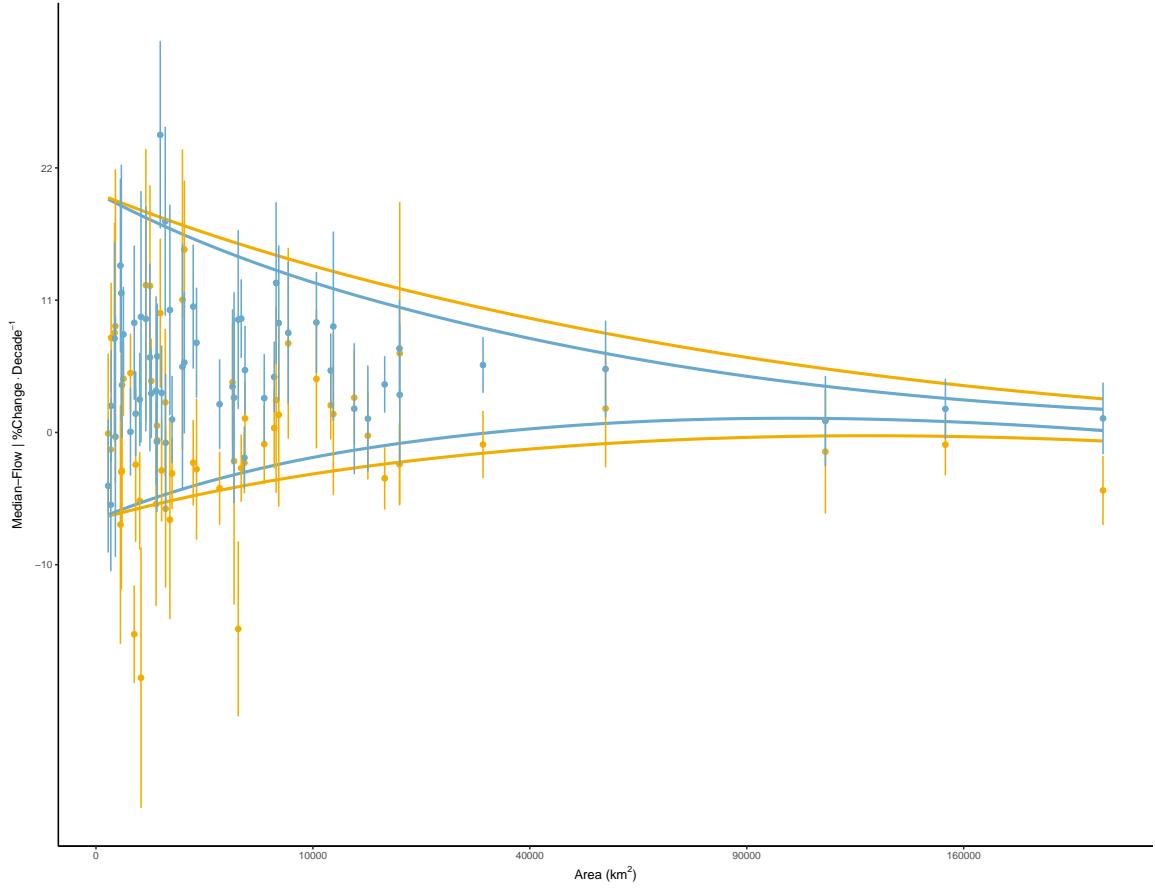


Figure 2.9: Example flow simulation results. Trend estimates (dots)  $\pm$ SE (vertical lines) of simulated null-model data (yellow) and the observed data upon which simulations were based (blue), plotted against watershed area. Curved lines represent the estimated variance function given the simulated data. This is a single example result of the 1000 null-model simulations produced from the data in supplemental figure 2.8.

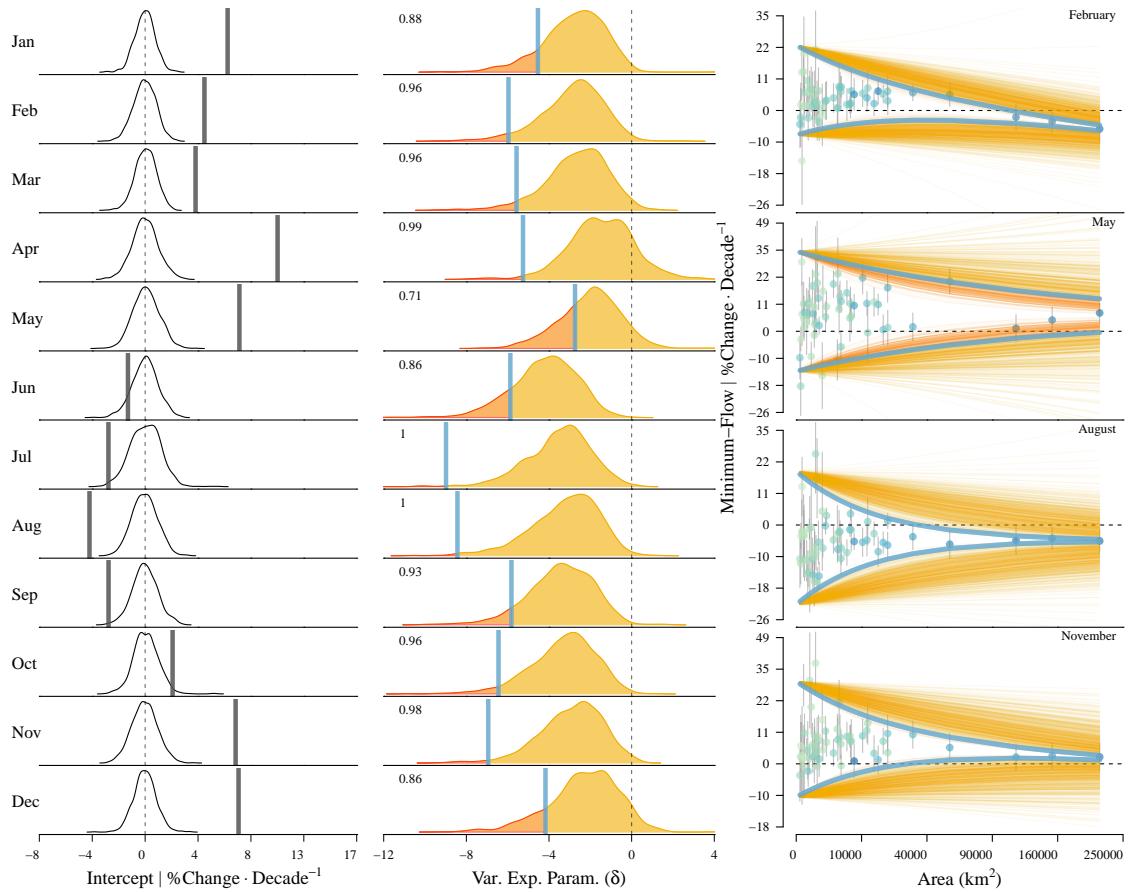


Figure 2.10: Monthly minimum flow trend attenuation within the Fraser River basin. **(Left)** Fraser River's basin-wide minimum-flow trend estimates (i.e., intercept = vertical grey lines) by month with density distributions of null-model simulations. Observed values falling further from the center of the density distribution suggest greater evidence for changes in minimum flow and a greater shift in magnitude. **(Center)** Observed monthly Fraser River minimum-flow variance exponent ( $\hat{\delta}$ , blue) and associated density distribution of simulated  $\hat{\delta}$  estimates. Decimal values represent the percent of simulated data exhibiting weaker attenuation (yellow) than observed. **(Right)** Trend estimates ( $\hat{b}_s$ )  $\pm$  one standard error (SE, grey) plotted against watershed area ( $\text{km}^2$ ), colored by climate portfolio strength (green = small, blue = large), for four seasonally representative months. These reflect months in the prior columns and describe the variation in percent change per decade of minimum flow among sites. Simulated lines ignore variance in the intercept and slope to focus visually on attenuation.

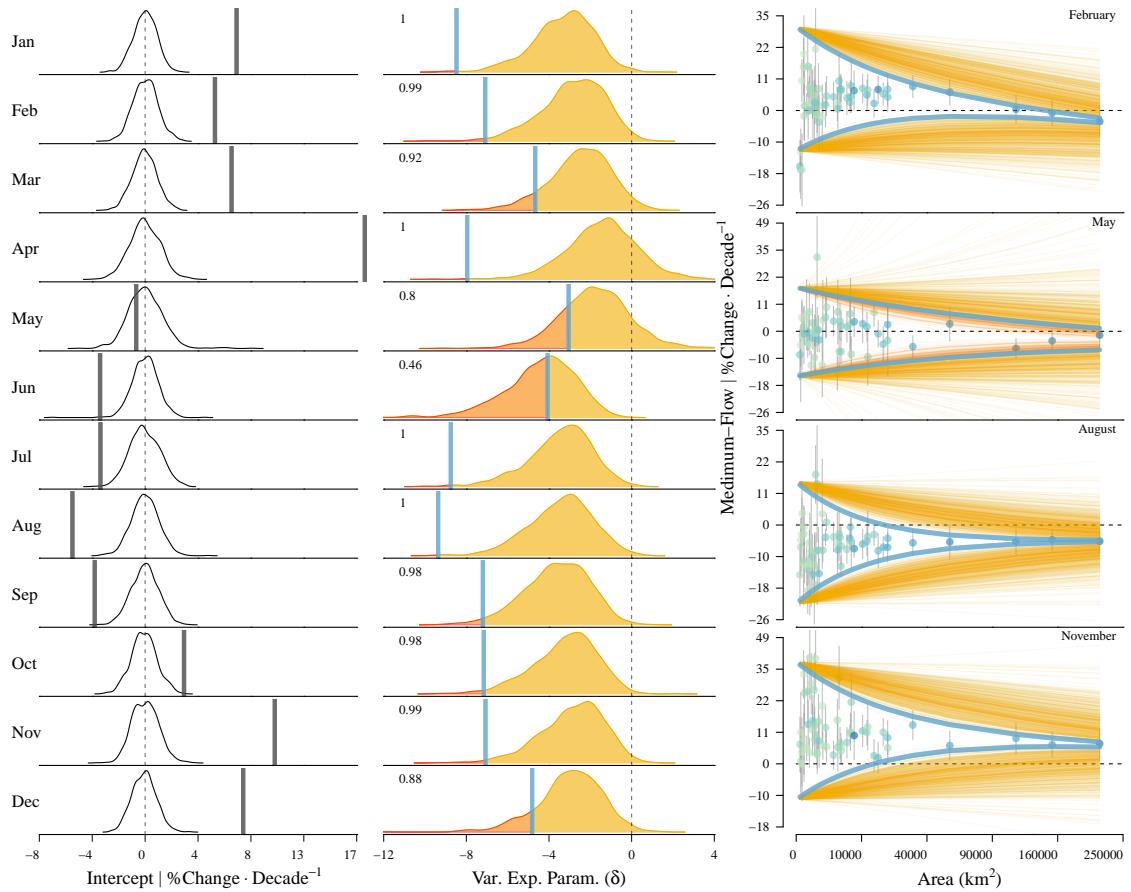


Figure 2.11: Monthly median flow trend attenuation within the Fraser River basin. (**Left**) Fraser River’s basin-wide median-flow trend estimates (i.e., intercept = vertical grey lines) by month with density distributions of null-model simulations. Observed values falling further from the center of the density distribution suggest greater evidence for changes in median flow and a greater shift in magnitude. (**Center**) Observed monthly Fraser River median-flow variance exponent ( $\hat{\delta}$ , blue) and associated density distribution of simulated  $\hat{\delta}$  estimates. Decimal values represent the percent of simulated data exhibiting weaker attenuation (yellow) than observed. (**Right**) Trend estimates ( $\hat{b}_s$ )  $\pm$  one standard error (SE, grey) plotted against watershed area ( $\text{km}^2$ ), colored by climate portfolio strength (green = small, blue = large), for four seasonally representative months. These reflect months in the prior columns and describe the variation in percent change per decade of median flow among sites. Simulated lines ignore variance in the intercept and slope to focus visually on attenuation.

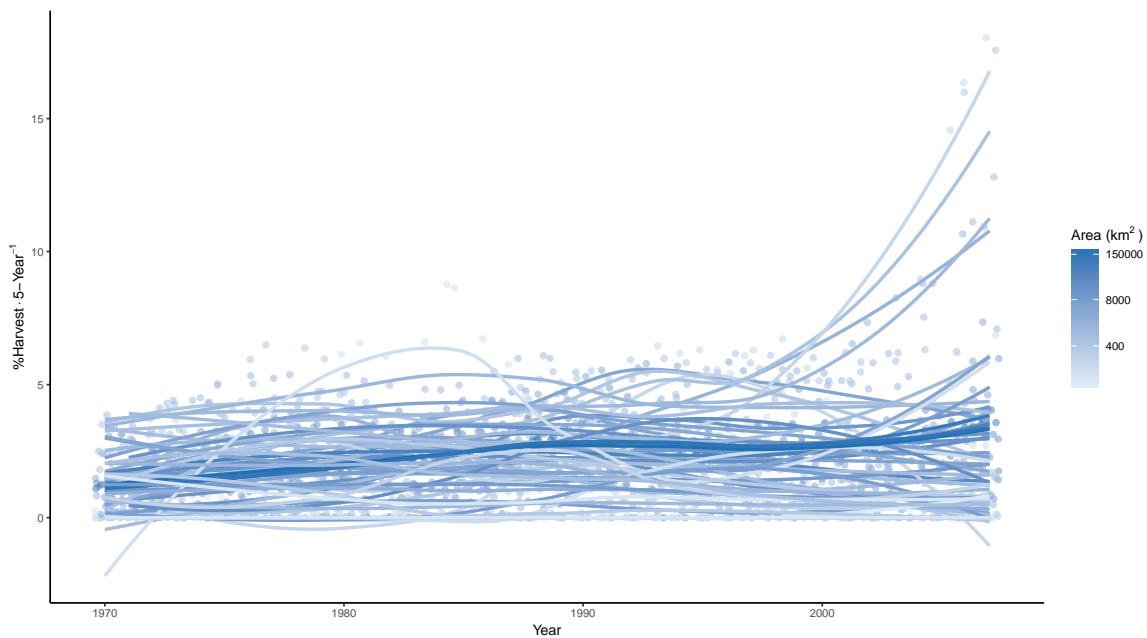


Figure 2.12: Timber harvest between 1970 and 2007 within 55 sub-basins of the Fraser River watershed in British Columbia, CA. Timber harvest is quantified as the percent of each sub-basins area harvested in the previous five years. Lines are loess smoothers indicating general trends over time for each sub-basin and are colored by sub-basin drainage area.

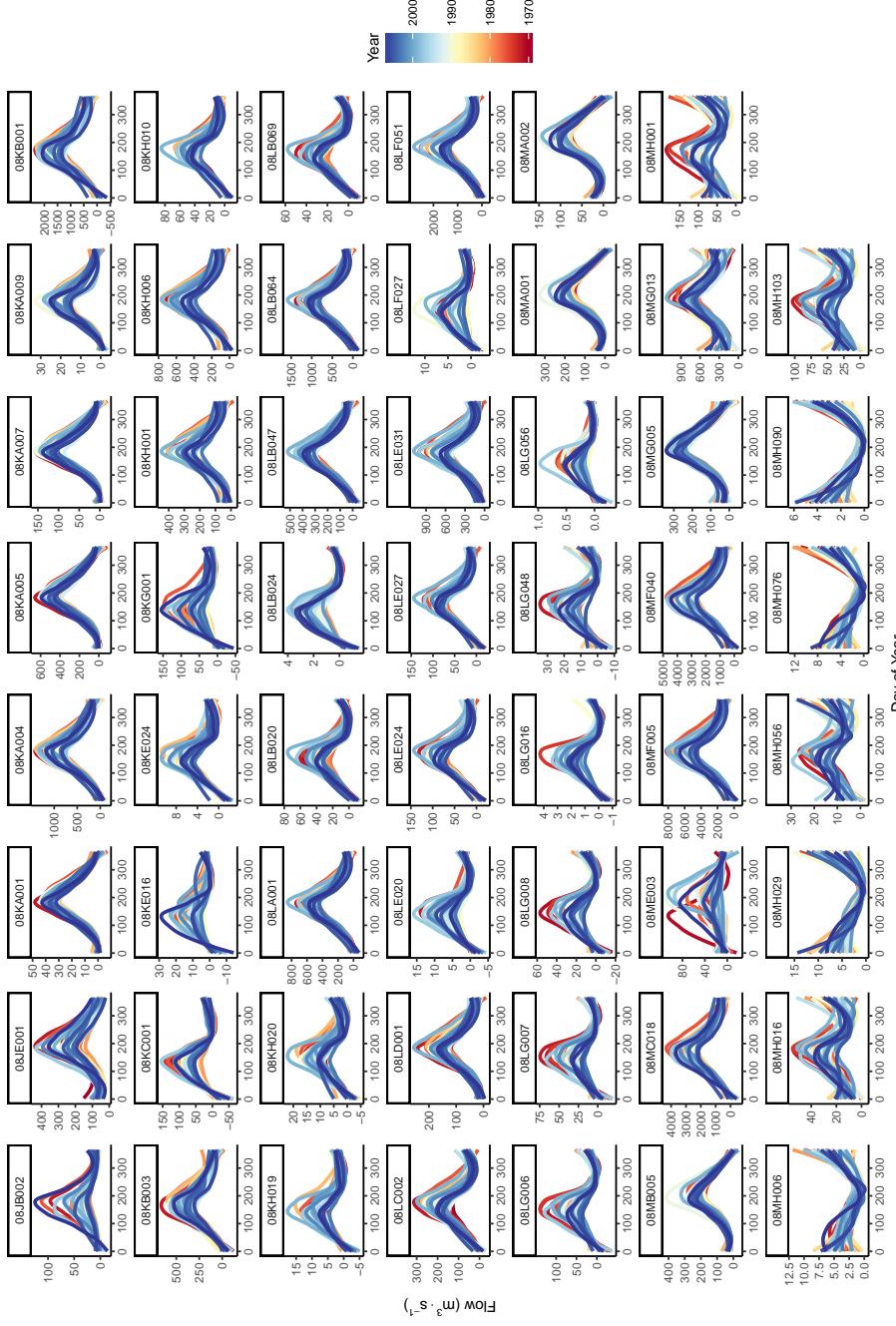


Figure 2.13: Smoothed, five-day rolling average flow rates for each day in each year between 1970 and 2007 at 55 flow gauge stations in the Fraser River watershed. Years are colored oldest (blue) to most recent (red). The distributions demonstrate that flow in the Fraser River basin is defined by a spring melt and typically low summer flows. Flow timing, duration and magnitude vary between sites but are largely grouped by coastal and interior locations. For instance, sites 08MH006, 08MH006, 08MH029, 08MH076 and 08MH090 experience an earlier freshet and higher fall and winter flows due to their coastal proximity.

## Chapter 3

# Spatial structuring of phenological match-mismatch in a highly migratory fish

### 3.1 Abstract

Climate change may be altering habitats unevenly over space, resulting in phenological mismatch between migratory predators and their prey. The severity of phenological mismatch may increase in migratory species as climates become increasingly dissimilar between origin and destination, which is likely a function of migration distance. Here we consider the relationship between migration distance, climate dissimilarity and consumer-prey mismatch for pink salmon (*Oncorhynchus gorbuscha*) in the Fraser River, an economically- and culturally-important migratory fish inhabiting the most productive salmon-producing watershed in Canada. Pink salmon are a migratory species that reproduce across climatically- and geomorphically-diverse habitats, and upon emerging swim downstream to a common estuary to feed on zooplankton. Arriving during the zooplankton bloom can increase juvenile salmon growth and survival. We used estimates of zooplankton bloom in the Strait of Georgia, observed and experimental spawn and emergence timing, physical stream temperature and discharge data, to predict pink salmon estuary arrival from 62 sites in 43 years across the Fraser River basin. Stream network models predicted life-history phenology over the river network and uncertainty was propagated throughout the analytical process. We discovered populations residing in the lower Fraser River basin experience more similar climates to that of the estuary, have shorter migration distances and are approximately 1.3 times more likely to arrive during the zooplankton bloom, than more climatically dissimilar populations originating upstream. Spatial structuring of mismatch was also related to the distribution shape of outmigration – more distant populations had broader distributions reducing the risk of complete mismatch. Increasingly distant populations with more dissimilar climates from that of the estuary had more uniform distributions with limited overlap potential (31% → 14%) nearly eliminating the opportunity for complete mismatch. These

findings suggest migratory distance and the spatial structuring of climate change are likely key factors controlling vulnerability to phenological match-mismatch.

### 3.2 Introduction

Climate change is disrupting the phenological synchrony of many ecological interactions (Durant et al., 2005; Visser et al., 2012; Thackeray et al., 2016; Deacy et al., 2017). As global temperatures rise many species are shifting the location and timing of life-history events (Parmesan & Yohe, 2003; Root et al., 2003). However, not all species are responding at the same rate thereby decoupling some predator and prey interactions (Cushing, 1990; Durant et al., 2007; Thackeray et al., 2016). For instance, pied flycatcher (*Ficedula hypoleuca*) migration timing is not tracking climate-driven shifts in caterpillar emergence, leading to mismatch in peak predator and prey phenology that is contributing to the observed population decline (Both et al., 2006). Thus, climate-change-driven phenological mismatch may pose challenges for predators, including those that are important for human economies and cultures.

Climate-driven changes in phenology do not always result in a phenological mismatch. Whether or not climate change effects species' phenological synchrony depends on the cues that predators and prey use (see 'cue hypothesis' Visser et al., 2012). For example species' phenological synchrony is unlikely to be affected when predators and prey use the same environmental cues for their phenology. However, if predators use static cues (e.g., photoperiod) while their prey use dynamic cues (e.g., temperature), changes in temperature are likely to cause phenological mismatch (Durant et al., 2007; Visser et al., 2012). Highly predictable prey phenology leads to tightly coupled predator and prey phenologies with low intra-annual variance (phenological breadth), while unpredictable prey phenology leads to high breadth in predator phenology (Baker et al., 1978). Subsequently, predators with higher phenological breadth are at decreased risk for a phenological mismatch (Durant et al., 2007). For example, Mertz & Myers (1994) found that cod (*Gadus morhua*) with more extended spawn timing (and thus extended larval emergence) have more consistent recruitment than populations with contracted spawn timing, presumably due to the higher phenology breadth buffering mismatch between cod larva and planktonic prey. Thus, both the type of cue (static/dynamic), as well as the breadth in phenology will influence the vulnerability of predators to a mismatch.

Migratory consumers may be particularly vulnerable to climate-induced mismatches with their prey, as migrants and prey are more likely to depend on different phenological cues (see 'cue hypothesis' Visser et al., 2012; Both et al., 2006; Visser & Both, 2005). Climate change is occurring unevenly over space with more asynchronous climatic dynamics occurring at greater distances (Moran, 1953), increasing match-mismatch challenges for migrants that travel greater distances (Durant et al., 2007). Indeed, Jones & Cresswell

(2010) observed that migrant birds that experience differential warming between breeding and wintering grounds are more likely to become mismatched with prey than migrants whose breeding and wintering grounds were changing at the same rate, an effect that was exacerbated in populations migrating greater distances. We hypothesize that populations inhabiting areas whose climates are similar to their destination habitat will be more phenologically matched than populations from more dissimilar climates. Given the relationship between distance and climate (Moran, 1953), we also predict that longer-distance migrants will be more vulnerable to climate mismatch.

Robust spatially and temporally diverse datasets describing predator and prey phenologies are rare (but see Burgess et al., 2018), likely explaining why potential spatial structure in match-mismatch is currently understudied. Yet, increasingly powerful spatial models provide an opportunity to leverage datasets with incomplete spatial or temporal coverage of phenology and environmental factors to create predictions of match-mismatch vulnerability across climatically diverse regions. For example, river network models provide a method for describing spatial relationships along a branching network allowing for a more robust statistical approach to the study of connected river systems (Peterson et al., 2013; Hoef & Peterson, 2010), which could be applied in the study of climate mismatch in migratory fishes in large rivers, such as anadromous salmon (*Oncorhynchus spp.*). Thus, there is an opportunity to use spatial models built on large datasets to provide empirically-based examinations of the hypothesis that long-distance migrants traversing climatically-dissimilar areas will be particularly vulnerable to phenological mismatch.

Anadromous Pacific (*Oncorhynchus spp.*) salmon are one culturally- and economically-important group of migratory consumers challenged by phenological mismatches. Young salmon migrate from freshwaters to the ocean where they rely upon zooplankton prey to grow large quickly, thereby increasing their odds of surviving to adulthood (Beamish & Mahnken, 2001). Zooplankton feed on phytoplankton whose abundance is dependent on nutrient and light availability near the ocean surface. Due to seasonal patterns in nutrients and light, phytoplankton blooms in mid- to high-latitude regions occur periodically, typically with a large bloom in the spring and a smaller bloom in the fall, subsequently resulting in zooplankton blooms (Friedland et al., 2017). Young salmon that migrate into the ocean during peak spring zooplankton bloom have higher growth and survival than salmon that migrate before or after peak prey availability (Chittenden et al., 2010; Mortensen et al., 2000; Scheuerell et al., 2009; Satterthwaite et al., 2014). Thus, the productivity of a given salmon population may be determined, in part, by the alignment of their migration phenology and the phenology of their ocean zooplankton prey. Yet, salmon spawn and rear in climatically- and geomorphically-diverse watersheds (Groot & Margolis, 1991). Their out-migration phenology is, at least for some species and populations, a plastic trait that is controlled by water temperatures (Crozier et al., 2008; Taylor, 2008). Thus, different populations of salmon from different freshwater locations may be differently able to track ocean

zooplankton phenology. We hypothesize that climate-driven phenological mismatch with the zooplankton bloom will be exacerbated for far-flung salmon populations from climatically different regions compared with those near the estuary.

There is emerging evidence that salmon migratory phenology is shifting and that climate mismatch may be causing population-level effects. For instance, Otero et al. (2014) found Atlantic salmon (*Salmo salar*) are outmigrating on average 2.5 days earlier per-decade since 1960 in the North Atlantic Ocean, similar to findings in the North Pacific where Taylor (2008) observed Alaskan pink salmon (*O. gorbuscha*) outmigrating 5 days earlier per-decade since 1970. These shifts in smolt outmigration appear to be occurring in different salmon species (Kovach et al., 2012, 2013), and are associated with increasing air, river and ocean temperatures (Taylor, 2008). Increasing sea surface temperatures are also shifting the timing (Richardson, 2008) and abundance (Boyce et al., 2010) of plankton blooms worldwide (Friedland et al., 2017). Although generally zooplankton are reaching peak bloom earlier, in some areas the bloom date has become more erratic (Allen & Wolfe, 2013). Regardless, the timing of salmon migration may not be tracking the shifting zooplankton bloom dynamics. The degree to which salmon phenology can track inter-annual variation in plankton bloom will likely contribute to which salmon populations are climate winners or climate losers (e.g., Chittenden et al., 2010; Malick et al., 2015). Thus, understanding potential spatial patterns in the risk of climate mismatch is of timely importance.

Here we examine whether there is spatial structure in phenological mismatch in a migratory fish of economic, ecological and cultural importance driven by spatial structure in climate dissimilarity. Locally-adapted salmon populations spawn throughout vast river networks that integrate diverse climate variability (Chezik et al., 2017), however migrants from all populations face the common challenge of matching up with the zooplankton bloom of a common estuary. Here we ask whether pink salmon populations closer to the estuary, with more similar climates, track inter-annual variation in zooplankton bloom better than populations that outmigrate from climatically dissimilar regions that are further upriver. Subsequently, we ask whether synchrony between pink salmon outmigration and their zooplankton prey has changed over time. By linking models that account for the river network spatial structure and propagating uncertainty over multiple life-history events, we estimate match-mismatch potential for 62 sites over four decades (1968-2010) within the most productive salmon watershed in Canada.

### 3.3 Materials and methods

We estimated juvenile pink salmon estuary arrival timing and phenological mismatch with their zooplankton prey over time for populations throughout the vast Fraser River of British Columbia, Canada. Long-term observational data of juvenile salmon outmigration timing and origin is limited in large rivers such as the Fraser River that contain many salmon

populations. Therefore our predictions were based on the following components. First, a model for pink salmon emergence timing was created through a quantitative synthesis of 104 experimental observations in studies from the region. This model was driven by water temperature during incubation and initiated by spawn timing. Pink salmon spawn timing estimates were derived from a fitted stream network model built on 379 unique site-year pink salmon spawn timing observations in the Fraser River watershed. Continuous daily water temperature records available at 62 sites since 1968 resulted 371 unique site-year emergence timing estimates. Thus, we predicted pink salmon emergence throughout the Fraser River watershed between 1970 and 2010, and by adding travel times to these phenologies, created estimates of the phenology of estuary arrival. These predictions of estuary arrival were compared to zooplankton phenology based on a long-term study of phytoplankton. Using these estimates of mismatch, we then asked whether the climate dissimilarity (yearly site-specific difference in climate index) was related to the amount of phenological mismatch, how this was spatially structured, and if the amount of mismatch was changing across years. All components included and propagated uncertainty. Thus, our study was based on observed life history data, environmental data, as well as experimental data. In the forthcoming methods we describe these models' structures, assumptions and limitations, their foundational observational data and our process of linking model estimates, their uncertainties and predictions in unobserved locations.

### 3.3.1 Study site

The Fraser River watershed drains approximately 217,000 km<sup>2</sup> of diverse terrain and climate (Figure 3.1). Originating in the western Cordillera of North America, the basin drains snow-laden Canadian Rockies and Coastal Mountains. Converging and running down the dry central plateau of British Columbia, the river eventually crashes through the coastal mountains by way of Hells Gate canyon and spills into the Strait of Georgia in the Pacific Ocean. We focus on the Fraser River watershed because the basin is large, topologically diverse, impacted by climate change (e.g., Kang et al., 2014, 2016; Déry et al., 2012; Schnorbus et al., 2010; Shrestha et al., 2012; Healey & Bradford, 2011) and home to hundreds of genetically and geographically distinct populations of five species of Pacific salmon as well as steelhead trout (*O. mykiss*). This salmon biodiversity supports commercial, recreational, and indigenous fisheries (Nesbitt & Moore, 2016). funneling such a wide diversity of climate and divergent climate shifts to the same ocean outlet (Chezik et al., 2017; Moore et al., 2015), salmon populations from throughout the Fraser River basin come from diverse freshwater habitats and yet experience similar ocean conditions.

### 3.3.2 Pink salmon

The Fraser River basin currently supports annual escapements of 2 to 20 million pink salmon during odd years (Ricker, 1989). Unlike other Pacific salmon species, juvenile pink salmon

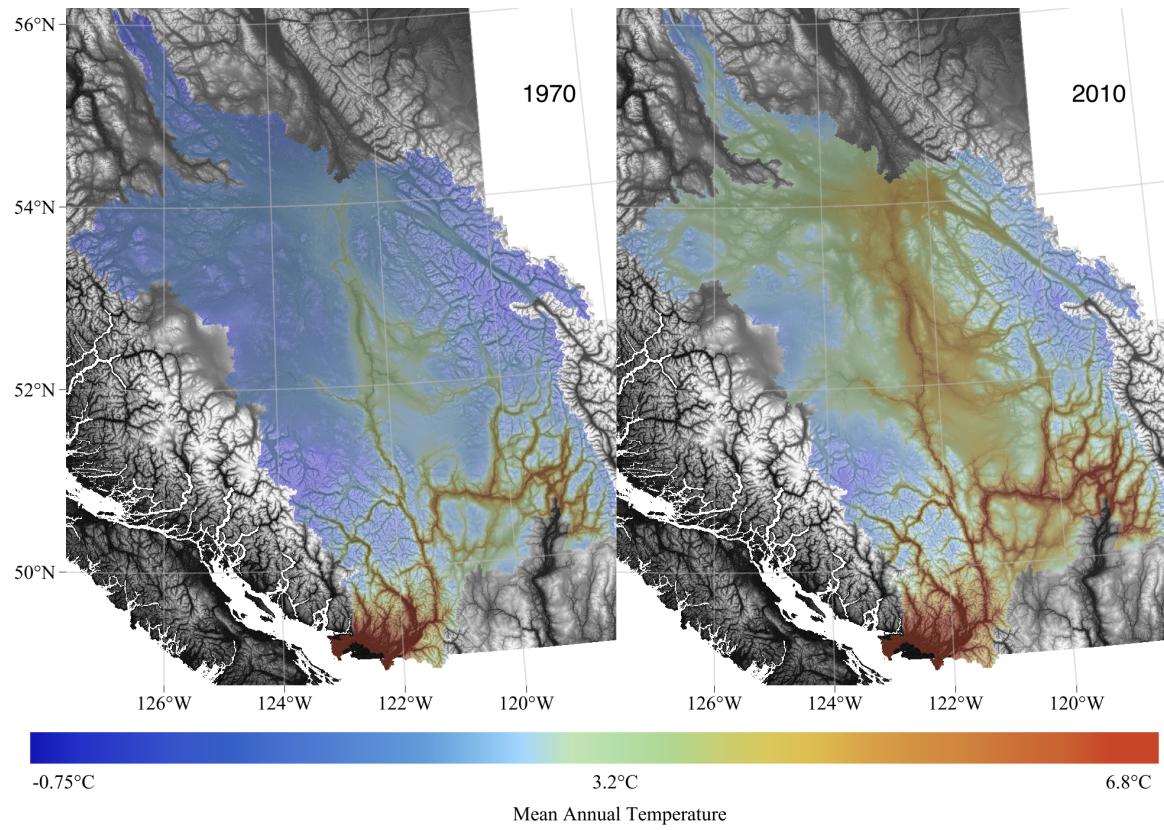


Figure 3.1: Fraser River watershed in British Columbia Canada, colored by the Climate WNA (Wang et al., 2016) estimated mean annual air temperature in 1970 (**left**) and 2010 (**right**).

migrate out to the ocean almost immediately upon emerging from the gravel. Thus, their migration timing will be largely driven by stream temperatures experienced by incubating eggs as well as the timing of spawning by their parents. There is extensive information on both spawn timing and river temperature from throughout the Fraser River watershed.

Pink salmon populations in the Fraser River are still recovering from a past extirpation event. A rockslide in 1913 in the Hells Gate area of the Fraser River, approximately 200 km upstream of the ocean, blocked migrations of upstream pink salmon populations. Fish passage was reinstated in 1947 (Roos, 1991) with the implementation of fishways. Since then, pink salmon populations have been gradually recolonizing upstream habitats (Pess et al., 2012). Given that pink salmon appear to still be recolonizing the Fraser River basin, we included potential locations that do not currently support pink salmon in our study.

### 3.3.3 Spawn timing

Spawn timing determines when thermal time begins accumulating thereby acting as an adult control on fry emergence and subsequently juvenile estuary arrival timing. Therefore, we must know when spawning occurred in order to estimate emergence. Spawn timing observations were not always coincident in time and space with recorded water temperature records (see below). In order to predict when spawning would occur where water temperature data are available, we built a spatial statistical stream-network linear mixed effects model (SSNM):

$$S_{s,t} = \mathbf{X}\boldsymbol{\beta} + \epsilon_{s,t}, \quad \epsilon \sim z_d + z_t + z_s + z_{nug}, \quad (3.1)$$

where spawn date  $S$  at site  $s$  in year  $t$  is predicted by a matrix of climate and landscape variables  $\mathbf{X}$ , where the relationship between spawn date and these predictors are described by a vector of  $\boldsymbol{\beta}$  coefficients. Residual error ( $\epsilon$ ) is decomposed into multiple random effects ( $z$ ) to capture the spatial autocovariance between populations. A "tail-down" autocovariance model ( $z_d$ ) describes the relationship among flow unconnected locations along the river network (Hoef & Peterson, 2010). We used year ( $z_t$ ) and site ( $z_s$ ) identifiers as random effects to account for year specific variation common across sites and repeated measures at individual sites respectively. The remaining independent and random error is captured in  $z_{nug}$ .

We gathered spawn timing data from the nuSEDS database owned and maintained by Fisheries and Oceans Canada (DFO). We selected pink salmon observations from throughout the Fraser River basin that included an estimate of the beginning and peak spawn date. To capture variation in spawn timing we assumed the spawning period to be approximated by a normal distribution and sampled distributions for each site and year. Standard deviation ( $\sigma$ ) estimates describing each distribution were approximated by searching for a value where the 99 and 50<sup>th</sup> percentile closely described the start and peak spawn dates. We sampled these distributions 22 times for each of the 379 unique site-year combinations,

resulting in 8734 spawn date estimates at 64 locations in odd years between 1957 and 1998. We limited our sub-sampling to reduce computational complexity and improve model fit times (Hoef et al., 2014).

