

INNOVATIVE STOCK ASSESSMENT METHODS AND MANAGEMENT SOLUTIONS
FOR SPATIALLY STRUCTURED FISH POPULATIONS

By

Yang Li

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Fisheries and Wildlife – Doctor of Philosophy

2018

ABSTRACT

INNOVATIVE STOCK ASSESSMENT METHODS AND MANAGEMENT SOLUTIONS FOR SPATIALLY STRUCTURED FISH POPULATIONS

By

Yang Li

Fish movement is a key characteristic of fish populations and is essential to account for from both conservation and management perspectives. Movement behavior can impact how fish are distributed, and whether their populations persist in face of ecosystem changes, and how stocks are assessed. This dissertation seeks to advance our knowledge in understanding of fish movement patterns in time and space, how those are related to other environmental variables, and how to best harness information on such movement as part of fishery assessment and management, in the case of overlapping fish populations where distinct or partially distinct spawning stocks mix during the harvest season. In Chapter 1, I developed a Bayesian variable selection framework for analyzing how factors impact movement intensity in large water areas. Based on the tag-recovery results of lake whitefish populations in Lake Huron from 2003 to 2011, I evaluated how different predictors influenced lake whitefish net movement distance in Lake Huron. These net movement distances were calculated from tagging results. By using a data-driven Bayesian variable selection method, results suggest that lake whitefish with greater total length had longer net distance, and fish started their annual spawning runs earlier in warmer years after acquiring and processing energy needed for spawning. Results also show that when relative *Diporeia* spp. density was high near the tagging site, lake whitefish tended to stay closer to their tagging site. In Chapter 2 I explored the use of spawning origin information of catch as a means for improving the stock assessments for overlapping fish populations. I also evaluated the influence of including annual recruitment penalties. Results suggested that incorporating

information on population-specific harvest age composition improved spawning stock biomass estimation throughout the years being assessed, and improved recruitment estimates only in the early assessment period. Including penalties on annual recruitment residuals improved recruitment estimates in terminal assessment years. In Chapter 3, I extended the spatial Brownie-Petersen tagging model for modeling multiyear tag-recovery data in a fishery context, and incorporated catch-at-age, and tag monitoring data jointly, for lake whitefish populations in Lake Huron. Previous studies of extending Brownie tagging models considering spatial structure were all based on a spatial assumption that fish start moving from where they were in the last time period, and did not recognize spawning site fidelity. We assumed 100% spawning site fidelity for lake whitefish in our model, and results suggested spawning populations in U.S. main basin has higher probability to overlap with other populations during fishing season compared to those in Canadian waters. In chapter 4, I extended the tagging model proposed in Chapter 3 to a more comprehensive framework that allowed for a continuum of spatial structures through modeling homing probability. Based on simulations, we explored how the degree of homing, the extent of spatial movements, and the types of data used, influenced estimability of parameters of interest. My results suggest that the model framework with only tag-recovery and fishing effort data had robust assessment performance in estimating movement rates, homing probability, natural mortality, and fully selectivity fishing mortality rates. With additional tag monitoring data, tag reporting rate can also be accurately estimated simultaneously. Including additional catch-at-age data did substantially improve the estimates of selectivity at age, slightly improved estimates of tag reporting and natural mortality rates, but the bias in estimating recruitment and spawning stock biomass can be high, especially for low productivity populations.

ACKNOWLEDGMENTS

I am grateful to many people for helping me throughout my doctoral program. First and foremost, I gladly acknowledge my advisor Dr. Jim Bence for his guidance and supervision when I was first fumbling with the original ideas for this research. His technical and editorial advice was essential to the completion of this dissertation, and has taught me innumerable lessons and insights on the workings of academic research in general. I am grateful to my co-advisor Dr. Travis Brenden for his gentle, yet insightful guidance and critique of both my analysis and writing. I thank other members of my graduate committee, Dr. William Taylor and Dr. Andy Finley, for their contributions and recommendations to this research. I sincerely thank Mark Ebener, Adam Cottrill, and Ji He for providing tag-recovery, catch-at-age, and tag monitoring data, Dr. Richard P. Barbiero for providing *Diporeia* abundances data, and Dr. Lacey Mason for providing surface water temperature data. I would also like to acknowledge Dr. Andrew Punt, and other three anonymous reviewers for commenting on earlier manuscript drafts for Chapter 1 and 2. I thank the Michigan DNR Fisheries Division, the Great Lakes Fishery Commission, and the Great Lakes Fishery Trust for financial support of this research. In addition, I acknowledge computational support and technical computing assistance from the High Performance Computing Center at the Institute for Cyber-Enabled Research at MSU. The modeling subcommittee of the Technical Fisheries Committee for 1836 Treaty Waters also provided input on my research. Lastly, I would like to thank all of my family and friends, whose loving kindness always bolstered my spirits during those periods of uncertainty when I did not believe I would finish the work. My parents, Hongping and Jiajun, have supported and inspired me in so many ways. My husband Zhen gave me a lot of help and support to my life and

research, and taught me many useful skills about programming and Bayesian. My baby Leo, provided me with countless hours of fun and laughter. Thank you all.

PREFACE

The chapters in this dissertation were drafted as stand-alone papers that will be submitted for publication in peer-reviewed journals. Chapter 1 was submitted and accepted for publication in the *Fisheries Research*. Chapter 2 was submitted and accepted for publication in the *Canadian Journal of Fisheries and Aquatic Sciences*. When submitted, Chapters 3 and 4 will include one or more co-authors. Consequently, chapters 1-4 are written with first person, plural narratives, even though I am listed as the sole author of the dissertation. All references are formatted in a style consistent with the *Canadian Journal of Fisheries and Aquatic Sciences*.

TABLE OF CONTENTS

LIST OF TABLES	x
LIST OF FIGURES	xiii
INTRODUCTION	1
Fish movement and Models of it	1
Fish Movement in Stock Assessment Models	4
Background on Lake Huron lake whitefish	7
Objectives	9
BIBLIOGRAPHY	11
CHAPTER 1 WHY DO LAKE WHITEFISH MOVE LONG DISTANCES IN LAKE HURON? BAYESIAN VARIABLE SELECTION OF FACTORS EXPLAINING FISH MOVEMENT DISTANCE	16
Abstract	16
Introduction	16
Methods	22
Data collection, selection, and calculation of net-movement distance	22
Explanatory variables	24
Case 1 (GDD hypothesis 1)	25
Case 2 (GDD hypothesis 2)	27
Model framework	28
Model description	29
Characterization of the posterior distribution using MCMC	31
Variable-wise summary	31
Model-wise summary	31
Model diagnosis, simulation study, and comparison with tree-based methods	32
Results	32
Variable-wise summary	33
GDD hypothesis 1 Case	33
GDD hypothesis 2 Case	34
Model-wise summary	37
Model diagnosis, simulation study, and comparison with tree-based methods	40
Model diagnosis	40
Simulation study	40
Comparison with tree-based methods	40
Discussion	41
Acknowledgements	46
APPENDICES	47
Appendix 1.A: Calculation of shortest water distance and GDD	49
Appendix 1.B: Model implementation	51
Appendix 1.C: Model diagnosis	56

Appendix 1.D: Simulation study.....	58
Appendix 1.E: Comparison with two tree-based methods.....	63
Appendix 1.F	65
BIBLIOGRAPHY.....	66
CHAPTER 2 CAN SPAWNING ORIGIN INFORMATION OF CATCH OR A RECRUITMENT PENALTY IMPROVE ASSESSMENT AND FISHERY MANAGEMENT PERFORMANCE FOR A SPATIALLY STRUCTURED STOCK ASSESSMENT MODEL?	71
Abstract.....	71
Introduction.....	72
Methods.....	76
Simulation framework	76
Operating model.....	79
Management procedure.....	86
Objective function.....	90
Application of the harvest control rule	92
Simulation Scenarios	93
Baseline scenario and alternative productivity and movement scenarios	94
Sensitivity Analyses.....	95
Data Quality	95
Uncertain Mixing Rates	96
Recruitment Variation.....	96
Target mortality	96
Performance statistics	97
Results.....	97
Baseline scenario	98
Alternative productivity and movement scenarios	101
Sensitivity Analyses.....	104
Discussion.....	106
Acknowledgements.....	110
APPENDIX.....	112
BIBLIOGRAPHY.....	120
CHAPTER 3 A BAYESIAN SPATIAL STRUCTURED TAGGING MODEL FOR ESTIMATING MOVEMENT PROBABILITIES, MORTALITY RATES, TAG REPORTING RATES AND ABUNDANCE FOR OVERLAPPING FISH POPULATIONS.....	124
Abstract.....	124
Introduction.....	124
Methods.....	129
Tagging model assumptions	129
Tagging model overview	130
Tag recovery dynamics.....	130
Population dynamics.....	138

Example application to lake whitefish	140
Design of Analysis	144
Model implementation	147
Performance statistics	148
Results	149
Model comparison	149
Summary of the best model: Model 14	150
Movement rates	150
Tag reporting rates	151
Instantaneous natural mortality rates	152
Instantaneous fishing mortality rates	152
Abundance	154
Assessing model fit	156
Sensitivity of parameter estimates to model parameterization	159
Comparison between best model and reduced model	161
Discussion	161
APPENDIX	167
BIBLIOGRAPHY	180
CHAPTER 4 CAN A TAGGING MODEL FRAMEWORK BE ADAPTED TO DIFFERENT SPATIAL STRUCTURE ASSUMPTIONS? A BAYESIAN SPATIALLY STRUCTURED ASSESSMENT MODEL FRAMEWORK INCORPORATING HOMING PROBABILITY	184
Abstract	184
Introduction	185
Methods	189
Simulation framework Overview	189
Operating model	196
Tagging model framework and its applications to different data availabilities	201
Tagging models	204
Estimated parameters	210
Selection of parameters for the operating model	211
Simulation scenarios	214
Baseline scenarios	214
Alternative homing and movement scenarios	215
Recruitment variation scenario	216
Data quality scenarios	216
Performance metrics	216
Results	217
Baseline scenarios	217
Alternative homing and movement assumptions	223
Sensitivity scenarios	230
Discussion	233
APPENDICES	240
Appendix 4.A: Reparameterization of the tagging dynamics of the ITCAAN model	241
Appendix 4.B: Life history parameters used in our simulations	244

Appendix 4.C: Prior Probability Distributions	246
Appendix 4.D: Likelihood component for models including tag monitoring data.....	247
Appendix 4.E: Likelihood components for models including catch at age data	249
BIBLIOGRAPHY.....	251
CONCLUSIONS AND RECOMMENDATIONS	255
Chapter 1 Conclusions and Recommendations.....	255
Chapter 2 Conclusions and Recommendations.....	256
Chapter 3 Conclusions and Recommendations.....	257
Chapter 4 Conclusions and Recommendations.....	258
Final thoughts.....	259
BIBLIOGRAPHY.....	262

LIST OF TABLES

<p>Table 1.1. Summary of candidate variables/terms and their interpretation and relationship to hypotheses. A. For continuous variables, “Hypothesis” (first column) states our a priori hypothesis associated with the variable, and the second column indicates sign of associated coefficient that would support that hypothesis. B. Similarly for interaction terms, but here a single hypothesis (our GDD Hypothesis 2) is associated with all interaction terms, and the second column describes the interpretation of coefficients and the pattern in their sign that would support the hypothesis. C. For categorical (dummy) variables we did not have explicit a priori hypothesis for the sign of coefficients but did hypothesize that these factors could influence net distance. For these variables one level of a factor is the baseline with coefficient fixed at zero, and this level (category) is given in the first column and interpretation of the sign of other coefficients in the second column. For A through C, “X” in the “GDD H1” column indicates that the variable was a candidate variable/term in our variable selection process for the GDD Hypothesis 1 Case, and the GDD H2 column likewise indicates if the variable/term was a candidate variable for the GDD Hypothesis 2 Case. The last row summarizes the total number of candidate variables for each GDD hypothesis.</p>	26
<p>Table 1.2. Posterior mean and 95% credible intervals for parameters of the highest posterior probability model.</p>	38
<p>Table 1.A1. Recoveries that were excluded in this study.</p>	48
<p>Table 1.A2. Summary of variable selection in the simulation study.</p>	61
<p>Table 1.A3. Summary of parameter estimation in simulation study.</p>	62
<p>Table 2.1. Composition of the assessment input data and objective function for the four assessment models we evaluated.</p>	78
<p>Table 2.2. Index variables and accents used in all equations.</p>	79
<p>Table 2.3. Coefficients for parameters used to generate different levels of productivity, data quality, recruitment variation, and annual-varying random generated rates in both operating and stock assessment models.</p>	81
<p>Table 2.4. Biomass calculation in the operating model.</p>	82
<p>Table 2.5. Simulation scenarios, including the baseline scenario and other combinations of productivity levels and stay rates, for four hypothetic populations used in the simulations.</p>	85
<p>Table 2.6. Scenarios for sensitivity analyses. In each sensitivity scenario, except for the change described below all other parameters are at their baseline levels.</p>	94
<p>Table 3.1. Symbols used in the model equations.</p>	131

Table 3.2. Equations for the tagging model framework.	134
Table 3.3. Exclusion criteria for fish recovered.....	143
Table 3.4. Assumptions that were used to generate different parameterizations of tag reporting rate and natural mortality in the different alternative parameterizations of the assessment models. Parameter dependency refers to what leads to different values for parameter.	145
Table 3.5. Model comparison results. The first four columns are four likelihood components, and the last four columns are D1, D2, p_D , and DIC for each model.	147
Table 3.6. Instantaneous rates of natural mortality (per year) for the best model.	152
Table 3.A1. Instantaneous rates of natural mortality (per year) for the best model without restriction on movement.	168
Table 3.A2. Instantaneous rates of natural mortality (per year) for the reduced best model without catch and age composition data.	168
Table 4.1. Symbols for all equations, except for estimated parameters that are defined in Table 4.4 and data sources that were described in Table 4.5.....	193
Table 4.2. Equations for the population dynamics in the operating model.	194
Table 4.3. Equations for the tagging model.....	195
Table 4.4. Description of estimated parameters (first and second column) and their inclusion in different tagging models (last three columns). The names of the last three columns “STR”, “STRO”, and “Full” indicate three assessment models: spatial tag recovery (STR) model, spatial tag recovery and observer (STRO) model, and the Full model. For each row, “X” indicates that the parameter was estimated in the assessment model; “O” indicates that parameter was assumed to be known without error in the assessment model; and “Flexible” indicates that the parameterization of the parameter was flexible and its estimability was evaluated in the assessment model.	202
Table 4.5. Description of data sources.....	203
Table 4.6. Inclusion of data set(s) for different tagging models.....	203
Table 4.7. Three levels of population productivity we assumed for six hypothetical management/harvest basins (Figure 4.1, high for NE, NW; medium for CE, CW, low for SE, SW), and two levels of recruitment variation we assumed in the baseline and sensitivity scenarios.....	213
Table 4.8. Different levels of data quality.	213
Table 4.9. Different levels of homing probability and movement assumptions we evaluated in a four (homing probability) by five (movement assumption) factorial experiment.	215

Table 4.A1. Weight- and maturity-at-age calculation and selectivity at age parameter values in the operating model.....	244
Table 4.A2. Equations for Ricker stock-recruitment parameterization in the operating model. .	245
Table 4.A3. Prior distributions for all estimated parameters.	246

LIST OF FIGURES

Figure 1.1. Map of the study area (Lake Huron) and seven tag release (spawning) sites. Of total 1368 recoveries, 659 were from Detour, 300 from Cheboygan, 243 from Burnt Island, 42 from Saginaw Bay, 43 from Sarnia, 56 from Alpena, and 25 from Fishing Islands.23

Figure 1.2. Posterior distributions for the number of selected variables (i.e., $q - 1$). The x-axis starts at 5 because all models selected at least five variables.33

Figure 1.3. Variable-wise summary results (posterior mean with 95% credible intervals) of the effect of variables (the β_j), with variables named on y-axis for the case with GDD hypothesis 1. Bars are highlighted by red color when the 95% credible interval does not cover 0, which is defined as a consistent effect. The number above each bar is the marginal inclusion probability.35

Figure 1.4. Variable-wise summary results (posterior mean with 95% credible interval of the effect β_j for the j th variable, as indicated in y-axis) for the case with GDD hypothesis 2. Bars are highlighted by red color when the 95% credible interval does not cover 0, which was defined as a consistent effect. The number above each bar is the marginal inclusion probability.36

Figure 1.5. Model-wise summary for top 12 models ranked according to their posterior probability mass, for the case of GDD hypothesis 2. Variables that were included in the top 12 models are given on the y-axis. Horizontal bar represents posterior 95% credible intervals and symbols on each bar the posterior mean for each coefficient included in a model, with the associated model given to the left of the bar. Thus when more bars are given for a variable it was included in more models.39

Figure 1.A1. χ^2 -discrepancy measures for both GDD hypothesis.57

Figure 1.A2. Simulated quadratic function of month effects.59

Figure 1.A3. Comparison of results from tree-based models for GDD hypothesis 1. Variables are ranked based on their importance according to the relative influence for boosted regression tree and out-of-bag prediction accuracy for random forests. The marginal inclusion probability from our Bayesian variable selection method is included in the parenthesis for comparison.63

Figure 1.A4. Investigation of variable GDD_Diff in GDD hypothesis 1 case: a) centered fitted function for GDD_Diff from the boosted regression trees; b) partial dependence for GDD_Diff from the random forests; c) the response variable versus GDD_Diff.64

Figure 1.A5. Comparison of results from tree-based models for GDD hypothesis 2. Variables are ranked based on their importance according to the relative influence for boosted regression tree (Friedman 2001) and out-of-bag prediction accuracy for random forests (Breiman 2001). The marginal inclusion probability from our Bayesian variable selection method is included in the parenthesis for comparison.64

Figure 2.1. The full closed-loop feedback simulation framework, which followed a management strategy evaluation approach.....	77
Figure 2.2. Ricker stock-recruitment relationships for populations with low, medium, and high level of productivity (Table 2.3). Two dashed lines represent the replacement lines for $F=0$ and target F and their intersections with stock-recruitment curves (dots) define equilibrium for low, baseline, and high productivity. Note that the target F is calculated based on the natural mortality rate and the status quo target total mortality ($A=0.65$).....	86
Figure 2.3. Simulation results (median \pm interquartile range) for population 1 (Table 2.5) in the baseline scenario. Full model names are in Table 2.1. (a) Relative error of estimating terminal assessment year SSB during simulation year 91 to 100. (b) In simulation year 100, relative error of estimating recruitment of the last ten assessment years. (c) Proportion of years SSB was lower than 20% of the unfished SSB level ($B_{20\%}$) over the last 25 years of simulations. (d) Mean annual yield for the fishing area surrounding spawning grounds of Pop1 over the last 25 years of simulations. (e) Mean interannual variation (IAV) in yield over the last 25 years of simulation. (f) Estimated autocorrelation for terminal year estimates of SSB during simulation years 75 to 100.....	100
Figure 2.4. Relative error in estimates of recruitment for the terminal assessment year during the simulation year 76 to 100 for an example simulation. Full model names are in Table 2.1	101
Figure 2.5. Simulation results (median \pm interquartile range) for populations 1 and 3 under scenarios 2 to 5 (Table 2.5). Full model names are in Table 2.1. Each column represents a different productivity and movement scenario, and each row presents a different performance statistic. The x-axis of each column indicates the productivity levels (L, A, H are low, average, and high productivity levels) and stay rates associated with the two populations results are presented for. For example, L70% means low productivity population with 70% stay rate. For each such productivity level and stay rate, results are given for the four different assessment methods, distinguished by different symbols. The second, fourth, and sixth rows represents the same performance statistics as for Figure 2.3c, 2.3e, and 2.3d. The first and third row are relative error of estimating terminal year SSB and recruitment in simulation year 100, respectively, with a 0 dashed line. The fifth row represents the average SSB over the last 25 years of simulation, and the dashed line is 20% of the unfished SSB.....	103
Figure 2.6. Simulation results (median \pm interquartile range) for Pop1 (Table 2.5) in sensitivity analyses. Full model names are in Table 2.1. Each column represents a sensitivity scenario, each row represents a performance metric (as described in Figure 2.5), and results in each panel are for the four assessment models.....	105
Figure 2.A1. Relative error of estimating recruitment for the terminal assessment year during the simulation year 75 to 100 in 20 example simulations.....	113
Figure 2.A2. Mean annual spawning stock biomass of each spawning population over the last 25 years of simulations. Each row of the panels represents different population productivity and stay rate scenarios (Table 2.5), and each column represents different scenarios of sensitivity analyses (Table 2.6). The x-axis of each panel represents the results of each spawning population. The	

dashed line is 20% of the unfished SSB. The symbols identify the median and the error bars identify the inter-quartile range of the simulations.....	114
Figure 2.A3. As for Figure 2.A2, except y-axis is mean annual yield for the fishing area surrounding spawning grounds of each spawning population (populations are intermixed with each other in each fishing area during harvest season) over the last 25 years of simulations done for each scenario.	115
Figure 2.A4. As for Figure 2.A2, except y-axis is the mean interannual variation (IAV) in yield over the last 25 years of simulation.	116
Figure 2.A5. As for Figure 2.A2, except y-axis is the relative error of estimating SSB of terminal assessment year.....	117
Figure 2.A6. Autocorrelation coefficient of estimating terminal year SSB during simulation year 75 to 100 (median \pm interquartile range) under all productivity and movement scenarios (include baseline scenario in the first row).	118
Figure 2.A7. Relative error (median \pm interquartile range) of estimating recruitment of assessment year 15 to 20 (i.e., year -4 to 0 relative to terminal assessment year) under all productivity and movement scenarios.	119
Figure 3.1. Map of Lake Huron with seven tag releasing and recovery basins.....	142
Figure 3.2. Posterior distribution of the movement probability that moving from each tag basin (panel name) to all recovery basins (the x-axis of each panel). Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Horizontal middle line: posterior median.	151
Figure 3.3. Posterior distribution of tag reporting rates by recovery basins. Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Middle dots: posterior median.....	152
Figure 3.4. Posterior distribution of fully selected fishing mortality by recovery years for each basin. Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Middle dots: posterior median.....	153
Figure 3.5. Posterior distribution of selectivity at age for all years, basins, and gear types. Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Middle dots: posterior median.....	154
Figure 3.6. Posterior distribution of abundance at age for each spawning population (row names) and recovery years (column names). Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Horizontal middle line: posterior median.....	155