Temperature has been loosely connected with spawn timing in pink salmon (Groot & Margolis, 1991) and may contribute to genetic controls on spawning and egg development (Hebert et al., 1998; Smoker et al., 1998). To account for temperature on spawn timing we used Climate WNA (Wang et al., 2016) to estimate mean fall (September - October) and winter (November - February) air temperature as well as mean fall precipitation at observed and prediction sites. We also included elevation to act as a relative estimate of migration strenuousness and to account for topographic interactions with precipitation. These four climate and landscape variables make up matrix  $\mathbf{X}$  in equation 3.1.

We chose a stream network model in part because pink salmon have only recently recolonized much of their historic range and in part because pink salmon have a strong straying tendency relative to other salmon species (Pess et al., 2012). Therefore we expect populations that are closer together along the river network to share more genetic information and respond more similarly to abiotic drivers than populations further apart. Accounting for network spatial relationships offers a way of directly addressing straying behaviour thereby reducing uncertainty in our spawn timing predictions.

### 3.3.4 Emergence timing model

Incubation temperatures control the rate of egg growth and development (Murray & Beacham, 1986), such that if we understand the relationship between temperature and development time, we can predict emergence timing using observed temperature data. In order to account for temperature fluctuations in the natural environment we took a degree-day approach where the accumulation of thermal time since fertilization predicts the timing of emergence. This model is appealing because stream temperature data is widely available throughout the Fraser River basin.

Studies that reported days to emergence during a controlled thermal regime were gathered to build a predictive model for emergence timing. For instance, Beacham & Murray (1986) incubated five stocks of fertilized pink salmon eggs at 4, 8 and 12°C and recorded the number of days to 50% emergence. By multiplying the days to emergence by the incubation temperature and summing these values, we get an estimate for the cumulative number of degree-days to emergence for each population under different thermal regimes. Many studies incubated eggs at thermal constants (Brannon, 1987; Beacham & Murray, 1988; Beacham, 1988; Murray & McPhail, 1988), while others used variable thermal regimes that mimic seasonal shifts (Murray & Beacham, 1986; Beacham & Murray, 1987). Temperature accumulation under variable thermal regimes were accounted for provided the details of each studies methods. Our literature search resulted in 104 estimates of cumulative degree-days

and days to emergence for 20 pink salmon populations in British Columbia, derived from seven studies.

To describe the relationship between thermal accumulation and days to emergence we used a linear regression model:

$$\text{logit}(\mathbf{E}_i) = b + m\text{CDD}_{T_0,i} + \eta_i, \quad \eta_i \sim \mathcal{N}(0, f(\text{CDD}_{T_0,i})), \quad (3.2)$$

where  $\mathbf{E}_i$  represents the number of days to emergence at each observation ( $i$ ) and  $\text{CDD}_{T_0,i}$  represents cumulative degree days per observation ( $i$ ) given a threshold temperature ( $T$ ) of 0°C (Chezik et al., 2014). We adjusted  $\text{CDD}_{T_0,i}$  by subtracting the mean  $\text{CDD}_{T_0}$ , thereby centering our predictor. To ensure the number of days to emergence didn't extend beyond one year, we logit transformed  $\mathbf{E}_i$  after scaling the data between 0 and 1 (i.e., dividing by 365). Fitted  $m$  and  $b$  parameters represent the mean effect of  $\text{CDD}_{T_0}$  on days to emergence and the mean number of days to emergence for the average number of  $\text{CDD}_{T_0}$ , respectively. In order to allow for heteroskedasticity in the data we modified our error structure using an exponential variance function:

$$f(\text{CDD}_{T_0,i}) = \sigma_E^2 \exp(2\delta\text{CDD}_{T_0,i}), \quad (3.3)$$

where we allowed for the variance in our error ( $\eta_i$ ) to increase with increasing  $\text{CDD}_{T_0}$  as defined by the estimated  $\delta$  parameter. This model was fit using RStan 2.16.2 (Stan Development Team, 2017) in a Bayesian framework using Hamiltonian Monte Carlo (HMC) sampling. We allowed four chains to burn in 4,000 iterations before sampling every third iteration 1,334 times, resulting in well mixed chains consistent with convergence and autocorrelation plots suggesting independence among samples. We used weakly informative normally distributed priors centered around zero with standard deviations of 1, 0.5 and 1 for  $b$ ,  $m$  and  $\delta$  respectively.

### 3.3.5 Emergence & arrival timing estimation

Pink salmon estimated emergence varies within and among populations. To capture parental and developmental contributions to variation in emergence timing we propagated error from our spawn timing model (eq. 3.1) to our emergence timing model (eq. 3.2). Defined by our spawn timing model's predicted mean spawn date and standard error estimates, we sampled prediction site student- $t$  distributions 1000 times, thereby capturing model uncertainty. Using these spawn date estimates, we then accumulated degree-days ( $\text{CDD}_{T_0,s}$ ) and days to emergence, calculating the probability of emergence each day.

Thermal accumulation ( $\text{CDD}_{T_0,s}$ ) was calculated by summing water temperatures greater than 0°C. Water temperature data were collected throughout the Fraser River basin by multiple agencies and is maintained by DFO in an unpublished database (for more information on temperature data, see Patterson et al., 2007; Thompson et al., 2010). Emergence proba-

bility was calculated by estimating  $E_s$  given posterior coefficient estimates of equation 3.2 and calculating the proportion less than or equal to the observed number of days of incubation. We retained emergence date estimates where the probability of emergence was greater than zero and less than one.

Immediate outmigration by juvenile pink salmon makes their estuary arrival timing a function of emergence timing and river distance (Neave, 1966). To estimate travel time between each population and the estuary we divided river distance by flow velocity at the time of emergence. Velocity was calculated by dividing discharge by the product of the rivers width and depth. Discharge and depth estimates were gathered from a flow gauge station in Hope, British Columbia, which is maintained by Environment Canada and made available in the HYDAT database. To estimate river width we used the statistically-derived relationships between discharge, depth and width described in Moody & Troutman (2002). We adjusted the intercept parameter that describes the relationship between flow and depth from a mean estimate of 0.27 to 0.52, well within the 95% CI range (0.12-0.63), in order to ensure unknown depth estimates agreed with the observed flow to depth relationship. Time was estimated in days and added to the day of emergence, resulting in distributions of estuary arrival dates for each site and year water temperature data were available.

### 3.3.6 Zooplankton bloom, match-mismatch and climate dissimilarity

Zooplankton are the critical food resource of juvenile pink salmon in the estuary (Neave, 1966). To determine if our arrival timing estimates coincide with food availability we need to know when past zooplankton abundances peaked. Allen & Wolfe (2013) developed a model of phytoplankton bloom estimates in the Strait of Georgia initialized with *in situ* observations and predicted by physical and biological observations. Using these modeled bloom date estimates and standard deviations, we estimated zooplankton bloom distributions for each year from 1968 to 2010.

Because zooplankton feed on phytoplankton, we approximated zooplankton bloom by applying a constant two week time lag to phytoplankton peak bloom date estimates (e.g., Chittenden et al., 2010). Multiplying the standard deviation around the mean ( $\sigma=5$ ) by three, we estimated a 30 day window capturing the most likely period over which the peak zooplankton bloom occurred. Because the peak bloom error estimate is assumed to be normally distributed, multiplying the standard deviation by three captures 99.6% of the variability around the mean estimate.

To quantify phenological match-mismatch between zooplankton bloom and pink salmon estuary arrival, we calculated the proportion of estimated arrival dates that coincide within the window of zooplankton bloom for a given year. Populations and years with greater overlap demonstrate greater opportunities for growth and survival, as previously documented by Malick et al. (2015) for pink salmon.

Ultimately, we wish to test whether similarity of climates between natal streams and the estuary is predictive of match-mismatch. To quantify how well climates across the basin track the estuary we calculated a relative measure of climatic dissimilarity in each year between each site and the estuary. Our climate dissimilarity index consists of calculating the absolute difference in temperature (mean, minimum, maximum) and precipitation (mean) between the estuary and each site, in each month of each year then standardizing these four climatic measures of similarity between zero and one for each spawn year and summing those standardized differences within each site and year. A spawn year spanned between July-01 and June-30 of the subsequent year, capturing direct impacts of climate on developing pink salmon embryos. The smaller the climate dissimilarity index value the greater the similarity between the estuary and the spawn site. Using these index values, we evaluate the relationship between climate similarity and pink salmon estuary arrival match-mismatch.

### 3.3.7 Spatial structure of climate dissimilarity and phenological match-mismatch

We tested the hypothesis that phenological mismatch was related with spatially-structured climate dissimilarity and examined whether there was a shift in predicted phenological mismatch over time. Specifically, to describe underlying trends and the spatial structure of timing mismatch we fit a SSNM:

$$\text{logit}(\hat{P}_{s,t}) = \text{CD}_{s,t}\beta + \epsilon_{s,t}, \quad \epsilon \sim z_u + z_t + z_{nug}, \quad (3.4)$$

where we explained observed phenological match, or overlap, between juvenile estuary arrival and zooplankton bloom ( $\hat{P}_{s,t}$ ) given each populations' natal stream climate dissimilarity ( $\text{CD}_{s,t}$ ) with the estuary. Similar to equation 3.1 our model includes random effects to account for year effects ( $z_t$ ) and the autocovariance among sites in the "tail-up" or flow-connected direction ( $z_u$ ), with remaining error ( $z_{nug}$ ) assumed to be independent and random. This model enables isolation of the effect of climate dissimilarity while accounting for any residual effect of spatial autocorrelation including distance ( $z_u$ ) with residual random noise captured in ( $z_{nug}$ ). We logit transformed our overlap estimates ( $\hat{P}$ ) in order to bound the data between 0 and 1. Using this model we predicted overlap given climate dissimilarity at regular intervals along the entire network thereby describing the mean effect in 1970 and 2010. To describe change in overlap (phenological match) over this 40 year period we calculated the difference in our prediction networks for these two years.

## 3.4 Results

### 3.4.1 Spawn timing spatial patterns

Modeled estimates suggest spawn timing has been trending later throughout the warming Fraser River basin (Figure 3.2) and demonstrates spatial patterns in both average timing and shifts in timing. Although our model demonstrated previously known patterns of spawning occurring earlier at higher latitude (Neave, 1966), our SSNM delineated more spatially explicit patterns of spawn timing by climatic region (Figure 3.2 & Supplementary Figure 3.8). For instance, spawn timing estimates in historically warmer southern-interior locations shifted later between 1970 and 2010 than historically cool locations in the north despite a greater degree of warming in the north (e.g., Figure 3.1). Therefore, while shifts in spawn timing among northern and western populations have been limited, populations in the southeast may have shifted by as much as two weeks. Standard error estimates of our predictions suggest the observed shifts in spawn timing may not constitute a significant linear change (Supplementary Figure 3.9). Also, we note that no population currently inhabits the northern portion of the Fraser River basin therefore those predictions are based on the stream network model of spawn timing informed by Climate WNA estimates of temperature and precipitation.

Populations closer on the river network shared more similar spawn timing than those further apart. Although a Torgogram (Hoef et al., 2014) suggested mixed evidence for flow-connected, -unconnected and linear autocorrelation (Supplementary Figure 3.10), when compared in an AIC framework the flow-unconnected network correlation structure outcompeted all other auto-correlation structures and their combinations. In other words, populations appeared more strongly correlated in the upstream direction than along flow-connected stream paths or in directions unconstrained by the network. Overall the random effects in equation 3.1 captured nearly all of the variation in spawn timing with the network ( $z_d$ ), year ( $z_y$ ) and site ( $z_s$ ) effects accounting for 51, 10 and 17% of the observed variation while climate and landscape covariates captured less than 1%. The limited explanatory power of our covariates corroborates the general consensus that abiotic factors are weakly informative of pink salmon spawn timing (Neave, 1966; Groot & Margolis, 1991).

### 3.4.2 Emergence timing model

Our synthesis of pink salmon incubation experiments confirmed strong relationships between water temperature and emergence timing. Intriguingly, the requisite number of days between fertilization and emergence declined with increasing cumulative degree-days (Figure 3.3). Therefore, locations with warmer temperatures required more thermal time but fewer days to emerge than cooler environments. For instance, a population experiencing a constant 5°C incubation temperature is expected to emerge 173 days after fertilization while a warmer 10°C location is expected to incubate for only 107 days before emerging. Although the

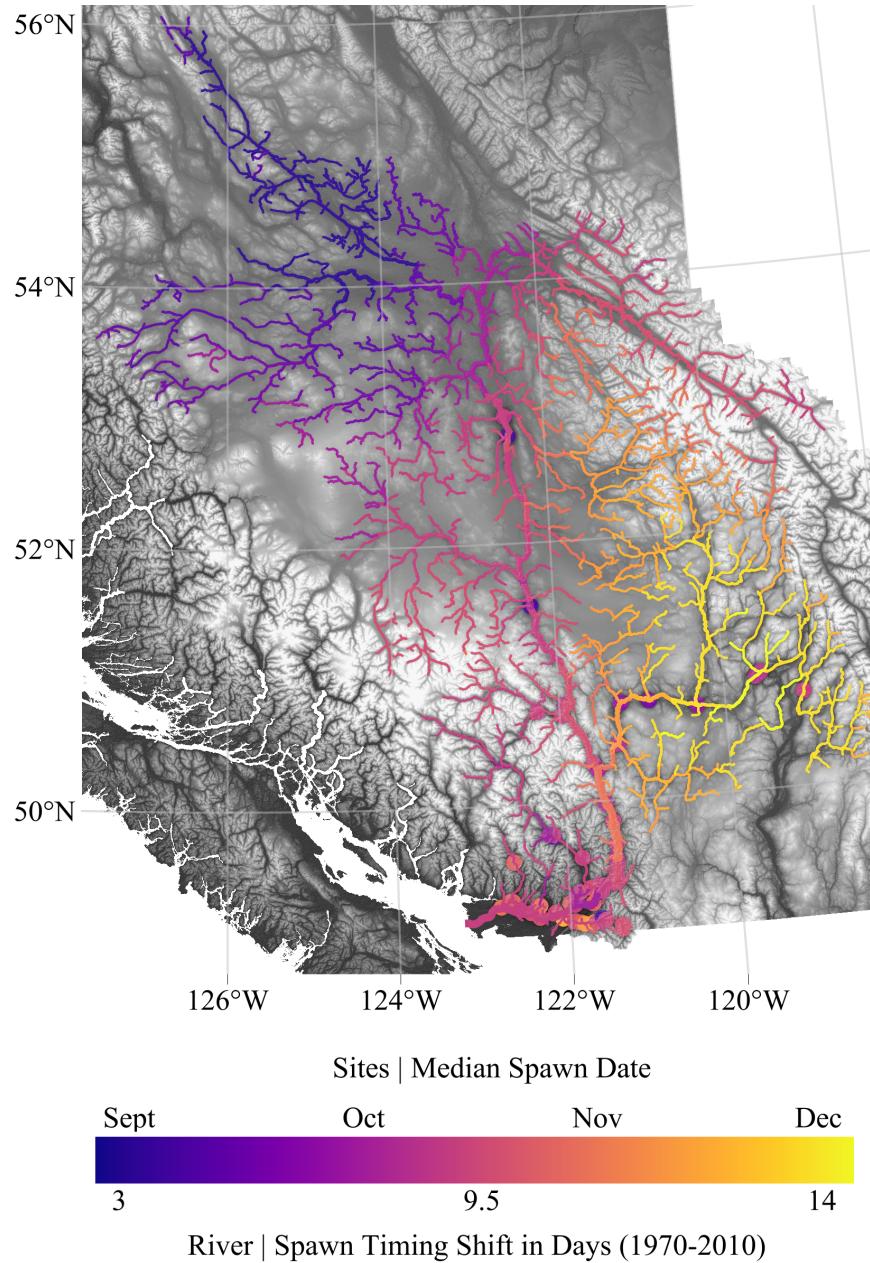


Figure 3.2: Fraser River basin pink salmon spawn locations, timing and projected shifts in timing. Points indicate the location of observed pink salmon spawning and are colored by the median spawn date across observed years. The Fraser River network is colored by the mean predicted shift in spawn date between 1970 and 2010 given our SSNM described in equation 3.1. Shifts in spawn date are expressed in days with positive values representing later seasonal spawning.

time it takes to emerge is nearly double at 5°C relative to 10°C, the number of cumulative degree-days is about 20% less. Thus, there is some degree of developmental compensation in emergence timing in pink salmon. Because the rate of degree-day accumulation is highly informative, we can effectively predict emergence timing simply by accumulating water incubation temperatures from the date of spawn.

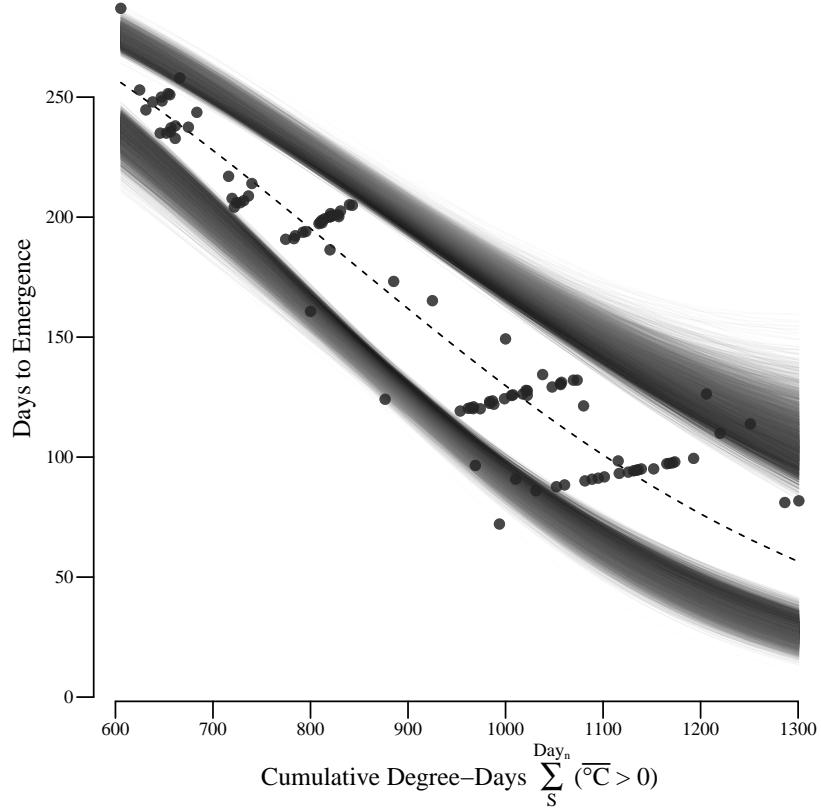


Figure 3.3: Days to emergence given the cumulative degree-days since fertilization. Cumulative degree-days are calculated as the summed mean daily temperatures above freezing ( $^{\circ}\text{C}_{T_0}$ ) from spawn (S) to emergence (day  $n$ ). Points represent observation estimates derived from published studies. The data were fit to equation 3.2 and the dotted line describes the average decline in days to emergence with increasing cumulative degree-days. The solid lines indicate the variance around the mean given all posterior estimates of  $\eta$ .

### 3.4.3 Estuary arrival, zooplankton bloom and climate dissimilarity

There was spatial structure in the predicted breadth and peak of pink salmon estuary arrival. By applying flow informed travel time estimates that ranged between less than a day to 2 weeks, to thermally-derived emergence predictions (i.e., eq. 3.2), initiated at 1000 normally distributed spawn dates (i.e., eq. 3.1), we predicted estuary arrival distributions at 62 sites across 43 years (1968-2010) (Figure 3.4). As the natal stream climate diverged from that of the estuary, these pink salmon estuary arrival distributions became increasingly uniform,

such that each arrival day shared increasingly similar probability (Figure 3.4). Concurrently, arrival distributions tended to be increasingly left skewed, spread over a wider range of dates and arriving predominately after the zooplankton bloom. Conversely, populations and years with relatively small climate dissimilarity values arrived in the estuary over a smaller range of dates, had a defined peak and were relatively well coincident with the zooplankton blooms. Overtime, inter-annual variation in zooplankton bloom has increased (Allen & Wolfe, 2013). Although pink salmon estuary arrival appears to be responsive to different years with different environmental conditions that are associated with differently-timed zooplankton blooms, there is often an overcorrection or a broadening of the distribution towards more uniform arrival (Supplementary Figures 3.11, 3.14).

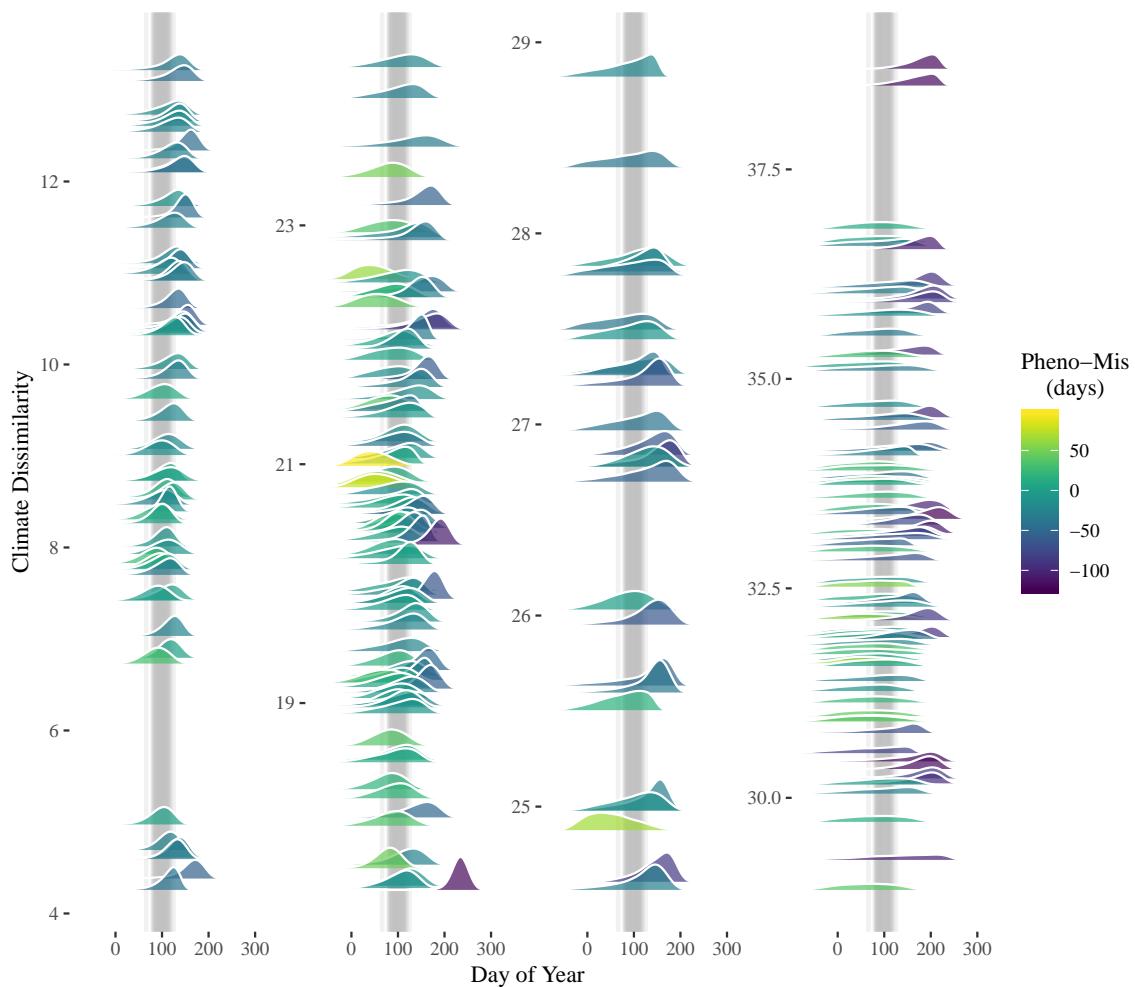


Figure 3.4: Estimated distributions of pink salmon estuary arrival timing overlaid on zooplankton bloom date estimates between 1968 and 2010, stratified by each populations natal stream climate dissimilarity with the estuary in a given year. Distributions are colored by the number of days the most common estuary arrival date (i.e., mode) missed the zooplankton bloom in the given year.

### 3.4.4 Phenological match-mismatch

On average, locations in the Fraser River basin closer to the estuary and whose climate was more similar to the estuary were estimated to have greater pink salmon estuary arrival overlap with the zooplankton bloom (Figures 3.5 & 3.6). For instance, pink salmon in the lower Fraser River basin are twice as likely on average to arrive during peak zooplankton bloom (31% overlap) than hypothetical populations migrating from near the Fraser River headwaters (14% overlap). Moreover, while climatically-similar locations lower in the watershed experienced zooplankton overlap as high as 56%, the least climatically similar regions never expressed overlap estimates greater than 26%. Similarly, the Lower Fraser River was unlikely to have low overlap values with only 6.5% of estimates falling in the lowest quartile compared with 25.2% of Upper Fraser River basin values.

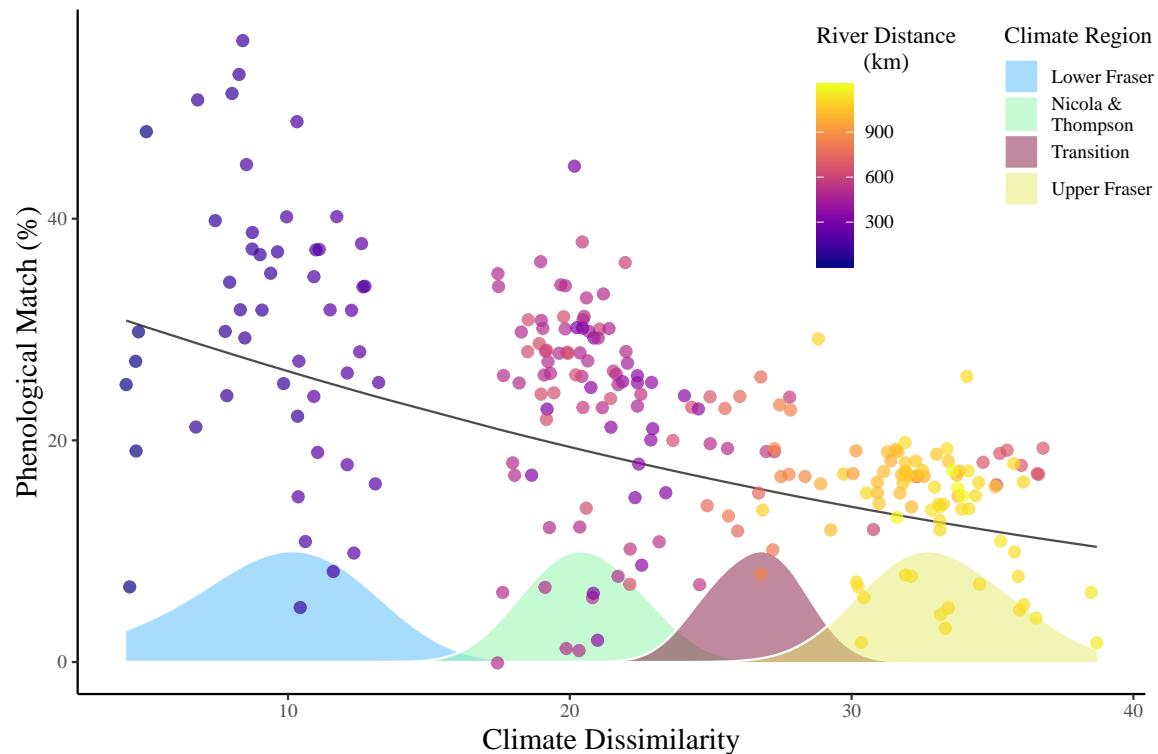


Figure 3.5: Percent phenological match between juvenile pink salmon arrival and the zooplankton bloom, given how dissimilar natal stream climates are from the estuary. Points describe the phenological match and climate dissimilarity estimates for each population in each year data were available and are colored by river distance to the estuary. The mean effect of climate dissimilarity on phenological match given a fitted estimate of equation 3.4 is described by the solid black line. Density distributions describe the spatial relationship of climate dissimilarity in the Fraser River basin (i.e., Supplementary Figure 3.8).

SSNM mapping (e.g., eq. 3.4) of pink salmon phenological match-mismatch with the zooplankton bloom across time revealed only small-magnitude shifts in the degree of climate

mismatch over the last 40 years (Figure 3.6). On average, the outmigration phenology of pink salmon as compared to zooplankton phenology became less similar in the lower Fraser River and main stem while the eastern and northwestern regions became more aligned (Figures 3.5, 3.6, & Supplementary Figure 3.8). Although our match-mismatch SSNM demonstrated a good fit and strong predictive power, we detected only minor directional shifts of no greater than 1.8% over the 40 years between 1970 and 2010, well within prediction standard error. Model results suggest climate dissimilarity does not capture all the variation (5%), especially among low dissimilarity values. The match-mismatch SSNM was able to capture much of the remaining variation (40%) in the  $z_u$  network spatial structure. LOOCV demonstrates that prediction standard error grows with observed overlap but rarely exceeded 15% (Supplementary Figure 3.12) and even captured extreme values extending well beyond the mean estimate (Supplementary Figure 3.13). Thus, while there is strong spatial structure in phenological mismatch in pink salmon, these spatial patterns do not appear to be shifting rapidly through time.

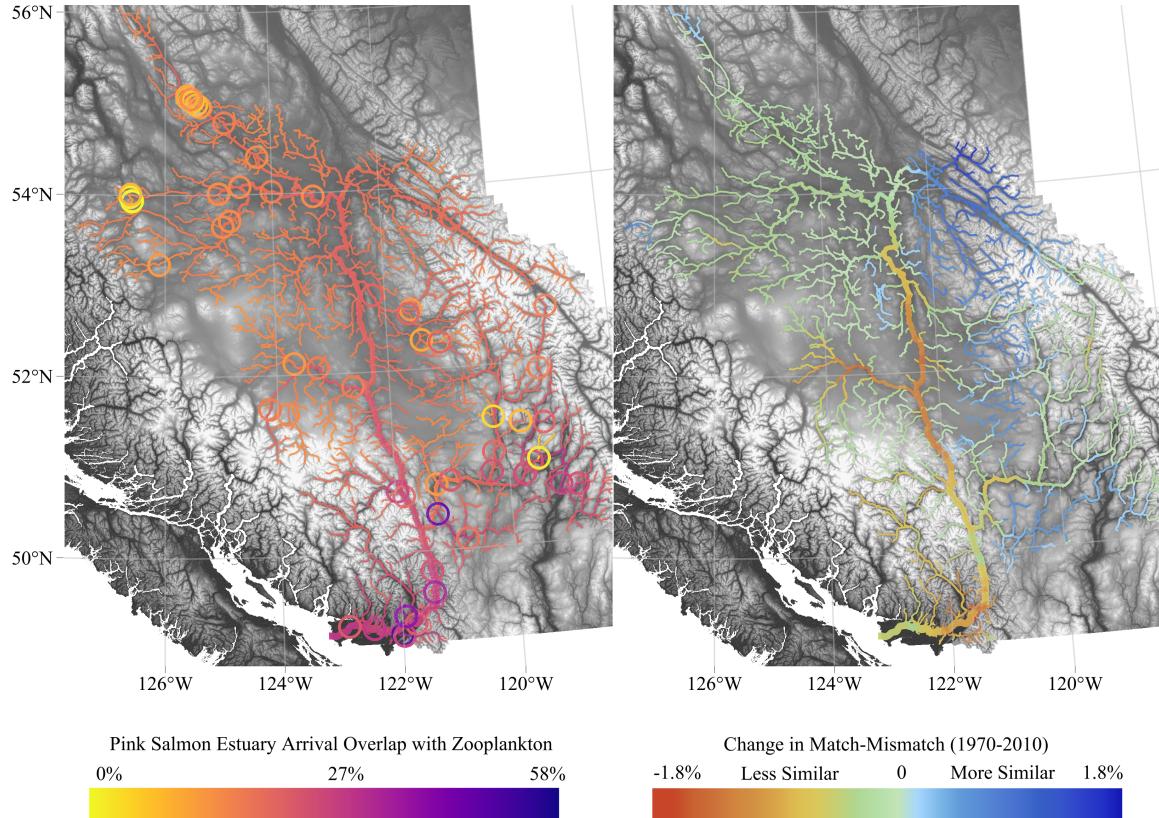


Figure 3.6: (left) Predicted phenological match (%) in estuary arrival with the 2010 zooplankton bloom throughout the Fraser River basin. Open circles indicate locations of temperature monitoring and are colored by the estimated median percent overlap across available years. (right) The difference between predicted 1970 and 2010 match-mismatch percentage across the Fraser River basin given equation 3.4.

Pink salmon populations were more likely present nearby if 20% or more of projected estuary arrival timing in the area was coincident with the zooplankton bloom (Figure 3.7). Climate dissimilarity naturally grouped into four regions of the Fraser River (Figure 3.5 & Supplementary Figure 3.8) and was strongly predictive of pink salmon presence and absence. The lower Fraser River basin constituted populations below Hells Gate Canyon while the Nicola River and Thompson River basins comprised populations above Hells Gate. A transition or edge region with few observed spawning populations was then evident before reaching the upper Fraser River basin where no known spawning populations currently exist. Thus, there is a regional structure in climate patterns, predictive of pink salmon-zooplankton match-mismatch, which covaries with the presence-absence of pink salmon populations.

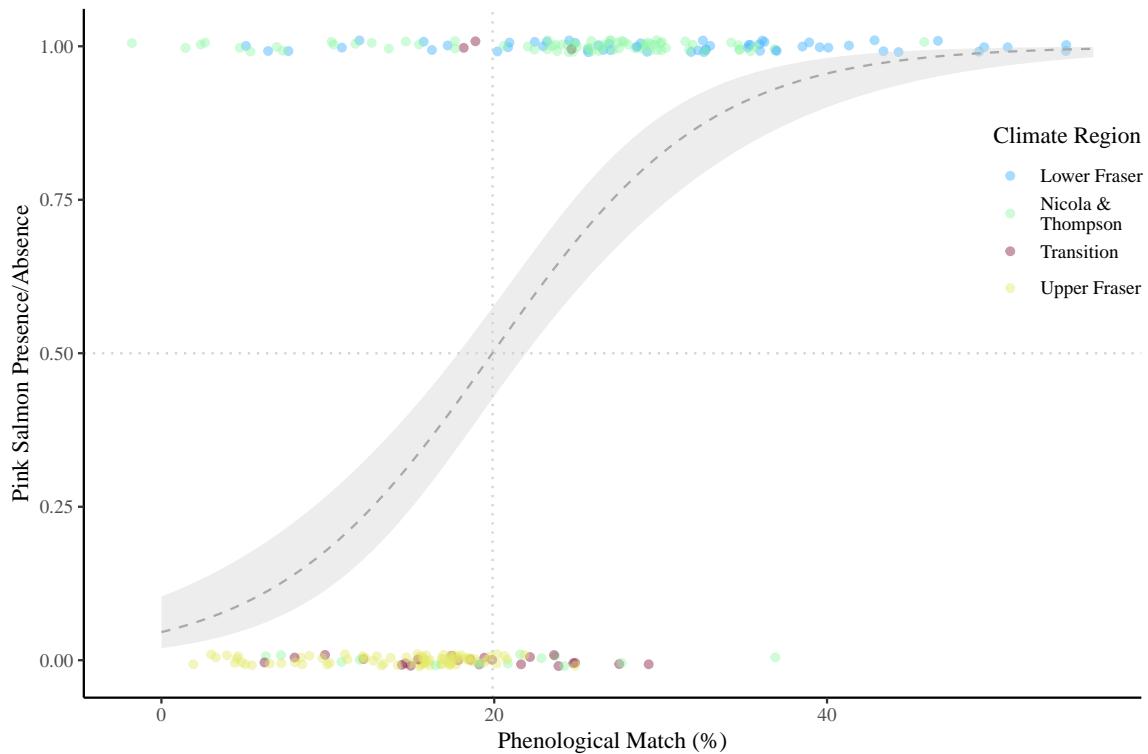


Figure 3.7: The presence (1) and absence (0) of pink salmon populations in the Fraser River basin given the phenological match (%) between estuary arrival of juvenile pink salmon and the zooplankton bloom. The points are colored by their natal stream climate dissimilarity region, (i.e., Figure 3.5 & Supplementary Figure 3.8) and are jittered both vertically and horizontally for visibility. The dotted grey line describes the mean effect of a binomial generalized linear model surrounded by a light grey standard error estimate.