Figure 3.7. Posterior distribution of mixed abundance at age for each year and basin. Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Horizontal middle line: posterior median.....	156
Figure 3.8. Comparison between the observed data and the posterior predictive distribution for the number of recovered fish, by recovery basin (x-axis) for each tag basin (panel). Error bar: 95% Bayesian highest posterior density interval (HPDI). Dot: Posterior median. Diamond: observed value.	157
Figure 3.9. Comparison between the observed data and the posterior predictive distribution for the catch from each recovery basin (x-axis) in each year (column) by each gear type (row). Error bar: 95% Bayesian highest posterior density interval (HPDI). Dot: Posterior median. Diamond: observed value.	158
Figure 3.10. Comparison between the observed data and the posterior predictive distribution for the age composition proportion in each recovery basin (x- column) in each year and gear type (row). Error bar: 95% Bayesian highest posterior density interval (HPDI). Dot: Posterior median. Diamond: observed value.	159
Figure 3.A1. Posterior distribution of the movement probability that moving from each tag basin (panel name) to all recovery basins (the x-axis of each panel). Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Horizontal middle line: posterior median.	168
Figure 3.A2. Posterior distribution of tag reporting rates by recovery basins. Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Middle dots: posterior median.....	169
Figure 3.A3. Posterior distribution of fully selected fishing mortality by recovery years for each basin. Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Middle dots: posterior median.....	169
Figure 3.A4. Posterior distribution of selectivity at age for all years, basins, and gear types. Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Middle dots: posterior median.	170
Figure 3.A5. Posterior distribution of abundance at age for each spawning population (row names) and recovery years (column names). Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Horizontal middle line: posterior median.	171
Figure 3.A6. Posterior distribution of mixed abundance at age for each year and basin. Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Horizontal middle line: posterior median.....	172
Figure 3.A7. Comparison between the observed data and the posterior predictive distribution for the number of recovered fish, by recovery basin (x-axis) for each tag basin (panel). Error bar:	

95% Bayesian highest posterior density interval (HPDI). Dot: Posterior median. Diamond: observed value.173

Figure 3.A8. Comparison between the observed data and the posterior predictive distribution for the catch from each recovery basin (x-axis) in each year (column) by each gear type (row). Error bar: 95% Bayesian highest posterior density interval (HPDI). Dot: Posterior median. Diamond: observed value.174

Figure 3.A9. Comparison between the observed data and the posterior predictive distribution for the age composition proportion in each recovery basin (x- column) in each year and gear type (row). Error bar: 95% Bayesian highest posterior density interval (HPDI). Dot: Posterior median. Diamond: observed value.175

Figure 3.A10. Posterior distribution of tag reporting rates by recovery basins for the reduced best model. Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Middle dots: posterior median.....176

Figure 3.A11. Posterior distribution of the movement probability that moving from each tag basin (panel name) to all recovery basins (the x-axis of each panel) for the reduced best model. Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Horizontal middle line: posterior median.177

Figure 3.A12. Posterior distribution of selectivity at age for all years, basins, and gear types for the reduced best model. Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Middle dots: posterior median.178

Figure 3.A13. Posterior distribution of fully selected fishing mortality by recovery years for each basin for the reduced best model. Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Middle dots: posterior median.178

Figure 3.A14. Comparison between the observed data and the posterior predictive distribution for the number of recovered fish, by recovery basin (x-axis) for each tag basin (panel) for the reduced best model. Error bar: 95% Bayesian highest posterior density interval (HPDI). Dot: Posterior median. Diamond: observed value.179

Figure 4.1. Map of the six hypothetical management/harvest basins (i.e., solid lines show the boundaries of each basin). The dotted circle inside of each basin is its spawning site (i.e., where tagged fish being released).....191

Figure 4.2. Flow chat for of the multi-state tag return probability for the tagged fish from cohort iak (tagged and released in year i at age a from basin k). Each rectangular box represents a state.204

Figure 4.3. Relative error of movement probability estimates for CE spawning population (medium productivity population) moving from their spawning site to all six recovery basins (the x-axis) under the baseline scenarios with alternative parameterizations of the natural mortality rate and homing probability in the assessment models (panel name). The triplet in each panel are

the results for each assessment model. Whiskers, boxes, and horizontal middle line are 95% and 50% inter quantile ranges, and median of 200 simulations.218

Figure 4.4. The natural mortality rate (a), homing probability (b), and tag reporting rate estimates under the baseline scenarios with alternative parameterizations of natural mortality rate and homing probability (panel name). The dashed line in (a) is the true value that was assumed in the operating model. The x-axis in each panel are the assessment models. Whiskers, boxes, and horizontal middle line are 95% and 50% inter quantile ranges, and median of 200 simulations.219

Figure 4.5. Relative error of selectivity at age, and fully selected fishing mortality estimates in CE basin under the baseline scenario in which natural mortality rate was estimated and homing probability was assumed to be known in the assessment models. The triplet in each panel are the results for each assessment model. Boxes, and horizontal middle line are 50% inter quantile ranges, and median of 200 simulations.221

Figure 4.6. Relative error of estimating recruitment (a) and SSB (b) in the Full model under the baseline scenario in which natural mortality rate was estimated and homing probability was assumed to be known in the assessment models. Whiskers, boxes, and horizontal middle line are 95% and 50% inter quantile ranges, and median of 200 simulations.222

Figure 4.7. The coverage probability of covering true value from the operating model in the 95% Bayesian credibility interval under the baseline scenarios with alternative parameterization of natural mortality rate and homing rate (row name). Each combination of shape and color represent results of an assessment model. Each panel represents an estimated parameter (group): f, fully selected fishing mortality rate; report: tag reporting rate; M, natural mortality rate; move: movement rate; rec, recruitment; natal, homing rate; ssb, spawning stock biomass; sel, selectivity at age. For some parameters that are varying by year, basin, and/or age, the coverage probability of each dimension/element was calculated. The middle point represents the median value, and the error bar represents the 50% interquartile ranges across all elements/dimensions.224

Figure 4.8. Relative error of recruitment estimates for the Full model under the scenarios with alternative homing assumptions. Each column represents a homing scenario, and each row represents a population. Whiskers, boxes, and horizontal middle line are 95% and 50% inter quantile ranges, and median of 200 simulations.225

Figure 4.9. Relative error of SSB estimates for the Full model under the scenarios with alternative homing assumptions. Each column represents a homing scenario, and each row represents a population. Whiskers, boxes, and horizontal middle line are 95% and 50% inter quantile ranges, and median of 200 simulations.226

Figure 4.10. Relative error of movement probability estimates for spawning populations of NE, CE, and SE basins (panel name) that moving from their spawning site to all six recovery basins (the x-axis) from the Full model under the scenarios with homing assumptions (row name). Whiskers, boxes, and horizontal middle line are 95% and 50% inter quantile ranges, and median of 200 simulations.227

Figure 4.11. Relative error of selectivity at age estimates for the Full model under the scenarios with alternative movement assumptions. Each panel represents a movement scenario. Whiskers, boxes, and horizontal middle line are 95% and 50% inter quantile ranges, and median of 200 simulations.228

Figure 4.12. Relative error of recruitment estimates for the Full model under the scenarios with alternative movement assumptions. Each column represents a movement scenario, and each row represents a population. Whiskers, boxes, and horizontal middle line are 95% and 50% inter quantile ranges, and median of 200 simulations.229

Figure 4.13. Relative error of SSB estimates for the Full model under the scenarios with alternative movement assumptions. Each column represents a movement scenario, and each row represents a population. Whiskers, boxes, and horizontal middle line are 95% and 50% inter quantile ranges, and median of 200 simulations.230

Figure 4.14. Assessment results for the Full model in sensitivity analysis. Shown are tag reporting rate estimates (a), natural mortality rate estimates (b), relative error of full selected fishing mortality estimates in all recovery years (c), and relative error of selectivity at age estimates (d). The panel name in all sections represents a sensitivity scenario. Whiskers, boxes, and horizontal middle line are 95% and 50% inter quantile ranges, and median of 200 simulations.231

Figure 4.15. Assessment results for the Full model in sensitivity analysis with high data quality. Shown are relative error of recruitment estimates (a) and relative error of SSB estimates (b) for all spawning populations (panel name) in all recovery years (x-axis). Whiskers, boxes, and horizontal middle line are 95% and 50% inter quantile ranges, and median of 200 simulations.232

Figure 4.16. Same as Figure 4.9, except that each row represents a scenario with alternative homing assumption or movement assumption, or a sensitivity scenario.233

INTRODUCTION

In this dissertation, I address fish movement through simulations and data analysis, both as a means to assessing movements and of incorporating movements into stock assessment models. Applications use data from lake whitefish in the Great Lakes, and simulations are based on life history and parameters for these lake whitefish stocks. Nevertheless, the results are intended to also provide general guidance and methods.

Fish movement and Models of it

Fish movement and population spatial structure have been increasingly investigated in the last two decades because of the development of tagging technologies, quantitative methods within the field, and computing power (Cadrin and Secor 2009). Many marine and fresh water fish species move for long distance at various times during their life cycle. Movement behavior not only allows fish to move away from unsuitable conditions and toward suitable conditions for survival and growth, but also permits migrations between habitats used by different life history stages (McMahon and Matter 2006). The behavior and fate of individual fish resulting from movements ultimately affect population-level processes. Movement performance of individual fish, including movement timing, intensity, direction, and distance, determines the spatial structure of their spawning populations. Movement behavior can influence the persistence of fish population in the face of ecosystem changes, which may further impact ecological interactions and evaluation in the long term (Gilliam and Fraser 2001). Management problems such as inaccurate assessment results, and inappropriate resulting management actions, can occur when actual fish movements do not agree with the spatial assumptions made in stock assessments and management decisions. Adverse outcomes can include local population depletion and

population collapse (Mitchell and Beauchamp 1988, Hutchings 1996, Fu and Fanning 2004, Rothschild 2007, Li et al. 2015). Thus, it is critically important to understand fish movement patterns in time and space from both individual and population perspectives.

Tagging studies have been widely used to explore movement behavior of individual fish and movement patterns of their population. Among different tagging technologies, such as conventional, electronic, acoustic, and archival tags, conventional tagging has the longest history and remains most widespread given the ratio of cost to information content and long-established methods for evaluating tag-recovery information (Lucas and Baras 2000). Conventional tagging can generate information on individual fish movements by recording the time and location of tag releasing and recovery. Although this method cannot provide information at fine temporal scales such as where fish go between tag and recovery, the relatively low cost allows conventional tagging program to cover large spatial areas, which makes it an explicitly useful approach for analyzing fish movement, for fishes that live in large water areas. Conventional tagging data can also be used to evaluate fish movement at the population level. For conventional tagging studies of adult fish, tagged fish are often tagged and released from their spawning sites during the spawning season (Durban et al. 2005, Heath et al. 2008, Ebener et al. 2010, Vincent et al. 2017). Thus, the tagging site also represents the spawning population identity of fish at the time of tag release. Tagged fish are frequently recovered during fishing years following the time of tagging. Thus conventional tagging data can be used for estimating the proportion of each spawning population occupying different regions during non-spawning periods (herein referred to as movement rates).

Previous research based on conventional tagging has explored how movement of individual fish was influenced by a variety of “triggering factors” (or cues) including life history

traits, time of tag and recovery, flow, and other environmental variables (Louca et al. 2009, Schlaff et al. 2014, Radinger and Wolter 2014, Drouineau et al. 2017). However, most of those research studies were limited to stream fish, and much less is known about movements of fish that live in large water areas. One substantial technical challenge faced by fish movement analysis in large water areas is the difficulty in calculating the shortest water distance between tag and recovery location (i.e., net movement distance) in large water areas with irregular shape, given no existing packages were available for such calculation. Net movement distance was used in this dissertation because the actual movement pathway is unknown and unpredictable based on conventional tagging data. Even without this challenge, pioneering studies of stream fish movement (where movement is generally linear) also had some limitations due to the analysis approaches. The pioneering studies of net movement distance used either ANOVA models or regression-tree based approaches to test if the net movement distance varied significantly in associate with factors they evaluated. However, the ANOVA models can provide only a rough picture of the continuous relationship between net movement distance and explanatory variables when jointly considering multiple factors and continuous covariates; while regression tree based approach recursively partitions the data points according to the categorization of the factors, which may lead to difficulty in interpreting the effects. Therefore, it is of general interest to develop a model framework for the analysis of conventional tagging data in a large water body to evaluate how factors impact fish movement distance based on sounder statistical approach. I addressed this topic in Chapter 1 of this dissertation, by analyzing individual net movement data for lake whitefish.

Population dynamic modelers and fisheries managers tend to be more interested in population-level process than in the behavior of individual fish, given populations have

demographic features and most fisheries are managed by management unit (McMahon and Matter 2006). When movement has been incorporated directly into the population dynamic models, movement rates between regions are generally used to model the overall proportions of each population moved to the different harvest regions and become additional key parameters to be estimated. Such movement is generally assumed to occur once a year right after the spawning period, and after movement fish that live in the same harvest region have the same survival and exploitation rate, until the next spawning period (Eveson et al. 2009, Vandergoot and Brenden 2014, Goethel et al. 2015a, Vincent et al. 2017). Depending upon the spawning site fidelity of the fish species, fish either would be assumed to return to their last spawning site to spawn or stay in the basin they had moved to and spawn there during the next spawning period. The first of these movement patterns is referred to as the overlap spatial structure (Porch et al. 2001, Goethel et al. 2011). The second is known as the diffusion (also called metapopulation) structure (Porch et al. 2001, Goethel et al. 2011). These two structures contrast in that spawning population identity is maintained with the overlap structure, whereas it can change each spawning season with the diffusion structure. Other movement patterns, such as seasonal movement (Fu and Fanning 2004), and movement that occurred at a different time of the year than right after spawning (Goethel et al. 2015b) have also occasionally allowed for in spatial movement models for some specific fish species.

Fish Movement in Stock Assessment Models

When prior knowledge of fish movement patterns and migration rates are available, how to incorporate such information into stock assessment models, and how it affects the long-term stock assessment and fisheries management performance are important topics that are attracting research. One known problem for assessment models, when applied to populations exhibiting

spatial structuring with moderate to high levels of intermixing, is that population-specific estimates of recruitment are uncertain or not estimable, and estimates of spawning stock biomass are unstable or biased, even when mixing rates are assumed known (Ying et al. 2011; Molton et al. 2012; Li et al. 2015). Li et al. (2015) proposed an overlap stock assessment model in which an integrated statistical catch-at-age (SCAA) assessment model was fit to overlapping fish populations by incorporating actual mixing rates in the model (rather than providing the model tagging data and estimating them as the often the case for models that use both tagging data and catch-at-age data). They found that mixing among areas caused problems in estimating population-specific annual recruitments, and this led to substantial uncertainty and bias in estimation of recruitment and biomass. They hypothesized that this problem could be resolved if additional population-specific data were provided to the assessment model, such that harvest data could be allocated to source populations. Hintzen et al. (2015) evaluated the influence of fishery-independent survey data on the performance of such an integrated SCAA method for intermixing fish populations, in which information on the classification of the catch to their spawning origin were used to inform survey indices (i.e., the proportions of survey sample to spawning populations). However, the catch data they used in the assessment model were not reallocated back to the spawning populations because their assessment model ignored spatial structure. Thus, mismatch between spatial structures in the assessment data and in the assessment model still existed. They found that spatially-explicit survey data marginally reduced bias in estimation of biomass, but when there were errors in classification rates, inaccuracies could actually increase. Chapter 2 addresses approaches to improving such spatial assessment models using data on spawning stock origin and through other means.

Conventional tagging data could be used to estimate the movement rates between regions. Many spatial structured tagging models based on conventional tagging data are now integrated with the analysis of catch-at-age data using one of the two tag integrated models: Spatial Brownie-Peterson tagging model (SBP) and integrated tagging and catch-at-age analysis (ITCAAN) (Eveson et al. 2009, Vandergoot and Brenden 2014, Goethel et al. 2015b, 2015c, Vincent et al. 2017). Both models make use of both conventional tagging and catch-at-age data, and allow estimates of movement rates, fishing and natural mortality rates, and abundance simultaneously. The underlying assumption for both types of models is that tagged fish and their spawning population at large experienced the same movement and survival dynamics. Thus, the subsequent recovery locations of tagged fish and harvest collected from different basins reveal where the cohort of the spawning population moved to from their spawning sites (often the tagging sites). The major apparent difference between SBP and ITCAAN models is how they treat the tag return process. The SBP model treats the tag return probability as a direct product of several conditional probabilities, such as movement, survival, exploitation, tag retention and tag reporting rates. In contrast, the ITCAAN model determines the tag return probabilities indirectly, by dividing the estimates of number of tag returns to the total number of tag release, and the number of tag returns is estimated by assuming tagged fish undergo the same dynamics as their spawning population. Chapter 3 extends the spatial Brownie-Petersen tagging model in the context of a spatial overlap-structure and emphasizes that the ITCAAN and SBP models make common assumptions, thus paving the way for unification of these approaches. Chapter 3 also attempts to apply this approach to lake whitefish in Lake Huron.

One limitation that SBP and ITCAAN models both have is that the spatial structure assumption of fish populations is assumed to be known in the assessment model, which has only

been either diffusion or overlap (Eveson et al. 2009, Vandergoot and Brenden 2014, Goethel et al. 2015b, 2015c). The diffusion and overlap structures assume either 0% or 100% homing, but actual fish populations will not generally perfectly match these categories. Tagging, genetic or hard structure chemical signature data can be used to estimate the degree of homing. When evaluated, even fish considered to have a high degree of homing have had some straying (Thorrold, 2001; Rooker et al., 2008; Ebener et al., 2010). In this dissertation, I refer to populations with some degree of homing that is less than 100% as having an “incomplete overlap structure”. Previous studies have revealed that ignorance or misspecification of spatial structure in stock assessment models can lead to biased estimates of population parameters and stock status, inappropriate harvest targets, and depletion of local populations, especially for low productivity populations (Ying et al. 2011, Molton et al. 2012, Li et al. 2015). Thus, it is potentially important to incorporate information on the degree of homing for populations with an incomplete overlap structure, rather than treating them as having either an overlap or diffusion structure in spatially structured tagging models or stock assessment models (Stewart et al. 2003). Chapter 4 develops and uses simulations to evaluate an approach for allowing for degree of homing in assessment models, thus generalizing and evaluating the approach of Chapter 3.

Background on Lake Huron lake whitefish

Lake whitefish (*Coregonus clupeaformis*), an ecologically, culturally, and economically important freshwater fish species in US, contributes to around 7 million kg commercial harvest annually from the Upper Great Lake regions, which equals \$16.6 million catch value (Lynch et al. 2015). In North America, Lake Huron had the largest commercial fishery for lake whitefish. However, in the recent decades, significant declines in abundance, body size, growth had been observed for lake whitefish in Lake Huron, and the potential explanations that have been

proposed include direct effects of invasive species, indirect effects of invasive species through displaced native prey, altered food web productivity, and harvest pressure (Ebener et al. 2008, 2010, Pothoven and Madenjian 2008, Rennie et al. 2009). According to past genetic and tagging studies (Pothoven and Madenjian 2008, Ebener et al. 2010, Eberts et al. 2017), lake whitefish also changed their diet, distribution, and movement in Lake Huron, which possibly was due to their needs of finding quality resources in the changed food web. Despite the changes in the spatial distribution of lake whitefish, they are still largely managed as “unit” stock in each management units and the movement between management units are ignored in current stock assessments. In Lake Huron, lake whitefish have been historically managed based on defined management districts, with different size, habitat, fishing pressure, and political jurisdictions. A total of 25 such districts were defined for Lake Huron, 17 in Canadian waters and 8 in U.S. waters. In recent years some of these units have been merged into larger assessment units and formal assessment and management activities have never been implemented for all the originally defined units. In each assessment unit within 1836 Treaty Waters, a statistical catch-a-age model have been used to estimate year- and age-specific abundance and mortality rates (Truesdell and Bence 2016). Similar assessments have been applied in a number of assessment areas in the U.S. outside of the treaty waters, and in Canadian waters (Adam Cottrill and Ji He personal communications). Simulation studies suggested that the pooling approach that has been adopted in Lake Huron can provide better management performance than separate assessment in the long term (Li et al. 2015). However, both separate and pooled approaches did not account for the actual spatial structure of populations in their stock assessment models. For lake whitefish in Upper Great Lakes region, both genetic and conventional tagging studies suggested that many lake whitefish populations intermixed during non-spawning period but have strong spawning site

fidelity so that most survived fish would move back to natal site to spawning during each year spawning season (Ebener et al. 2010, Stott et al. 2010). Only very limited studies have been done to understand lake whitefish movement patterns in upper Great Lakes, and movement rates between populations are unknown and have never been incorporated into current stock assessment models. Thus, it is of general interest to understand lake whitefish movement patterns from both individual and population perspectives, and the long-term effects of incorporating population spatial structure into current stock assessment models. Thus, my dissertation not only seeks to develop and apply general methodology to better understand fish movement and how to use this information in assessment models, but to also to make specific application to lake whitefish in Lake Huron.