## 3.5 Discussion

Here we discovered spatially-structured climate dissimilarity was linked to phenological mismatch between ocean zooplankton in the Strait of Georgia and migratory pink salmon

inhabiting the vast Fraser River watershed in British Columbia, Canada. Based on modeled pink salmon outmigration estimates from 62 locations across 43 years, salmon populations originating close to the estuary are twice as likely to match with their zooplankton prey than those in the farthest reaches of the basin. Using spatial modeling approaches, we created predictive maps of climate-driven phenological mismatch throughout the Fraser River basin (Figure 3.6). Indeed, distant pink salmon populations had more dissimilar climates ( $corr = 0.91$  for relationship between distance and climate dissimilarity) from the estuary than more proximal populations. These findings suggest climate may drive strong spatial patterning in mismatch between migratory consumers and their prey, modified by spatial gradients in environmental conditions and phenotypic plasticity. Thus, we provide rare empirical evidence that longer-distant migrants from dissimilar climates may be more vulnerable to match-mismatch.

Large river networks provide spatial structure for both environmental conditions as well as populations (e.g., Isaak et al., 2014; Rodriguez-Iturbe et al., 2009; Yeakel et al., 2014). In a watershed the size of the United Kingdom, we created predictive models of river spawn timing, emergence, and climate-driven mismatch. In more northern and interior locations where the climate is on average cooler (Figure 3.1), we found spawn timing was generally earlier (Figure 3.2) but because of the colder water temperatures during embryo-larval development, pink salmon outmigration was on average 21 days later than their downstream counterparts in coastal, southern British Columbia. By comparison these upstream populations miss the peak abundance of their zooplankton prey more often and by a larger margin (Figure 3.4). Typically river network models are used to predict abiotic variables such as temperature (e.g., Isaak et al., 2017), here we uniquely apply these models to linked physical and biological dimensions of a watershed to describe the structural architecture of life-history timing and variation. As large river networks like the Fraser River system become increasingly impacted by climate change, local climate manifestations will aggregate downstream (e.g., Chezik et al., 2017) and confer spatial patterning to the timing and variation of life history thereby organizing climate-driven match-mismatch across the basin.

The phenology of juvenile pink salmon is controlled by the spawn timing of their parents and the temperature of the stream in which they are incubating. We built a robust predictive model of pink salmon emergence timing as a function of stream temperature through a quantitative synthesis of published studies. This synthesis demonstrated that pink salmon exhibit some level of compensatory dynamics in their emergence timing (Branstrom, 1987). While we expected growth and development to scale with temperature (Brett, 1969) such that emergence occurs upon accumulating a fixed number of degree-days, instead we found that individuals emerged faster and slower under cooler and warmer environments respectively, presumably in order to maximize a life history and fitness benefit (Angilletta, 2009). Such developmental compensation among pink salmon populations may act as an unaccounted-for mechanism to stabilize spring emergence timing despite variable climates

thereby increasing the opportunity for juvenile pink salmon to arrive in the estuary during the zooplankton bloom.

Climate change will impact salmon populations through multiple interacting pathways across the salmon life cycle (Crozier et al., 2008). For instance, warming temperatures (Figure 3.1) may be delaying spawn timing of adult pink salmon by 8.5 days on average throughout the Fraser River basin with interior locations shifting by as much as two weeks (Figure 3.2). Delayed spawning not only reduces adult salmon exposure to increasingly warm river temperatures but also decreases thermal exposure for embryo development (Figure 3.3). Warming rivers should drive earlier outmigration of pink salmon, according to the relationship between emergence timing and temperature. Thus, warming rivers would benefit more inland and northern pink salmon populations that currently can't emerge early enough to match their zooplankton prey. Yet, in other locations later adult spawn timing in warming streams may actually keep pink salmon emergence relatively consistent through time. Indeed, there have been little directional shifts in predicted match-mismatch throughout the Fraser River watershed (Figure 3.6). Pink salmon in Alaska have demonstrated micro-evolutionary shifts favoring early out-migration that are hypothesized to be due to warming spring marine environments (Kovach et al., 2012; Manhard et al., 2017), and have allowed tracking of the zooplankton bloom which is advancing at a rate of 14 days per decade in the Gulf of Alaska (Richardson, 2008). In the Strait of Georgia, directionality of the shifting bloom is less clear, rather the phenology of ocean plankton suggest that the bloom is becoming more variable (Allen & Wolfe, 2013). The potential tensions between climate impacts on different life stages will govern the population-level impacts of climate change on species (Crozier et al., 2008; Mangel & Satterthwaite, 2008). Previous studies (e.g, Malick et al., 2015) have found that pink salmon mismatch with zooplankton prey can influence their population productivity. Thus, given our findings here, we speculate that the spatially-structured mismatch that we predict may contribute to differential population dynamics and productivities across Fraser River pink salmon. Indeed, it has proved challenging to connect match-mismatch dynamics with population dynamics (McLean et al., 2016; Miller-Rushing et al., 2010).

It is even possible that climate-driven phenological mismatches may contribute to limiting pink salmon expansion into the upper Fraser River watershed. Currently, despite thermal compensatory egg development (Figure 3.3), we expect that much of the Fraser River watershed is too cold to promote rapid enough egg development to arrive in the estuary during the zooplankton bloom. By comparing sites in currently inhabited and uninhabited regions, we found overlap with the zooplankton bloom above 20% was strongly indicative of sites near established populations (Figure 3.7). Furthermore, regions of the Fraser River network predicted to have less than 20% overlap with the zooplankton bloom are strongly coincident with cold temperatures (Figure 3.1 & Supplementary Figure 3.8). Thus, cold temperatures towards the headwaters may be one potential factor limiting further expan-

sion of pink salmon. Climate warming of these cold regions, all other things equal, may actually enable pink salmon to have life-cycle phenology that better aligns with the ocean food web. To date, historic pink salmon abundances have yet to be realized since the Hells Gate slide of 1913 (Pess et al., 2012; Ricker, 1989), hypothesized to be due in part to the loss of large-bodied individuals capable of swimming the extreme canyon flows (Ricker, 1989), and accompanying genetic and phenotypic potential. Although small by historic pre-slide measures, populations above Hells Gate continue to be larger in body size than those populations below the canyon (Ricker, 1989). Regardless, our research suggests that climate warming is not only providing new challenges for salmon but also may provide new opportunities in some colder habitats (e.g., Dunmall et al., 2016; Nielsen et al., 2013).

We discovered strong spatial structure in the breadth of phenology of pink salmon migration. Specifically, more upstream populations with greater climate dissimilarity from that of the estuary, had broader phenological distributions than more downstream pink salmon populations (Figure 3.4). In this case, the broadening may be proximally driven by population specific non-linear relationships between temperature and development (Figure 3.3) (Brett, 1969). The result of this phenological breadth is that upstream populations are buffered from completely missing the peak zooplankton bloom window (Figure 3.4). Thus, spatial structuring of mismatch will not only be driven by spatial patterns in phenological peaks but also spatial patterns of phenological variability (e.g., Mertz & Myers, 1994).

Our methods utilized a series of large-datasets and relational analyses to make general predictions on spatial match-mismatch between pink salmon and their estuarine zooplankton prey. We note that our study did not directly observe salmon migration – Fraser River watershed pink salmon populations are genetically indistinguishable with traditional genetic stock identification approaches (Holtby & Ciruna, 2008), such that reconstructing population-specific historic outmigrations is impossible. To conservatively address and reduce uncertainty in our estuarine arrival predictions we, 1. built our models using large amounts of observational data, 2. propagated uncertainty across life-history events, 3. utilized models that account for structure in correlation and variance, and 4. built life-stage models from first principles which relate the environment to growth and behaviour. This does not mean we have captured all forms of variation, for instance populations may have unaccounted-for locally adapted temperature-emergence relationships, but given the genetic similarity among populations across geographically distinct regions, variation is likely very small. That said, pink salmon are believed to be adjusting to climate change more rapidly than other Pacific salmon species (Crozier, 2015), possibly expressing rapid evolution (Kovach et al., 2012). As such we might expect future populations inhabiting currently uninhabited locations to express different life histories in order to cope with local challenges. Ultimately we feel these methods largely capture the underlying processes that describe variation in estuary arrival among populations and fundamentally allows for prediction at unobserved locations extending spatially beyond the observed data.

Here we provide evidence, through a study of an economically- and culturally-important migratory fish, that consumers migrating farther distances can be more vulnerable to phenological mismatches with their prey. Understanding the vulnerability and responses of populations, species, and regions to climate change is a major research direction (e.g., Comte & Olden, 2017). For example, currently climate change envelope models rely heavily on physiological limits of species in their current environment (Lawler et al., 2011), or their behavioural capacity to cope (e.g., Sinervo et al., 2010) to extrapolate range limitations in the future. Here we build predictive maps of match-mismatch for a migratory fish, and discovered strong spatial structure such that longer-distance migrants from climatically dissimilar regions are more vulnerable to mismatch with their prey. If climate change further disrupts environmental predictability across space and time, long-distance migrants will increasingly find their phenologies out of sync with their prey. Thus, migratory distance and how it is linked to climate dissimilarity may be a key factor in contributing to climate change vulnerability.

### 3.6 Acknowledgements

This study relied on decades of physical and biological data collected by those at Environment Canada (EC) and Fisheries and Oceans Canada (DFO). Raw HYDAT and NUSeds data can be obtained at the EC and DFO websites respectively. Emergence, spawn, flow and stream temperature data limited to this study can be obtained on GitHub, along with all code ([http://github.com/kchezik/Timing-Mismatch\\_Pink-Salmon.git](http://github.com/kchezik/Timing-Mismatch_Pink-Salmon.git)). To analyze these data we used R v3.4.3 (<https://www.r-project.org/>), QGIS v2.18 and the ArcGIS dependent FLoWS and STARS toolboxes along with the SSN R-package developed collaboratively at the US Geological Survey, US Forest Service and US National Oceanic and Atmospheric Administration. K.A. Chezik and J.W. Moore were supported by the Liber Ero Chair of Coastal Science and Management and the Natural Sciences and Engineering Research Council of Canada. K.A Chezik was also supported by the KEY big data graduate scholarship. S.M. Wilson was supported by the Vanier Canada Graduate Scholarship. All authors were supported by Simon Fraser University. Earlier versions of this manuscript were improved by comments from David Patterson at DFO and Dr. Jonathan Armstrong at Oregon State University.

### 3.7 Supplementary material

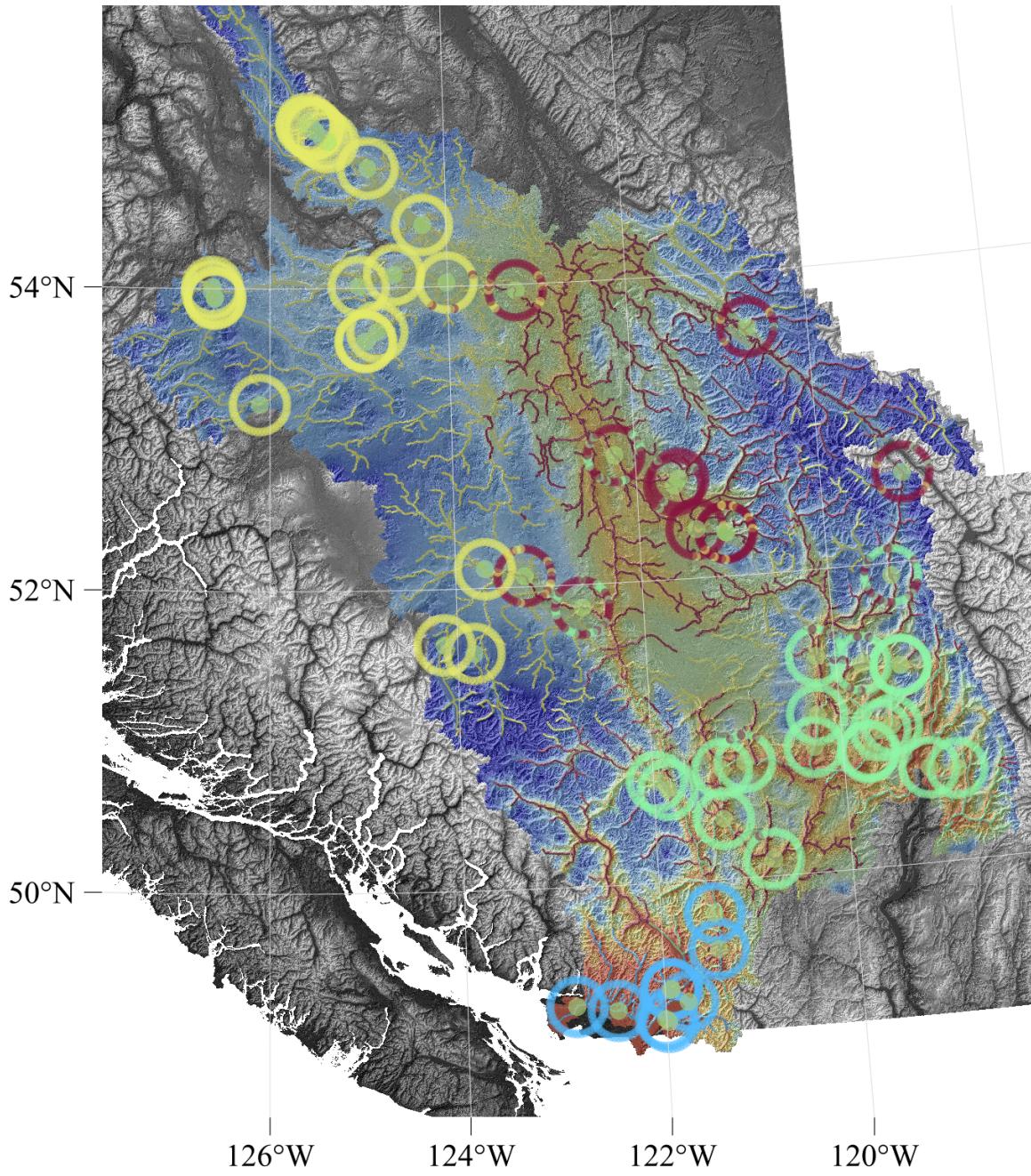


Figure 3.8: Climate regions by year among emergence timing prediction sites in the Fraser River basin in British Columbia Canada. Points (green) indicate temperature monitoring locations where emergence timing can be estimated. The ring around monitoring locations are made up of climate divergence estimates (points) for each year between 1970 and 2010 colored by climate divergence group. Four distinct climate divergence groups are estimated, located in the lower Fraser basin (blue), the Nicola and Thompson region (green), the central and northeast region (red) and the west and northwest region (yellow). Climate group boundaries vary from year to year within edge sites. The river network is colored by mean climate group and the Fraser River watershed is colored by mean annual temperature across years.

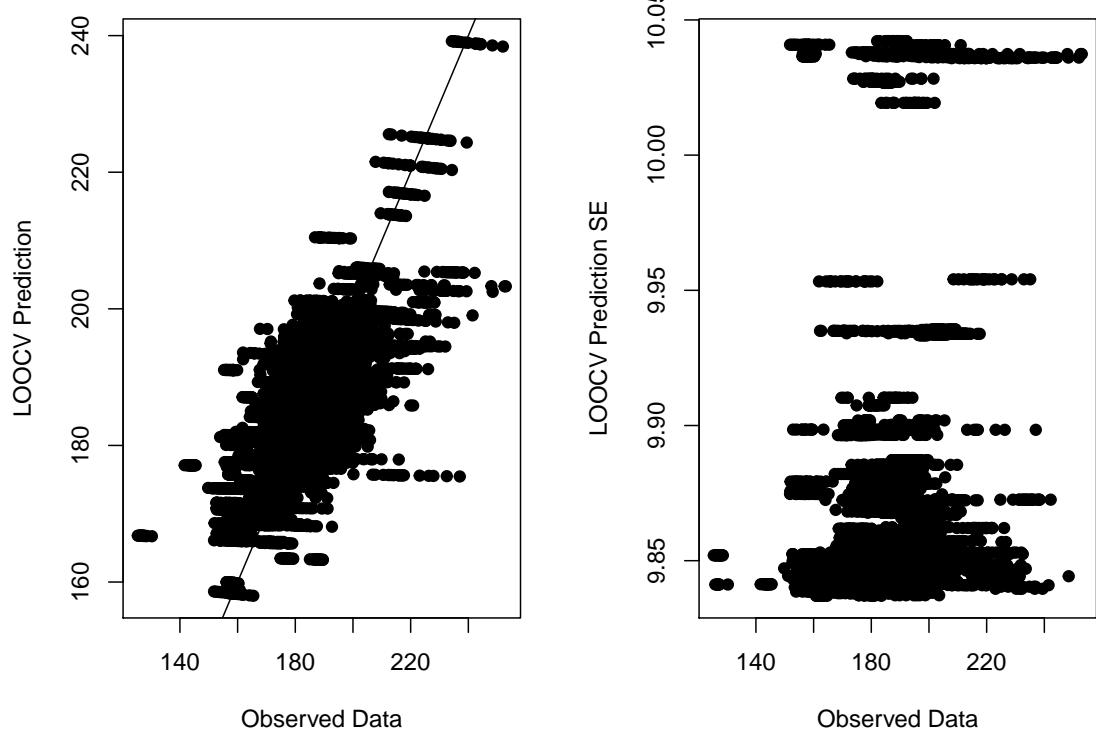


Figure 3.9: (**Left**) Leave one out cross validation (LOOCV) given the observed spawn timing estimates (solid line 1:1). (**Right**) LOOCV prediction standard error estimates in days given the observed data.

### Estimation Method: MethMoment

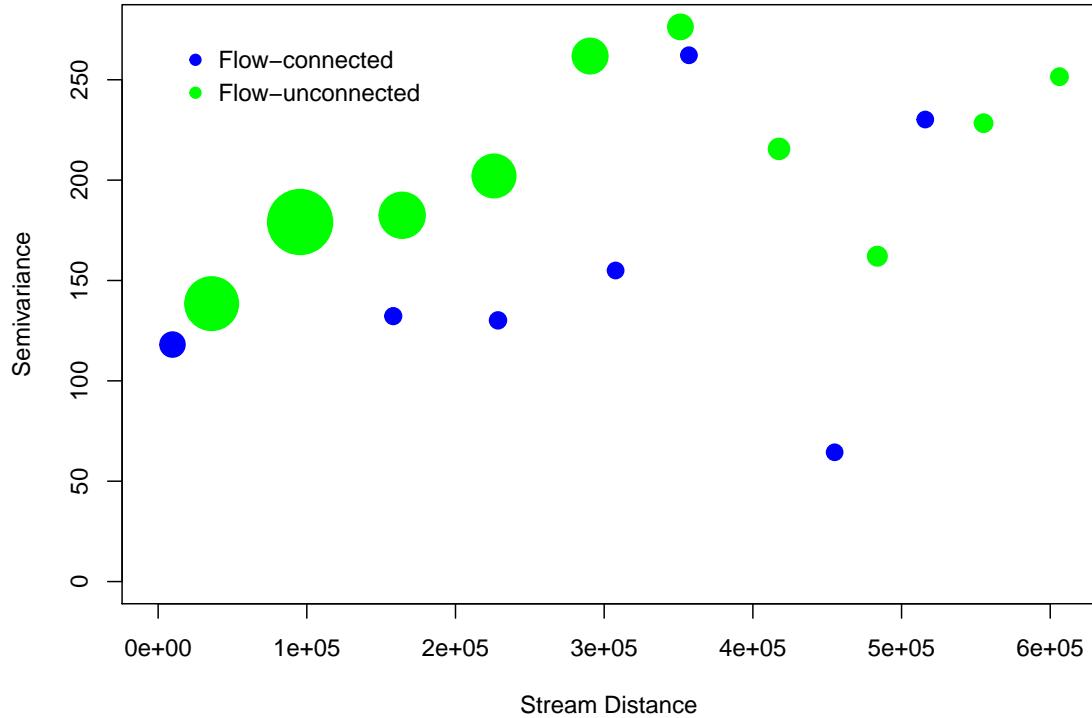


Figure 3.10: Torgogram describing autocorrelation in spawn timing among flow connected and unconnected locations in the Fraser River basin where lower semivariance indicates higher autocorrelation. The size of the circles are proportional to the number of pairs binned in the given stream distance (Hoef et al., 2014).

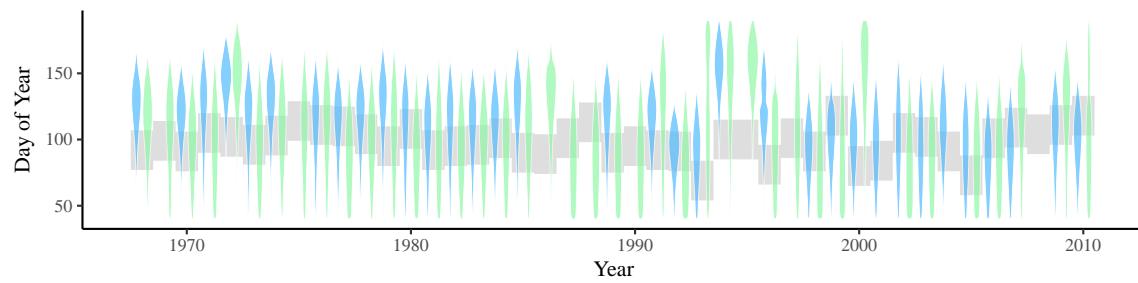


Figure 3.11: Overlap between pink salmon estuary arrival and zooplankton peak bloom window in the Straight of Georgia from 1968 to 2010. Pink salmon arrival distributions represent populations residing in the lower Fraser River below Hells Gate (blue) and above Hells Gate in the Thompson and Nicola region (green). Zooplankton bloom date range (grey ribbon) is derived from model output in Allen & Wolfe (2013).

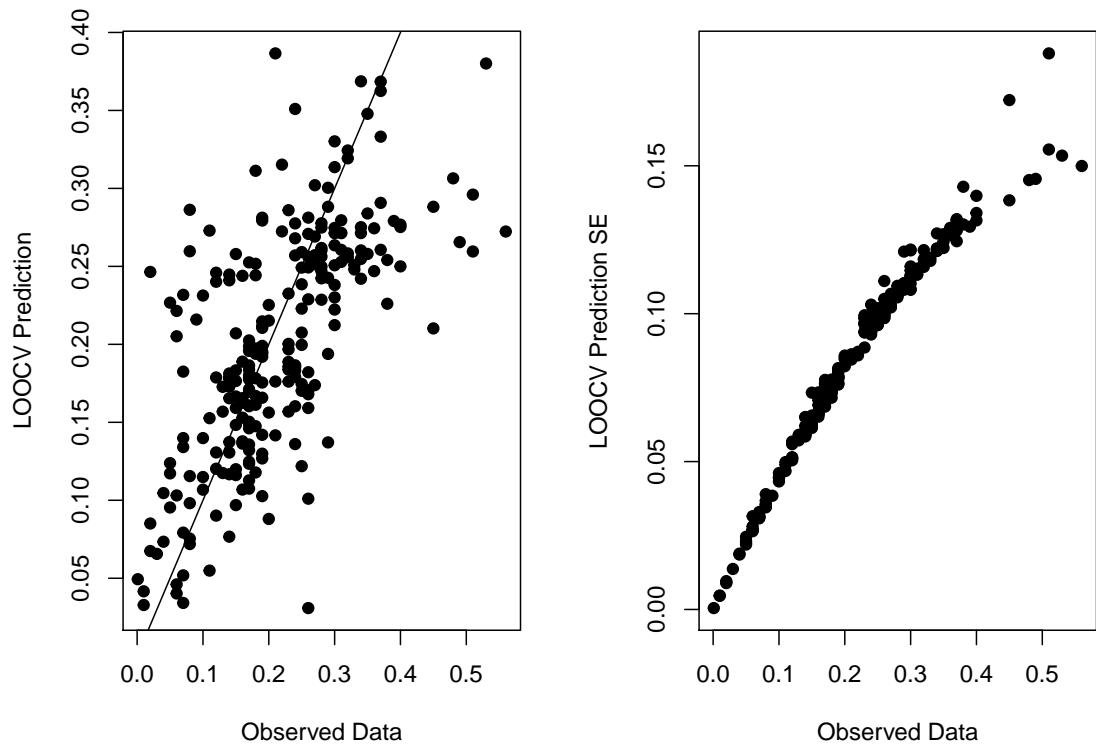


Figure 3.12: (**Left**) Leave one out cross validation (LOOCV) given the observed phenological match (%) between estuary arrival and zooplankton bloom among sites and years in the Fraser River watershed in British Columbia Canada (solid line 1:1). (**Right**) LOOCV prediction standard error estimate (%) given the observed data.

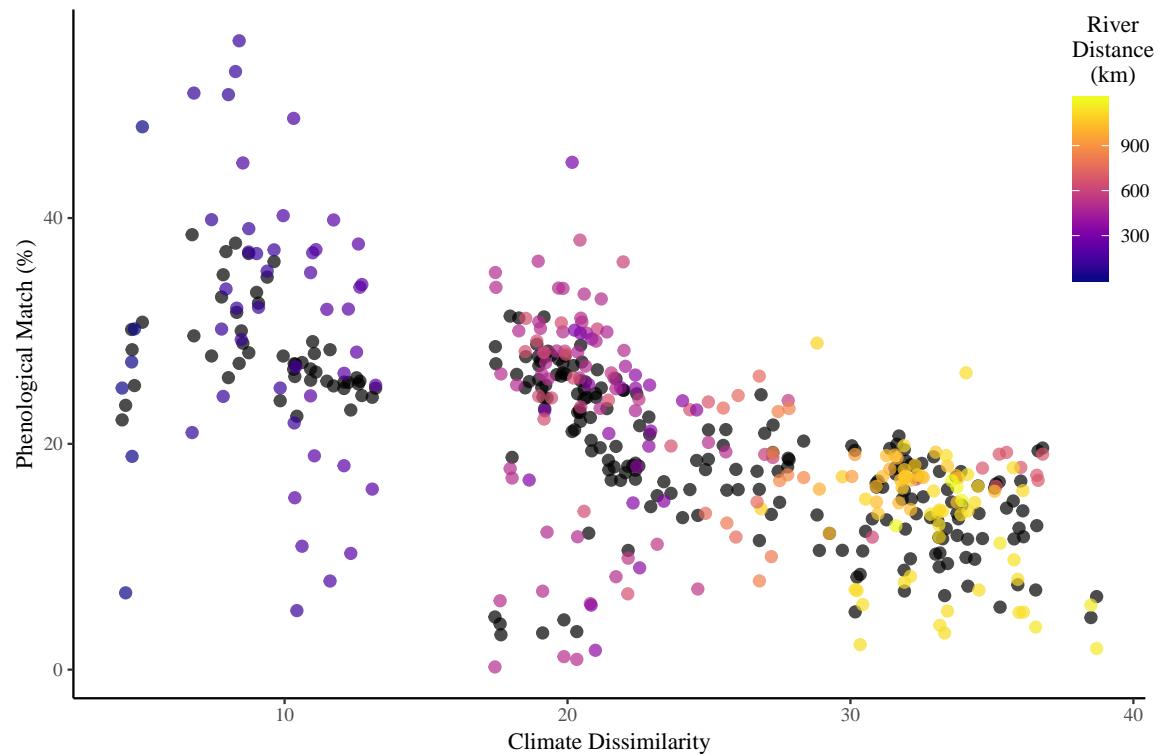


Figure 3.13: Phenological match (%) between juvenile pink salmon arrival and the zooplankton bloom, given the standardized degree of climate dissimilarity between each natal stream and the estuary in the given year. Colored points describe the overlap and climate divergence estimates for each population in each year data were available and are colored by river distance to the estuary. Black points represent prediction estimates given leave one out cross validation per a fitted equation 4 in the main text.

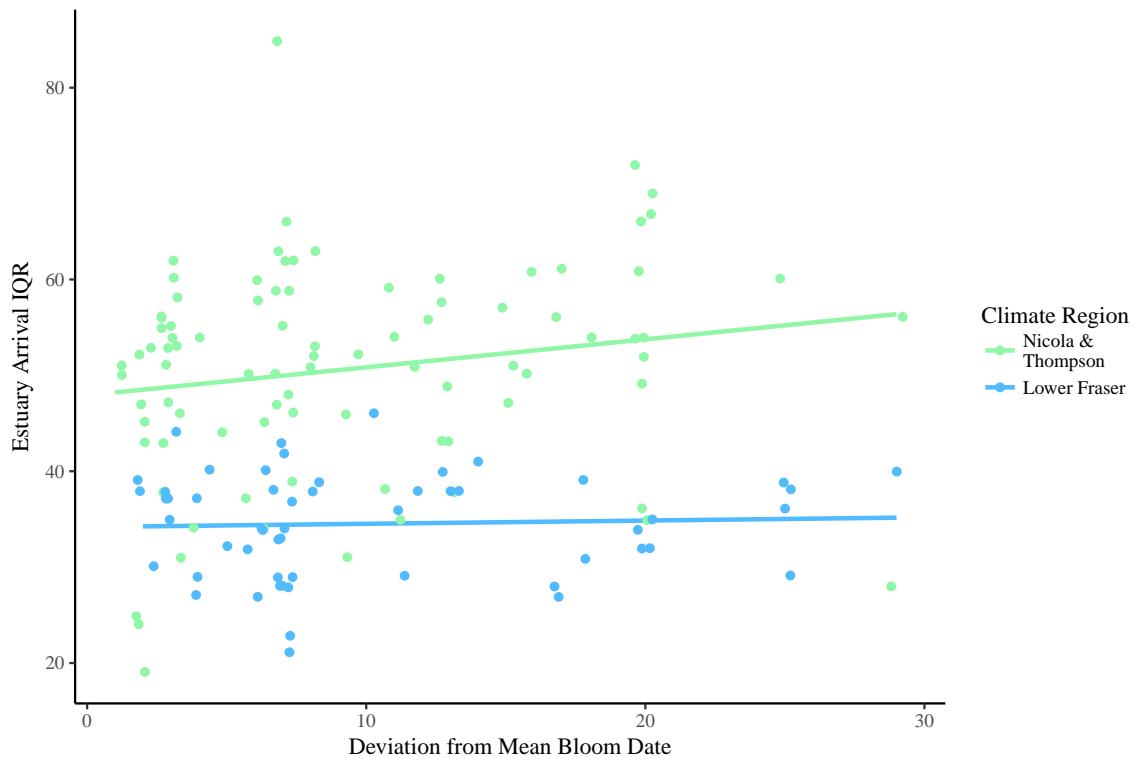


Figure 3.14: Changing estuary arrival timing inter-quartile range given the strength of the deviation from the mean zooplankton bloom date as described in Allen & Wolfe (2013). The points are colored by climate region above (green) and below (blue) Hells Gate. Lines describe the mean trend line given a simple linear model.

# Chapter 4

## Cleaning stream temperature data with hidden markov models

### 4.1 Abstract

Data cleaning is increasingly a chokepoint in the research process as remote sensing technology facilitates mass-monitoring efforts that return prodigious data volumes. Stream temperature is essential to assessing ecological processes in the lotic environment and has become relatively trivial to monitor as sensors have become small and relatively inexpensive. Here I present a semi-automated method of cleaning stream temperature data using Hidden Markov modeling (HMM), an unsupervised learning methodology that assess observation error in a probabilistic framework. Receiving raw observational data and limited prior constraints, this HMM assess the likelihood of data arising from a stream or air temperature model and bins them according to user-defined certainty criteria. Using stream temperature data collected at 10 locations over 4 years in the Thompson River watershed in central British Columbia Canada, I assess this model by comparing manual efforts with HMM results at a variety of certainty thresholds. I demonstrate strong model efficacy when the air model is constrained under high certainty requirements, with as little as 2% disagreement between manual and automated methods at some sites. Significant disagreement (27%) occurred when stream temperature data exhibited strong spring hysteresis and overlap between air and water state estimates. In total, this model exhibited 84% agreement with manual methods, suggesting it could prove to greatly simplify stream temperature cleaning while retaining high levels of data quality.

### 4.2 Introduction

Researchers are increasingly relying on digital remote sensing technologies to capture ecological processes over large spatial extents and at higher temporal resolution than has historically been possible (see ZSL Journal: Remote Sensing in Ecology and Conservation). For example, LiDAR is a remote sensing tool used to collect three dimensional images (e.g.,

Washington State Department of Natural Resources, 2017). A single flight over a forest can map individual trees and be used to measure forest growth (e.g., Caughlin et al., 2016), canopy density (Lee & Lucas, 2007), terrain topology (e.g., Washington State Department of Natural Resources, 2017), photosynthetic rate and even species composition (e.g., Barbosa et al., 2017; Asner et al., 2017). Thus, high volumes of simple physical environmental measurements such as light and reflectance, as collected by LiDAR, can shed light on a variety of ecological patterns and processes. Ongoing cost reductions and miniaturization of sensors measuring fundamental physical data (e.g., temperature, light, oxygen) have made mass implementation easy and relatively affordable resulting in temporally-resolved monitoring arrays (Isaak et al., 2017). This era of big data ecology is improving our fundamental understanding of ecosystems and their management (Hampton et al., 2013).

A current challenge in ecology is processing the massive influx of data. Although we can now collect large quantities of data over space and time, which can return refined understandings of ecosystem processes, data errors are more difficult to detect without rigorous and time-consuming analysis. If errors in big data were strictly random occurrences we could assume they are subsumed by the data volumes themselves but because sensor technology is systematic by nature, often so are their errors, possibly leading to spurious artificial patterns. Inaccurate findings due to a lack of due diligence undermines the study, mis-informs future research and damages the credibility of science when discovered (Liu et al., 2016). Data cleaning is thus requiring increasingly sophisticated approaches, with some disciplines developing clever post-processing tools that use a variety of rule-based algorithms to eliminate many common errors (e.g., Habib et al., 2009). Rule-based error correction or "knowledge engineering" (Domingos, 2015), is wholly dependent on humans to identify unique error types and construct a recipe that cleans data by rigid logic. Programming logic to identify each new source of erroneous data is laborious and often leads to complexity limitations as the number of influential variables compound.

Increasingly the solution to the problem of compounding complexity in big data is machine learning. Neural networks, support vector machines (SVM), naive Bayes, random forest, etc., are common tools for grouping data by similar features. Rather than being hamstrung by rigid rules or parametric functions, these methods infer underlying processes and group data through feedback training, which in theory become more defined as data volumes increase. Erroneous data could constitute its own group but often these methods need a labeled dataset that is representative of the broader population to discern clear boundaries among groups. These broadly representative datasets are not easily compiled and are often incomplete. Tools such as PCA and Isomap can group raw data without the supervision of a labeled dataset but these tools still require humans to identify groups that constitute errors. Moreover, without considering context, erroneous data can be mislabeled and computational complexity often plagues these methods as data dimensionality and volume increases. To identify errors in data, which are inherently unlabeled, there is a

need for unsupervised learning methods that allow the data to be probabilistically grouped under known physical constraints. Hidden Markov Models (HMM) offer a construct that is at the intersection of human and machine intelligence, where humans provide potential conceptual models from an *a priori* understanding of the processes underlying the data and the computer bins the data under these constraints.

Water temperature governs aquatic biological processes and is a fundamental and essential component to understanding aquatic ecology (Angilletta, 2009). Flowing water is especially thermally dynamic and essential to understanding lotic freshwater ecosystems. Monitoring stream water temperature has become a relatively trivial and routine task due to the development of cheap, long-lived, small and environmentally robust temperature sensors (Isaak et al., 2017). As a result, sensors are being deployed at increasingly high spatial and temporal resolution, resulting in exponentially growing volumes of data even for the simplest of studies (Webb et al., 2008). The dynamic nature of flowing water leads to these sensors experiencing many types of error. For example, snow-derived extreme spring flows (i.e., freshet) in northern latitudes can lead to sensors being buried in sediment, pushed onto river banks, tangled up in low hanging tree limbs or simply elevated in the water column due to shifts in the river bed. These scenarios are common, often go unidentified until retrieval and lead to mixed air, water and ground temperature data. These errors can be reduced by expert deployment but cannot altogether be avoided. Furthermore, human-derived errors are also common and can be systematic or random. For instance, air temperature is often recorded during deployment and retrieval of sensors. Alternatively, sensors may be found by curious public who remove the logger from the water column resulting in random air temperature readings of various duration. Finally, the sensors themselves can experience errors such as battery failure during extreme cold periods that become resolved during warmer periods. Thus, there may be many potential sources of errors that may or may not be obvious in a temperature time series. Currently, it is recommended that scientists undertake a rigorous yet subjective manual data-cleaning process that involves cross comparing nearby and concurrently monitored stream and air temperatures as well as flow data (Sowder & Steel, 2012). A machine-learning approach to error detection could reduce the subjectivity inherent in traditional methods of data cleaning.

Here we build a Hidden Markov Model for stream temperature data with the primary goal of separating water from erroneous air temperature measurements in an effort to reduce post-processing time and human subjectivity. Constructed under a probabilistic Bayesian framework, this model assigns probabilities to data as to whether it is water temperature or erroneous air temperature. Additionally, HMMs leverage the power of autocorrelation in time series by incorporating temporally-weighted support of adjacent data to inform categorization at the current time point. Moreover, our model leverages known physical properties of seasonal temperature cycles and modeled air temperature estimates readily available to constrain the model and improve certainty and accuracy. Combined, this model

provides a framework for leveraging 'big data' to identify errors in stream temperature records and automate the increasingly laborious and time-consuming process of cleaning these data.