Objectives

The specific objectives for each dissertation chapter are as follows:

Chapter 1. Develop a Bayesian variable selection framework for analyzing how factors affect net movement distance of individual fish in large water areas. Fit the model using conventional tagging and other environmental data that are available for Lake Huron lake whitefish to better understand lake whitefish movements in Lake Huron.

Chapter 2. Use simulations to explore the use of spawning origin information of catch and the influence of including annual recruitment penalties as a means for improving integrated age-structured stock assessments for overlapping fish populations.

Chapter 3. Extend the spatial Brownie-Petersen tagging model for modeling multiyear tag-recovery data in a fishery context, incorporating catch-at-age, and tag monitoring data jointly, and demonstrate the method for lake whitefish populations in Lake Huron, in which an overlap spatial structure was assumed.

Chapter 4. Extend the tagging model proposed in Chapter 3 to a more comprehensive framework that allowed for a continuum of spatial structures (incomplete overlap) through modeling homing probability. Use simulations to explore how the degree of homing, the extent of spatial movements, and which types of data are used, influenced estimability of parameters of the proposed model framework.

BIBLIOGRAPHY

BIBLIOGRAPHY

- Cadrin, S.X., and Secor, D.H. 2009. Accounting for Spatial Population Structure in Stock Assessment: Past, Present, and Future. *In* *The Future of Fisheries Science in North America*. Springer Netherlands, Dordrecht. pp. 405–426. doi:10.1007/978-1-4020-9210-7_22.
- Drouineau, H., Bau, F., Alric, A., Deligne, N., Gomes, P., and Sagnes, P. 2017. Silver eel downstream migration in fragmented rivers: use of a Bayesian model to track movements triggering and duration. *Aquat. Living Resour.* **30**. doi:10.1051/alr/2017003.
- Durban, J.W., Elston, D.A., Ellifrit, D.K., Dickson, E., Hammond, P.S., and Thompson, P.M. 2005. Multisite Mark-Recapture for Cetaceans: Population Estimates With Bayesian Model Averaging. *Mar. Mammal Sci.* **21**(1): 80–92. doi:10.1111/j.1748-7692.2005.tb01209.x.
- Ebener, M.P., Brenden, T.O., Wright, G.M., Jones, M.L., and Faisal, M. 2010. Spatial and temporal distributions of lake whitefish spawning stocks in Northern lakes Michigan and Huron, 2003–2008. *J. Great Lakes Res.* **36**: 38–51. doi:10.1016/j.jglr.2010.02.002.
- Ebener, M.P., Kinnunen, R.E., Schneeberger, P.J., Mohr, L.C., Hoyle, J.A., and Peeters, P. 2008. Management of Commercial Fisheries for Lake Whitefish in the Laurentian Great Lakes of North America. *Int. Gov. Fish. Ecosyst. Learn. from Past, Find. Solut. Futur.*: 99–143. Available from <http://www.miseagrant.umich.edu/downloads/research/journals/08-310.pdf>.
- Eberts, R.L., Wissel, B., Simpson, G.L., Crawford, S.S., Stott, W., Hanner, R.H., Manzon, R.G., Wilson, J.Y., Boreham, D.R., and Somers, C.M. 2017. Isotopic Structure of Lake Whitefish in Lake Huron: Evidence for Regional and Local Populations Based on Resource Use. *North Am. J. Fish. Manag.* **37**(1): 133–148. doi:10.1080/02755947.2016.1245225.
- Eveson, J.P., Laslett, G.M., and Polacheck, T. 2009. A spatial model for estimating mortality rates, abundance and movement probabilities from fishery tag-recovery data. *In* *Modeling Demographic Processes In Marked Populations*. doi:10.1007/978-0-387-78151-8.
- Fu, C., and Fanning, L.P. 2004. Spatial considerations in the management of Atlantic cod off Nova Scotia, Canada. *North Am. J. Fish. Manag.* **24**(3): 775–784. doi:10.1577/M03-134.1.
- Gilliam, J.F., and Fraser, D.F. 2001. MOVEMENT IN CORRIDORS: ENHANCEMENT BY PREDATION THREAT, DISTURBANCE, AND HABITAT STRUCTURE. *Ecology* **82**(1): 258–273. doi:10.1890/0012-9658(2001)082[0258:MICEBP]2.0.CO;2.

- Goethel, D.R., Legault, C.M., and Cadrin, S.X. 2015a. Testing the performance of a spatially explicit tag-integrated stock assessment model of yellowtail flounder (*Limanda ferruginea*) through simulation analysis. *Can. J. Fish. Aquat. Sci.* **72**(4): 582–601. doi:10.1139/cjfas-2014-0244.
- Goethel, D.R., Legault, C.M., and Cadrin, S.X. 2015b. Demonstration of a spatially explicit, tag-integrated stock assessment model with application to three interconnected stocks of yellowtail flounder off of New England. *ICES J. Mar. Sci.* **72**(1): 164–177. doi:10.1093/icesjms/fsu014.
- Goethel, D.R., Legault, C.M., and Cadrin, S.X. 2015c. Testing the performance of a spatially explicit tag-integrated stock assessment model of yellowtail flounder (*Limanda ferruginea*) through simulation analysis. *Can. J. Fish. Aquat. Sci.* **72**(4): 582–601. doi:10.1139/cjfas-2014-0244.
- Goethel, D.R., Quinn, T.J., and Cadrin, S.X. 2011. Incorporating spatial structure in stock assessment: Movement modeling in marine fish population dynamics. *Rev. Fish. Sci.* **19**(2): 119–136. doi:10.1080/10641262.2011.557451.
- Heath, M.R., Kunzlik, P.A., Gallego, A., Holmes, S.J., and Wright, P.J. 2008. A model of meta-population dynamics for North Sea and West of Scotland cod—The dynamic consequences of natal fidelity. *Fish. Res.* **93**(1–2): 92–116. doi:10.1016/j.fishres.2008.02.014.
- Hutchings, J.A. 1996. Spatial and temporal variation in the density of northern cod and a review of hypotheses for the stock's collapse. *Can. J. Fish. Aquat. Sci.* **53**(5): 943–962. doi:10.1139/f96-097.
- Li, Y., Bence, J.R., and Brenden, T.O. 2015. An evaluation of alternative assessment approaches for intermixing fish populations: a case study with Great Lakes lake whitefish. *ICES J. Mar. Sci.* **72**(1): 70–81. doi:10.1093/icesjms/fsu057.
- Louca, V., Lindsay, S.W., and Lucas, M.C. 2009. Factors triggering floodplain fish emigration: Importance of fish density and food availability. *Ecol. Freshw. Fish* **18**(1): 60–64. doi:10.1111/j.1600-0633.2008.00323.x.
- Lucas, M.C., and Baras, E. 2000. Methods for studying spatial behaviour of freshwater fishes in the natural environment. *Fish Fish.* **1**(4): 283–316. doi:10.1046/j.1467-2979.2000.00028.x.
- Lynch, A.J., Taylor, W.W., Beard, T.D., and Lofgren, B.M. 2015. Climate change projections for lake whitefish (*Coregonus clupeaformis*) recruitment in the 1836 Treaty Waters of the Upper Great Lakes. *J. Great Lakes Res.* **41**(2): 415–422. Elsevier B.V. doi:10.1016/j.jglr.2015.03.015.
- McMahon, T.E., and Matter, W.J. 2006. Linking habitat selection, emigration and population dynamics of freshwater fishes: A synthesis of ideas and approaches. *Ecol. Freshw. Fish* **15**(2): 200–210. doi:10.1111/j.1600-0633.2006.00130.x.
- Mitchell, T.J., and Beauchamp, J.J. 1988. Bayesian Variable Selection in Linear Regression. *J. Am. Stat. Assoc.* **83**(404): 1023–1032. Available from

<http://www.jstor.org/stable/2290129>.

- Molton, K.J., Brenden, T.O., and Bence, J.R. 2012. Control rule performance for intermixing lake whitefish populations in the 1836 Treaty waters of the Great Lakes: A simulation-based evaluation. *J. Great Lakes Res.* **38**(4): 686–698. doi:10.1016/j.jglr.2012.09.014.
- Porch, C.E., Turner, S.C., and Powers, J.E. 2001. Virtual population analyses of atlantic bluefin tuna with alternative models of transatlantic migration: 1970–1997. *Int. Comm. Conserv. Atl. Tunas Collect. Vol. Sci. Pap.* **52**(3): 1022–1045.
- Pothoven, S.A., and Madenjian, C.P. 2008. Changes in Consumption by Alewives and Lake Whitefish after Dreissenid Mussel Invasions in Lakes Michigan and Huron. *North Am. J. Fish. Manag.* **28**(1): 308–320. doi:10.1577/M07-022.1.
- Radinger, J., and Wolter, C. 2014. Patterns and predictors of fish dispersal in rivers. *Fish Fish.* **15**(3): 456–473. doi:10.1111/faf.12028.
- Rennie, M.D., Sprules, W.G., and Johnson, T.B. 2009. Factors affecting the growth and condition of lake whitefish (*Coregonus clupeaformis*). *Can. J. Fish. Aquat. Sci.* **66**(12): 2096–2108. doi:10.1139/F09-139.
- Rothschild, B.J. 2007. Coherence of Atlantic cod stock dynamics in the Northwest Atlantic Ocean. *Trans. Am. Fish. Soc.* **136**(3): 858–874. doi:10.1577/T06-213.1.
- Schlaff, A.M., Heupel, M.R., and Simpfendorfer, C.A. 2014. Influence of environmental factors on shark and ray movement, behaviour and habitat use: a review. *Rev. Fish Biol. Fish.* **24**(4): 1089–1103. doi:10.1007/s11160-014-9364-8.
- Stewart, I.J., Quinn, T.P., and Bentzen, P. 2003. Evidence for fine-scale natal homing among island beach spawning sockeye salmon, *Oncorhynchus nerka*. *Environ. Biol. Fishes* **67**(1): 77–85. doi:10.1023/A:1024436632183.
- Stott, W., VanDeHey, J.A., and Sloss, B.L. 2010. Genetic diversity of lake whitefish in lakes Michigan and Huron; sampling, standardization, and research priorities. *J. Great Lakes Res.* **36**(SUPPL. 1): 59–65. doi:10.1016/j.jglr.2010.01.004.
- Truesdell, S.B., and Bence, J.R. 2016. A Review of Stock Assessment Methods for Lake Trout and Lake Whitefish in 1836 Treaty Waters of Lake Huron , Lake Michigan and Lake Superior Quantitative Fisheries Center Technical Report T2016-01. (March). doi:10.6084/m9.figshare.3123949.
- Vandergoot, C.S., and Brenden, T.O. 2014. Spatially varying population demographics and fishery characteristics of Lake Erie walleyes inferred from a long-term tag recovery study. *Trans. Am. Fish. Soc.* **143**(1): 188–204. doi:10.1080/00028487.2013.837095.
- Vincent, M.T., Brenden, T.O., and Bence, J.R. 2017. Simulation testing the robustness of a multiregion, tag-integrated assessment model that exhibits natal homing and estimates natural mortality and reporting rate. *Can. J. Fish. Aquat. Sci.* **74**(11): 1930–1949. doi:10.1139/cjfas-2016-0297.

Ying, Y., Chen, Y., Lin, L., and Gao, T. 2011. Risks of ignoring fish population spatial structure in fisheries management. *Can. J. Fish. Aquat. Sci.* **68**(12): 2101–2120. doi:10.1139/f2011-116.

CHAPTER 1

WHY DO LAKE WHITEFISH MOVE LONG DISTANCES IN LAKE HURON? BAYESIAN VARIABLE SELECTION OF FACTORS EXPLAINING FISH MOVEMENT DISTANCE

Abstract

Understanding fish movement patterns is vital for stock assessment and fishery management. We used a variable selection procedure in a Bayesian framework to understand what factors most likely affect the net movement distance of individual fish based on a conventional tag-recovery study of lake whitefish populations in Lake Huron during 2003-2011, where fish of this species with spawning site fidelity were tagged during the spawning season and recovered throughout the year. We found that fish with greater total length, and those that were tagged and released from tagging sites near Cheboygan and Alpena, Michigan, moved longer net distances than fish from other tagging sites. Habitat conditions also had a profound effect on net movement distance. We found that shorter movement distances by lake whitefish can be expected if the relative density of the benthic amphipod *Diporeia* spp. was higher near the tagging site during the recovery year. We also found evidence that lake whitefish may start their annual spawning migration runs earlier in warmer years. More generally, our Bayesian framework for analysis of conventional tagging data has potential for wide applicability, and detailed model details and our code are provided to facilitate this.

Introduction

Many fish species move for long distances at various times during their life cycle, and movements made by individuals vary from regular and predictable migration to less-predictable resource driven nomadism (Runge et al. 2014). Most previous research that evaluated changes in fish spatial locations focused on either the triggering factors or distance between initial and final

fish location (e.g., Albanese et al., 2004; Radinger and Wolter, 2014), or on estimating net movement/migration rates of populations (Polacheck et al. 2006, Vandergoot and Brenden 2014).

Fish movement is essential from both conservation and management perspectives. Movement behavior can influence how fish are distributed, whether their populations persist in the face of ecosystem changes, and how stocks are assessed. Fish movement can further influence ecological interactions and evolution (Lidicker and Stenseth 1992). Management problems such as inaccurate assessment results, or inappropriate catch limits, can occur when actual fish movements do not agree with the spatial assumptions made in stock assessments and management decisions, which can result in local population depletion and population collapse (Mitchell and Beauchamp 1988, Hutchings 1996, Fu and Fanning 2004, Rothschild 2007, Li et al. 2015).

Despite its ecological and management importance, understanding of fish movement patterns in time and space, and how movements are related to environmental variables, is still limited. Moreover, most previous research that focused on the triggering factors (i.e., factors causing the initiation of movement) and net fish movement distance were limited to stream fish, given the easy calculation of net distance moved from conventional tagging data. Much less is known about movement of fish that live in large water areas. Most of which is known has been derived from electronic tagging data, although there are many long-term conventional tagging programs. While technological advances make the use of acoustic or pop-up tags increasingly useful, conventional tags are still more widely used for estimating population size, mortality, and tracking individual growth, given their lower price. Conventional tagging data can also provide information on the location at tag release and tag recovery, which could be used for the

estimation of movement route and intensity (e.g., net fish movement distance) (e.g., Albanese et al., 2004; Gilliam and Fraser, 2001).

The goal of this study was to develop a model framework for analysis of how factors impact the distance fish move from when they are tagged until they are recovered ('net fish movement distance' hereafter) in a larger water body, based on conventional tag-recovery results. We based our research on several lake whitefish (*Coregonus clupeaformis*) spawning stocks in Lake Huron of the Laurentian Great Lakes of North America. As an ecological and economically important fish species in the Great Lakes, lake whitefish have been found to move freely among multiple management units during the non-spawning period, but show a high degree of natal homing, so nearly all mature fish return to spawn at the same location each year (Ebener et al. 2010b). Previous research on lake whitefish movement patterns provides a useful platform for us to derive a priori hypotheses about the potential factors that influence movement. Since the establishment of dreissenid mussels in the early 1990s, the ecosystem of four of the five Great Lakes have changed substantially, including an overall decrease in the density of lake whitefish's preferred food- *Diporeia* spp. (Mohr and Nalepa 2005, McNickle et al. 2006, Barbiero et al. 2011). In this context, Rennie et al. (2012) evaluated the relationship between lake whitefish migration distance and growth rate, and found that the least mobile population of lake whitefish was supported by a remnant *Diporeia* spp. population. Ebener et al. (2010b) found that stock identity and season of recapture affected net movement distance most strongly, while the influence of variables such as sex, year, fish total length, and time at large was weaker. Although the role of temperature has not been directly implicated in explaining patterns in the fish movement, the association between lake whitefish harvest and surface water temperature suggested that such a connection may exist (Price et al. 2003).

The pioneering studies of net movement distance used either a regression-tree based approach or ANOVA models to test whether net movement distance varied significantly in association with the factors they evaluated (e.g., Albanese et al., 2004; Ebener et al., 2010b; Radinger and Wolter, 2014). Because some studies estimated the effects of different factors as additive (i.e., causing a given distance change rather than a percentage change in net movement distances), it is hard to generalize the results from studies with different spatial and temporal scales. When jointly considering multiple factors and continuous covariates, the ANOVA approach can provide only a rough picture of the continuous relationship between net movement distance and explanatory factors. Thus, a more thorough regression analysis is needed. The regression-tree based approach seeks to approximate nonlinearity and interactions in the relationships between the net movement distances and multiple factors by recursively partitioning the data points according to the categorization of the factors (Ebener et al. 2010b). Such partitioning may have difficulty in interpreting the effects, if the observations from the same tag or recovery area happen to be separated into different branches of the tree. Some regression-tree applications have partitioned data by site (i.e., different sites on different branches), and this can make it difficult to develop a general understanding of movement (Ebener et al. 2010b). In addition, although it is possible for regression-tree based approaches to rank or select variables based on variable importance measures, they do not provide any further insight of the uncertainty associated with their rankings or selections. Also information criteria, such as Akaike's information criterion and the Bayesian information criterion, commonly used as penalization terms for the number of parameters in model, are not applicable for nonparametric tree-based models (Claeskens and Hjort 2008).

We therefore considered a global linear regression model that accounts for joint effects of multiple factors and the heterogeneity among sites, to study the relationship between the net movement distance and individual factors. We further conducted a variable selection procedure under a Bayesian framework to explore the plausibility of alternative regression models that include various explanatory variables, and assess the associated uncertainty. Bayesian variable selection treats the regression model itself as random among all possible models with different sets of variables. Thus, it accounts for model uncertainty in the overall assessment of uncertainty by making inferences on how probable alternative models are after consideration of the data. The implementation of Bayesian variable selection via the reversible jump Markov chain Monte Carlo (rjMCMC) (Green 1995) procedure is substantially more efficient in exploring the model space than the traditional approaches such as all-subsets-regression (Woznicki et al. 2016). While we believe our approach has substantial advantages over regression-tree approaches, it could miss some nonlinear effects that could be identified by regression-trees. Thus, as a check on robustness we compared our results with those from regression-tree methods.

We considered how net distance moved from tagging to recapture locations changed monthly and over years, and how this net movement pattern depended upon tagging location. In addition, we considered how life history traits, namely total length, and sex, and habitat features, namely *Diporeia* spp. density and water temperature, played a role in these net movement patterns. Thus, the variables we considered as potential explanatory factors in this study were tagging year, recovery year, recovery month, year(s) between tag and recovery, fish total length, sex, tagging (spawning) site, and the habitat variables based on *Diporeia* spp. density and growing degree days.

Our goal was to provide not only insight on how those factors influenced lake whitefish movement in Lake Huron, but also a model framework for analyzing movement mechanism based on conventional tagging data. Although Bayesian variable selection in linear regression is a long-established approach (Mitchell and Beauchamp 1988), it was rarely used in ecology or more specifically for uncovering explanations for movements (Drouineau et al. 2017, Ethier et al. 2017). Drouineau et al. (2017) used a Bayesian state-space model to analyze the effects of different environmental factors in triggering migration of silver eel in fragmented rivers. Ethier et al. (2017) used Bayesian models and variable selection to evaluate how environmental variables influenced regional variation in population trends of Bobolink. Both studies used a mixture distribution of priors (i.e., normal plus zero-inflation), which were estimated using a Gibbs sampler. However, their variable selection procedure did not introduce a penalty such as BIC for increasing number of selected variables. Also the Gibbs sampler usually involves scanning all variables at each iteration, which could be computational expensive, especially when the number of candidate variables is large.

To the best of our knowledge, this study is the first to apply the Bayesian variable selection approach to compare the effects of various factors on fish net movement distance by introducing an explicit prior penalty on model complexity, and the most comprehensive to date in terms of the range of factors affecting whitefish movement. To avoid sampling all indicators within a Gibbs sampler circle as in Drouineau et al. (2017) and Ethier et al. (2017), we adopt the reversible jump MCMC algorithm for model exploration that mimics stepwise selection and subsets regression technique, which is more computationally efficient. Thus our research introduces an approach to fish movement studies, which has the potential to be much more effectively interrogate a large number of predictor variables. To facilitate usage of our approach,

we provide the open-source code for MATLAB program which is online available at to implement the method.