## 4.3 Methods

### 4.3.1 Hidden markov models

Hidden Markov Models (HMM) are a type of state-space model that offers a probabilistic framework governing the transition between states through time (Hamilton, 2016). In stream temperature data primarily two states exist, the desired state where the data exhibit stream temperature dynamics and an erroneous state where the data exhibit air temperature dynamics. The probability of being in either state is described by the first order Markov chain,

$$p(z_t|z_{t-1}), \quad (4.1)$$

where the state probabilities of air and water ( $k = a, w$ ) at time point  $t$  are informed by the state probabilities in the previous time point. This Markovian structure makes explicit an expectation that the current time point will likely remain in the same state as the previous. It is not possible to observe the true state directly so it is inferred  $z_t$  by considering the probability of the observed data ( $y_t$ ) in either state as,

$$p(y_t|z_t = k, \theta), \quad (4.2)$$

where coefficients describing each state and a transition matrix describing the global propensity to change between states are captured in  $\theta$ . The joint distribution of the state model  $p(z_t|z_{t-1})$  and observation model  $p(y_t|z_t)$  described by,

$$p(z_{1:T}, y_{1:T}) = [p(z_1) \prod_{t=2}^T p(z_t|z_{t-1})] [\prod_{t=1}^T p(y_t|z_t)], \quad (4.3)$$

returns the full probability of each state given the relative support of the data in the current time step and the systems susceptibility to a state change, weighted by the state probabilities in the previous time step (Damiano et al., 2017). A more in-depth discussion of the HMM model and algorithm can be found in Hamilton (2016) and Damiano et al. (2017).

### 4.3.2 States: water & air temperature models

Water and air temperatures follow the seasonal dynamics derived from changing levels of solar radiation related to the earth's orbit and tilt of its axis. The shared ultimate cause of temperature and the passive exchange of energy between the two mediums results in stream and water temperature cycles being closely correlated. But due to differences in the

thermal dynamics of gases and liquids, the temperature cycles' of air and water do diverge in magnitude and timing. Therefore, I used the same cyclical model (e.g., Shumway & Stoffer, 2000) for both water and air temperature states described by,

$$y_t = \alpha_k + A_k \cos(2\pi\omega + \tau_k\pi) + \eta_k, \quad (4.4)$$

where a cosine curve capturing the seasonal oscillation of temperature ( $y_t$ ) is modified by state coefficients that describe the annual dynamics of air and water temperature ( $z_{w,a}$ ). The mean annual temperature is captured in  $\alpha$  while the range of temperature values around  $\alpha$  (i.e., amplitude) are captured in  $A$ . The frequency of the temperature cycle ( $\omega$ ) is described by,

$$\omega = d_{1:\gamma}/\gamma, \quad (4.5)$$

where  $\gamma$  is the number of observations per cycle and  $d_t$  the integer location of observation  $t$  in the cycle. I include a seasonal adjustment ( $\tau$ ) to shift the cosine curve such that  $d_1$  coincides with the point in the annual temperature cycle best described by the first of January, which is typically approximated by an adjustment of  $\pi$  in the northern hemisphere, beginning the curve near its lowest point.

The variance around the mean temperature is a distinguishing feature between annual air and water temperature cycles. Here I assume the variance is normally distributed in both states and captured in sigma.

$$\eta_k \sim N(0, \sigma_k) \quad (4.6)$$

I do not expect errors to always be independent and identically distributed because auto-correlation between data-points at the sub-daily level will not be captured by the water state model. As long as the primary components of the two states are well described, compromising the normality assumption should not impact the efficacy of state estimation.

I expect  $A_k$ ,  $\tau_k$  (eq. 4.4) and  $\sigma_k$  (eq. 4.6) to be ordered (i.e.,  $w \prec a$ ) because water's high thermal capacity delays the response time to seasonal shifts in solar radiation and reduces the amplitude and variance of water's annual temperature cycle. Therefore, I expect the seasonality of water temperatures to lag behind air and express depressed amplitude and variance relative to the air temperature state. Ordering of variables not only captures the differential dynamics of the states but increases the stability of estimation by reducing the potential for state inversion (i.e., label switching) during the multi-chain MCMC process.

Although neither air or stream temperature profiles are perfectly described by a cosine curve, stream temperatures are particularly prone to hysteresis where spring and fall temperatures are not symmetrical around summer's peak temperatures (Letcher et al., 2016). Hysteresis in the annual stream temperature curve occurs for many reasons such as the high specific energy of water and snowmelt and rain (e.g., Lisi et al., 2015). In northern latitudes

spring snow melt depresses stream temperatures initially but wanes in the summer leading to low flows and warmer stream temperatures, only to return to a cooler state as fall rains increase river discharge. To allow our stream temperature curves a modicum of seasonal flexibility, I include a double cycle cosine curve that modulates the annual curve (eq. 4.4),

$$\text{season} = \phi \cos(2\pi\omega\frac{1}{2} + (\tau_w + c)\pi) \quad (4.7)$$

such that winter and summer water temperatures are elevated, while spring and fall temperatures are depressed. The season effect is tied to the  $\tau_w$  value in eq. 4.4 and shifted by a constant ( $c = .5$  or  $1.5$ ) to align the coldest day of the year with a seasonal shift of zero. I also include a seasonal expansion factor ( $\phi$ ) that describes the strength of the seasonal effect where zero suggests no effect of season. I limited this effect to be no greater than 3 degrees so as to limit its flexibility and avoid 'reaching' by the model to capture seasonally correlated air temperature readings.

### 4.3.3 Temperature HMM

By evaluating the likelihood of the data ( $y_t$ ) given the state models (eqs. 4.4 & 4.7), weighted by the previous time step's state probabilities (eq. 4.1), it is now possible to make inference about which state ( $k = a, w$ ) the data are in at each time point. Iteratively evaluating eq. 4.3 returns state probabilities that indicate whether the temperature data represent an air or water source as well as the models certainty of that estimate. I implemented this process in the probabilistic programming language Stan (Stan Development Team, 2017) with one modification. Given water temperature is the desired state, I expect most datasets will have limited air temperature records likely leading to the inaccurate coincidence of state parameter estimates during the fitting process. Furthermore, tools are currently available to approximate  $\alpha_a$  and  $A_a$  with relatively high accuracy over space and time (e.g., Wang et al., 2012). Therefore, to avoid coincidence in state parameter estimates which would result in unclear state probabilities, I provide the air state model with known and modelled  $\alpha$  and  $A$  estimates. Priors indicating estimated parameter value expectations were also incorporated into the model thereby providing coefficient guidance to states with limited data at a given site (see *Model Priors and Known Quantities* for values).

### 4.3.4 Observed data

To demonstrate a real-world application, I applied our HMM to raw stream temperature data collected at 10 sites in the Thompson River basin in central British Columbia Canada (Figure 4.1). Collected for the purpose of understanding geomorphic impacts on stream temperature, these data are not paired with air temperature data and represent a real-world example of the challenges surrounding stream temperature data cleaning. These stream temperature data were collected between July of 2014 and August of 2017 and likely con-

tain examples of all aforementioned error types. Although unconfirmed, I know errors are likely due to evidence upon retrieval but the exact location of those errors in the time series is imprecisely unknown. To compare HMM and traditional methods of data cleaning I manually inspected and labelled data as either air or water using prior knowledge and personal expertise (e.g., Sowder & Steel, 2012). I acknowledge that the true state of temperature is unknown. Eighteen weather stations owned and operated by Environment Canada (Figure 4.1), collecting air temperatures during the same period, were consolidated to represent eleven regions in the Thompson River basin and used as a visual reference to compare and contrast modeled source estimates with direct observations.

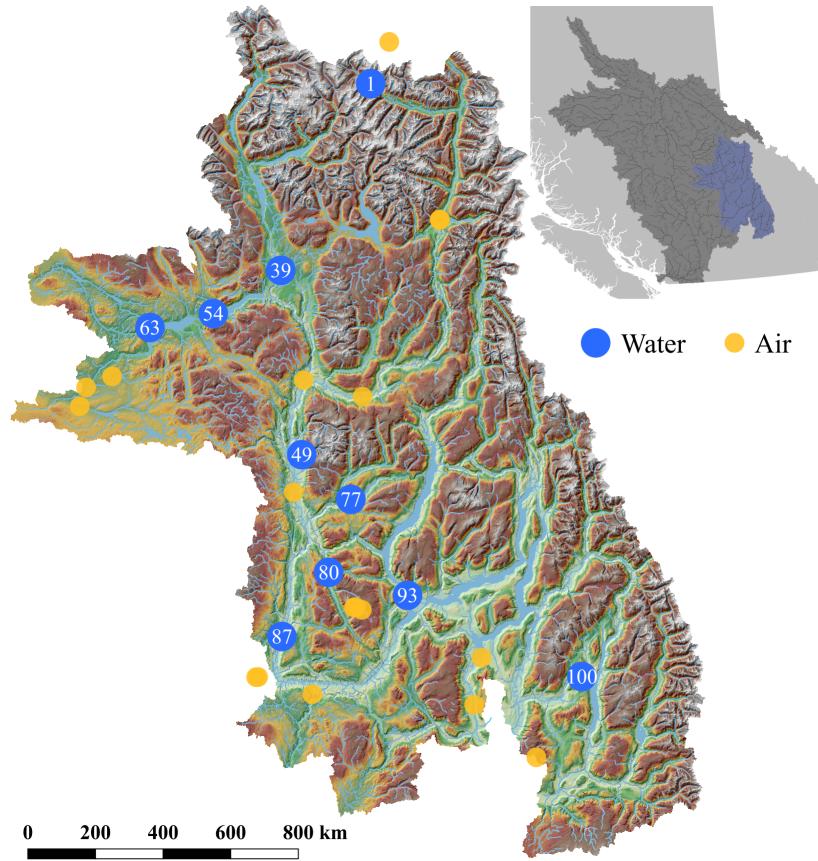


Figure 4.1: Observed water and air temperature locations within the Thompson River watershed, nested within the Fraser River watershed (dark grey) in central British Columbia Canada (light grey). Water temperature locations (blue) were monitored between July 2014 and August 2017 concurrently with air temperature stations (yellow) maintained by Environment Canada. Water temperature locations are labeled by their site number indicating the location of model results shown in figure 4.3.

#### 4.3.5 Model priors & known quantities

I expect temperature data in a given state to be more likely to remain in that state than transition to another state. Therefore, I applied weakly informative beta priors with shape parameters  $\alpha = 10$  and  $\beta = 2$  (median=0.85) for the probability of remaining in the current state and the inverse for the probability of transitioning between states. Moderately informative priors and fixed values were applied to some parameters where more is known about their values. By extracting mean annual air temperatures and associated annual temperature range estimates from the climate tool ClimateBC (Wang et al., 2012), I was able to provide locally-adjusted (e.g., elevation, latitude)  $\alpha_a$  and  $A_a$  parameter estimates. I used ClimateBC air parameter estimates to provide the mean water temperature ( $\alpha_w$ ) and amplitude ( $A_w$ ) parameters moderately strong priors ( $\sigma = 5$ ). As the coldest temperatures are widely understood to occur near the first of the year in the northern hemisphere, I estimated  $\tau$  to coincide with a cosine curve adjustment of  $\pi$ .

I was able to directly calculate the parameter values in eq. 4.5 because temperature cycles occur at regular intervals. Each year typically contains 365 days with 366 in leap years. Thus  $\gamma$  is typically some multiplication of 365 depending on the temperature sampling frequency. In this study I collected data every other hour resulting in 4380 observations per cycle. Each observations'  $d$  value was calculated as the hour of the year starting with January-01 00:00:00 (M-D HH:MM:SS).

#### 4.3.6 Model post-hoc calculations

Manual cleaning of the data suggest air temperatures constitute approximately 4.3% of the 121,764 data points used in this analysis. This imbalance of data increases the probability of type II errors where water temperature data is inaccurately labeled as air, a problem that is amplified by the HMMs temporal structure. While using previous time-steps' probabilities to inform subsequent probabilities, the model takes advantage of temporal auto-correlation and at times overcomes model imprecision, it also increases the potential for perpetuating a state belief beyond local evidence of a state change. This bias can be partially overcome by smoothing the data using the forward-backward algorithm described in Damiano et al. (2017), which considers the state probabilities in the future as well as the past. Furthermore, rather than use a state certainty break-point probability of 50%, it's possible to adjust this value to reflect the inherent probabilities of each state in the data. Because water data is expected much more frequently than air, these state certainty break-points should be asymmetric – break points should be more strict when labeling data air and less strict when labeling data water. Here I tested disagreements with manual efforts at varied break-points to identify values that limit disagreement and approach manual efforts of cleaning.

Type I errors are most likely to occur when air data overlap with the water state during daily temperature oscillations. Assuming strong evidence of air temperatures in a day con-

fers the air state to the entire day, it's possible to improve on smoothed estimates and avoid some of these type I errors by considering the totality of the daily data probabilities. By exponentiating the log-mean water state probabilities within each day I calculated a weighted daily state probability. By log transforming the water state probabilities before calculating the mean, I down-weighted the high likelihood water state values within a given day and upweighted high likelihood air state probabilities. In this way, I return a state probability for each day where high likelihood air state probabilities disproportionately contribute and confer the air state to the entire day. When the daily probability suggests an air state on that day, we attribute that state to all values within a given day regardless of individual observations suggesting otherwise. In considering observational and daily level probabilities, we can overcome some of the common transition errors that occur when data states overlap.

## 4.4 Results

Of the 121,764 observations, the HMM identified 95,514 that had greater than 50% probability of being water leaving 26,250 observations labeled air. This assignment agreed with the manual cleaning process 80% of the time but ranged wildly by site and year. Among sites, disagreement ranged between 6 and 39% while at the yearly level the range widened to between 1 (site 39 in 2014) and 61% (site 93 in 2016). Divergence between manual and automated methods was not greater in data-limited years such as 2014 and 2017 where only part of the year was recorded nor in site 93 where data was lost after deployment (Figure 4.3).

By relaxing the certainty required to label data in the water state and inversely tightening the requirement for the air state, the overall disagreement between manual and HMM methods were reduced to approximately 7% (Figure 4.2). Thus, shifting the certainty break points can dramatically reduced type II disagreements, while only marginally increasing type I disagreements. The break-point required to reach disagreement minimization is approaching certitude in the air state and nearing complete relaxation of certainty in the water state.

Using separate break-points of 1% for individual observations and  $1 \times 10^{-7}\%$  for daily probability averages, the model does not minimize disagreements between manual and HMM methods but rather operates in a middle ground that does not eliminate water data unnecessarily while identifying the most extreme air temperature records. Using human judgment to identify these break-points, HMM disagreement among sites was reduced to as low as 2% and importantly, only as high as 27% (Figure 4.3). At the annual level this resulted in one site having less than 1% disagreement (site 49 in 2017) and only one year above 50% (site 93 in 2016). General agreement between manual and HMM methods increased 4% to 84% under these conditions, compared with the default 50% break-point.

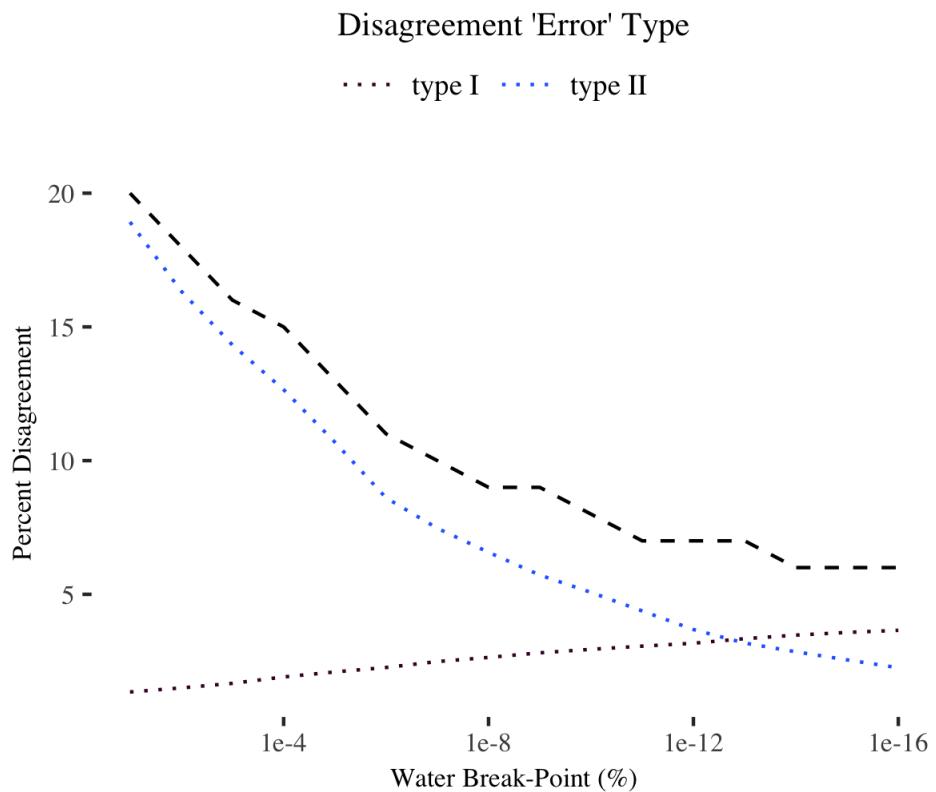


Figure 4.2: Shifting disagreement between manual and automated cleaning of stream temperature data as the water certainty break-point decreases. The 'type' of disagreement assumes manual cleaning to be the 'true' state, where a type I error indicates when the HMM labels data water while manual methods indicate air to be the 'true' source. Type II 'errors' occur when the HMM labels data air while the data were manually labeled water. The total percentage of data in disagreement between these data cleaning methods is indicated by the long-dashed line.

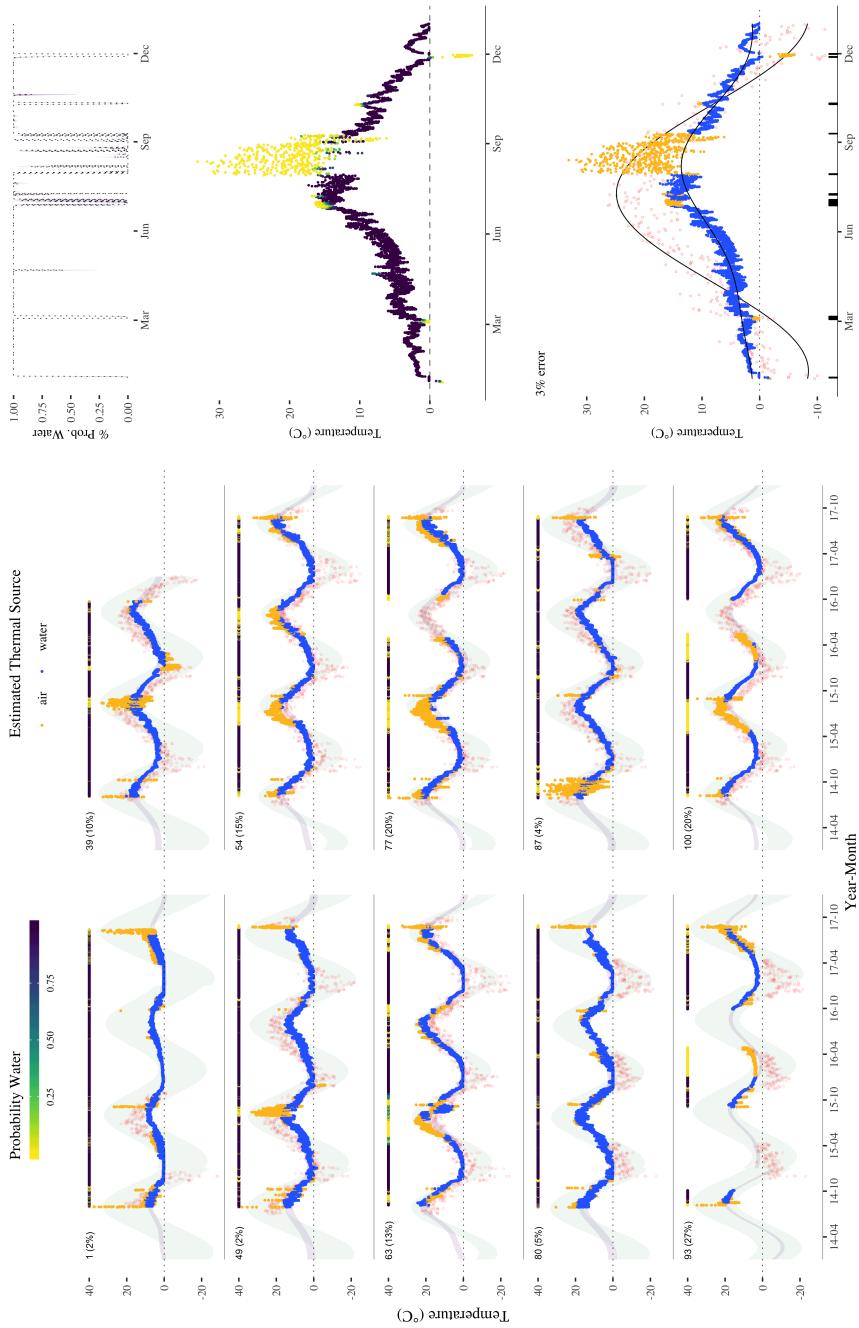


Figure 4.3: **(Left)** Source estimates for temperature observations at 10 locations within the Thompson River watershed in central British Columbia, Canada . Air (yellow) and water (blue) source estimates were calculated using break-points of greater than 1% probability for water at the observation level and greater than  $1 \times 10^{-7}\%$  when averaged to the daily level. HMM state estimate probabilities are described by the color bar above each time series. State model 95% confidence intervals describe the probability density window for air (green) and water (purple). Panels are labeled by site number, associated with locations in figure 4.1 and total disagreement (%) . For visual consideration, mean daily air temperature observations at the nearest monitoring station are displayed in red. **(Right)** Temperature observations for site 49 in 2015 and associated source probability estimates. A rug plot in the lower panel identifies where manual and HMM models disagree and solid lines describe the mean estimate of the annual air and water profiles.

## 4.5 Discussion

Here I explore the potential for unsupervised machine learning to clean remotely-sensed data. This approach arguably reduces potential subjectivity and time. Using a well-defined HMM the model probabilistically separates water and air temperature signals within a time series using raw temperature observations and a limited number of easily-obtained annual air temperature statistics (Figure 4.3). The HMM agreed with manual-cleaning protocols in some sites and years (e.g., sites: 1, 39, 49, 80, 87), suggesting that with some additional improvements it could become a reliable tool requiring minimal user intervention. The flexibility of the model to accommodate variation among datasets while remaining constrained enough to limit mistakes offers a fair and consistent method for identifying errors. As such, the model may offer a greater degree of objectivity than human pattern recognition. Improving this model to be consistently indistinguishable from high-quality manual efforts not only reduces the need for hours of monotonous manual data-cleaning but also promises reduced variation in data quality.

Despite the power and objectivity of this model, there are limitations to this approach and I do not believe humans should be altogether eliminated from the data cleaning process. As demonstrated, 'error' rates change with the certainty threshold (Figure 4.2). This suggests a mechanism by which researchers and data managers can contribute to the data cleaning processes in a meaningful but less work intensive way. By adjusting the water's certainty requirements below 50%, the frequency of type II 'errors' rapidly decrease with only small gains in type I 'errors', resulting in an overall reduction in discrepancies between HMM and manual cleaning (Figure 4.2). The cost of incidentally eliminating water data versus retaining air data is largely subjective and dependent on the needs of the research. Often temperature data are summarized such that incidental air data are largely obfuscated in subsequent analysis while the gains of having more complete data that fully describe a period of time lead to larger  $n$  and greater power. Moreover, the impact of type I 'errors' depends in part on whether or not the data happen to be seasonally coincident. If the expected water temperature is nearly identical to air temperature, the ultimate source may be irrelevant. By building our HMM in a Bayesian framework, uncertainty is fully acknowledged, offering a tool for researchers to assess their studies needs and select the data with sufficient accuracy.

The efficacy of stream temperature state estimation is dependent on the degree of coincidence with the approximated air state, degree of seasonal change in water temperature variation and the degree of hysteresis in the annual temperature cycle. Notably, the HMM struggled to properly estimate a convincing water model when air and water temperatures were strongly coincident. Sites 54, 63, 77, 93 and 100 were particularly afflicted, consistently under-estimating late spring and early summer stream temperatures, resulting in poor agreement with manual data cleaning (Figure 4.3). An exploration of these sites revealed they are

directly downstream of lakes and are therefore likely reflecting lake temperature dynamics (Figure 4.1). The influence of lakes is evident by sudden drops in temperature, indicative of lake turning events where cool water from the lake bottom is mixed due to wind disturbance and propagated downstream (Wetzel, 2001; Lisi & Schindler, 2015). When the lake is thermally stratified, the stream is fed by lake surface waters which strongly reflect low variance air temperatures during the snow free period of the year. As such, the primary distinguishing component between air and water temperature is their variance. Future work could include additional model complexity to capture such dynamics.

The variance between states becomes the only distinguishing characteristic when air and water temperatures overlap, suggesting improved state variance estimates could increase model accuracy. Because many sites have little or no temperature data to inform the air state error term (e.g., site 54), the variance is estimated to be only marginally larger than the water state and only because the variance is ordered (see eq. 4.6). By applying a global model to the air state error term, the HMM may improve its ability to distinguish between states, leveraging estimates at sites where air data are more prevalent. This could also be achieved at the site level by applying a strong and well defined prior on the variance term.

Hysteresis in the annual temperature cycle is difficult to fit with deterministic trigonometric functions but their rigidity is also necessary to limit model flexibility when HMM states are strongly correlated and coincident. Here seasonal and annual cosine curves capture much of the site-specific hysteresis in water temperatures due to seasonal shifts in weather patterns but do not capture correlations at finer time scales. For instance, site 49 in figure 4.3 demonstrates a lag in spring temperatures consistent with snowmelt cooling the stream despite increased solar radiation (Lisi et al., 2015). Once the spring freshet concludes in late June and early July the stream rapidly warms to more closely follow air temperatures. This short period of momentum can not be captured by the rigid underlying water model and as such extends beyond the model's likelihood resulting in data being incorrectly labeled. Typically, short periods of momentum in time series models would be captured in either an auto-regressive (AR) or moving-average (MA) component where a proportion of the observation or error in the previous time step(s) would be included in the subsequent mean estimate (Shumway & Stoffer, 2000). We were unable to introduce an AR or MA component to our HMM as it made the state models excessively flexible and unable to detect state changes. A possible solution is to allow the water state's error term to fluctuate seasonally (e.g., eq. 4.7) and be described by a distribution that gives higher probabilities to more extreme values (e.g., student-*t* distribution). This would better accommodate momentum driven auto-correlation, improve water state mean estimates and reduce the likelihood of mislabeling water data (i.e., type II errors).

Big data promises improved accuracy, precision and understanding with increased data volumes. The true power though is through the 'wisdom of the crowds' (Surowiecki, 2004) or drawing upon information across sites, years and days to inform parameter estimates when

data is limited. As discussed earlier with variance estimates in the data impoverished air state models, drawing on sites that are data rich to inform data poor sites could extend the model to smaller datasets capturing little of the annual cycle and improve state estimation accuracy. For instance, the parameter  $\tau$  adjusts the cosine curve such that the lowest point coincides with the coldest part of the year, a value which is shared among sites with minor variation. By assuming  $\tau$  estimates are described by some global distribution, it's possible to improve  $\tau$  estimates at sites where data are noisy or missing by considering sites with higher certainty. This could improve the off-set estimate in  $\tau$  between the air and water state and better align air state models with the unseen air data (e.g., Figure 4.3 site 54). This method could be extended to the water state's annual mean and amplitude estimates or possibly relate these values to their air temperature, such as those derived from the ClimateBC tool. A potential short-coming of global models informing parameters with substantial variation is the propensity for the local estimate to be pulled towards the global mean. If drawn away from the data sufficiently, the state model is no longer coincident and therefore that state is believed less likely by the HMM. Exploring and tuning global models could prove extremely powerful and extend the HMMs ability beyond the conventional analytical scope of traditional manual methods.

As the volumes of digital data increase so do the opportunities to ask novel, broad-scale and impactful questions that shouldn't be encumbered by the labor costs of ensuring data quality. Here I describe the specification of a Hidden Markov Model that separates air from stream temperature data in order to reduce post-processing QAQC and human subjectivity. The model's construction aims to keep inputs no more complicated than what is typically output by stream temperature loggers and gathered by GPS at the time of deployment. Geo-spatial coordinates can then inform climate models such as ClimateBC or PRISM and approximate air state parameters. The simplicity of the model inputs aim to improve the models appeal and facilitate wide usership. Further developments of this model could improve its accuracy such that humans need only adjust the certainty break-point at which we label data. A simple software-tool could then be developed that allows a user to adjust the breakpoint and observe the changing state of the data until local user knowledge and HMM approximation best align. Subsequently, a *post-hoc* analysis could be fit that fills data-gaps and provides uncertainty estimates resulting in complete datasets useful to a wide variety of ecological studies and monitoring efforts. In these ways, data cleaning tools such as this can facilitate our ability to grow databases, leverage big data in ecological research and tackle increasingly global ecological challenges.

## 4.6 Acknowledgements

This project relied heavily on a tutorial by Damiano et al. (2017), using the code therein as a framework. Built in the statistical coding languages R v3.5.1 and rStan v2.17.4, all

code is available on GitHub at <https://github.com/kchezik/rTDataScrub>. K.A Chezik was supported by the Liber Ero Chair of Coastal Sciences and Management, the SFU KEY Big Data Fellowship, and Simon Fraser University. I also wish to acknowledge David A. Patterson at the Department of Fisheries and Ocean Canada for support in monitoring the Thompson River watershed.

# Chapter 5

## Impacts of network structure and landscape complexity on the cumulative thermal exposure of migrating salmon

### 5.1 Abstract

Warming water temperatures may challenge migrating adult salmon. Elevated and persistent thermal stress during upstream migration is linked to a variety of negative effects on salmon survival including increased disease, lower energy density and curtailed aerobic scope. Yet linking landscape controls to temperature dynamics across space and time and quantifying the migratory thermal exposure of salmon remains a challenge. In this study we combine a large in-stream temperature collection effort with improved analytical approaches to assess how landscape features control temperature regimes and vary cumulative levels of thermal exposure experienced by salmon migrating throughout the large and diverse Thompson River watershed in central British Columbia Canada. We used temperature data collected continuously between 2014 and 2017 at major river confluences in two neighboring and topographically distinct sub-basins to describe annual temperature dynamics resulting in 290 unique site-year observations. By quantifying landscape and climate contributions to temporal stream temperature dynamics, we demonstrate the importance of landscape and network interactions that lead to remarkably different temperature dynamics in the two neighboring sub-basins. For instance, high elevation lakes and glaciers were associated with cooler stream temperatures and low thermal exposure probability while low elevation lakes and wildfire exhibited the inverse. Integrated by the river network, dominant warming features were shown to lead to higher exposure potential along common migratory corridors (i.e., mainstem), leading to higher thermal exposure for populations with longer migrations. Importantly, in the four years of observation, high inter-annual variation in migratory thermal exposure was observed, suggesting systems like the Thompson River watershed, which

are currently exhibiting near prohibitive temperatures for migrating salmon, may become more persistently troublesome as global temperatures rise. By linking watershed attributes to the thermal challenges of upstream migration, our study identifies differential exposure potentials among salmon populations that may be useful to conservation assessment and improve management prioritization.

## 5.2 Introduction

Freshwater temperatures reflect climate and landscape interactions (Webb et al., 2008) and control ecological processes (Angilletta, 2009). Freshwater temperatures and their impacts on biota are of increasing concern with rising global temperatures (Comte & Olden, 2017). A filtered manifestation of climate (Griffiths et al., 2014), the contribution of climate change to shifting freshwater temperatures is complexified by landscape dynamics (e.g., Lisi et al., 2013; MacDonald et al., 2003). Moreover, as catchment area increases, a greater variety of landscape features contribute to water's thermal profile and their relative importance shift in magnitude. The hierarchical structuring of the river's diverse and changing thermalscape organizes the freshwater ecosystem (Brown & Hannah, 2008), dampening variation (Moore et al., 2015; Steel et al., 2016) and buffering against the changing climate (Chezik et al., 2017). As a result, thermal extremes and the associated stress on organisms will vary depending on network location and the composition of the upslope landscape.

Different landscape features and land-uses influence river temperatures. For example, lakes are known to act as heat sinks, warming and stabilize the thermal regimes of downstream rivers (Wetzel, 2001; Moore, 2006), while glaciers and steep slopes often contribute to stream cooling during the summer through the contribution of melt- and groundwater (Cadbury et al., 2008; Lisi et al., 2015; Stahl & Moore, 2006). Human interventions such as urbanization (Nelson & Palmer, 2007), logging (MacDonald et al., 2003) and stream regulation (Moore, 2006) generally warm riverine temperatures although dams exhibit distinctively mixed effects (Moore, 2006; Todd et al., 2005). Thus, different landscape features will impact downstream water temperatures differently with relative influences shifting seasonally. Accordingly, downstream water temperature regimes are a hierarchical aggregate of the contributing watershed (Vannote et al., 1980; Isaak et al., 2014). Tributaries that drain basins of different composition and size will add their distinct thermal signature resulting in a constantly shifting thermal mosaic over space and time (Palmer et al., 2009; Peterson et al., 2007).

Organisms that migrate through river networks to access essential habitat may experience temperatures that at times exceed their thermal limits depending on their timing and route through the network (Steel et al., 2016; Keefer et al., 2015). Renowned for their extraordinary migrations from the Pacific Ocean to the interior of Western North America, adult Pacific Salmon (*Oncorhynchus spp.*) may transit diverse habitats of varied tempera-

tures in order to reproduce in their natal streams (Quinn, 2011). A coldwater species, adult salmon may be close to or beyond their thermal optimum during their migration (Isaak et al., 2015; Keefer et al., 2008, 2018) leading to thermal stress that can result in pre-spawn mortality (e.g., Hinch et al., 2012). For instance, in especially warm years en route mortality has exceeded 90% in some stocks of Fraser River sockeye (*O. nerka*) (Cooke et al., 2004; Macdonald, 2000). Spawning throughout river systems west of the Rocky Mountain Range at different times throughout the year, salmon species and populations experience varied levels of thermal exposure depending on their migration timing and route (Keefer & Caudill, 2015). If water temperatures increase as expected under climate change, adult migration may become increasingly strenuous, possibly leading to loss of suitable habitat (Ruesch et al., 2012) or abundance declines (Rand et al., 2006; Hague et al., 2011) if rapid evolution can't keep pace with climate change (Reed et al., 2011).

Estimating salmon migratory thermal exposure is challenging as the river thermalscape is quite variable over space and time (Fullerton et al., 2018). Historically, the thermal challenge of migration has typically been assessed low in the watershed (e.g., Wild Fish Conservancy, 2015). Such an approach is pragmatic because temperatures lower in the basin are typically warmer (Steel et al., 2016) and constitute a common migratory corridor to many populations, thereby greatly simplifying the monitoring effort. However, the accumulation of migratory sub-lethal effects on salmon survival has become increasingly concerning (Healey & Bradford, 2011), necessitating improved broad-scale temperature monitoring and modeling (e.g., Isaak et al., 2015). Statistical inference facilitates prediction at unobserved locations, improving estimation of network-wide thermal exposure. To date, these approaches have often greatly simplified temporal (e.g., Isaak et al., 2015) or spatial (e.g., Quinn et al., 1997) thermal exposure contributions. However, studies that have placed internal tags in salmon have found that thermal exposure is dependent on migration timing and route along the network (Keefer et al., 2015). Thus, there is a need for further incorporation of temperature dynamics across space and time towards understanding thermal exposure of migratory salmon.

The Thompson River watershed in central British Columbia Canada is an important salmon catchment where temperatures may be challenging these migratory fishes. The Thompson is a tributary to the greater Fraser River, a system which has exhibited substantial warming over the last half century (Morrison et al., 2002; Ferrari et al., 2007). The Thompson River is home to steelhead (*Oncorhynchus mykiss*), pink (*O. gorbuscha*), Chinook (*O. tshawytscha*), coho (*O. kisutch*) and sockeye (*O. nerka*) salmon, including the Adams River migration of sockeye salmon. Yet, the Thompson River has exhibited many population declines across species due to degraded ocean conditions and freshwater habitat (Labelle, 2009). Sockeye salmon migrating to the South Thompson River watershed have exhibited an increase in pre-spawn mortality (Young et al., 2006; Hinch et al., 2012). Linked to a shift towards earlier adult migration, energy depletion has been cited as a cause

of decreased survival with warmer migration temperatures speculated to be a fundamental contributor (Crossin et al., 2008). Similarly, increasing temperatures is believed to be contributing to the decline of coho, steelhead and Chinook abundances (Bradford & Irvine, 2000). Although direct links of temperature on these species is to date unclear, sub-lethal effects have been shown to exacerbate the impacts of disease and parasites (Gilhousen, 1992) on energy density and curtail aerobic scope (Eliason et al., 2011), making migration more strenuous and survival less likely (Martins et al., 2011). These findings suggest that as water temperatures warm, persistent thermal stress during migration is likely to increase mortality rates through multiple ultimate mechanisms.