Methods

Data collection, selection, and calculation of net-movement distance

Lake whitefish were tagged and released in a study coordinated by one of us (Mark P. Ebener) at 21 individual tagging sites from nine spawning stocks in Lake Huron from late October through December (i.e., spawning season) of 2003-2006. Total length (mm) of all 35,285 tagged fish were measured before release, spatial coordinates of the tagging and release location, and date of release were recorded for each fish. Lake whitefish were tagged on or very near the spawning grounds and subsequently killed when recovered by the commercial or recreational fishery. The commercial fishing season for lake whitefish is not closed in Ontario waters during the spawning season, but it is closed in Michigan waters. Thus, fish tagged and released at Detour, Cheboygan, Alpena, and Saginaw Bay (Figure 1.1) were extant 1-4 weeks before being subjected to fishing and tag recovery. At Burnt Island, the Fishing Islands, and Sarnia fish were also tagged during the spawning season, but commercial fishing was occurring simultaneously during tagging so they had little time to be extant prior to tag recovery. Recovery happened from December 2003 until December 2012, with the majority being recovered by commercial fishermen, and the rest recovered during fishery surveys. Subsets of the data used here were previously reported by Ebener et al. (2010a, 2010b), and details of the tagging methodology are given by Ebener et al. (2010a).

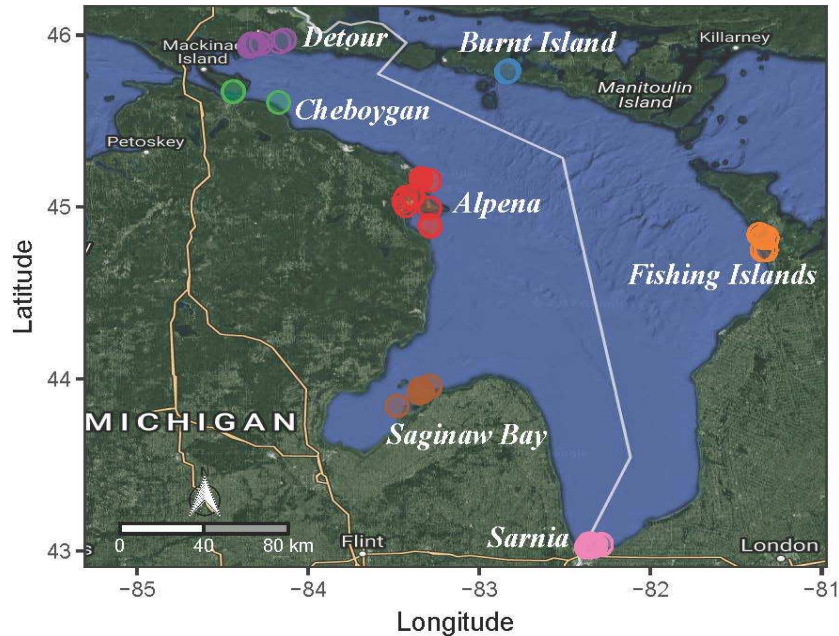


Figure 1.1. Map of the study area (Lake Huron) and seven tag release (spawning) sites. Of total 1368 recoveries, 659 were from Detour, 300 from Cheboygan, 243 from Burnt Island, 42 from Saginaw Bay, 43 from Sarnia, 56 from Alpena, and 25 from Fishing Islands.

Our analysis focused on drivers of net movement distance of lake whitefish tagged and recovered in Lake Huron. We thus restricted attention to recoveries for which net distance movement could be calculated and for which explanatory variable data were available. Only recoveries that had location information recorded (either by latitude and longitude or by 10-minute by 10-minute statistical grid, treated as though recovered at the grid center) were considered. In addition, we excluded observations from fish that were recovered within two days of release, as well as those without their recapture date, sex, or total length recorded (i.e., explanatory variables). We also removed fish that were recovered from Lake Michigan because of our focus on movement within Lake Huron and because our explanatory variables were from Lake Huron. We further excluded recoveries from two tagging sites that each produced only two total recoveries, and the two fish recovered during 2012. Thus of the total of 2,098 reported lake whitefish recoveries, 1,368 recoveries were used in this study. Details of data exclusion are

described in Supplementary Table 1.A1. These recovered fish had total lengths between 375-667 mm at the time of tagging, and were tagged and released from seven spawning sites (Figure 1.1).

We used log-transformed net movement distance as a response variable because net movement distances were highly skewed. We calculated net movement distance based on the shortest water distance between tagging and recovery locations, using a Dijkstra type shortest path algorithm (Vincenty, 1975; online Appendix 1.A). We standardized log-transformed net movement distance by subtracting the mean and dividing by standard deviation prior to analysis.

Explanatory variables

We hypothesized that net movement distance for lake whitefish in Lake Huron would be influenced by 1) life history traits, which included total length, and sex; 2) temporal factors, which included tagging year (tag_Y), recovery year (rec_Y), recovery month (rec_M), and year(s) between tagging and recovery (year_lag); and 3) habitat condition, which included *Diporeia* spp. density, and growing degree days; and 4) tagging (spawning) sites. These hypotheses, related variables, and the expected sign of the associated coefficients, if hypotheses were supported, are in Table 1.1. Due to the strong spawning site fidelity of lake whitefish (i.e., nearly all lake whitefish move back to where they born each year during the spawning season), we only considered the habitat conditions during the recovery year as a predictor. That is, the net movement is in actuality the net movement since the prior spawning season. We used relative *Diporeia* spp. density, which was the *Diporeia* spp. density of the release location divided by the mean of all sampled stations in Lake Huron for that year. The U.S. EPA Great Lakes National Program Office collected *Diporeia* samples every August since 1999 at 12 Lake Huron stations (Barbiero et al. 2011). The release location density was defined as the density at

the sampled location closest to the release location. Our hypothesis was that lake whitefish tended to stay near their tagging locations when *Diporeia* density was higher in that vicinity.

We proposed two alternative hypotheses for the relationship between growing degree days (GDDs) (i.e., also known as thermal time, a weather-based indicator about heat acumination for assessing fish growth; e.g., Chezik et al., 2014) and lake whitefish net movement distance, and these led to two distinct sets of GDD variables. These two sets were used in two alternative analyses. We calculated GDDs based on mean daily (daytime) surface temperatures from the Great Lakes Surface Environmental Analysis (GLSEA) remote sensing surface water temperature data (See online Appendix 1.A).

Case 1 (GDD hypothesis 1)—Lake whitefish respond to growing conditions they had experienced during the current year. Thus, they would tend to be closer to their tagging (spawning) site when the growing degree days (GDD) at the tagging location was greater than the lake average GDD during that same time period. This led us to define the explanatory variable relative GDD difference (“GDD_Diff”), calculated as: $GDD_Diff = (GDD_{tag} - GDD_{lake}) / GDD_{lake}$, where GDD_{tag} and GDD_{lake} are the cumulated non negative degree days (°C. days) that exceeded 5°C (Rennie et al. 2009) at the tagging location or for the lake-average, respectively, from the first day of the recovery year to the day of recovery.

Table 1.1. Summary of candidate variables/terms and their interpretation and relationship to hypotheses. A. For continuous variables, “Hypothesis” (first column) states our a priori hypothesis associated with the variable, and the second column indicates sign of associated coefficient that would support that hypothesis. B. Similarly for interaction terms, but here a single hypothesis (our GDD Hypothesis 2) is associated with all interaction terms, and the second column describes the interpretation of coefficients and the pattern in their sign that would support the hypothesis. C. For categorical (dummy) variables we did not have explicit a priori hypothesis for the sign of coefficients but did hypothesize that these factors could influence net distance. For these variables one level of a factor is the baseline with coefficient fixed at zero, and this level (category) is given in the first column and interpretation of the sign of other coefficients in the second column. For A through C, “X” in the “GDD H1” column indicates that the variable was a candidate variable/term in our variable selection process for the GDD Hypothesis 1 Case, and the GDD H2 column likewise indicates if the variable/term was a candidate variable for the GDD Hypothesis 2 Case. The last row summarizes the total number of candidate variables for each GDD hypothesis.

A. Continuous Variables				
Variable Name	Hypothesis	If support, sign of covariate	GDD H1	GDD H2
Length	Greater total length, fish range further from tagging site.	>0	X	X
years_lag	Longer lag between tagging and recovery year, recoveries tend to be further from tagging site.	>0	X	X
<i>Diporeia</i>	Higher relative <i>Diporeia</i> spp. density near the tagging site, fish stay closer to their tagging site.	<0	X	X
GDD_Diff	Greater GDD at the tagging location than the lake average, fish stay closer to their tagging site.	<0	X	
B. Interaction Terms				
Names	Hypothesis	Sign of coefficient	GDD H1	GDD H2
Sep× GDD _{lake} Oct× GDD _{lake} Nov× GDD _{lake} Dec× GDD _{lake}	In years when lake average GDD is higher there is a shift in spawning timing. This is reflected in shorter net distances in one or more adjacent spawning months, and longer net distances in later months.	If <0, fish are closer to tagging site with higher GDD _{lake} during that month, and if >0 further away. Support for hypothesis would be >0 coefficient for one or more adjacent months of Sep – Nov, and <0 coefficient for later months.		X

Table 1.1. (cont'd)

C. Categorical (Dummy) Variables				
Variable Names	Baseline category (effect was 0)	Interpretation of coefficient	GDD H1	GDD H2
tag_site: Cheboygan tag_site: Burnt_Island tag_site: Alpena tag_site: Sarnia tag_site: Saginaw_Bay tag_site: Fish_Islands	Fish tagged and released from Detour (Figure 1)	If >0, larger net distance than baseline; if <0, shorter net distance than baseline	X	X
sex: Female	Male tagged fish	Same as above	X	X
rec_Y: 2003 rec_Y: 2004 rec_Y: 2005 rec_Y: 2007 rec_Y: 2008 rec_Y: 2009 rec_Y: 2010 rec_Y: 2011	Tagged fish recovered from 2006	Same as above	X	X
rec_M:7 rec_M:8 rec_M:9 rec_M:10 rec_M:11 rec_M:12 rec_M:1 rec_M:2 rec_M:3 rec_M:4 rec_M:5	Tagged fish recovered in June of each year	Same as above	X	X
tag_Y: 2003 tag_Y: 2005 tag_Y: 2006	Fish tagged and released from 2004	Same as above	X	X
Total number of candidate variables for each case (without intercept)			33	36

Case 2 (GDD hypothesis 2)—The spawning season of lake whitefish would be shifted earlier in the year, in years for which GDDs accumulated faster, because individual fish would reach a physiological status allowing spawning earlier under such conditions. Preliminary model

fits without a GDD effect indicated that lake whitefish were generally closer to the spawning location during September through December, than at other times of the year. We therefore assumed that GDD might potentially influence net movement distance (to varying degrees) only during these months. Thus, we added four additional interaction variables (recovery month \times GDD_{lake}) for September through December recoveries. We used GDD_{lake} because fish would be living and feeding away from their spawning/tagging sites until moving to those sites for spawning.