Thermal exposure depends on salmon river entry timing and migration rate in concert with the aforementioned climate and landscape spatial controls on the migratory river thermalscape. Here we ask: 1. In what capacity do climate and landscape features contribute to the dynamics of the thermalscape of the Thompson River system? 2. What is the thermal exposure posed to salmon with different migration timing and routes? To address these questions we model climate and landscape effects on seasonal water temperature dynamics using a statistical stream network framework (Hoef & Peterson, 2010; Peterson et al., 2013). Previous work has generally used such network models to predict average mean August water temperature or some other summary statistic as a function of watershed characteristics (e.g., Isaak et al., 2015): here we use these spatial models to actually model the temporal dynamics of water temperatures. This modeling advance allows for the re-creation of thermal dynamics along any migratory route while accounting for changing seasonal dynamics and upslope landscape effects. Understanding the various contributors to stream temperature during migration should inform routes and populations at greater risk of thermal stress and inform how watershed landuse and climate change alter such risks.

## 5.3 Methods

### 5.3.1 Thompson River watershed

To capture the thermal heterogeneity of the Thompson River watershed, Hobo Pendant data loggers were deployed at 103 locations between 2014 and 2017, with the goal of capturing confluence interactions and specific landscape effects (Figure 5.1). By placing loggers in triads, upstream and downstream of major confluences, we aimed to capture the relative contribution of each stream to the subsequent downstream temperatures (e.g., Marsha et al., 2018). We focused logger deployment in the North Thompson River watershed because the basin contains more distinct regions where the contributions of specific landscape features may be more pronounced. In its northern reaches, the North Thompson River watershed is dominated by steep slopes, snow-capped mountains and small glaciers. The west side of the watershed is perched on a high plateau, is relatively flat and dominated by lakes connected by relatively slow-moving streams. In the transition between the northern and western regions

is a large inactive, volcanic field, largely captured within Wells Grey Provincial Park, and distinguished by canyons, rapid flows and dramatic waterfalls. The southern portion of the North Thompson River watershed is narrow, largely composed of the mainstem with many small creeks and streams flowing out of increasingly arid hills. The South Thompson River watershed is dominated by lakes, which are fed by small streams tumbling out of the Canadian Rocky Mountains. The north and south basins of the Thompson River merge in the southwest before continuing to the mainstem of the Fraser River and ultimately to the Pacific Ocean. The Thompson River region is dominated by conifers, is actively logged and frequently experiences wildfires. Agriculture and ranching occur in the watershed but the topology of the region limit these industries largely to the floodplain along the mainstem with the exception of some more extensive orchard agriculture in the South Thompson River watershed's Okanagan Valley and low density ranching in the western portion of the North Thompson River watershed.

### **5.3.2 Stream temperature data**

We monitored stream temperatures at two-hour intervals from July of 2014 to August of 2017, with routine annual data collection and redeployment. By returning each year we ensured the loggers remained in the water column and that depleted batteries and lost loggers were replaced. We affixed the loggers to the landscape by attaching cable to boulders, bridge pylons or large trees expected to remain constant and unyielding to extreme river flows. White PVC shields were attached to the cable to protect the loggers from debris and solar radiation that may bias temperature readings (e.g., Isaak & Horan, 2011). To prevent de-watering events, we sunk loggers attached to weights in deep pools protected from the direct influence of the rivers thalweg when possible. We cleaned the temperature data via manual inspection and flagged data believed to be associated with de-watering events (e.g., Sowder & Steel, 2012). Concerned with the most likely thermal experience, we averaged raw temperature readings to the daily level.

### **5.3.3 Statistical models**

Interested in how temperatures change during adult salmon migration we require linked spatial and temporal models that describe thermal conditions in any given place at any given time on the river network. To do this we first characterize temperature's temporal pattern and then subsequently model those descriptive coefficients given climatic and landscape covariates in a spatial statistical stream network model (SSNM). Combining process and statistical approaches this unique methodology parses landscape, network and seasonal dynamics that then allow for prediction of temperatures across time at any place on the network.

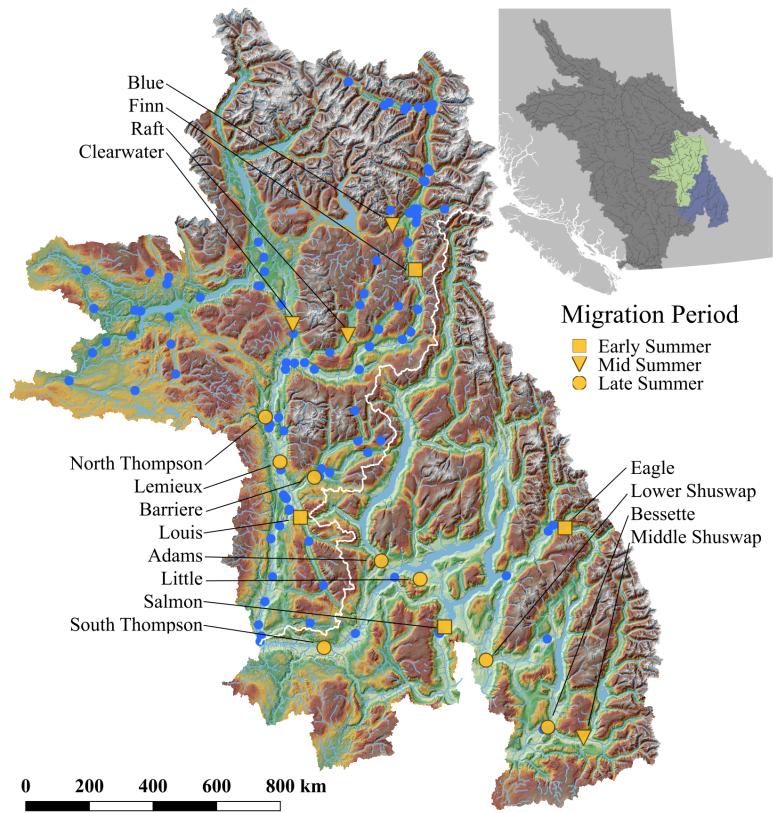


Figure 5.1: Stream temperature monitoring locations (blue) and Chinook salmon populations (yellow) within the Thompson River watershed. The inset indicates the North (green) and South Thompson's (purple) watershed locations within the Fraser River watershed (dark grey) in British Columbia Canada (light grey). The shape of the Chinook salmon population locations indicate the timing of migration through the Albion Test Fishery in early (June), mid (July) and late (August) summer, with the exception of Louis which is considered to migrate in the spring (March-May) but was grouped into the early summer period for this study (see Parken et al., 2008).

## Temporal stream temperature model

To characterize the seasonal dynamics of stream temperatures we used two cyclical models in linear combination,

$$y_{s,t} = \text{Annual}_{s,\text{yr}} + \text{Season}_{s,\text{yr}} + \eta_s, \quad \eta_s \sim N(0, \sigma_s) \quad (5.1)$$

where mean daily ( $t$ ) temperatures ( $y$ ) at each site ( $s$ ) are described primarily by broad seasonal dynamics (*Annual*) and modified by local seasonal conditions (*Season*). The remaining variance ( $\sigma$ ) not captured by these models is assumed to be independent and random (i.e. *i.i.d.*).

Dominated by seasonal shifts in solar radiation due to the tilt of the earth's axis and its orbit around the sun, we describe the primary periodic dynamics of stream temperature using a cyclical model (e.g., Shumway & Stoffer, 2000),

$$\text{Annual}_{s,\text{yr}} = \alpha_{s,\text{yr}} + A_{s,\text{yr}} \cos(2\pi\omega + \tau_{1,s}\pi) \quad (5.2)$$

where the annual oscillation of temperature is described by a modified cosine curve. Shifting the cosine curve vertically captures the mean annual water temperature ( $\alpha$ ) while amplifying the curve captures the annual range of temperatures around the mean ( $A$ ). The frequency of the temperature cycle is captured in ( $\omega$ ) and described by,

$$\omega = d_{1:\gamma}/\gamma, \quad (5.3)$$

where  $\gamma$  is the number of observations per year and  $d_t$  is the integer location of observation  $t$  in the year. Because we averaged our data to the daily level, we used the day of year as a common value of  $d_t$  across sites, and set  $\gamma$  to 365 (or 366 in leap years). Although the solstices should vaguely coincide with the extremes in annual temperature, shifts due to local climatic variation are allowed in  $\tau_1$ . Because  $d_1$  is equivalent to January 1<sup>st</sup>,  $\tau_1$  is approximately described by a value of 1 in the Northern Hemisphere or a shift in the cosine curve of  $\pi$  such that the cosine curve begins near its lowest point.

Hysteresis is a common feature of annual stream temperatures, where rising spring and declining fall temperatures are not symmetrical around summer's peak temperature (Letcher et al., 2016), rendering a simple cosine curve a poor descriptive model on its own. Divergences from the expected pattern of temperature given changing levels of solar radiation are typically a result of seasonal precipitation, where higher stream flows depress stream temperatures (Lisi et al., 2015). In the northern hemisphere this is commonly observed in the spring when snow melt drives a basin-wide freshet. To account for seasonal patterns in stream temperature driven by local climate, we include a double cycle cosine curve that acts to modulate the annual curve (eq. 5.2),

$$Season_{s,yr} = \phi_{s,yr} \cos(2\pi\omega\frac{1}{2} + \tau_{2,s}\pi) \quad (5.4)$$

where  $\phi$  is an expansion factor that alters the strength of the hysteresis and just as in equation 5.2,  $\tau_2$  shifts the curve to align with local seasonal effects. In the Thompson River watershed,  $\tau_2$  typically corresponds with a value of 1.5 which results in warmer winter and summer stream temperatures and cooler spring and fall temperatures. This is because low flows in the summer and winter lead to a greater influence of solar radiation and ground water respectively, while snow-melt and rain dominate in the spring and fall in the Thompson River region.

Implementing this model in the probabilistic programming language Stan (Stan Development Team, 2017) allowed us to explore the entire probability space of these parameters and propagate that variability to the subsequent spatial model. Furthermore, using a Bayesian framework allowed us to provide reasonable prior estimates and limit the parameter space over which the model must search. For instance,  $\alpha$  and  $A$  parameters are not reasonably expected to go below 0°C, therefore we limited those parameters to above zero and provided weak lognormal priors centered on 2 with a variance of 1. Similarly, given the cyclical nature of our model,  $\tau$  values can take on reasonable estimates at regular intervals *ad infinitum*. Running a multi-chain (MCMC) analysis requires the chains agree on an estimate which can be difficult when multiple values lead to the same fit. To avoid disagreement, we limited  $\tau$  to positive values and provided weakly normal priors with a variance of 0.25 centered on 0.9 for  $\tau_1$  and 1.5 for  $\tau_2$ . The expansion factor  $\phi$  was limited to between 0 and 3, maximizing the hysteresis adjustment to no more than 3°C in either direction thereby limiting the models flexibility and influence. After fitting equation 5.1 to the data, we visually inspected the fits at each site and year and consulted diagnostic statistics accepting  $\hat{R}$  value less than 1.1 and effective sampling size ratios of greater than 0.001 (Gelman et al., 2013). We excluded sites and years that had poor visual fits even when other diagnostics were not problematic. Strict data requirement cut-offs were not used as data near the seasonal extremes play a disproportionate role in fitting equation 5.1 parameters. Using a more subjective approach allowed us to include more sites despite varying degrees of missing data.

### Spatial stream network model

Advances in stream network modeling, that account for branching architecture, directed flow, longitudinal connectivity and weighted confluence contributions, have greatly enhanced our ability to account for network spatial relationships and proportion landscape and climate contributions to observations on the network (Hoef & Peterson, 2010; Peterson et al., 2007, 2013; Isaak et al., 2014). Applications of these models to stream temperature data have largely been limited to temporally static summary statistics – for example, mean summer temperature – that attempt to understand network mediated landscape contributions

to biologically meaningful thresholds (Isaak et al., 2010). Here we advance this approach by using these spatial statistical stream network models on the actual parameters that describe the entire seasonal dynamics. This approach allowed us to deconstruct how different watershed features affect the dynamics of temperature regions – for example, watershed features may differently affect the average temperature or the amplitude or seasonality of the streams annual temperature profile.

Currently, multi-variate methods that link predictors to a variety of responses simultaneously are not available within the stream-network framework. Therefore, we built individual models for each parameter resulting in five stream network models with the general construction,

$$\text{std\_coef}_{s,yr} = \mathbf{X}\boldsymbol{\beta} + \epsilon_{s,yr}, \quad \epsilon \sim z_d + z_{yr} + z_s + z_{nug}, \quad (5.5)$$

where standardized equation 5.1 coefficients (std\_coef, see table 5.1) at site  $s$  in year  $yr$  are predicted by a matrix of climate and landscape variables  $\mathbf{X}$ , where the relationship between the temporal model's coefficients and these predictors are described by a vector of  $\boldsymbol{\beta}$  coefficients. The error term is decomposed into random effects ( $z$ ) accounting for basin wide year effects ( $z_{yr}$ ), site effects ( $z_s$ ), exponentially weighted flow connected autocovariance ( $z_d$ ) and residual *i.i.d* error ( $z_{nug}$ ) (Hoef & Peterson, 2010).

In order to calculate predictor variables at each site, we first built a landscape network describing the branching architecture of our streams (Theobald et al., 2006) and delineated reach contributing areas (RCA) for each downstream confluence in the network (Peterson & Hoef, 2014). Geo-referenced stream data were gathered from the British Columbia Data Catalogue made available by the Ministry of Citizens Services. Streams were subset to greater than second order, cleaned of network braiding, complex confluences, pseudo-nodes and all vectors were directed towards the outlet. Sites were snapped to the network (Figure 5.1) and center-points of each stream segment were used as prediction locations. Reach contributing areas were delineated using a 25 metre digital elevation model (DEM). Prior to delineation, the DEM was cleaned of topological errors and the known stream network was burned into the DEM at a 5 metre depth. All predictor variables were summarized by RCA using the *zonal\_stats* tool in the Python package *rasterstats*.

Climate and landscape predictor variables known to significantly contribute to stream temperatures (e.g., Isaak et al., 2010), were gathered from multiple sources. Climate data were calculated using the open source tool ClimateBC (Wang et al., 2016), which extracts and downscals PRISM climate normal data and extrapolates to any location within British Columbia (Daly et al., 2008). Sampling at 1 km resolution across the Thompson River watershed, we used ClimateBC to estimate the mean annual air temperature in each year between 2014 and 2017. We calculated the amplitude around the mean as the difference between the average maximum summer air temperature (June-Aug.) and the average minimum winter

air temperature (Dec.-Feb.). Static landscape variables included glacial coverage gathered from the Rudolph Glacier Inventory (v5.0) (Arendt et al., 2015), lake area calculated from polygons in the BC Freshwater Atlas and elevation extracted from the aforementioned DEM. To characterize dominant changes to the landscape in recent decades, we extracted high resolution forest change data from the National Forest Information System. Derived from Landsat images, these data identify logging and wildfire for each year between 1985 and 2010 at a 30 metre resolution (White et al., 2017). Summing these data within each RCA, we determined the total loss of forest in each category over the 25 year period as a general measure of impact. Climate and elevation data were averaged by RCA whereas landscape variables were summarized by total km<sup>2</sup> in each RCA.

Ultimately, we are interested in the contribution of each predictor variable from the entire watershed to each observation and prediction site. Using the spatial relationships defined in our landscape network, we calculated the upslope area average of each predictor for each site. To characterize the contribution of glaciers, lakes and forest-change relative to the size of each sites catchment, we divided these variables by their catchment area, thereby calculating their proportional area coverage. Visual inspection suggested a logit transformation of these percentages in order to spread the data more evenly and limit the influence of extreme data points. Climate, elevation and catchment area were all centered and scaled in order to limit the influence of the y-intercept and allow effect size comparison across variables (Table 5.1).

Not all predictor variables display strong relationships with each temporal coefficient, therefore we limited predictors and interactions in each model to those that resulted in lower Akaike Information Criterion (AIC), significant gains in explained variance and improved leave one out cross validation results. Once SSNM models were constructed for each parameter, we iteratively sampled coefficient estimates from the posteriors of equation 5.1, fit our SSNM models to those estimates and predicted temporal coefficients at our prediction sites. This process resulted in 250 estimates of each coefficient for each year (2014-2017) at 4376 prediction locations throughout the Thompson River watershed.

To approximate the thermal exposure of migrating salmon, we need to define when salmon might arrive in the Thompson River, how quickly they might move through the river network, and what temperatures are considered stressful. Previously, migration timing of Chinook populations throughout the Thompson River watershed were estimated in 2000 and 2001 by Parken et al. (2008) using coded wire tags and genetic sampling. Measured at the Albion Test Fishery, approximately 50 km upstream from the Fraser River mouth, populations occurred in three run-timing groups, early (June), mid (July) and late (August) summer. Using an estimated migration rate of 36 km per day (Salinger & Anderson, 2006), we calculated the number of days ( 13 days) required to travel the 447 km between the Albion Test Fishery and the confluence of the North and South Thompson Rivers. Although we do not account for temperature-driven changes in migration speed (e.g., Salinger & Anderson,

	Variable	Mean	SD
Response	$\ln(\alpha)$	0.29	
	$\ln(A)$	0.26	
	$\phi$	0.72	
	$\tau_1$	0.04	
	$\tau_2$	0.17	
	$\sigma$	0.20	
Predictor	Air Amplitude $H_2O$	29.6	3.55
	Air Mean $H_2O$	3.11	1.94
	$\ln(\text{CA}) H_2O$	2.75	2.21
	Elevation $H_2O$ (m)	1431	349

Table 5.1: Mean and standard deviation (SD) values of predictor and response variables. Only the standard deviation value is reported for the response variable as these values were only standardized but not scaled. Moreover, these SD values are averages from the 250 samples drawn from equation 5.1 parameter posteriors. Predictor variables were both centered and scaled and the respective mean and SD values were calculated across all observed and predicted sites. Transformations of some variables by the natural log ( $\ln$ ) was done prior to calculating the mean and SD.  $H_2O$  indicates predictors that constitute the average value among values averaging over the contributing area. For instance, the average contributing area elevation among predicted and observed sites was 1431 which is similar to, but not the same as, the Thompson's overall mean elevation of 1264 metres as sites are nested.

2006), we do sample migration dates uniformly across the migration windows which should consume individual variation and generalize our findings more widely. Moreover, although we are using migration timing data for Chinook salmon, we note that these different run-timing groups also apply to sockeye salmon and were chosen by Parken et al. (2008) to correspond with contemporary conservation requirements for salmon more broadly (Bailey & Canadian Science Advisory Secretariat, 2001).

Upon estimating salmon arrival in the Thompson River system, we used our predicted SSNM model estimates of equation 5.1 parameters to calculate expected daily temperature probability distributions at each prediction site on the expected day of passage. To do this we calculated the day of passage by dividing the upstream distance of any prediction site by the migration speed (36 km/day) and added the result to the date of arrival at the outlet, giving the expected  $d$  value at each prediction site in the network, dependent on the day of passage through the Albion Test Fishery. Using the expected  $d$  value in combination with all 250 site predicted coefficient estimates we calculated the full distribution of mean temperature and standard deviation estimates at each prediction point (e.g. Figure 5.2). These distributions were then used to calculate three metrics of thermal exposure that describe the difficulty of migrating to and spawning at any location on the network.

*Thermal Exposure Probability:* Temperature sensors capture a single point estimate within a thermally heterogeneous environment. We are interested in the most likely (i.e.,

mean) thermal experience and the probability of experiencing stressful temperatures above a known threshold at a given place and time. Assuming thermal heterogeneity of the environment is normally distributed around the mean temperature estimate, we can calculate the area under the curve equal to or exceeding any threshold temperature that may be stressful to salmon and divided this value by the total area, returning the probability of experiencing temperatures as warm or warmer than the given threshold at that point in time. For example, if our mean temperature estimate was 17°C with a standard deviation estimate of 2, the probability of a temperature above 19°C would be 0.16. Using all 250 mean and standard deviation estimates at a given site, we calculate a distribution of probabilities and return the median probability of experiencing an above threshold temperature at that location. Thus, the thermal exposure probability is the median probability of experiencing a temperature greater than the given threshold at a given site.

*Cumulative Migratory Thermal Exposure:* Following the flow connected path from each site to the outlet, we summed all above threshold mean temperature estimates. These values indicate temperatures with at least 50% probability of exceeding a given threshold. A greater number of these high thermal exposure probability estimates result in greater cumulative migratory thermal exposure, indicating persistent or acute thermal challenges during migration. For instance, locations with extremely high thermal exposure probability will result in a greater proportion of the 250 mean temperature estimates being above the threshold temperature, resulting in an acute increase in cumulative migratory thermal exposure. Whereas persistent thermal exposure would occur when many locations along the migratory route have mean estimates above threshold. Thus, higher cumulative thermal exposures indicate fish migration routes with greater potential of experiencing deleterious temperatures overall. We reported the median value for each run-timing group to capture the typical cumulative thermal exposure throughout the migration window.

*Average Migratory Thermal Exposure Probability:* Finally, we divided the cumulative migratory thermal exposure by the total number of temperature estimates along the migration route ( $n = \#sites \cdot 250$ ), giving the average migratory thermal exposure probability. While the migratory exposure strictly increases during upstream migration, the exposure probability can both increase and decrease, thereby depicting changing temporal and landscape effects on temperature during migration and separating persistent and acute thermal challenges on the network. This metric can also be thought of as the proportion of temperature estimates with a greater than 50% chance of exceeding the temperature threshold, over the course of the migration. We again reported the median value for each run-timing group to characterize the typical cumulative migratory thermal exposure probability.

This flexible analytical approach can calculate thermal exposure at a variety of thresholds and spatial extents as determined by species thermal limits and behaviours. Here we use a lower threshold of 19°C to indicate significant but manageable stress and a value of 22°C to indicate temperatures blocking migration per values identified for sockeye, steelhead and

Chinook salmon (Richter & Kolmes, 2005). We emphasize that these temperature thresholds are for considering temperature – in order to consider risk we would also need to consider the sensitivities of different populations, which can vary due to different local adaptations (e.g., Eliason et al., 2011; Anttila et al., 2018). Ultimately, this process produces median thermal exposure probabilities by migration group, in any year and at any temperature threshold of interest, at any point in the network. As such, we calculated the median cumulative migratory thermal exposure and median spawn site thermal exposure probability for the 16 populations of Chinook salmon identified by citetParken:2008 in the Thompson River watershed (Figure 5.1) per each populations migration route and run-timing.

## 5.4 Results

### 5.4.1 Temporal stream temperature model

Of the 408 unique site-year combinations, we selected 290 that displayed strong fits, capturing the maximum temperatures and hysteresis present in the data (e.g., Figure 5.2). Of those the data were spread fairly evenly between years with 61 estimates in 2014, 91 in 2015, 67 in 2016 and 71 in 2017. Many sites displayed poor fits when temperature data near the annual extremes were limited and thus were eliminated from the spatial analysis. Frequently, data were limited due to dewatering events, late (i.e., fall) deployment in 2014 or logger loss. The selected sites in figure 5.2 exhibit many of the characteristics that either constituted acceptance or rejection. For instance, the North Thompson (NT) Headwaters and Barrièr River lower in the North Thompson River watershed displayed four quality years of data and acceptable fits. Meanwhile, Bridge Creek and the Shuswap River each exhibited one year, 2014 (late deployment) and 2016 (logger loss) respectively, that did not capture enough data to constitute convincing estimates of all parameters. Thus, these site-years were removed from the spatial analysis. Overall fits were strong with no apparent bias seasonally except for when peak summer temperatures were extremely 'sharp' but our estimates of variance largely captured these values at 95% confidence (e.g., Bridge Creek, 2015 Figure 5.2).

### 5.4.2 Spatial stream network models

Stream network (SSNM) model selection resulted in primarily strong fits that accounted for much of the variance in the temporal model parameters (eq. 5.1). Leave-one-out cross validation demonstrates the predictive capability of our models with largely strong 1:1 relationships between observed and predicted parameter values (Figure 5.3). Variance ( $\sigma$ ) and timing ( $\tau$ ) coefficient models exhibited particularly strong predictive relationships with limited exception, likely due to their minimal variance among sites and shared values across years (i.e., eq. 5.2). Elevation was the only covariate included in these three SSNM models (i.e.,  $\sigma$  and  $\tau_{1,2}$ ), with higher average catchments leading to lower values across parameters.

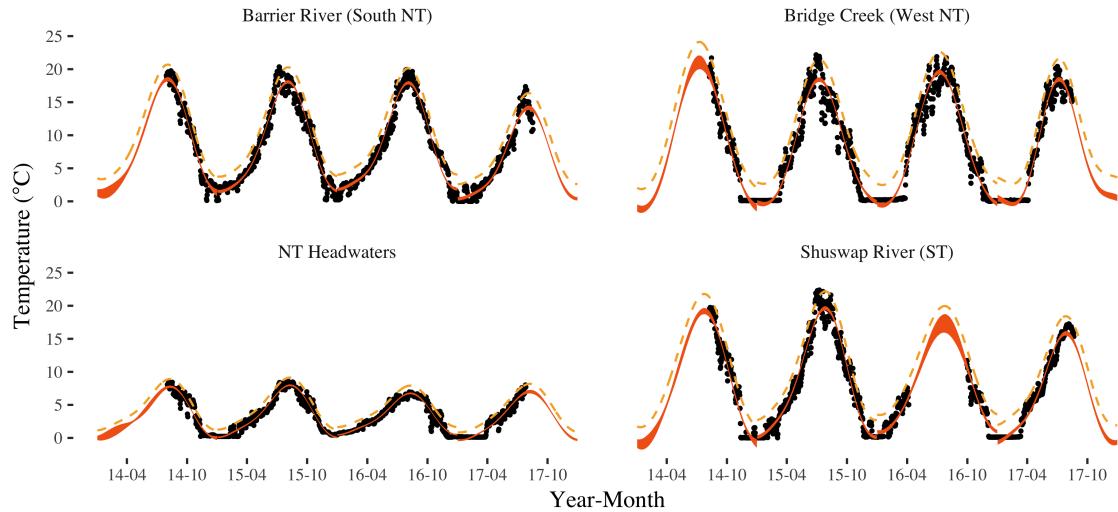


Figure 5.2: Examples of model fits to equation 5.1 at 4 sites in Thompson River watershed. The red ribbon describes the 95% credibility interval of the mean temperature estimate, while the orange dashed-line describes the average 95% variance window above the mean. The Barrière River and Bridge Creek are tributaries to the North Thompson River (NT), while the Shuswap River drains Mabel Lake and flows into the South Thompson River (ST).

In other words, higher elevation was associated with less thermal variation ( $\sigma$ ), a later onset of spring ( $\tau_1$ ) and delayed seasonal effects of snow and rain ( $\tau_2$ ). A key difference between these three models is the contribution of the covariate relative to the correlation structures. Elevation dominated in the  $\sigma$  model (46%), while only capturing 10% of the variation in the  $\tau$  models. Site-level correlation ( $z_s$ ) accounted for much of the remaining variance in all three models but  $\tau_1$  demonstrated a particularly strong effect of site (62%) while variance in the  $\tau_2$  parameter was primarily captured through autocorrelation ( $z_d$ ) along the flow connected network (52%). All three of these SSNM models had very little unaccounted for residual variance ( $\ll 1\%$ ) (Figure 5.3).

Covariates that contributed to the efficacy of  $\alpha$ ,  $A$  and  $\phi$  SSNM models varied with the exception of mean annual air temperature which was the dominant predictor across models (Figure 5.4). Water's mean annual temperature ( $\alpha$ ) was additionally positively affected by the annual range of temperatures averaged across the contributing area (Air Amplitude), the percent catchment burned (WildFire), the catchment's mean elevation (Elevation) and the percentage of lake area (Lakes). The interaction of elevation and lake predictors was notably the only cooling influence on  $\alpha$  while counter-intuitively elevation alone had a slight warming effect. Overall, these covariates captured on average 37% of the variation in the observed  $\alpha$  values while auto-correlation along the flow connected network captured 29%. Much of the remainder was accounted for by correlation within years (25%) and within sites (4%) with 5% of the variance in  $\alpha$  remaining unaccounted.

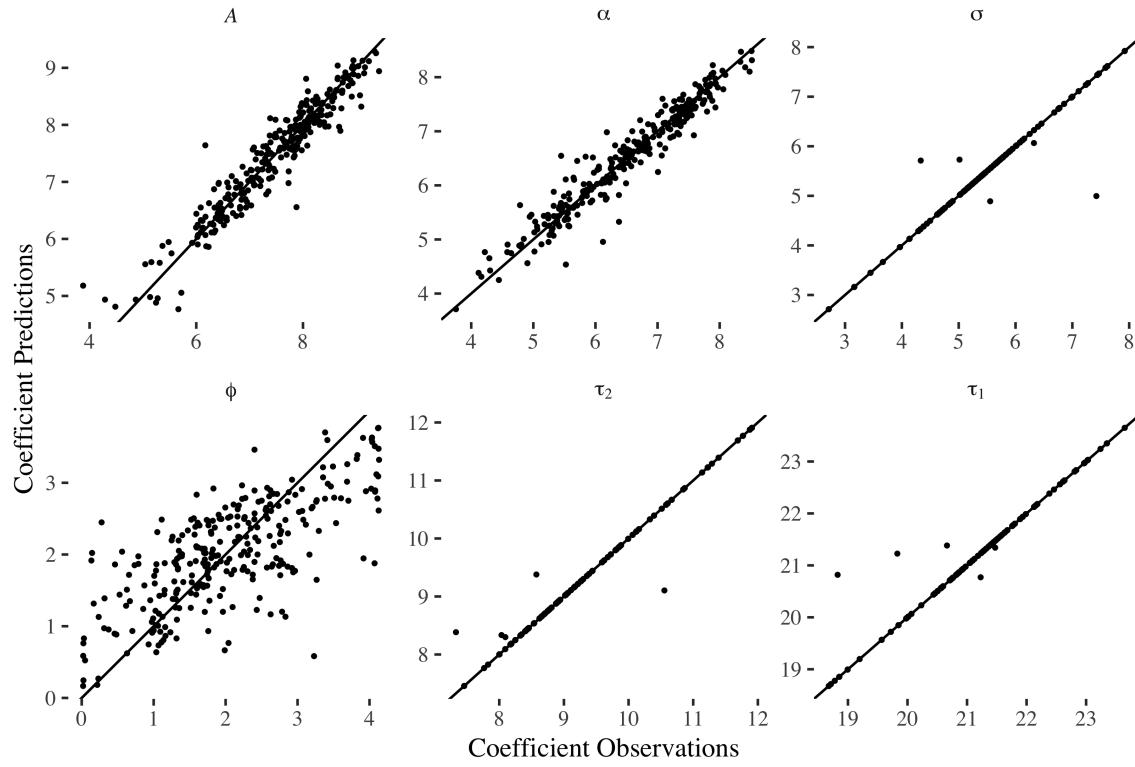


Figure 5.3: Evaluation of SSNM model predictive power using leave-one-out cross validation for each temporal model parameter. A 1:1 line describes a perfect fit between observed and SSNM predicted values. Parameter estimates were randomly drawn from equation 5.1 parameter posteriors and predicted by SSNM models (see Figure 5.4 for covariates).

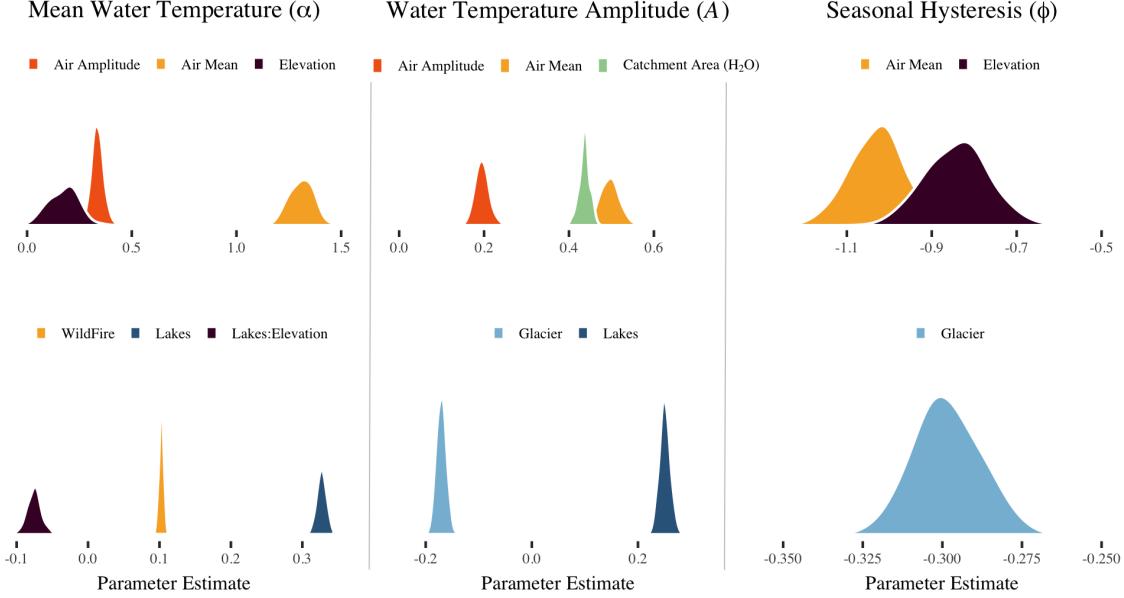


Figure 5.4: Coefficient estimates for mean annual water temperature ( $\alpha$ ), water temperature amplitude ( $A$ ) and seasonal hysteresis ( $\phi$ ) SSNMs by column respectively. Top row coefficients have been scaled while the bottom row have been logit transformed making parameter estimate effect sizes comparable by row within models.

The SSNM describing the amplitude ( $A$ ) of water temperatures around the mean showed positive relationships with the contributing basins mean air amplitude, the catchment area, its percentage lake area and glacial coverage. The percent catchment area covered by glaciers reduced the range of annual temperatures around the mean (center, Figure 5.4). Thus, glaciers dampen temperature variation. The covariates in this SSNM captured the largest amount of variance in the data (46%) compared with covariates in all other SSNMs. The remainder was largely evenly distributed between the flow connected correlation structure ( $z_d = 19\%$ ) and correlation within sites ( $z_s = 18\%$ ).

The  $\phi$  parameter controlling the strength of seasonal hysteresis was notably under-estimated by the SSNM at high values and over-estimated at low values, resulting in relatively even attribution of this parameter across sites (Figure 5.3). Neither the covariates (10%) nor the autocorrelation structures ( $z_d$ ,  $z_{yr}$ ,  $z_s$ ) dominated in accounting for variance in this parameter (20, 25 and 7% respectively). The little variance captured by covariates included negative effects of mean annual air temperatures, mean catchment elevation and percent catchment area covered by glaciers (Figure 5.4). In other words, seasonal hysteresis is reduced in colder regions, a characteristic typical of higher elevations, and further dampened by the presence of glaciers at high elevations.

### 5.4.3 Projected network thermal exposure

Spatial projections of thermal exposure probabilities highlight the thermal challenge of the South Thompson River's mainstem and the North Thompson River's western high plateau (Figure 5.5 & Supplementary Figure 5.7). The potential for stressful conditions above 19°C were relatively elevated during the mid-summer migration and least likely during the late-summer migration (Supplementary Figure 5.7). Furthermore, the likelihood of stressful conditions varied across years with a general decline between 2014 and 2017 across migratory periods. Thermal exposure at 22°C was extremely unlikely across migrations and years (Supplementary Figure 5.8). Of those regions that had elevated stress potentials above 19°C, large lakes in low elevation basins constituted the most common landscape characteristic, such as in the Okanagan Valley region of the South Thompson River watershed.

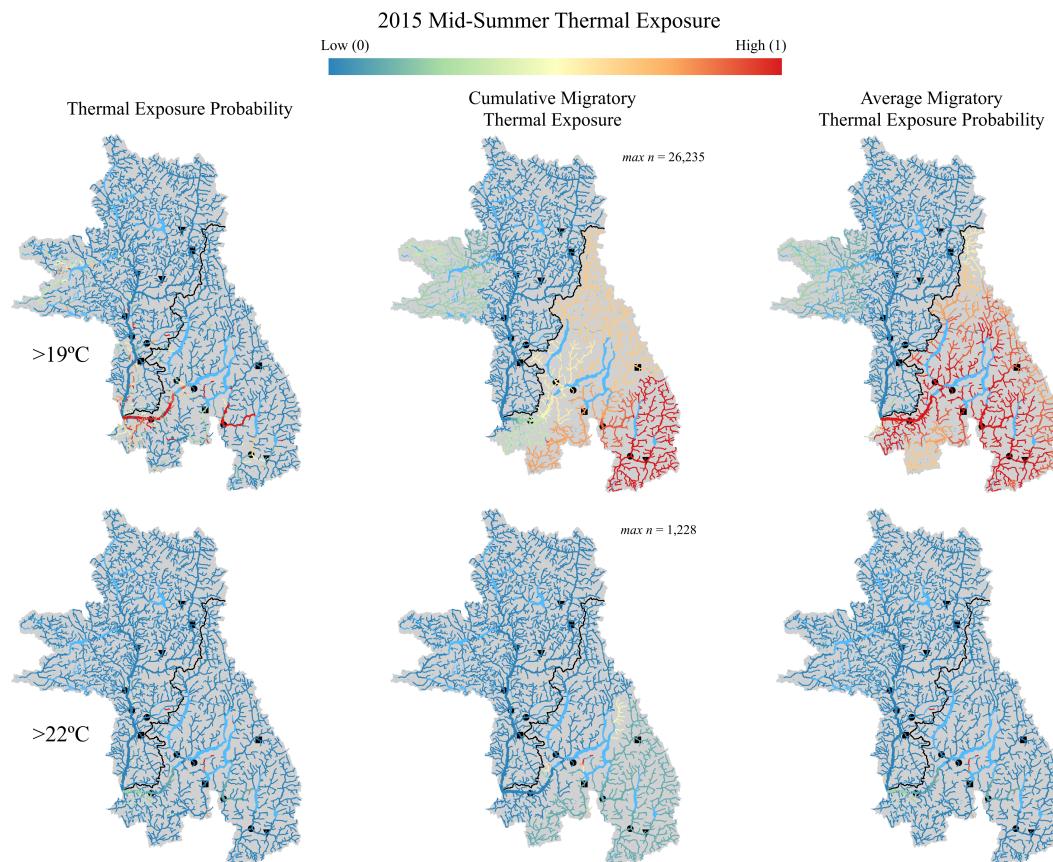


Figure 5.5: Thermal exposure above 19 and 22°C in the North and South Thompson River watersheds as exemplified by the mid-summer migration in 2015. (**left**) The median probability of experiencing above threshold temperatures at each point in the network. (**center**) The cumulative migratory thermal exposure from the outlet to each point on the network. The color bar ranges from 0 (low) to  $\text{max } n$  (high) rather than depicting a probability in this column. (**right**) The median value of the average migratory thermal exposure probability.

Reflecting regionally elevated thermal exposure probabilities, cumulative migratory thermal exposure generally increased with migratory distance and was much higher in the South Thompson River watershed (Figures 5.5, 5.6 & Supplementary Figure 5.9). In other words, South Thompson River watershed residents are more likely to experience stressful temperature than those in the North Thompson River watershed, and the opportunity of experiencing these temperatures grows the further these southern populations migrate. This finding was particularly true in 2014, early and mid-summer of 2015 and mid-summer of 2016, while elevated thermal exposure was relatively unlikely in 2017. Cumulative migratory thermal exposure above 22°C, was nearly non-existent in either basin across sites, years and migratory periods (Supplementary Figure 5.10).

Contrary to the cumulative thermal exposure, the average thermal exposure probability declined with migratory distance in the South Thompson River watershed (Figure 5.5 & Supplementary Figure 5.11). In the North Thompson River watershed, the probability of thermal stress was notably highest in the western portion of the watershed and did not follow the migratory distance patterns of the South Thompson. These findings also reflect the thermal exposure probability dynamics of the mainstem versus the tributaries (Figure 5.5 & Supplementary Figure 5.11). Moving further up the watershed and away from the mainstem reduces the local probability of thermal stress leading to an overall decline in the migratory probabilities of thermal stress. Thus, once migratory salmon make it through the warm lower mainstem portions of these rivers, they will generally encounter lower temperatures. There was little to no evidence that the 22°C threshold is likely among any migration route (Supplementary Figure 5.12).

#### **5.4.4 Chinook salmon thermal exposure**

Different Chinook salmon populations with different migration timing and routes had different exposure to excessively high temperatures. South Thompson River watershed populations are much more likely to experience thermal exposure than those in the North Thompson River watershed (Figure 5.6 & Supplementary Figure 5.13). In many years the potential for thermal exposure ( $>19^{\circ}\text{C}$ ) among southern populations persisted after diverging off the mainstem and arriving on the spawning grounds but rarely above 50% probability. At the 22°C threshold, only the mid-summer migrants in the South Thompson River watershed exhibited any chance ( $>50\%$ ) of experiencing these temperatures during migration and they were extremely unlikely across populations on the spawning grounds (Figure 5.6). Notably, longer distance migrations were exposed to a greater number of likely stressful temperatures in the southern watershed. Broadly, thermal exposure was consistent across run-timing groups and years within watersheds though fewer populations are observed during the more stressful mid-summer migration window in the South Thompson (Supplementary Figure 5.13). Across populations, thermal exposure potential declined from 2014 to 2017.

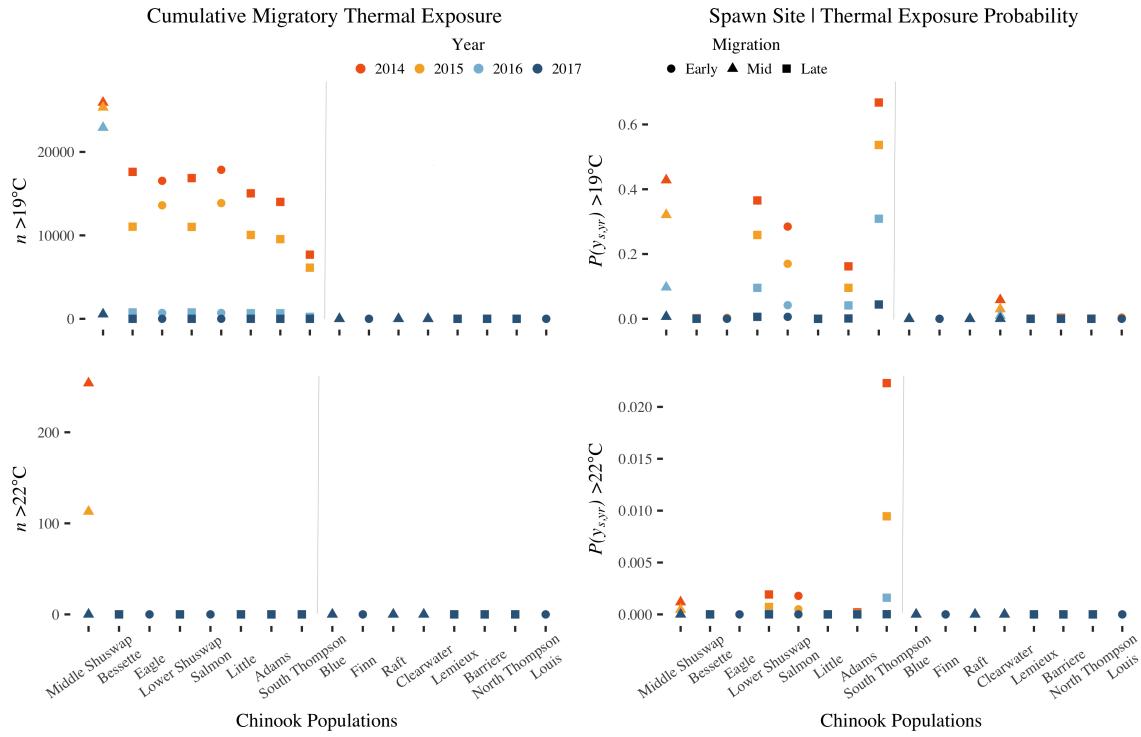


Figure 5.6: Thermal exposure above 19 and 22°C for 16 Chinook salmon populations in the South and North Thompson River watersheds from 2014 to 2017. (**left**) The cumulative migratory thermal exposure along the populations migratory route. (**right**) The thermal exposure probability upon arriving on the spawning grounds. All values are calculated given the populations run-timing and populations are ordered from furthest to shortest migration within their respective watershed. Migration refers to run-timing through the Albion Test Fishery near the outlet of the Fraser River in the early (June:Circle), mid (July:Triangle) and late (August:Square) summer.

## 5.5 Discussion

This study linked a large field temperature collection effort with new analytical approaches to model stream temperatures throughout a large watershed towards predicting thermal exposure of different migratory salmon populations. We focused on the 38,000 km<sup>2</sup> Thompson River watershed of British Columbia, a climatically-diverse watershed of significant topographical complexity, home to imperiled salmon populations. We decomposed the different ways that climate, land-use, and watershed features contribute to the temperature regimes of this region – for example, wildfire and lakes lower in the watershed were associated with warmer average temperatures, whereas glaciers damped temperature variation. Accordingly, there was great spatial variation in river temperatures yet this heterogeneity demonstrated significant spatial organization by the river network. Spatial and temporal thermal variation along migratory corridors result in different degrees of thermal exposure among migrating Chinook salmon populations, destined for locations throughout the basin and returning at varied times throughout the summer.

Watershed features and their interactions demonstrated diverse controls on the different aspects of temperature regimes. Lakes were a particularly dominant landscape feature contributing strongly to stream temperatures (Figure 5.4). Although the North and South Thompson River watersheds share nearly the same total lake area (726 and 756 km<sup>2</sup> respectively), these two basins exhibit dramatically different lake effects on temperature. Unlike the North Thompson River watershed, 47% of the South Thompson River's two longest and largest rivers' are consumed by large lakes. The impoundment of tributary waters in these lakes allows for surface waters to warm from increased solar radiation over the summer. The subsequent outflow is then dominated by these warm surface temperatures which do not reflect colder inputs from upstream (Wetzel, 2001). In contrast the North Thompson River watershed exhibits a cooling effect of lakes due to the mediating effect of elevation. In this watershed many of the lakes are found at high elevation and near the headwaters where snow and glacial melt dominate, leading to a cooling of downstream temperatures (Figure 5.4). Although this pattern is somewhat confounded by a strong latitudinal gradient, the western portion of the North Thompson stands in contrast with greater thermal exposure potential despite its northern location. In this western region, lakes are characteristically more similar to the South Thompson River basin in network location and size, exerting a similar, albeit milder, influence. Together these findings demonstrate the complex mixture of landscape interactions filtering the climate and leading to heterogeneous stream temperature dynamics, organized by the network.

Warming global temperatures could alter the effects of current landscape features. As we found lakes at higher elevation were associated with cooler stream temperatures, possibly due to short residence times and the water's cyrospheric origin, making these lakes cooling features for downstream habitat conditions (Dorava & Milner, 2000; Slemmons et al., 2013).

As global temperatures rise, the greater Fraser River basin is transitioning away from a snow-dominated precipitation regime towards an increasingly snow-rain hybrid (Kang et al., 2016), common to similarly temperate regions around the globe (Hartmann et al., 2013). With less snow and ice on the landscape, our models suggest the thermal character of many locations will become warmer as the cooling effect of high elevation lakes are lost (Figure 5.4). Moreover, just as these cool-water reservoirs are depleting, climate change is producing hotter and persistently dry periods that are encouraging a greater risk of wildfire (Jolly et al., 2015). The loss of canopy cover along stream banks allows a greater degree of direct solar radiation on the waters surface, increasing stream temperatures (Pettit & Naiman, 2007). Although our study already identifies wildfire's warming contribution to stream temperatures, we might expect an increasing degree of influence as wildfires become larger and more frequent on the landscape (Heyck-Williams et al., 2017), leading to a greater impact of fire on stream temperature and thus cold-water fishes (Williams et al., 2015).

The analytical approach we employed advances the application of spatial stream network models so that they recreate the seasonal dynamics of temperatures on the river network, rather than just time-invariant summary statistics of temperature. Complementary temporal and spatial models may correct for poor fits and result in more accurate and precise estimates in unknown locations. It's largely understood that considering the spatial linkages between observations in a stream network improves prediction in unknown locations (Peterson et al., 2013; Hoef & Peterson, 2010). By leveraging shared information between estimates along the network and by refitting the model to all posterior values, we not only explore the full range of possible values at a given site but also draw more extreme values towards the mean. As a result, spurious temporal model estimates are properly down-weighted towards a broader consensus among data points, improving temperature estimates and subsequently migratory thermal exposure estimates. In this way, we not only provide fully probabilistic estimates of thermal exposure to extreme temperatures but also narrow the variance and improve certainty. Unfortunately, we build these models separately and were required to eliminate temporal parameter estimates in years where data was limited, but uncertainty was propagated across modeling steps. To avoid data loss, it might be possible to fit the temporal model simultaneously under a hierarchical spatial model, such that data informing temporal model parameters could be shared along the network, leading to better temporal fits and estimates in data poor years. Ultimately, our data collection and statistical methodology generalize the spatial heterogeneity of the thermalscape, providing insights into the most likely thermal experience and the relative challenge of different migratory routes. More spatially resolved data might improve the accuracy and precision of the model but would still not capture distinct cold-water refuges (e.g. ground water seeps), that salmon might use as thermal oases, leap-frogging through challenging areas to more suitable habitat (Fullerton et al., 2018). More complex models with much greater geomor-

phic resolution or direct observation studies utilizing internal tagging methodologies (e.g., Keefer & Caudill, 2015) could be useful for identifying these habitat features.

We discovered that different Chinook salmon populations with different migration routes and timing had different thermal exposures. Many locations were consistently likely to reach 19°C, a temperature that can cause stress, and some locations and times exhibited elevated probabilities of 22°C, which can cause mortality in some populations (Richter & Kolmes, 2005). Salmon migrating through different parts of the watersheds were exposed to greatly different water temperatures. In the southern watershed, the downstream rivers that salmon must migrate through reached the highest temperatures of the whole watershed, such that these migratory corridors present one of the biggest thermal challenges to these adult salmon. In contrast populations in the northern watershed exhibited limited and localized thermal challenges most broadly isolated to the watersheds uninhabited western region (Figure 5.6). In addition, different run-timing groups of salmon were also exposed to different water temperatures with the mid-summer run in the South Thompson River watershed demonstrating the greatest exposure (Figure 5.6 & Supplementary Figure 5.13). These data are a key step in understanding the thermal risks posed to different salmon populations at present day and the future. It is important to acknowledge that different populations of salmon have different thermal tolerances due to local adaptation (e.g., Eliason et al., 2011). It is also possible that rapid evolution or phenotypic plasticity in run-timing may allow salmon to keep up with climate change (Reed et al., 2011). Migrating salmon also may behaviorally thermoregulate by using coldwater refuges not capture in this study. Regardless, these systems appear extremely close to upper thermal thresholds for salmon suggesting small increases in summer temperature could lead to much higher thermal stress and poorer outcomes for these populations. Currently, thermal stress is primarily considered at the common migration corridor near the ocean outlet and on the spawning grounds as this is a particularly sensitive life history (Richter & Kolmes, 2005), but our work shows that thermal challenges exist in many locations of the river network. By considering the cumulative thermal challenge of migration, now and in the future, we may more effectively identify populations of greater thermal exposure and associated conservation concern than previously recognized.

The observed inter-annual variation of thermal exposure in the Thompson River basin suggest limited climatic shifts could result in increasingly persistently challenging stream temperature conditions for salmon in the future. Climate projections in the Fraser River basin predict a shift in precipitation from a snow to rain dominated regime, leading to earlier spring run-off and lower summer flows (Shrestha et al., 2012). Warming stream temperatures and reduced river flows have already been observed in the Fraser Basin (Morrison et al., 2002). During this study, the particularly dry winter of 2014 produced at least 100" less snow pack than the subsequent three years according to local ski resort records, and exhibited the warmest stream temperatures and the greatest probability of >22°C exposure (see

Supplementary Figures). Meanwhile increased snowfall in 2015 and 2016 exhibited lower stream thermal exposure despite approximately 1.5°C warmer basin wide mean annual air temperatures. In combination, cool climates and normal winter snow fall in 2017 resulted in the lowest thermal exposure in this study. These observations suggest that punctuated climate extremes from inter-annual variation, driven by regional climate processes such as the Pacific Decadal Oscillation or El Niño, will drive inter-annual variation in thermal exposure conditions. As climate change progresses, rising mean annual air temperatures will reduce snowpack and increase glacial retreat thereby weakening the thermal buffering capacity of the watershed. Under such conditions it is reasonable to expect variability will lead to particularly acute stressful conditions in some years (Hague et al., 2011) that will systematically and abruptly act on some salmon populations more than others, possibly negatively affecting diversity in the watershed (Ward et al., 2015; Beechie et al., 2006).

Our study has two main implications for the management of salmon. First, we identify populations that have higher exposure to potentially dangerously high, water temperatures. Such information can inform on-going risk assessment processes and conservation prioritization. For example, 7 Chinook salmon populations in the South Thompson River watershed were likely to experience persistent temperatures at or above 19°C during migration in some years (Supplementary Figure 5.13), with stressful conditions continuing onto the spawning grounds for 5 of those population (Figure 5.6). Potential management actions include precautionary intervention in both commercial and recreational fisheries for stocks that may be suffering from high thermal exposure (e.g., Macdonald et al., 2010). Second, we develop a predictive framework that connects watershed characteristics to temperature regimes. Thus, information such as this can help identify watershed features that may predispose systems to thermal risk. For example, while lakes are foundationally important habitats to salmon (e.g., sockeye salmon), they may render watersheds more vulnerable to excessively high temperatures. Moreover, it may illuminate regions where climate affects could exacerbate thermal exposure such as forests with increasing wildfire risk or the loss of glacial effects.

## 5.6 Acknowledgements

This study relied on many open source tools including R v3.5.1 (<https://www.r-project.org/>), rStan v2.17.4, Python v2.7 and QGIS v2.18. All R code can be found at <http://github.com/kchezik/River-Network-Temperature.git>. Data processing was facilitated greatly by GNU Parallel (Tange, 2011). We also wish to acknowledge David A. Patterson at the Department of Fisheries and Ocean Canada for support with, and continued monitoring of, the Thompson River watershed. K.A. Chezik was financially supported by the SFU KEY Big Data Fellowship. K.A Chezik and J.W. Moore were also supported by the Liber Ero Chair of Coastal Sciences and Management and Simon Fraser University.

## 5.7 Supplementary material

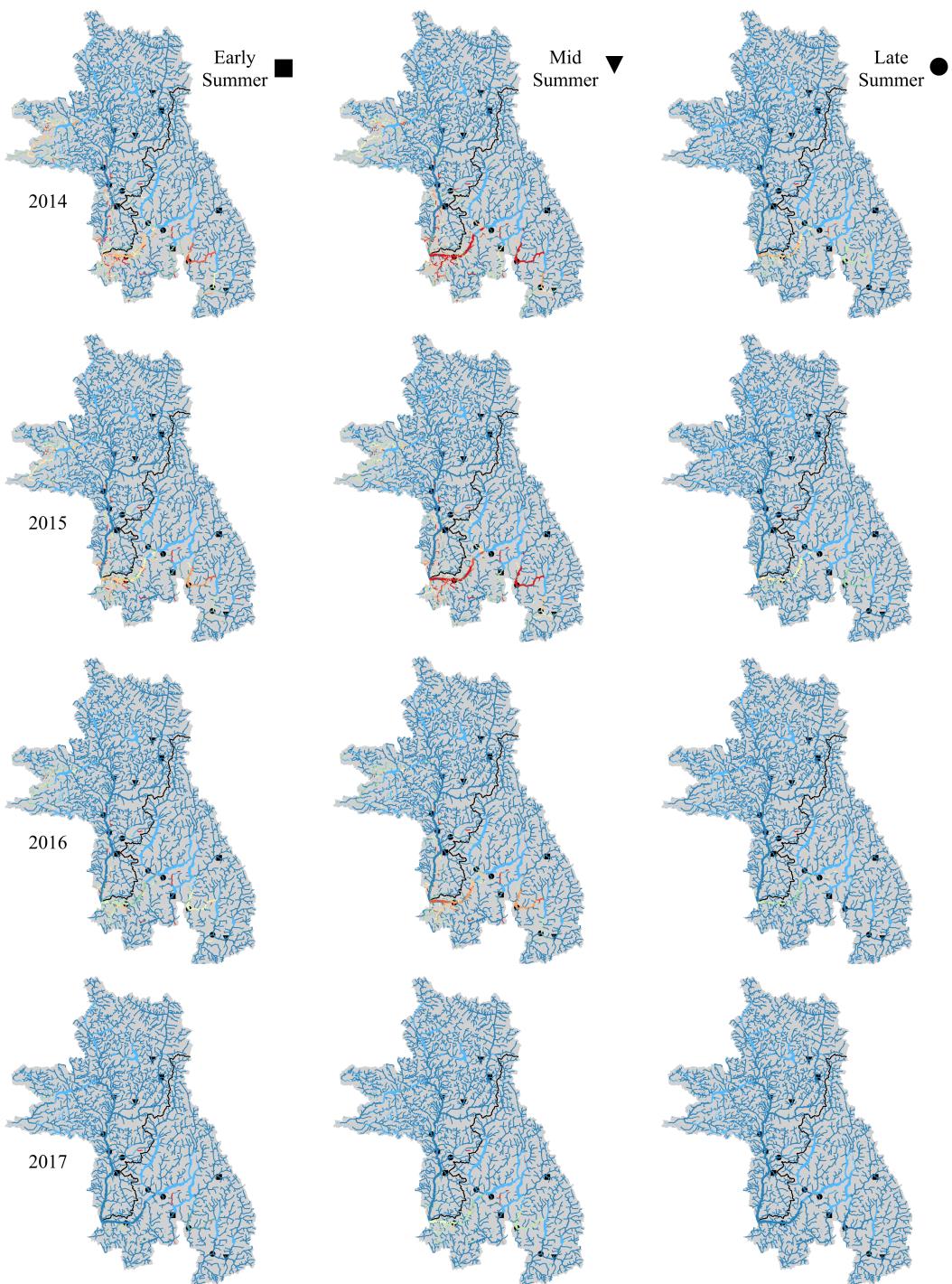


Figure 5.7: Probability (0 = blue, 1 = red) of thermal exposure above 19°C in the North and South Thompson River watersheds. Median migratory period estimates projected during early (June), mid (July) and late (August) summer between 2014 and 2017. Points indicate spawning locations of Chinook populations in the watersheds and their shape indicates their migration timing through the Albion Test Fishery near the mouth of the Fraser River.

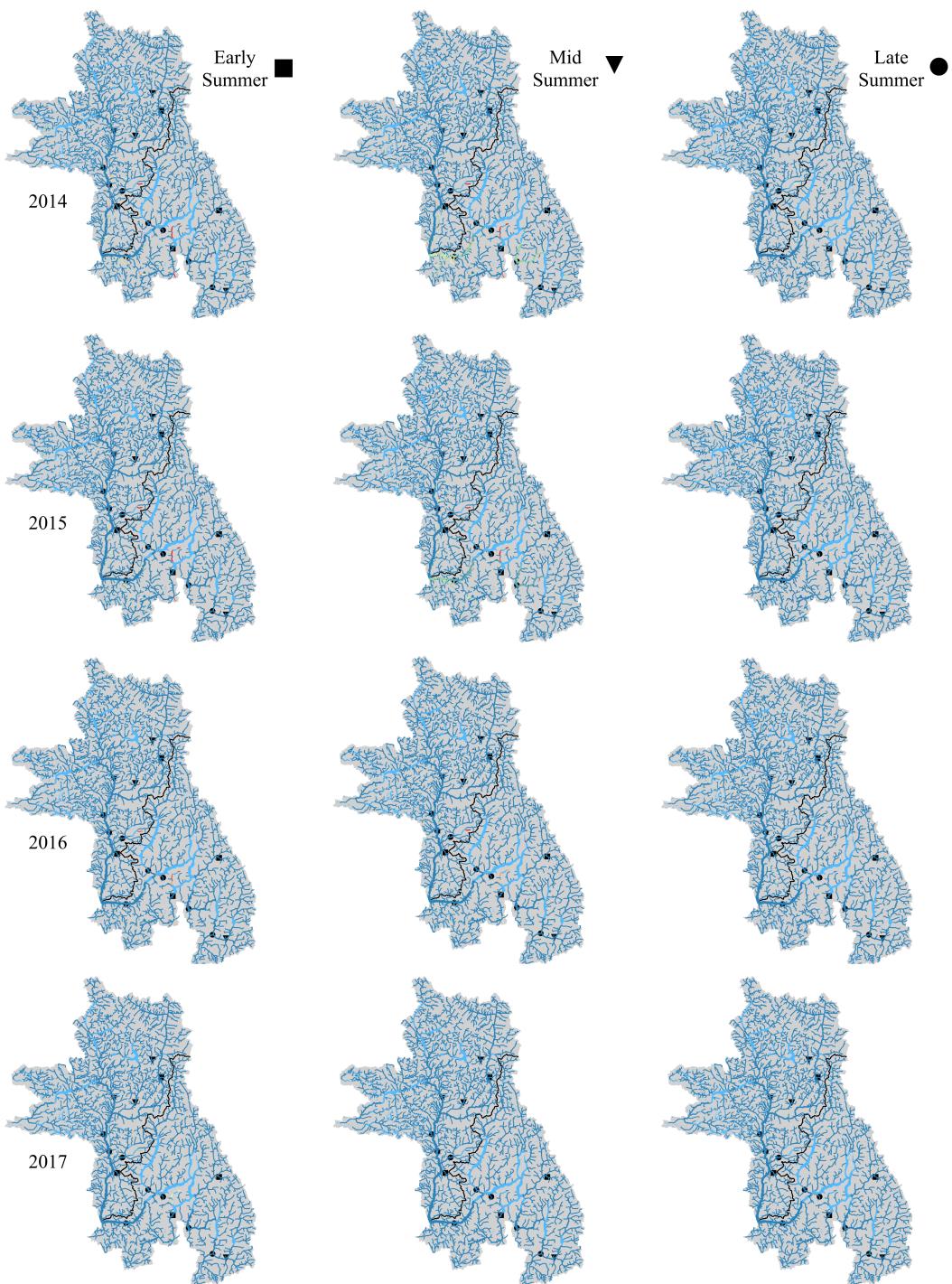


Figure 5.8: Probability (0 = blue, 1 = red) of thermal exposure above 22°C in the North and South Thompson River watersheds. Median migratory period estimates projected during early (June), mid (July) and late (August) summer between 2014 and 2017. Points indicate spawning locations of Chinook populations in the watersheds and their shape indicates their migration timing through the Albion Test Fishery near the mouth of the Fraser River.

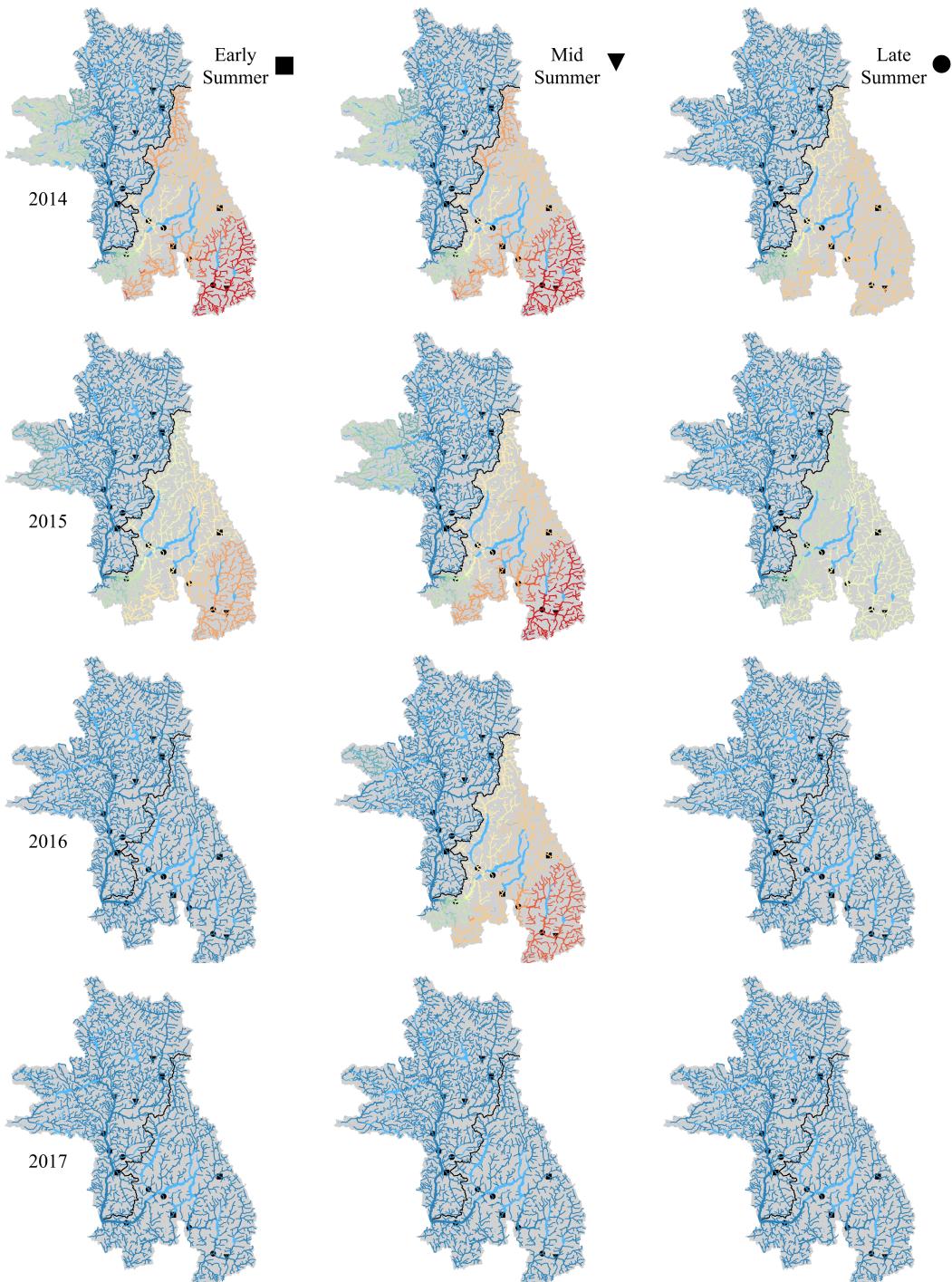


Figure 5.9: The median cumulative number of likely ( $>50\%$  probability) thermal exposures above  $19^{\circ}\text{C}$  during migration up the North and South Thompson River watersheds. Values are scaled between 0 (blue) and 1 (red) by the largest value across migratory periods and years ( $\max n = 26,235$ ). Migratory period estimates projected during early (June), mid (July) and late (August) summer between 2014 and 2017. Points indicate spawning locations of Chinook populations in the watersheds and their shape indicates their migration timing through the Albion Test Fishery near the mouth of the Fraser River.

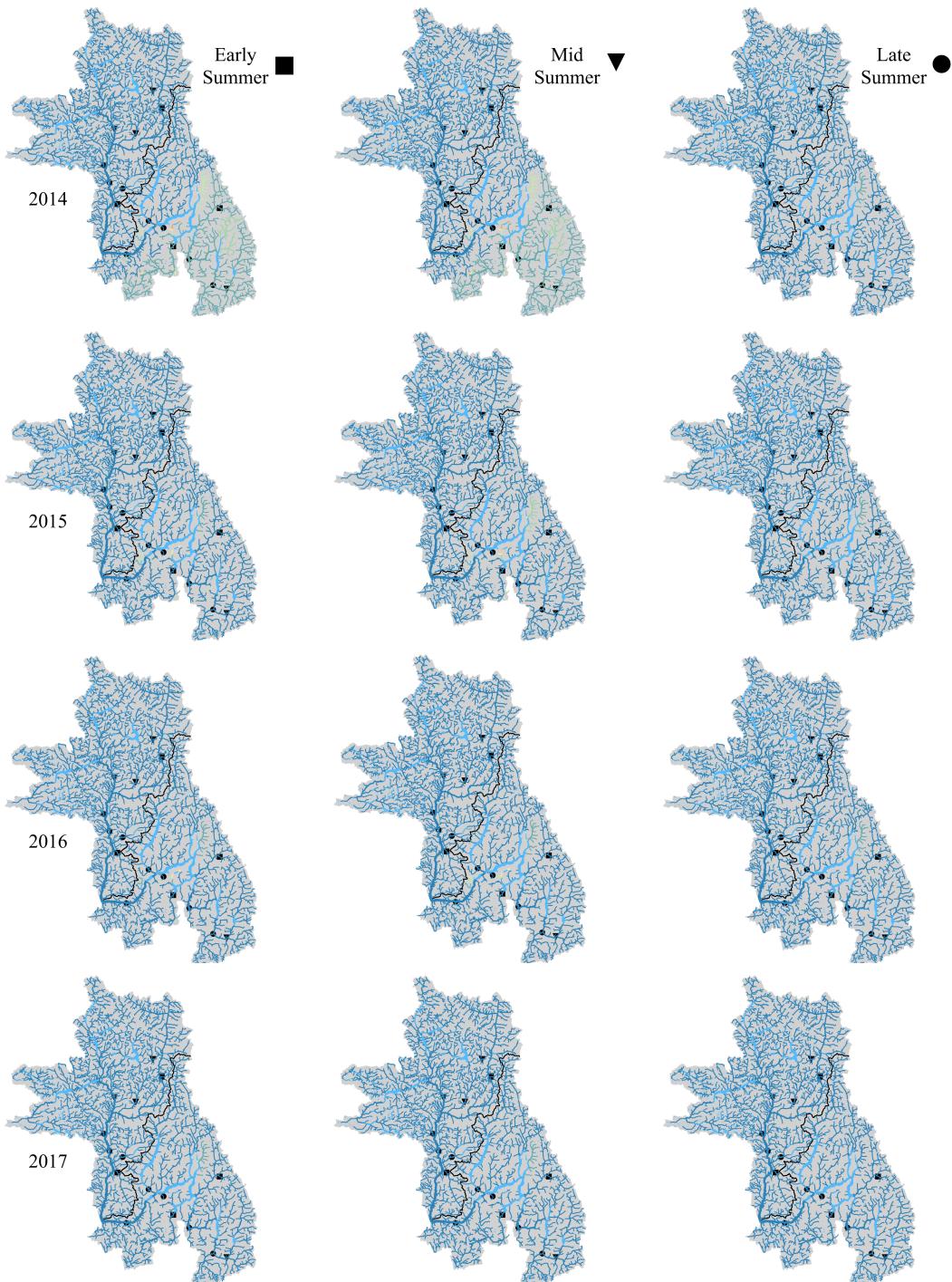


Figure 5.10: The median cumulative number of likely ( $>50\%$  probability) thermal exposures above  $22^{\circ}\text{C}$  during migration up the North and South Thompson River watersheds. Values are scaled between 0 (blue) and 1 (red) by the largest value across migratory periods and years ( $\max n = 1,271$ ). Migratory period estimates projected during early (June), mid (July) and late (August) summer between 2014 and 2017. Points indicate spawning locations of Chinook populations in the watersheds and their shape indicates their migration timing through the Albion Test Fishery near the mouth of the Fraser River.

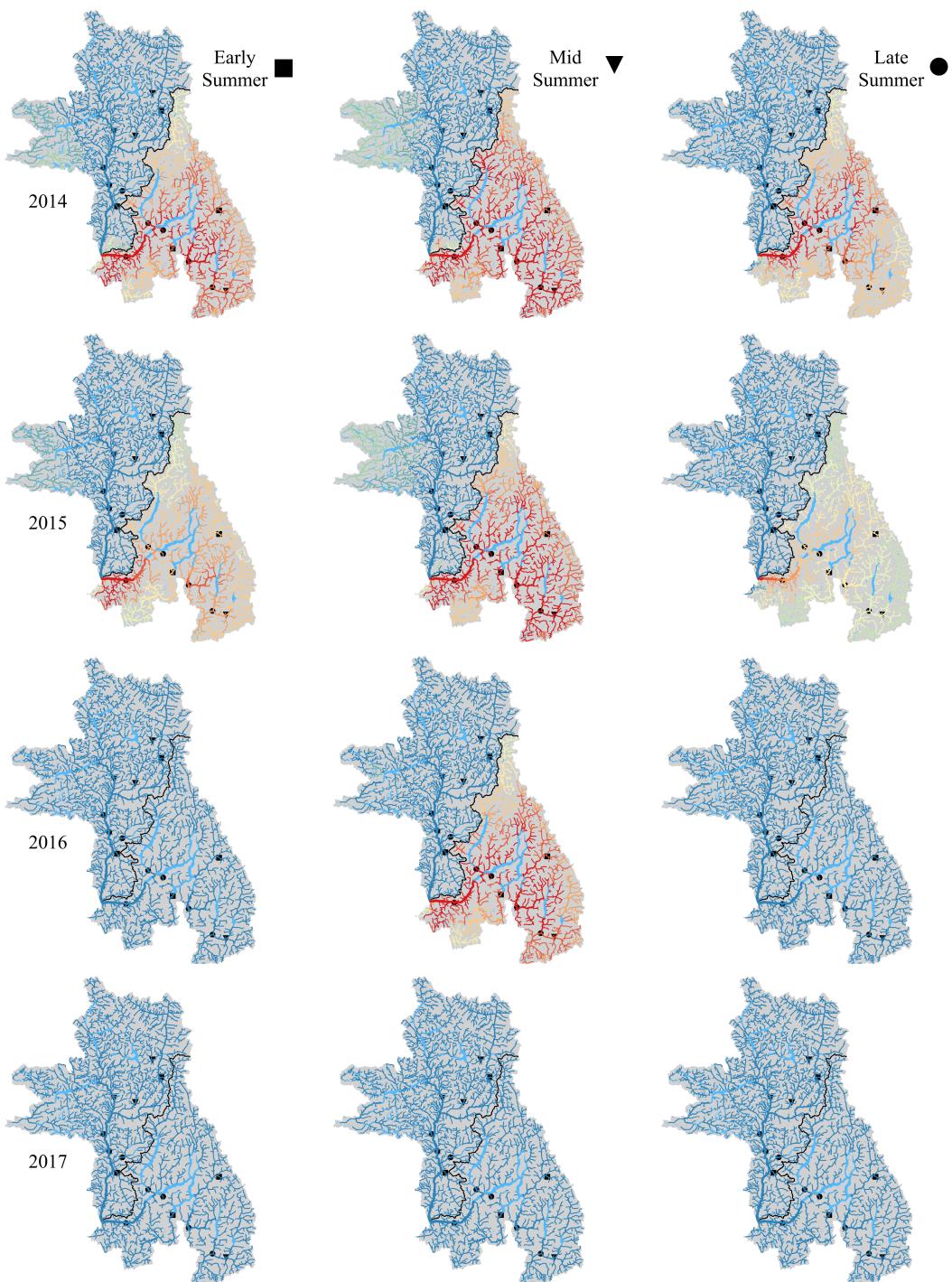


Figure 5.11: The average migratory thermal exposure probability (0 = blue, 1 = red) above 19°C during upstream migration in the North and South Thompson River watersheds. Run-timing estimates projected during early (June), mid (July) and late (August) summer between 2014 and 2017. Points indicate spawning locations of Chinook populations in the watersheds and their shape indicates their migration timing through the Albion Test Fishery near the mouth of the Fraser River.

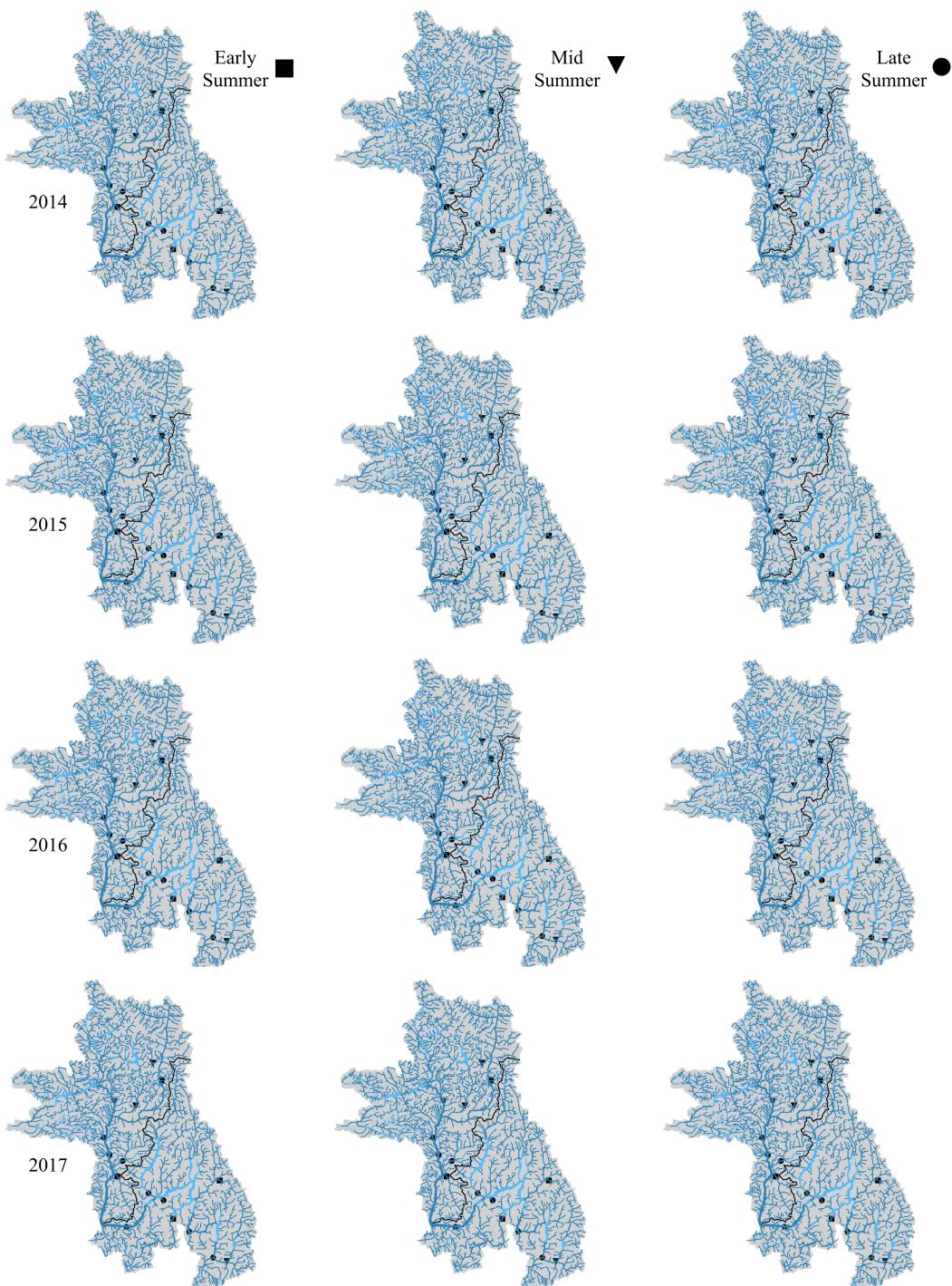


Figure 5.12: The average migratory thermal exposure probability (0 = blue, 1 = red) above 22°C during upstream migration in the North and South Thompson River watersheds. Run-timing estimates projected during early (June), mid (July) and late (August) summer between 2014 and 2017. Points indicate spawning locations of Chinook populations in the watersheds and their shape indicates their migration timing through the Albion Test Fishery near the mouth of the Fraser River.

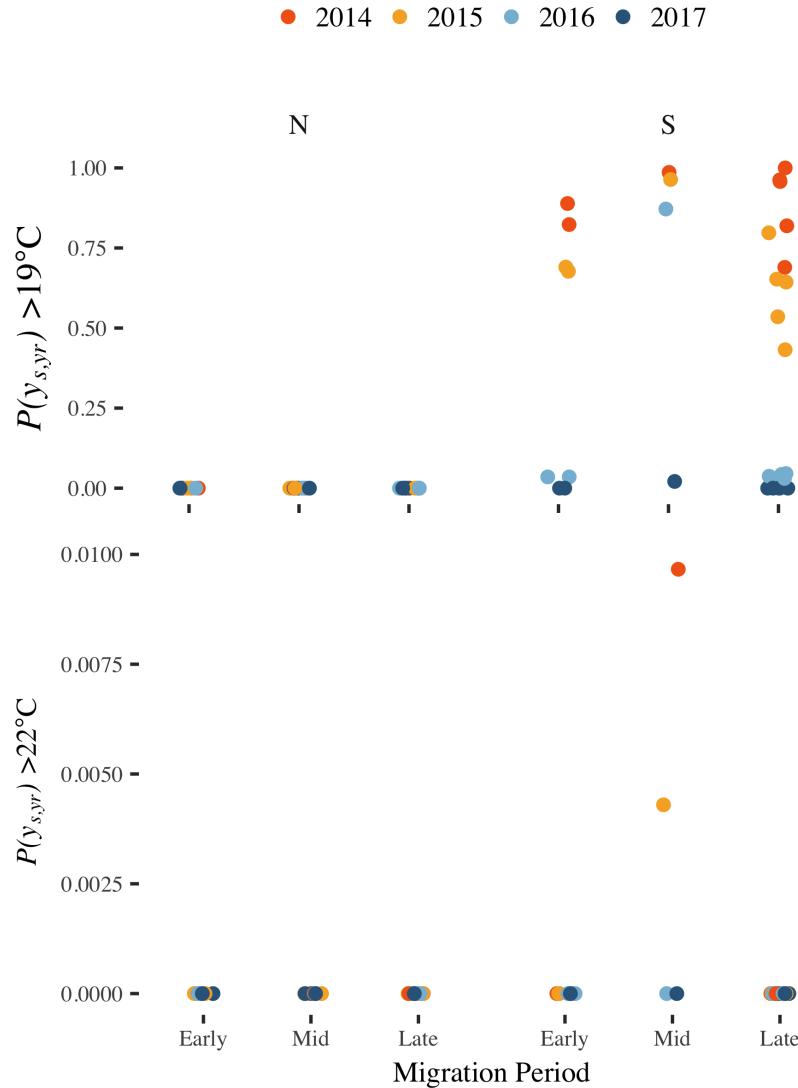


Figure 5.13: The average migratory thermal exposure probability of experiencing above threshold temperatures for each Chinook population within each run-timing period in the North (N) and South (S) Thompson River watersheds between 2014 and 2017. Migration period represents the run-timing of Chinook populations through the Albion Test Fishery in the lower Fraser River in early (June), mid (July) and late (August) summer. Points are horizontally jittered to improve visibility with 2(N), 3(N), 3(N), 2(S), 1(S) and 5(S) populations in the early, mid and late migration respectively.

# Chapter 6

## General discussion

River networks are complex and adaptive systems, continuously communicating with their watersheds topographic, geomorphic and climatic character. Structured connectivity underlies the conversation and organizes the on-network dynamics and evolution of the system. Although the physical formation of river networks has received significant theoretical and empirical treatment (e.g., Horton, 1945; Leopold, 1994; Rice et al., 2008), only recently have network dynamics been extended to the biotic sphere (e.g., Benda et al., 2004; Campbell Grant et al., 2007; Rodriguez-Iturbe et al., 2009).

The primary aim of my thesis was to understand the implications of river network mediated landscape and climatic interactions on riverine biota. Specifically, chapter two discovered that river networks dampen the local hydrologic effects of climate change by integrating diverse climate portfolios. Chapter three found that juvenile pink salmon (*Oncorhynchus gorbuscha*) who migrate further on the network from their natal stream to the ocean, are more likely to miss the zooplankton bloom in the estuary. Chapter four developed an unsupervised machine-learning tool that facilitates the cleaning of increasingly abundant stream temperature data. Finally, chapter five analyzed a large stream-temperature data-set to reveal differential thermal exposure probabilities among adult Chinook salmon (*O. tshawytscha*) migratory routes. Exposure probability was found to depend on migration timing and interacting climate, landscape and network dynamics. Collectively, I am hopeful that this thesis contributes positively to the: 1. general conceptual understanding of climate change in watersheds, 2. advancement of analytical approaches in river temperature research, and 3. specific understanding of the dynamics of my focal system – the Fraser River – and its culturally-, economically-, and ecologically-important salmon.

In the remainder of this general discussion, I raise several broad conceptual ideas and discuss their potential future directions.

## 6.1 Integrating the climate mosaic

Pervasive throughout my thesis is the concept that watersheds consist of a climate mosaic which mediated by the landscape becomes manifest in water and integrated by the network. The dendritic architecture that links and integrates different climatic regions underpins the river networks' emergent stability. More than a simple averaging of variability (Doak et al., 1998), river networks dynamically adjust to pervasive climate patterns (Seybold et al., 2017) to maximize downstream energy dampening (Rinaldo et al., 2014) leading to increasingly stable downstream flow (Asano & Uchida, 2010; Sanford et al., 2007) and temperature dynamics (Moore et al., 2015). The duality maintained by the network of dynamic-stability in these two ecological master variables establishes the conditions that promote portfolio processes that dampen meta-population dynamics (Carrara et al., 2012; Yeakel et al., 2014; Terui et al., 2018). Thus, from the physical processes that develop out of pairing network structure with climatic heterogeneity emerge biotic stability processes. Indeed, the climate mosaic is implicated here in chapters two, three and five, driving 1. downstream hydrologic trend stability, 2. life-history diversity that organizes phenological match-mismatch potential, and 3. migratory thermal exposure potential, respectively. Segregating and organizing fundamental biological processes (e.g., Fullerton et al., 2017), river networks clearly play a central role in understanding how species and populations will fare under future climate conditions, landuse practices and water management regimes. Furthering our understanding of how networks stitch together the climate tapestry at increasingly resolved scales deserves further attention as it will inform conservation strategies (Lowe et al., 2006).

## 6.2 Considering shifting interactions

The physical nature of many landscape features are expected to remain relatively constant on human time-scales, but their fundamental interactions with the climate may change. For instance, precipitation patterns in temperate regions have been and are expected to continue to shift from snow to rain dominate regimes as winter temperatures warm (Melillo et al., 2014; Kang et al., 2016). A loss of winter snow pack (Kang et al., 2014) and glacier mass (Huss & Hock, 2018) is a direct loss of water storage that leads to reduced stream discharge over the summer (Moore et al., 2009). In some regions, the melting cyrosphere may initially cool stream temperatures through temporarily increased flows but eventually these reserves will dry up leading to longer surface water residence times, lower heating capacity and subsequently warmer water temperatures (Caissie, 2006). Under such a radical shift in climate, watershed geomorphic features that once were indicative of cooler water temperatures may change due to shifting interaction strengths with climate and other landscape variables (Ward, 1985; Griffiths et al., 2014). These interactions are complex and depend on spatial and temporal relationships of network location, relative size, and changing seasonal dynamics such that a broad and temporally static understanding of landscape effects may

be insufficient at the local scale (Steel et al., 2016). Looking forward it will be important to assess and account for temporal shifts and their effects on biota, especially as transitions could occur abruptly, leading to alternate state dynamics (Scheffer et al., 2001).

### 6.3 Shifting climate portfolios

Downstream dampening dynamics may be vulnerable to climate change. Although chapter two demonstrates flow trend diversity dampens downstream flow trends and inter-annual variability, upstream trend diversity may be a function of landscape dynamics that will not persist indefinitely. As discussed, many flow trends are likely derived from the melting cyrosphere (Stahl & Moore, 2006), a reservoir of diversity that is rapidly depleting in some temperate watersheds (Huss & Hock, 2018). The decline of snow dominant flow regimes could lead to increased basin-wide homogeneity resulting in more dramatic downstream trends and possibly greater variability (Moore et al., 2009; Shrestha et al., 2017). Currently it's unclear how basin shape, topographic composition, landuse, flow management and climate combine to contribute to flow trend diversity. Quantifying the climate mediating landscape attributes that drive the observed flow trends in chapter two may explain variation in climate buffering capacity among sub-basins. This in turn may reveal basin compositions vulnerable to a loss of climate buffering capacity under future climate scenarios. Applying these findings to other river systems we might explore changing river network buffering dynamics more widely and discover potential climate tipping points where rivers become suddenly more dynamic due a loss of climate portfolio diversity.

### 6.4 Towards linking local landscape features and biology

Despite a general understanding of climate and landscape interactions that drive the climate mosaic, spatially and temporally refined understandings still stand to improve future predictions of change and their impacts on biota. For instance, Lisi et al. (2013) demonstrated spawn timing in sockeye salmon (*O. nerka*) to be closely linked to average summer stream temperatures, a habitat characteristic shown to be well predicted by watershed size, steepness and lake area. I discovered similar findings in my fifth chapter where larger lakes at lower elevation had an exceptional warming effect on stream temperatures. Adult salmon compelled through these migration corridors had much higher thermal exposure potential than those whose upstream migration was primarily free-flowing and lake limited. In both cases, landscape heterogeneity and climate interact to produce fine-scale habitat variation that differentially act on populations at the local level. This has implications for management as ecosystems and resources are often managed at a high level. For instance, salmon are managed as stocks composed of many populations spread throughout a watershed. It's becoming increasingly apparent that local level processes in aggregate will dictate the higher level dynamics at which we currently manage. As such continued efforts to refine the linkages

between the climate and landscape to varied life-histories will provide necessary insights to the long-term conservation and preservation of ecosystems.

## 6.5 River network structured life-history risk

Risks to aquatic organisms may be structured by river networks. For instance, in my third chapter I discovered the potential for juvenile pink salmon to miss the zooplankton bloom in the estuary depended on their emergence location on the network. Those that emerged further upriver, where the climate was less similar to that of the estuary, found it more difficult to precisely track inter-annual variation in zooplankton bloom timing. My fifth chapter also grappled with the idea of structured risk potential by demonstrating diverse yet organized thermal exposure opportunity, impacting some Chinook salmon population more than others. These findings suggest the risk potential of changing climate or landuse may be spatially and temporally explicit, disproportionately affecting some populations.

If the spatial organization of risk potential is coincident with the same forces that shape diversity, some life-histories may be disproportionately affected by perturbation. For instance, my second chapter demonstrates diverse flow trends in the Fraser River watershed indicative of shifting precipitation regimes. Such changes to climate and flow have already been shown to be negatively affecting some life-history types more than others. For example, Beechie et al. (2006) examined the spatial segregation of stream- and ocean-type Chinook populations in the Pacific Northwest Puget Sound region and demonstrated stream-type disproportionately utilize snowmelt-dominated habitat which as previously discussed is declining in prevalence in many regions (Bindoff et al., 2013; Melillo et al., 2014). Salmon populations throughout the Pacific rim have exhibited similar geographic life-history segregation (Beacham et al., 2003; Crozier & Zabel, 2006; Eliason et al., 2011; Rogers & Schindler, 2008; Carr-Harris et al., 2018), suggesting certain life-histories across species could be imperiled.

Life-history diversity is widely understood to confer stability and resilience to river network ecosystems through asynchrony dynamics (e.g., Schindler et al., 2010; Terui et al., 2018). As some life-histories may be more threatened than others, diversity within river system meta-populations may be declining, leading to greater response synchrony and risk potential (Moore et al., 2010). The river network analysis framework should facilitate the identification of such systematic loss of diversity and the potential impacts to the stability dynamics of meta-populations as it may highlight patterns of organization not apparent in traditional spatial analysis.

## 6.6 Monitoring at scale

Recent statistical model developments incorporating river network spatial dynamics has greatly improved predictions of essential habitat characteristics on the river network (Pe-

terson et al., 2013). In this thesis, I used a statistical stream network model (SSNM) (Hoef & Peterson, 2010) to directionally integrate landscape and climate covariates within the network and uniquely predict life-history and temporal stream temperature dynamics on the network. Despite structural improvements in river network spatial statistics, significant variability remained in these models as many stream habitat dynamics happen at highly resolved spatial and temporal scales. I acknowledge that my thesis does not focus on local level variation in river temperatures that could potentially be important to the biology of these systems. Improving the precision and accuracy of predictions at the local level requires highly resolved monitoring efforts. For instance, the concern that increasingly common and persistent extreme water temperatures may be negatively impacting stream biota depends on the opportunity for thermal refuge. Contributions of hyporheic flow and ground-water inputs that create cool pockets of stream habitat are only observable at fine-grain spatial resolutions (Fullerton et al., 2018). Monitoring at spatial and temporal resolutions that capture essential river network dynamics will greatly improve model assessment and development, facilitate fundamental knowledge and help assess present habitat conditions while probabilistically informing future expectations under varied climate and landscape conditions (Steel et al., 2017).

## 6.7 In conclusion

I hope that my thesis has contributed to thinking about large rivers systems in this era of climate change. As global temperatures rise and precipitation patterns shift, stream temperatures and flow dynamics that underpin lotic ecosystems will change in non-uniform ways over space and time. Only by considering the network structure are we to understand how such spatial heterogeneity will be organized and how organisms on different branches of the network might respond. In considering the broad-scale dynamics over the network, drivers of local dynamics might become more clear and inform management approaches that are more widely applicable to the entire system. River systems are naturally robust and resilient, taking advantage of these attributes offers a path for meeting the challenge of climate change.

# Bibliography

- Allen, S. E., & Wolfe, M. (2013). Hindcast of the timing of the spring phytoplankton bloom in the Strait of Georgia, 1968–2010. *Progress in Oceanography*, 115, 6–13.
- Anderson, S. C., Moore, J. W., McClure, M. M., Dulvy, N. K., & Cooper, A. B. (2015). Portfolio conservation of metapopulations under climate change. *Ecological Applications*, 25(2), 559–572.
- Angilletta, M. J. (2009). *Thermal adaptation: a theoretical and empirical synthesis*. Oxford University Press.
- Anttila, K., Farrell, A. P., Patterson, D. A., Hinch, S. G., & Eliason, E. J. (2018). Cardiac SERCA activity in sockeye salmon populations: an adaptive response to migration conditions. *Canadian Journal of Fisheries and Aquatic Sciences*, 999, 1–5.
- Arendt, A., Bliss, A., Bolch, T., Cogley, J., Gardner, A., Hagen, J., Hock, R., Huss, M., Kaser, G., Kienholz, C., et al. (2015). Randolph glacier inventory—a dataset of global glacier outlines: Version 5.0. *Global Land Ice Measurements from Space, Boulder Colorado, USA*.
- Arkema, K. K., Guannel, G., Verutes, G., Wood, S. A., Guerry, A., Ruckelshaus, M., Kareiva, P., Lacayo, M., & Silver, J. M. (2013). Coastal habitats shield people and property from sea-level rise and storms. *Nature Climate Change*, 3(10), 913–918.
- Asano, Y., & Uchida, T. (2010). Is representative elementary area defined by a simple mixing of variable small streams in headwater catchments? *Hydrological Processes: An International Journal*, 24(5), 666–671.
- Asner, G. P., Martin, R. E., Knapp, D. E., Tupayachi, R., Anderson, C. B., Sinca, F., Vaughn, N. R., & Llactayo, W. (2017). Airborne laser-guided imaging spectroscopy to map forest trait diversity and guide conservation. *Science*, 355(6323), 385–389.
- Bailey, R. E., & Canadian Science Advisory Secretariat (2001). *Summary of stock assessment information for selected early returning Chinook salmon populations of the Fraser River watershed*. Canadian Science Advisory Secretariat.
- Baker, R., et al. (1978). *Evolutionary ecology of animal migration*. Holmes & Meier Publishers.
- Barbosa, J. M., Asner, G. P., Hughes, R. F., & Johnson, M. T. (2017). Landscape-scale GPP and carbon density inform patterns and impacts of an invasive tree across wet forests of Hawaii. *Ecological Applications*, 27(2), 403–415.

- Beacham, T. D. (1988). A genetic analysis of early development in pink (*Oncorhynchus gorbuscha*) and chum salmon (*Oncorhynchus keta*) at three different temperatures. *Genome*, 30(1), 89–96.
- Beacham, T. D., & Murray, C. B. (1986). Comparative developmental biology of pink salmon, *Oncorhynchus gorbuscha*, in southern British Columbia. *Journal of Fish Biology*, 28(2), 233–246.
- Beacham, T. D., & Murray, C. B. (1987). Effects of transferring pink (*Oncorhynchus gorbuscha*) and chum (*Oncorhynchus keta*) salmon embryos at different developmental stages to a low incubation temperature. *Canadian Journal of Zoology*, 65(1), 96–105.
- Beacham, T. D., & Murray, C. B. (1988). Variation in developmental biology of pink salmon (*Oncorhynchus gorbuscha*) in British Columbia. *Canadian Journal of Zoology*, 66(12), 2634–2648.
- Beacham, T. D., Supernault, K. J., Wetklo, M., Deagle, B., Labaree, K., Irvine, J. R., Candy, J. R., Miller, K. M., Nelson, R. J., & Withler, R. E. (2003). The geographic basis for population structure in Fraser River Chinook salmon (*Oncorhynchus tshawytscha*). *Fishery Bulletin*, 101(2), 229–243.
- Beamish, R., & Mahnken, C. (2001). A critical size and period hypothesis to explain natural regulation of salmon abundance and the linkage to climate and climate change. *Progress in Oceanography*, 49(1), 423 – 437. Pacific climate variability and marine ecosystem impacts.
- Beechie, T., Buhle, E., Ruckelshaus, M., Fullerton, A., & Holsinger, L. (2006). Hydrologic regime and the conservation of salmon life history diversity. *Biological Conservation*, 130(4), 560–572.
- Benda, L., Poff, N. L., Miller, D., Dunne, T., Reeves, G., Pess, G., & Pollock, M. (2004). The network dynamics hypothesis: how channel networks structure riverine habitats. *BioScience*, 54(5), 413–427.
- Bindoff, N. L., Stott, P. A., AchutaRao, M., Allen, M. R., Gillett, N., Gutzler, D., Hansingo, K., Hegerl, G., Hu, Y., Jain, S., Mokhov, I., Overland, J., Perlitz, J., Sebbari, R., & Zhang, X. (2013). *Detection and attribution of climate change: From global to regional*. In: *Climate Change 2013: The Physical Science basis. Contribution of Working group I to Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press.
- Both, C., Bouwhuis, S., Lessells, C., & Visser, M. E. (2006). Climate change and population declines in a long-distance migratory bird. *Nature*, 441(7089), 81–83.
- Boyce, D. G., Lewis, M. R., & Worm, B. (2010). Global phytoplankton decline over the past century. *Nature*, 466(7306), 591.
- Bradford, M. J., & Irvine, J. R. (2000). Land use, fishing, climate change, and the decline of Thompson River, British Columbia, coho salmon. *Canadian Journal of Fisheries and Aquatic Sciences*, 57(1), 13–16.

- Brannon, E. L. (1987). Mechanisms stabilizing salmonid fry emergence timing. *Canadian Special Publication of Fisheries and Aquatic Sciences*, 96, 120–124.
- Brett, J. (1969). Temperature and fish. *Chesapeake Science*, 10(3-4), 275–276.
- Brown, L. E., & Hannah, D. M. (2008). Spatial heterogeneity of water temperature across an alpine river basin. *Hydrological Processes: An International Journal*, 22(7), 954–967.
- Burgess, M. D., Smith, K. W., Evans, K. L., Leech, D., Pearce-Higgins, J. W., Branston, C. J., Briggs, K., Clark, J. R., du Feu, C. R., Lewthwaite, K., et al. (2018). Tritrophic phenological match–mismatch in space and time. *Nature ecology & evolution*, 2(6), 970–975.
- Cadbury, S., Hannah, D., Milner, A., Pearson, C., & Brown, L. (2008). Stream temperature dynamics within a New Zealand glacierized river basin. *River Research and Applications*, 24(1), 68–89.
- Caissie, D. (2006). The thermal regime of rivers: a review. *Freshwater biology*, 51(8), 1389–1406.
- Campbell Grant, E. H., Lowe, W. H., & Fagan, W. F. (2007). Living in the branches: population dynamics and ecological processes in dendritic networks. *Ecology letters*, 10(2), 165–175.
- Carr-Harris, C. N., Moore, J. W., Gottesfeld, A. S., Gordon, J. A., Shepert, W. M., Henry Jr, J. D., Russell, H. J., Helin, W. N., Doolan, D. J., & Beacham, T. D. (2018). Phenological diversity of salmon smolt migration timing within a large watershed. *Transactions of the American Fisheries Society*, 147(5), 775–790.
- Carrara, F., Altermatt, F., Rodriguez-Iturbe, I., & Rinaldo, A. (2012). Dendritic connectivity controls biodiversity patterns in experimental metacommunities. *Proceedings of the National Academy of Sciences*.
- Caughlin, T. T., Rifai, S. W., Graves, S. J., Asner, G. P., & Bohlman, S. A. (2016). Integrating lidar-derived tree height and landsat satellite reflectance to estimate forest regrowth in a tropical agricultural landscape. *Remote Sensing in Ecology and Conservation*, 2(4), 190–203.
- Chezik, K. A., Anderson, S. C., & Moore, J. W. (2017). River networks dampen long-term hydrological signals of climate change. *Geophysical Research Letters*, 44(14), 7256–7264.
- Chezik, K. A., Lester, N. P., & Venturelli, P. A. (2014). Fish growth and degree-days I: selecting a base temperature for a within-population study. *Canadian Journal of Fisheries and Aquatic Sciences*, 71(1), 47–55.
- Chittenden, C. M., Jensen, J. L., Ewart, D., Anderson, S., Balfry, S., Downey, E., Eaves, A., Saksida, S., Smith, B., Vincent, S., et al. (2010). Recent salmon declines: a result of lost feeding opportunities due to bad timing? *PLoS One*, 5(8).
- Comte, L., & Olden, J. D. (2017). Climatic vulnerability of the world's freshwater and marine fishes. *Nature Climate Change*, 7(10), 718.

- Cooke, S. J., Hinch, S. G., Farrell, A. P., Lapointe, M. F., Jones, S. R., Macdonald, J. S., Patterson, D. A., Healey, M. C., & Van Der Kraak, G. (2004). Abnormal migration timing and high en route mortality of sockeye salmon in the Fraser River, British Columbia. *Fisheries*, 29(2), 22–33.
- Crossin, G. T., Hinch, S., Cooke, S., Welch, D., Patterson, D., Jones, S., Lotto, A., Leggatt, R., Mathes, M., Shrimpton, J., et al. (2008). Exposure to high temperature influences the behaviour, physiology, and survival of sockeye salmon during spawning migration. *Canadian Journal of Zoology*, 86(2), 127–140.
- Crozier, L. G. (2015). Impacts of climate change on salmon of the Pacific Northwest. *Technical Review, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Seattle, Washington*, (p. 46).
- Crozier, L. G., Hendry, A., Lawson, P. W., Quinn, T., Mantua, N., Battin, J., Shaw, R., & Huey, R. (2008). Potential responses to climate change in organisms with complex life histories: evolution and plasticity in Pacific salmon. *Evolutionary Applications*, 1(2), 252–270.
- Crozier, L. G., & Zabel, R. W. (2006). Climate impacts at multiple scales: evidence for differential population responses in juvenile Chinook salmon. *Journal of Animal Ecology*, 75(5), 1100–1109.
- Cushing, D. (1990). Plankton production and year-class strength in fish populations: an update of the match/mismatch hypothesis. *Advances in Marine Biology*, 26, 249–293.
- Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., Curtis, J., & Pasteris, P. P. (2008). Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *International Journal of Climatology: a Journal of the Royal Meteorological Society*, 28(15), 2031–2064.
- Damiano, L., Peterson, B., & Michael, W. (2017). *A Tutorial on Hidden Markov Models using Stan*. A Case Study submitted as a candidate for a contributed talk in StanCon 2018.  
URL <https://github.com/luisdamiano/stancon18>
- Deacy, W. W., Armstrong, J. B., Leacock, W. B., Robbins, C. T., Gustine, D. D., Ward, E. J., Erlenbach, J. A., & Stanford, J. A. (2017). Phenological synchronization disrupts trophic interactions between Kodiak brown bears and salmon. *Proceedings of the National Academy of Sciences*, 114(39), 10432–10437.  
URL <http://www.pnas.org/content/114/39/10432>
- Déry, S. J., Hernández-Henríquez, M. A., Owens, P. N., Parkes, M. W., & Petticrew, E. L. (2012). A century of hydrological variability and trends in the Fraser River basin. *Environmental Research Letters*, 7(2).
- Déry, S. J., Stahl, K., Moore, R., Whitfield, P., Menounos, B., & Burford, J. E. (2009). Detection of runoff timing changes in pluvial, nival, and glacial rivers of western Canada. *Water Resources Research*, 45(4).

- Doak, D., Bigger, D., Harding, E., Marvier, M., RE, O., & Thomson, D. (1998). The statistical inevitability of stability-diversity relationships in community ecology. *Am Nat*, *151*(3), 264–276.
- Domingos, P. (2015). *The master algorithm: How the quest for the ultimate learning machine will remake our world*. Basic Books.
- Donat, M. G., Lowry, A. L., Alexander, L. V., O’Gorman, P. A., & Maher, N. (2016). More extreme precipitation in the world’s dry and wet regions. *Nature Climate Change*, *6*(5), 508–513.
- Dorava, J. M., & Milner, A. M. (2000). Role of lake regulation on glacier-fed rivers in enhancing salmon productivity: the Cook Inlet watershed, south-central Alaska, USA. *Hydrological Processes*, *14*(16-17), 3149–3159.
- Dunmall, K. M., Mochnacz, N. J., Zimmerman, C. E., Lean, C., & Reist, J. D. (2016). Using thermal limits to assess establishment of fish dispersing to high-latitude and high-elevation watersheds. *Canadian Journal of Fisheries and Aquatic Sciences*, *73*(12), 1750–1758.
- Durant, J. M., Hjermann, D. Ø., Anker-Nilssen, T., Beaugrand, G., Mysterud, A., Pettorelli, N., & Stenseth, N. C. (2005). Timing and abundance as key mechanisms affecting trophic interactions in variable environments. *Ecology Letters*, *8*(9), 952–958.
- Durant, J. M., Hjermann, D. Ø., Ottersen, G., & Stenseth, N. C. (2007). Climate and the match or mismatch between predator requirements and resource availability. *Climate research*, *33*(3), 271–283.
- Eliason, E. J., Clark, T. D., Hague, M. J., Hanson, L. M., Gallagher, Z. S., Jeffries, K. M., Gale, M. K., Patterson, D. A., Hinch, S. G., & Farrell, A. P. (2011). Differences in thermal tolerance among sockeye salmon populations. *Science*, *332*(6025), 109–112.
- Fagan, W. F. (2002). Connectivity, fragmentation, and extinction risk in dendritic metapopulations. *Ecology*, *83*(12), 3243–3249.
- Ferrari, M. R., Miller, J. R., & Russell, G. L. (2007). Modeling changes in summer temperature of the Fraser River during the next century. *Journal of Hydrology*, *342*(3-4), 336–346.
- Ficklin, D. L., Abatzoglou, J. T., Robeson, S. M., Null, S. E., & Knouft, J. H. (2018). Natural and managed watersheds show similar responses to recent climate change. *Proceedings of the National Academy of Sciences*, *115*(34), 8553–8557.
- Filmon, G. (2003). *Firestorm 2003 Provincial Review*. British Columbia.
- Fortuna, M. A., Gómez-Rodríguez, C., & Bascompte, J. (2006). Spatial network structure and amphibian persistence in stochastic environments. *Proceedings of the Royal Society of London B: Biological Sciences*, *273*(1592), 1429–1434.
- Friedland, K. D., Mouw, C. B., Asch, R. G., Ferreira, A. S. A., Henson, S., Hyde, K. J., Morse, R. E., Thomas, A. C., & Brady, D. C. (2017). Phenology and time series trends of the dominant seasonal phytoplankton bloom across global scales. *Global Ecology and Biogeography*, *27*(5), 551–569.

- Fullerton, A., Torgersen, C., Lawler, J. J., Steel, A. E., Ebersole, J., & Lee, S. (2018). Longitudinal thermal heterogeneity in rivers and refugia for coldwater species: effects of scale and climate change. *Aquatic sciences*, 80(1), 3.
- Fullerton, A. H., Burke, B. J., Lawler, J. J., Torgersen, C. E., Ebersole, J. L., & Leibowitz, S. G. (2017). Simulated juvenile salmon growth and phenology respond to altered thermal regimes and stream network shape. *Ecosphere*, 8(12).
- Gelman, A., & Hill, J. (2008). *Data analysis using regression and multilevel hierarchical models*. Cambridge University Press.
- Gelman, A., Stern, H. S., Carlin, J. B., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. Chapman and Hall/CRC.
- Gilhouse, P. (1992). *Prespawning mortalities of sockeye salmon in the Fraser River system and possible causal factors*. International Pacific Salmon Fisheries Commission.
- Griffiths, J. R., Schindler, D. E., Ruggerone, G. T., & Bumgarner, J. D. (2014). Climate variation is filtered differently among lakes to influence growth of juvenile sockeye salmon in an Alaskan watershed. *Oikos*, 123(6), 687–698.
- Groot, C., & Margolis, L. (1991). *Pacific salmon life histories*. UBC Press.
- Guerry, A. D., Polasky, S., Lubchenco, J., Rebecca, C., Daily, G. C., Griffin, R., Ruckelshaus, M., Bateman, I. J., Duraiappah, A., Elmquist, T., Feldman, M. W., Folke, C., Hoekstra, J., Kareiva, P. M., Keeler, B. L., Li, S., Emily, M., Ouyang, Z., Reyers, B., Ricketts, T. H., Rockström, J., Tallis, H., & Vira, B. (2015). Natural capital and ecosystem services informing decisions: From promise to practice. *Proc Natl Acad Sci*, 112(24), 7348–7355.
- Habib, A., Bang, K. I., Kersting, A. P., & Lee, D.-C. (2009). Error budget of LiDAR systems and quality control of the derived data. *Photogrammetric Engineering & Remote Sensing*, 75(9), 1093–1108.
- Hague, M., Ferrari, M., Miller, J., Patterson, D., Russell, G., Farrell, A., & Hinch, S. (2011). Modelling the future hydroclimatology of the lower Fraser River and its impacts on the spawning migration survival of sockeye salmon. *Global Change Biology*, 17(1), 87–98.
- Hamilton, J. D. (2016). *Regime Switching Models*, (pp. 1–7). London: Palgrave Macmillan UK.
- Hampton, S., Strasser, C., & Tewksbury, J. (2013). Big data and the future of ecology. *Frontiers in Ecology and the Environment*, 11(3), 156–162.
- Hartmann, D., Klein Tank, A., Rusticucci, M., Alexander, L., Brönnimann, S., Charabi, Y., Dentener, F., Dlugokencky, E., Easterling, D., Kaplan, A., Soden, B., Thorne, P., Wild, M., & Zhai, P. (2013). *Observations: Atmosphere and Surface*. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press.

- Healey, M., & Bradford, M. (2011). The cumulative impacts of climate change on Fraser River sockeye salmon (*Oncorhynchus nerka*) and implications for management. *Canadian Journal of Fisheries and Aquatic Sciences*, 68(4), 718–737.
- Hebert, K., Goddard, P., Smoker, W. W., & Gharrett, A. (1998). Quantitative genetic variation and genotype by environment interaction of embryo development rate in pink salmon (*Oncorhynchus gorbuscha*). *Canadian Journal of Fisheries and Aquatic Sciences*, 55(9), 2048–2057.
- Heyck-Williams, S., Anderson, L., & Stein, B. (2017). *Megafires: The Growing Risk to America's Forests, Communities, and Wildlife*. Washington, DC: National Wildlife Federation.
- Hinch, S., Cooke, S., Farrell, A., Miller, K., Lapointe, M., & Patterson, D. (2012). Dead fish swimming: a review of research on the early migration and high premature mortality in adult Fraser River sockeye salmon *Oncorhynchus nerka*. *Journal of Fish Biology*, 81(2), 576–599.
- Hirabayashi, Y., Mahendran, R., Koitala, S., Konoshima, L., Yamazaki, D., Watanabe, S., Kim, H., & Kanae, S. (2013). Global flood risk under climate change. *Nature Climate Change*, 3(9), 816–821.
- Hoef, J. V., Peterson, E., Clifford, D., & Shah, R. (2014). SSN: An R package for spatial statistical modeling on stream networks. *Journal of Statistical Software*, 56(3), 1–45.
- Hoef, J. V., & Peterson, E. E. (2010). A moving average approach for spatial statistical models of stream networks. *Journal of the American Statistical Association*, 105(489), 6–18.
- Holtby, L. B., & Ciruna, K. A. (2008). Conservation units for pacific salmon under the wild salmon policy. Tech. Rep. 2007/070, Department of Fisheries and Oceans, Ottawa, ON(Canada).
- Horton, R. E. (1945). Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology. *Geological Society of America bulletin*, 56(3), 275–370.
- Hulme, M. (2016). 1.5 °C and climate research after the Paris Agreement. *Nature Climate Change*, 6(3), 222–224.
- Huss, M., & Hock, R. (2018). Global-scale hydrological response to future glacier mass loss. *Nature Climate Change*, 8(2), 135–140.
- Isaak, D. J., & Horan, D. L. (2011). An evaluation of underwater epoxies to permanently install temperature sensors in mountain streams. *North American Journal of Fisheries Management*, 31(1), 134–137.
- Isaak, D. J., Luce, C. H., Rieman, B. E., Nagel, D. E., Peterson, E. E., Horan, D. L., Parkes, S., & Chandler, G. L. (2010). Effects of climate change and wildfire on stream temperatures and salmonid thermal habitat in a mountain river network. *Ecological Applications*, 20(5), 1350–1371.

- Isaak, D. J., Peterson, E. E., Ver Hoef, J. M., Wenger, S. J., Falke, J. A., Torgersen, C. E., Sowder, C., Steel, E. A., Fortin, M.-J., Jordan, C. E., Ruesch, A. S., Som, N., & Monestiez, P. (2014). Applications of spatial statistical network models to stream data. *Wiley Interdisciplinary Reviews: Water*, 1(3), 277–294.
- Isaak, D. J., Wenger, S. J., Peterson, E. E., Ver Hoef, J. M., Nagel, D. E., Luce, C. H., Hostetler, S. W., Dunham, J. B., Roper, B. B., Wollrab, S. P., Chandler, G. L., Horan, D. L., & Parkes-Payne, S. (2017). The NorWeST summer stream temperature model and scenarios for the Western U.S.: A crowd-sourced database and new geospatial tools foster a user community and predict broad climate warming of rivers and streams. *Water Resources Research*, 53(11), 9181–9205.
- Isaak, D. J., Young, M. K., Nagel, D. E., Horan, D. L., & Groce, M. C. (2015). The cold-water climate shield: delineating refugia for preserving salmonid fishes through the 21<sup>st</sup> century. *Global Change Biology*, 21(7), 2540–2553.
- Jolly, W. M., Cochrane, M. A., Freeborn, P. H., Holden, Z. A., Brown, T. J., Williamson, G. J., & Bowman, D. M. J. S. (2015). Climate-induced variations in global wildfire danger from 1979 to 2013. *Nature Communications*, 6(7537).
- Jones, H. P., Hole, D. G., & Zavaleta, E. S. (2012). Harnessing nature to help people adapt to climate change. *Nature Climate Change*, 2(7), 504–509.
- Jones, T., & Cresswell, W. (2010). The phenology mismatch hypothesis: are declines of migrant birds linked to uneven global climate change? *Journal of Animal Ecology*, 79(1), 98–108.
- Kang, D., Gao, H., Shi, X., ul Islam, S., & Déry, S. J. (2016). Impacts of a rapidly declining mountain snowpack on streamflow timing in Canada's Fraser River basin. *Scientific reports*, 6(19299).
- Kang, D., Shi, X., Gao, H., & Déry, S. J. (2014). On the changing contribution of snow to the hydrology of the Fraser River Basin, Canada. *Journal of Hydrometeorology*, 15(4), 1344–1365.
- Keefer, M. L., & Caudill, C. C. (2015). Estimating thermal exposure of adult summer steelhead and fall Chinook salmon migrating in a warm impounded river. *Ecology of Freshwater Fish*, 25(4), 599–611.
- Keefer, M. L., Clabough, T. S., Jepson, M. A., Johnson, E. L., Peery, C. A., & Caudill, C. C. (2018). Thermal exposure of adult Chinook salmon and steelhead: Diverse behavioral strategies in a large and warming river system. *PloS One*, 13(9), e0204274.
- Keefer, M. L., Clabough, T. S., Jepson, M. A., Naughton, G. P., Blubaugh, T. J., Joosten, D. C., & Caudill, C. C. (2015). Thermal exposure of adult Chinook salmon in the Willamette River basin. *Journal of Thermal Biology*, 48, 11–20.
- Keefer, M. L., Peery, C. A., & Heinrich, M. J. (2008). Temperature-mediated en route migration mortality and travel rates of endangered snake river sockeye salmon. *Ecology of Freshwater Fish*, 17(1), 136–145.

- Kingsolver, J. G. (2009). The well-temperatured biologist: (American Society of Naturalists presidential address). *The American Naturalist*, 174(6), 755–768.
- Kovach, R. P., Gharrett, A. J., & Tallmon, D. A. (2012). Genetic change for earlier migration timing in a pink salmon population. *Proceedings of the Royal Society of London B: Biological Sciences*, 279(1743), 3870–3878.
- Kovach, R. P., Joyce, J. E., Echave, J. D., Lindberg, M. S., & Tallmon, D. A. (2013). Earlier migration timing, decreasing phenotypic variation, and biocomplexity in multiple salmonid species. *PloS One*, 8(1), e53807.
- Labelle, M. (2009). Status of Pacific salmon resources in southern British Columbia and the Fraser River basin. *Pacific Fisheries Resource Conservation Council, Vancouver, BC*, 91p.
- Lawler, J. J., Wiersma, Y. F., & Huettmann, F. (2011). *Using species distribution models for conservation planning and ecological forecasting* In: *Predictive Species and Habitat Modeling in Landscape Ecology* [Ashton, D.C. and Wiersma, Y.F. and Huettmann, F. (eds.)]. Springer.
- Lee, A. C., & Lucas, R. M. (2007). A LiDAR-derived canopy density model for tree stem and crown mapping in Australian forests. *Remote Sensing of Environment*, 111(4), 493–518.
- Leopold, L. B. (1994). *A View of the River*. Harvard University Press.
- Letcher, B. H., Hocking, D. J., O’Neil, K., Whiteley, A. R., Nislow, K. H., & O’Donnell, M. J. (2016). A hierarchical model of daily stream temperature using air-water temperature synchronization, autocorrelation, and time lags. *PeerJ*, 4, e1727.
- Lindsay, J. (2016). Whitebox GAT: a case study in geomorphometric analysis. *Computer Geoscience*, 95, 75–84.
- Lisi, P. J., & Schindler, D. E. (2015). Wind-driven upwelling in lakes destabilizes thermal regimes of downstream rivers. *Limnology and Oceanography*, 60(1), 169–180.
- Lisi, P. J., Schindler, D. E., Bentley, K. T., & Pess, G. R. (2013). Association between geomorphic attributes of watersheds, water temperature, and salmon spawn timing in Alaskan streams. *Geomorphology*, 185, 78–86.
- Lisi, P. J., Schindler, D. E., Cline, T. J., Scheuerell, M. D., & Walsh, P. B. (2015). Watershed geomorphology and snowmelt control stream thermal sensitivity to air temperature. *Geophysical Research Letters*, 42(9), 3380–3388.
- Liu, J., Li, J., Li, W., & Wu, J. (2016). Rethinking big data: A review on the data quality and usage issues. *ISPRS Journal of Photogrammetry and Remote Sensing*, 115, 134 – 142.
- Louisiana Department of Natural Resources (1998). *Coast 2050: Toward a Sustainable Coastal Louisiana*. Baton Rouge, La: Louisiana Department of Natural Resources.
- Lowe, W. H., Likens, G. E., & Power, M. E. (2006). Linking scales in stream ecology. *AIBS Bulletin*, 56(7), 591–597.

- MacArthur, R. H. (1984). *Geographical ecology: patterns in the distribution of species*. Princeton University Press.
- Macdonald, J. (2000). Mortality during the migration of Fraser River sockeye salmon (*Oncorhynchus nerka*): a study of the effects of ocean and river environmental conditions in 1997. *Can. Tech. Rep. Fish. Aquat. Sci.*, 12, 2315.
- MacDonald, J., Herunter, H., & Moore, R. (2003). Forestry impacts on fish habitat in the northern interior of British Columbia: A compendium of research from the Stuart-Takla fish-forestry interaction study. Canadian Technical Report of Fisheries and Aquatic Sciences 2509, Fisheries and Oceans Canada.
- Macdonald, J. S., Patterson, D. A., Hague, M. J., & Guthrie, I. C. (2010). Modeling the influence of environmental factors on spawning migration mortality for sockeye salmon fisheries management in the Fraser River, British Columbia. *Transactions of the American Fisheries Society*, 139(3), 768–782.
- Macklin, M. G., & Lewin, J. (2015). The rivers of civilization. *Quaternary Science Reviews*, 114, 228–244.
- Malick, M. J., Cox, S. P., Mueter, F. J., & Peterman, R. M. (2015). Linking phytoplankton phenology to salmon productivity along a north–south gradient in the Northeast Pacific Ocean. *Canadian Journal of Fisheries and Aquatic Sciences*, 72(5), 697–708.
- Maness, H., Kushner, P. J., & Fung, I. (2013). Summertime climate response to mountain pine beetle disturbance in British Columbia. *Nature Geoscience*, 6(1), 65–70.
- Mangel, M., & Satterthwaite, W. H. (2008). Combining proximate and ultimate approaches to understand life history variation in salmonids with application to fisheries, conservation, and aquaculture. *Bulletin of Marine Science*, 83(1), 107–130.
- Manhard, C. V., Joyce, J. E., & Gharrett, A. J. (2017). Evolution of phenology in a salmonid population: a potential adaptive response to climate change. *Canadian Journal of Fisheries and Aquatic Sciences*, 74(10), 1519–1527.
- Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, 7(1), 77–91.
- Marsha, A., Steel, E. A., Fullerton, A. H., & Sowder, C. (2018). Monitoring riverine thermal regimes on stream networks: Insights into spatial sampling designs from the Snoqualmie River, WA. *Ecological Indicators*, 84, 11–26.
- Martins, E. G., Hinch, S. G., Patterson, D. A., Hague, M. J., Cooke, S. J., Miller, K. M., Lapointe, M. F., English, K. K., & Farrell, A. P. (2011). Effects of river temperature and climate warming on stock-specific survival of adult migrating Fraser River sockeye salmon (*Oncorhynchus nerka*). *Global Change Biology*, 17(1), 99–114.
- McCabe, G. J., & Wolock, D. M. (2014). Spatial and temporal patterns in conterminous United States streamflow characteristics. *Geophysical Research Letters*, 41(19), 6889–6897.

- McCluney, K. E., Poff, N. L., Palmer, M. A., Thorp, J. H., Poole, G. C., Williams, B. S., Williams, M. R., & Baron, J. S. (2014). Riverine macrosystems ecology: sensitivity, resistance, and resilience of whole river basins with human alterations. *Frontiers in Ecology and the Environment*, 12(1), 48–58.
- McLean, N., Lawson, C. R., Leech, D. I., & Pol, M. (2016). Predicting when climate-driven phenotypic change affects population dynamics. *Ecology Letters*, 19(6), 595–608.
- Melillo, J., Richmond, T., & Yohe, G. W. (2014). *Climate change impacts in the United States: The Third National Climate Assessment..* U.S. Government Printing Office.
- Mertz, G., & Myers, R. (1994). Match/mismatch predictions of spawning duration versus recruitment variability. *Fisheries Oceanography*, 3(4), 236–245.
- Miller-Rushing, A. J., Høye, T. T., Inouye, D. W., & Post, E. (2010). The effects of phenological mismatches on demography. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 365(1555), 3177–3186.
- Moody, J. A., & Troutman, B. M. (2002). Characterization of the spatial variability of channel morphology. *Earth Surface Processes and Landforms*, 27(12), 1251–1266.
- Moore, J. W., Beakes, M. P., Nesbitt, H. K., Yeakel, J. D., Patterson, D. A., Thompson, L. A., Phillis, C. C., Braun, D. C., Favaro, C., Scott, D., Charmaine, C., & Atlas, W. I. (2015). Emergent stability in a large, free-flowing watershed. *Ecology*, 96(2), 340–347.
- Moore, J. W., McClure, M., Rogers, L. A., & Schindler, D. E. (2010). Synchronization and portfolio performance of threatened salmon. *Conservation Letters*, 3(5), 340–348.
- Moore, R., Fleming, S., Menounos, B., Wheate, R., Fountain, A., Stahl, K., Holm, K., & Jakob, M. (2009). Glacier change in western North America: influences on hydrology, geomorphic hazards and water quality. *Hydrological Processes*, 23(1), 42–61.
- Moore, R. D. (2006). Stream temperature patterns in British Columbia, Canada, based on routine spot measurements. *Canadian Water Resources Journal*, 31(1), 41–56.
- Moran, P. (1953). The statistical analysis of the Canadian Lynx cycle. *Australian Journal of Zoology*, 1(3), 291–298.
- Morrison, J., Quick, M., & Foreman, M. (2002). Climate change in the Fraser River watershed: flow and temperature projections. *Journal of Hydrology*, 263, 230–244.
- Mortensen, D., Wertheimer, A., Taylor, C., & Landingham, J. (2000). The relation between early marine growth of pink salmon, *Oncorhynchus gorbuscha*, and marine water temperature, secondary production, and survival to adulthood. *Fishery Bulletin*, 98(2), 319–335.
- Murray, C. B., & Beacham, T. D. (1986). Effect of varying temperature regimes on the development of pink salmon (*Oncorhynchus gorbuscha*) eggs and alevins. *Canadian Journal of Zoology*, 64(3), 670–676.
- Murray, C. B., & McPhail, J. (1988). Effect of incubation temperature on the development of five species of pacific salmon (*Oncorhynchus*) embryos and alevins. *Canadian Journal of Zoology*, 66(1), 266–273.

- Narain, U., Margulis, S., & Essam, T. (2011). Estimating costs of adaptation to climate change. *Climate Policy*, 11(3), 1001–1019.
- Neave, F. (1966). Pink salmon in British Columbia. *Bulletin – International North Pacific Fisheries Commission*, 18.
- Nelson, K. C., & Palmer, M. A. (2007). Stream temperature surges under urbanization and climate change: Data, models, and responses 1. *Journal of the American Water Resources Association*, 43(2), 440–452.
- Nesbitt, H. K., & Moore, J. W. (2016). Species and population diversity in Pacific salmon fisheries underpin indigenous food security. *Journal of Applied Ecology*, 53(5), 1489–1499.
- Nielsen, J. L., Ruggerone, G. T., & Zimmerman, C. E. (2013). Adaptive strategies and life history characteristics in a warming climate: Salmon in the arctic? *Environmental Biology of Fishes*, 96(10-11), 1187–1226.
- Nijssen, B., O'Donnell, G., Hamlet, A., & Lettenmaier, D. (2001). Hydrologic sensitivity of global rivers to climate change. *Climate Change*, 50, 143–175.
- Otero, J., L'Abée-Lund, J. H., Castro-Santos, T., Leonardsson, K., Storvik, G. O., Jonsson, B., Dempson, B., Russell, I. C., Jensen, A. J., Baglinière, J.-L., et al. (2014). Basin-scale phenology and effects of climate variability on global timing of initial seaward migration of atlantic salmon (*Salmo salar*). *Global Change Biology*, 20(1), 61–75.
- Palmer, M. A., Lettenmaier, D. P., Poff, N. L., Postel, S. L., Richter, B., & Warner, R. (2009). Climate change and river ecosystems: Protection and adaptation options. *Environmental Management*, 44(6), 1053–1068.
- Palmer, M. A., & Richardson, D. C. (2009). *The Princeton guide to ecology In: Provisioning services: a focus on freshwater*, (p. 625). Princeton University Press.
- Parken, C. K., Candy, J. R., Irvine, J. R., & Beacham, T. D. (2008). Genetic and coded wire tag results combine to allow more-precise management of a complex Chinook salmon aggregate. *North American Journal of Fisheries Management*, 28(1), 328–340.
- Parmesan, C., & Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, 421(6918), 37.
- Patterson, D., Macdonald, J., Skibo, K., Barnes, D., Guthrie, I., & Hills, J. (2007). *Reconstructing the summer thermal history for the lower Fraser River, 1941 to 2006, and implications for adult sockeye salmon (*Oncorhynchus nerka*) spawning migration*. Fisheries and Oceans Canada Cultus Lake, BC.
- Pecl, G. T., Araújo, M. B., Bell, J. D., Blanchard, J., Bonebrake, T. C., Chen, I.-C., Clark, T. D., Colwell, R. K., Danielsen, F., Evengård, B., Falconi, L., Ferrier, S., Frusher, S., Garcia, R. A., Griffis, R. B., Hobday, A. J., Janion-Scheepers, C., Jarzyna, M. A., Jennings, S., Lenoir, J., Linnetved, H. I., Martin, V. Y., McCormack, P. C., McDonald, J., Mitchell, N. J., Mustonen, T., Pandolfi, J. M., Pettorelli, N., Popova, E., Robinson, S. A., Scheffers, B. R., Shaw, J. D., Sorte, C. J. B., Strugnell, J. M., Sunday, J. M., Tuanmu, M.-N., Vergés, A., Villanueva, C., Wernberg, T., Wapstra, E., & Williams, S. E. (2017). Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. *Science*, 355(6332).

- Pess, G., Hilborn, R., Kloehn, K., Quinn, T., & Bradford, M. (2012). The influence of population dynamics and environmental conditions on pink salmon (*Oncorhynchus gorbuscha*) recolonization after barrier removal in the Fraser River, British Columbia, Canada. *Canadian Journal of Fisheries and Aquatic Sciences*, 69(5), 970–982.
- Peterson, E., & Hoef, J. V. (2014). Stars: An arcgis toolset used to calculate the spatial information needed to fit spatial statistical models to stream network data. *Journal of Statistical Software*, 56(2), 1–17.
- Peterson, E. E., Hoef, J. M., Isaak, D. J., Falke, J. A., Fortin, M., Jordan, C. E., Kristina, M., Monestiez, P., Ruesch, A. S., Sengupta, A., Som, N., Steel, A. E., Theobald, D. M., Torgersen, C. E., & Wenger, S. J. (2013). Modelling dendritic ecological networks in space: an integrated network perspective. *Ecology Letters*, 16(5), 707–719.
- Peterson, E. E., Theobald, D. M., & ver Hoef, J. M. (2007). Geostatistical modelling on stream networks: developing valid covariance matrices based on hydrologic distance and stream flow. *Freshwater biology*, 52(2), 267–279.
- Pettit, N. E., & Naiman, R. J. (2007). Fire in the riparian zone: Characteristics and ecological consequences. *Ecosystems*, 10(5), 673–687.
- Pinheiro, J., & Bates, D. (2000). *Mixed-Effects Models in S and S-PLUS*. Springer.
- Poff, L. N., Olden, J. D., Merritt, D. M., & Pepin, D. M. (2007). Homogenization of regional river dynamics by dams and global biodiversity implications. *Proceedings of the National Academy of Sciences*, 104(14), 5732–5737.
- Quinn, T. P. (2011). *The behavior and ecology of Pacific salmon and trout*. UBC press.
- Quinn, T. P., Hodgson, S., & Peven, C. (1997). Temperature, flow, and the migration of adult sockeye salmon (*Oncorhynchus nerka*) in the Columbia River. *Canadian Journal of Fisheries and Aquatic Sciences*, 54(6), 1349–1360.
- Rand, P., Hinch, S., Morrison, J., Foreman, M., MacNutt, M., Macdonald, J., Healey, M., Farrell, A., & Higgs, D. (2006). Effects of river discharge, temperature, and future climates on energetics and mortality of adult migrating Fraser River sockeye salmon. *Transactions of the American Fisheries Society*, 135(3), 655–667.
- Rauscher, S. A., Pal, J. S., Diffenbaugh, N. S., & Benedetti, M. M. (2008). Future changes in snowmelt-driven runoff timing over the western US. *Geophysical Research Letters*, 35(16).
- Reed, T. E., Schindler, D. E., Hague, M. J., Patterson, D. A., Meir, E., Waples, R. S., & Hinch, S. G. (2011). Time to evolve? potential evolutionary responses of fraser river sockeye salmon to climate change and effects on persistence. *PLoS One*, 6(6), e20380.
- Rice, S., Roy, A., & Rhoads, B. (2008). *River confluences, tributaries and the fluvial network*. John Wiley & Sons.
- Richardson, A. J. (2008). In hot water: zooplankton and climate change. *ICES Journal of Marine Science*, 65(3), 279–295.

- Richter, A., & Kolmes, S. A. (2005). Maximum temperature limits for chinook, coho, and chum salmon, and steelhead trout in the Pacific Northwest. *Reviews in Fisheries Science*, 13(1), 23–49.
- Ricker, W. E. (1989). History and present state of the odd-year-pink salmon runs of the Fraser River region. Tech. rep., Department of Fisheries and Oceans Canada.
- Rinaldo, A., Rigon, R., Banavar, J. R., Maritan, A., & Rodriguez-Iturbe, I. (2014). Evolution and selection of river networks: Statics, dynamics, and complexity. *Proceedings of the National Academy of Sciences*, 111(7), 2417–2424.
- Rodriguez-Iturbe, I., Muneepakul, R., Bertuzzo, E., Levin, S. A., & Rinaldo, A. (2009). River networks as ecological corridors: A complex systems perspective for integrating hydrologic, geomorphologic, and ecologic dynamics. *Water Resources Research*, 45(1).
- Rogers, L. A., & Schindler, D. E. (2008). Asynchrony in population dynamics of sockeye salmon in southwest Alaska. *Oikos*, 117(10), 1578–1586.
- Roos, J. F. (1991). *Restoring Fraser River Salmon: A History of the International Pacific Salmon Fisheries Commission 1937-1985*. Vancouver, British Columbia (Canada) Pacific Salmon Commission.
- Root, T. L., Price, J. T., Hall, K. R., Schneider, S. H., et al. (2003). Fingerprints of global warming on wild animals and plants. *Nature*, 421(6918), 57.
- Ruesch, A. S., Torgersen, C. E., Lawler, J. J., Olden, J. D., Peterson, E. E., Volk, C. J., & Lawrence, D. J. (2012). Projected climate-induced habitat loss for salmonids in the John Day river network, Oregon, USA. *Conservation Biology*, 26(5), 873–882.
- Salinger, D. H., & Anderson, J. J. (2006). Effects of water temperature and flow on adult salmon migration swim speed and delay. *Transactions of the American Fisheries Society*, 135(1), 188–199.
- Sanford, S., Creed, I., Tague, C., Beall, F., & Buttle, J. (2007). Scale-dependence of natural variability of flow regimes in a forested landscape. *Water Resources Research*, 43(8).
- Satterthwaite, W. H., Carlson, S. M., Allen-Moran, S. D., Vincenzi, S., Bograd, S. J., & Wells, B. K. (2014). Match-mismatch dynamics and the relationship between ocean-entry timing and relative ocean recoveries of Central Valley fall run Chinook salmon. *Marine Ecology Progress Series*, 511, 237–248.
- Scheffer, M., Carpenter, S., Foley, J. A., Folke, C., & Walker, B. (2001). Catastrophic shifts in ecosystems. *Nature*, 413(6856), 591.
- Scheuerell, M. D., Zabel, R. W., & Sandford, B. P. (2009). Relating juvenile migration timing and survival to adulthood in two species of threatened Pacific salmon (*Oncorhynchus spp.*). *Journal of Applied Ecology*, 46(5), 983–990.
- Schindler, D. E., Hilborn, R., Chasco, B., Boatright, C. P., Quinn, T. P., Rogers, L. A., & Webster, M. S. (2010). Population diversity and the portfolio effect in an exploited species. *Nature*, 465(7298), 609.

- Schnorbus, M. A., Werner, A., & Bennett, K. (2010). *Quantifying the water resource impacts of mountain pine beetle and associated salvage harvest operations across a range of watershed scales: Hydrologic modeling of the Fraser River Basin*, vol. 423. Pacific Forestry Centre.
- Seybold, H., Rothman, D. H., & Kirchner, J. W. (2017). Climate's watermark in the geometry of stream networks. *Geophysical Research Letters*, 44(5), 2272–2280.
- Shrestha, R., Cannon, A., Schnorbus, M., & Change, Z. F. (2017). Projecting future non-stationary extreme streamflow for the Fraser River, Canada. *Climatic Change*, 145(3), 289–303.
- Shrestha, R. R., Schnorbus, M. A., Werner, A. T., & Berland, A. J. (2012). Modelling spatial and temporal variability of hydrologic impacts of climate change in the Fraser River basin, British Columbia, Canada. *Hydrological Processes*, 26(12), 1840–1860.
- Shumway, R. H., & Stoffer, D. S. (2000). *ARIMA Models In: Time series analysis and its applications*, (p. 67). Springer, 3 ed.
- Simard, S. (2015). Conversations in the forest: The roots of nature's equanimity. *SGI Quarterly*, 79, 8–9.
- Sinervo, B., Mendez-De-La-Cruz, F., Miles, D. B., Heulin, B., Bastiaans, E., Villagrán-Santa Cruz, M., Lara-Resendiz, R., Martínez-Méndez, N., Calderón-Espinosa, M. L., Meza-Lázaro, R. N., et al. (2010). Erosion of lizard diversity by climate change and altered thermal niches. *Science*, 328(5980), 894–899.
- Slemmons, K. E., Saros, J. E., & Simon, K. (2013). The influence of glacial meltwater on alpine aquatic ecosystems: a review. *Environmental Science: Processes & Impacts*, 15(10), 1794–1806.
- Smoker, W. W., Gharrett, A. J., & Stekoll, M. S. (1998). Genetic variation of return date in a population of pink salmon: a consequence of fluctuating environment and dispersive selection. *Alaska Fishery Research Bulletin*, 5(1), 46–54.
- Sowder, C., & Steel, E. A. (2012). A note on the collection and cleaning of water temperature data. *Water*, 4(3), 597–606.
- Stahl, K., & Moore, R. (2006). Influence of watershed glacier coverage on summer streamflow in British Columbia, Canada. *Water Resources Research*, 42(6).
- Stan Development Team (2017). Stan modeling language users guide and reference manual. *Version 2.16.0*, <http://mc-stan.org>.
- Steel, A. E., Sowder, C., & Peterson, E. E. (2016). Spatial and temporal variation of water temperature regimes on the Snoqualmie River network. *Journal of the American Water Resources Association*, 52(3), 769–787.
- Steel, E. A., Beechie, T. J., Torgersen, C. E., & Fullerton, A. H. (2017). Envisioning, quantifying, and managing thermal regimes on river networks. *BioScience*, 67(6), 506–522.

- Stralher, A. (1964). Quantitative geomorphology of drainage basins and channel network. *Handbook of Applied Hydrology*, (pp. 4–76).
- Strayer, D. L., & Dudgeon, D. (2010). Freshwater biodiversity conservation: recent progress and future challenges. *Journal of the North American Benthological Society*, 29(1), 344–358.
- Surowiecki, J. (2004). *The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business*. Doubleday & Co. New York, NY, US.
- Tange, O. (2011). GNU Parallel - the command-line power tool. *The USENIX Magazine*, 36(1), 42–47.  
URL <http://www.gnu.org/s/parallel>
- Taylor, S. G. (2008). Climate warming causes phenological shift in pink salmon, *Oncorhynchus gorbuscha*, behavior at Auke Creek, Alaska. *Global Change Biology*, 14(2), 229–235.
- Terui, A., Ishiyama, N., Urabe, H., Ono, S., Finlay, J. C., & Nakamura, F. (2018). Metapopulation stability in branching river networks. *Proceedings of the National Academy of Sciences*.
- Thackeray, S. J., Henrys, P. A., Hemming, D., Bell, J. R., Botham, M. S., Burthe, S., Helaouet, P., Johns, D. G., Jones, I. D., Leech, D. I., et al. (2016). Phenological sensitivity to climate across taxa and trophic levels. *Nature*, 535(7611), 241.
- Theobald, D. M., Norman, J. B., Peterson, E., Ferraz, S., Wade, A., & Sherburne, M. (2006). Functional linkage of water basins and streams (FLoWS) v1 user's guide: ArcGIS tools for network-based analysis of freshwater ecosystems. *Natural Resource Ecology Lab, Colorado State University, Fort Collins, CO*, 43.
- Thompson, L., Baudry, V., Moore, B., Hague, M., Senciall, D., Mercer, S., Patterson, D., of Fisheries, D., & Oceans, B. S. B., Burnaby (2010). A standardised process for the rescue, archival and quality control of historic water temperature data for the Fraser River Watershed, British Columbia. Tech. rep., DFO, Burnaby, BC Canada.
- Todd, C. R., Ryan, T., Nicol, S. J., & Bearlin, A. R. (2005). The impact of cold water releases on the critical period of post-spawning survival and its implications for murray cod (*Maccullochella peelii peelii*): a case study of the Mitta Mitta River, southeastern Australia. *River Research and Applications*, 21(9), 1035–1052.
- Tonolla, D., Wolter, C., Ruhtz, T., & Tockner, K. (2012). Linking fish assemblages and spatiotemporal thermal heterogeneity in a river-floodplain landscape using high-resolution airborne thermal infrared remote sensing and in-situ measurements. *Remote Sensing of Environment*, 125, 134–146.
- Vannote, R. L., Minshall, G. W., Cummins, K. W., Sedell, J. R., & Cushing, C. E. (1980). The river continuum concept. *Canadian Journal of Fisheries and Aquatic Sciences*, 37(1), 130–137.

- Visser, M. E., & Both, C. (2005). Shifts in phenology due to global climate change: the need for a yardstick. *Proceedings of the Royal Society of London B: Biological Sciences*, 272(1581), 2561–2569.
- Visser, M. E., Holleman, L. J., & Gienapp, P. (2006). Shifts in caterpillar biomass phenology due to climate change and its impact on the breeding biology of an insectivorous bird. *Oecologia*, 147(1), 164–172.
- Visser, M. E., te Marvelde, L., & Lof, M. E. (2012). Adaptive phenological mismatches of birds and their food in a warming world. *Journal of Ornithology*, 153(1), 75–84.
- Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S. E., Sullivan, C. A., Liermann, C. R., et al. (2010). Global threats to human water security and river biodiversity. *Nature*, 467(7315), 555–561.
- Wang, T., Hamann, A., Spittlehouse, D., & Carroll, C. (2016). Locally downscaled and spatially customizable climate data for historical and future periods for North America. *PLoS One*, 11(6), e0156720.
- Wang, T., Hamann, A., Spittlehouse, D. L., & Murdock, T. Q. (2012). ClimateWNA—high-resolution spatial climate data for western North America. *Journal of Applied Meteorology and Climatology*, 51(1), 16–29.
- Ward, E. J., Anderson, J. H., Beechie, T. J., Pess, G. R., & Ford, M. J. (2015). Increasing hydrologic variability threatens depleted anadromous fish populations. *Global Change Biology*, 21(7), 2500–2509.
- Ward, J. (1985). Thermal characteristics of running waters. In *Perspectives in Southern Hemisphere Limnology*, (pp. 31–46). Springer.
- Ward, J. (1989). The four-dimensional nature of lotic ecosystems. *Journal of the North American Benthological Society*, 8(1), 2–8.
- Washington State Department of Natural Resources (2017). The Bare Earth: How LiDAR in Washington State exposes geology and natural hazards.  
URL <https://arcg.is/1DGeqL>
- Webb, B. W., Hannah, D. M., Moore, R. D., Brown, L. E., & Nobilis, F. (2008). Recent advances in stream and river temperature research. *Hydrological Processes: An International Journal*, 22(7), 902–918.
- Wetzel, R. G. (2001). *Limnology: lake and river ecosystems*. Gulf Professional Publishing.
- White, J. C., Wulder, M. A., Hermosilla, T., Coops, N. C., & Hobart, G. W. (2017). A nationwide annual characterization of 25 years of forest disturbance and recovery for Canada using Landsat time series. *Remote Sensing of Environment*, 194, 303–321.
- Wild Fish Conservancy (2015). 2015 temperature and flow conditions of Pacific Northwest rivers: A quality and quantity crisis and the need for fishery closures and the development of a NOAA approved drought management plan for the protection of ESA-listed salmonids. Tech. rep., Wild Fish Conservancy.

- Williams, J., Isaak, D., Imhof, J., Hendrickson, D., & McMillan, J. (2015). Cold-water fishes and climate change in North America. *Reference Module in Earth Systems and Environmental Sciences*.
- Wood, E. F., Sivapalan, M., & Beven, K. (1990). Similarity and scale in catchment storm response. *Reviews of Geophysics*, 28(1), 1–18.
- Woodrow, K., Lindsay, J. B., & Berg, A. A. (2016). Evaluating DEM conditioning techniques, elevation source data, and grid resolution for field-scale hydrological parameter extraction. *Journal of Hydrology*, 540, 1022–1029.
- Xenopoulos, M. A., & Lodge, D. M. (2006). Going with the flow: using species-discharge relationships to forecast losses in fish biodiversity. *Ecology*, 87(8), 1907–14.
- Yeakel, J., Moore, J., Guimarães, P., & Aguiar, M. (2014). Synchronisation and stability in river metapopulation networks. *Ecology Letters*, 17(3), 273–283.
- Young, J. L., Hinch, S. G., Cooke, S. J., Crossin, G. T., Patterson, D. A., Farrell, A. P., Kraak, G. V. D., Lotto, A. G., Lister, A., Healey, M. C., et al. (2006). Physiological and energetic correlates of en route mortality for abnormally early migrating adult sockeye salmon (*Oncorhynchus nerka*) in the Thompson River, British Columbia. *Canadian Journal of Fisheries and Aquatic Sciences*, 63(5), 1067–1077.
- Zhang, M., & Wei, X. (2014). Alteration of flow regimes caused by large-scale forest disturbance: a case study from a large watershed in the interior of British Columbia, Canada. *Ecohydrology*, 7(2), 544–556.