After creating dummy variables and choosing the category with the largest number of observations as the baseline category for each factor, we have a total of 34 (for GDD hypothesis 1 case) and 37 (for GDD hypothesis 2 case) candidate variables including the intercept (Table 1.1). Note that there was no dummy variable created for the baseline category (i.e., tagging site: Detour, recovery month: June, tagging year: 2004, recovery year: 2006, or sex: Male), because it was defined as zero for all other categories for that factor. All explanatory variables were standardized like net movement distance.

Model framework

We used Bayesian variable selection to identify the highly probable subsets of predictors for the linear regression and, given a set of predictors, we assessed likely parameter values. Given the Bayesian approach we used, inferences were based on a posterior distribution, which depends jointly on assumed prior distributions and the likelihood of the data. Model components (i.e., regression model, prior distributions, and likelihood) are described in Section 2.31 (Model Description) and how we used Markov chain Monte Carlo (MCMC) techniques to derive

posterior distributions in Section 2.3.2. A separate model selection process was conducted for the two cases (GDD hypotheses).

Model description

Each possible model is of the form:

$$Y = X_J \beta_J + \epsilon, \epsilon \sim N(0, \tau^2 I_N) \quad (1.1)$$

where Y is the response variable (i.e., log-transformed net movement distance) with N observations, X_J is the $N \times q_J$ design matrix (containing data for the predictors included for that regression), β_J is a vector of parameter coefficients (an intercept included in every model plus $q_J - 1$ additional coefficients for the predictor variables included for that model) and ϵ is the residual error. We assumed here homogenous, normal and independent residual errors, with variance τ^2 . We assumed independent errors given the relatively large distances between tagging sites (Figure 1.1) and because tagging and recovery spatial factors were included as potential explanatory variables. As described in online Appendix 1.B, I_N could be replaced with a selected correlation matrix. A model with a specific subset of selected variables is represented by J , which formally is an index set, that maps the q variables in the selected model to the larger set of p possible variables.

The $\beta_J = (\beta_1, \beta_2, \dots, \beta_q)'$ had a normal prior $\beta_J \sim N(0, \tau^2 \Lambda)$, where $\Lambda = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_q\}$. The λ_j s were modeled as arising from a higher level inverse-gamma prior distribution ('hyperprior') with shape parameter a_λ and scale parameter b_λ . We assumed a hyperprior with inverse-gamma density for τ^2 with shape a_τ and scale b_τ . The hyperparameters were set to the values $a_\lambda = a_\tau = 2$ and $b_\lambda = b_\tau = 0.001$, which correspond to a rather dispersed prior distribution.

The normal prior with diagonal variance-covariance matrix for the λ_j s represents a decision to use a Bayesian counterpart to Ridge regression. The λ_j s represent the signal to noise ratio of the effects in the model, and their magnitude played a role in whether an effect was included and the size of selected models. Modeling them as arising from a hyperprior (rather than specifying their values) allowed for adaptive learning on which variables to include during the model search process.

We included an intercept in all models to account for the grand mean level of Y , as is often done for variable selection. There are a total of 2^{p-1} possible models (i.e., an intercept-only model, all possible models with one additional variable, all possible models with two additional variables, etc.). We specified the prior probability of each model as arising from the product of a prior probability for a model of a given size (i.e., $\pi(q)$), multiplied by the probability of a specific model given its size:

$$\pi(\mathcal{J}) = \pi(\mathcal{J}_q, q) = \pi(\mathcal{J}_q|q)\pi(q) \quad (1.2)$$

We let $\pi(q) \propto \exp\{-\kappa q\}$ for integer q from $\{1, 2, \dots, p\}$. Here \propto means “proportional to” up to a constant that is irrelevant in making inferences about the hyper-parameter κ . This placed higher prior probability on models with smaller size, as is consistent with common practice in variable selection, and the rate at which the prior probability falls as model size increases was determined by κ . We set $\kappa = \log(N)/2$, which is analogous to a BIC-type penalty on the number of selected variables (Schwarz 1978). Conditional on q , each model \mathcal{J}_q had an equal chance of being selected, i.e., $\pi(\mathcal{J}_q|q) = 1/\binom{p-1}{q-1}$ for $q > 1$, and for $q = 1$ no selection is needed.

Characterization of the posterior Distribution using MCMC

We used Markov chain Monte Carlo methods to determine the posterior distribution. We used a hybrid reversible jump technique (rjMCMC), because it performs well when selecting among different sets of variables, which involved trans-dimensional states of Markov chain (Green 1995, Woznicki et al. 2016). Our procedure involved running multiple chains and combining converged portions of these into one set of "retained samples." The retained samples were summarized to highlight desired properties of the posterior distributions.

Details on the implementation of the hybrid of rjMCMC for model search and Gibbs sampler for parameters given the model, as well as procedures for evaluating MCMC convergence and producing the retained samples are given in online Appendix 1.B. We summarized the posterior distributions for regression model parameters in two ways:

Variable-wise summary— This provided a summary conditional on the j -th variable being selected. This was based on summarizing all samples included in the final MCMC chains for a model that included the j^{th} variable. For the corresponding β_j , the posterior mean and 95% (equal probability tail) credible intervals were constructed from these samples. As a measure of the importance of each variable we also calculated the marginal inclusion probability (Barbieri and Berger 2004), as the proportion of all retained MCMC samples that included the j^{th} variable in the model.

Model-wise summary— This was conditional on one specific model \mathcal{J} in the posterior samples, and thus was based only on retained MCMC samples for that model. We provide such summaries for the 12 "top" models. Here models are ranked based on the posterior probability, calculated as the proportion of all retained MCMC samples that were model \mathcal{J} . For the top

models, we summarized the posterior distributions of the β_j s for all variables in \mathcal{J} , again in terms of the posterior mean and 95% credible interval.

Model diagnosis, simulation study, and comparison with tree-based methods

We used the posterior predictive assessment of model fitness using the χ^2 -discrepancy (Gelman et al. 1996), based upon which we calculated the Bayesian p-value for the top models in both GDD hypothesis cases. We also conducted simulations to evaluate how well our Bayesian variable selection procedure can discover the true set of important variables and estimate the corresponding effects, under five different scenarios with varying combinations of true predictor variable effects. We also applied two tree-based methods to our data, and compared the top variables from tree-based methods, the gradient boosting regression tree method (Ethier et al. 2017) and the random forests approach (Breiman 2001), to the selected variables from our variable-wise summary. Detailed methods for our diagnostic procedures, simulations, and tree-based applications are given in online appendices C, D, and E respectively, and performance statistics resulting from the simulations and tree-based methods are also presented in the appendices.

Results

The posterior distributions of the number of selected variables were similar for the two GDD cases and suggested that the most probable model sizes had 6 and 7 variables including an intercept (Figure 1.2). However, the selected variables were quite different (see Section 3.1).

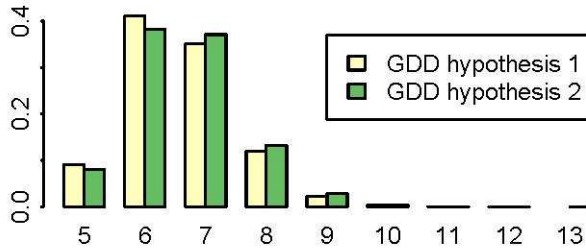


Figure 1.2. Posterior distributions for the number of selected variables (i.e., $q - 1$). The x-axis starts at 5 because all models selected at least five variables.

Variable-wise summary

GDD hypothesis 1 Case: There were 10 variables with 95% credible intervals that did not cover 0, which we define as "consistent effects" (Figure 1.3). Variables that had consistent effects generally had high marginal inclusion probability, and more generally variables with higher probability of inclusion tend to have more of their posterior distribution on one side of zero (Figure 1.3). The six top variables (length, tagging site: Cheboygan and Alpena, *Diporeia*, and recovery months October and November) had marginal inclusion probability above 0.75 (i.e., they are selected by more than 75% of the total posterior samples). The variable recovery month September also had a relatively large marginal inclusion probability (0.40). The other variables that were detected as consistent effects had substantially lower marginal inclusion probability (<0.07) are: years lag, tagging site Fishing islands, and recovery month December. According to the posterior mean of those 10 variables with consistent effects, fish with greater length, longer lag between the tagging and recovery years, released at tagging site Cheboygan, Alpena, and Fishing Islands, and recovered in December had greater net movement distance, while fish released at the tagging site with higher density of *Diporeia*, and recovered during September, October, and November had shorter net movement distance. Our first GDD hypothesis was not supported by the variable selection results because the 95% credible interval of the associated effect covered 0, and had a marginal inclusion probability of only 0.004.

GDD hypothesis 2 Case: As was true for the previous case, consistency of effects and the marginal probability of inclusion were positively associated (Figure 1.4). The six top variables in terms of marginal inclusion probability (length, tagging site: Cheboygan and Alpena, *Diporeia*, recovery month November, interaction effect between lake-average GDD and recovery month October) were similar to the top variables for GDD hypothesis 1. The major exceptions were for the recovery month October and the GDD associated variables (Figure 1.4). Consistent with the results for GDD hypothesis 1, fish with greater length, longer lag between the tagging and recovery years, and released at tagging site Cheboygan, Alpena, and Fishing islands had greater net movement distance, while fish released at the tagging site with higher density of *Diporeia*, and recovered during September, and November had shorter net movement distance. The effect of recovery month October had a similar negative posterior mean, although the effect was less consistent. The less consistent effect of October is likely associated with the inclusion of the GDD associated variables for Hypothesis 2. Our second hypothesis of GDD was well supported by the variable selection results. The interaction effect between lake-average GDD and recovery month October, and the interaction effect between lake-average GDD and recovery month November were both consistent, with a negative posterior mean and the former was smaller than the latter. That is, fish tended to have smaller net movement distance in October and November if the lake-average GDD was greater, and the effect was larger in October than that in November. On the other hand, the interaction between lake-average GDD and recovery month December was also consistent, but with a positive posterior mean, which suggested that fish tended to have greater net movement distance in December if the lake-average GDD was greater. An overall interpretation of these effects is a shift in the spawning season in association with

GDD, with more fish close to the spawning grounds by October and having moved away by December, when GDD was higher.

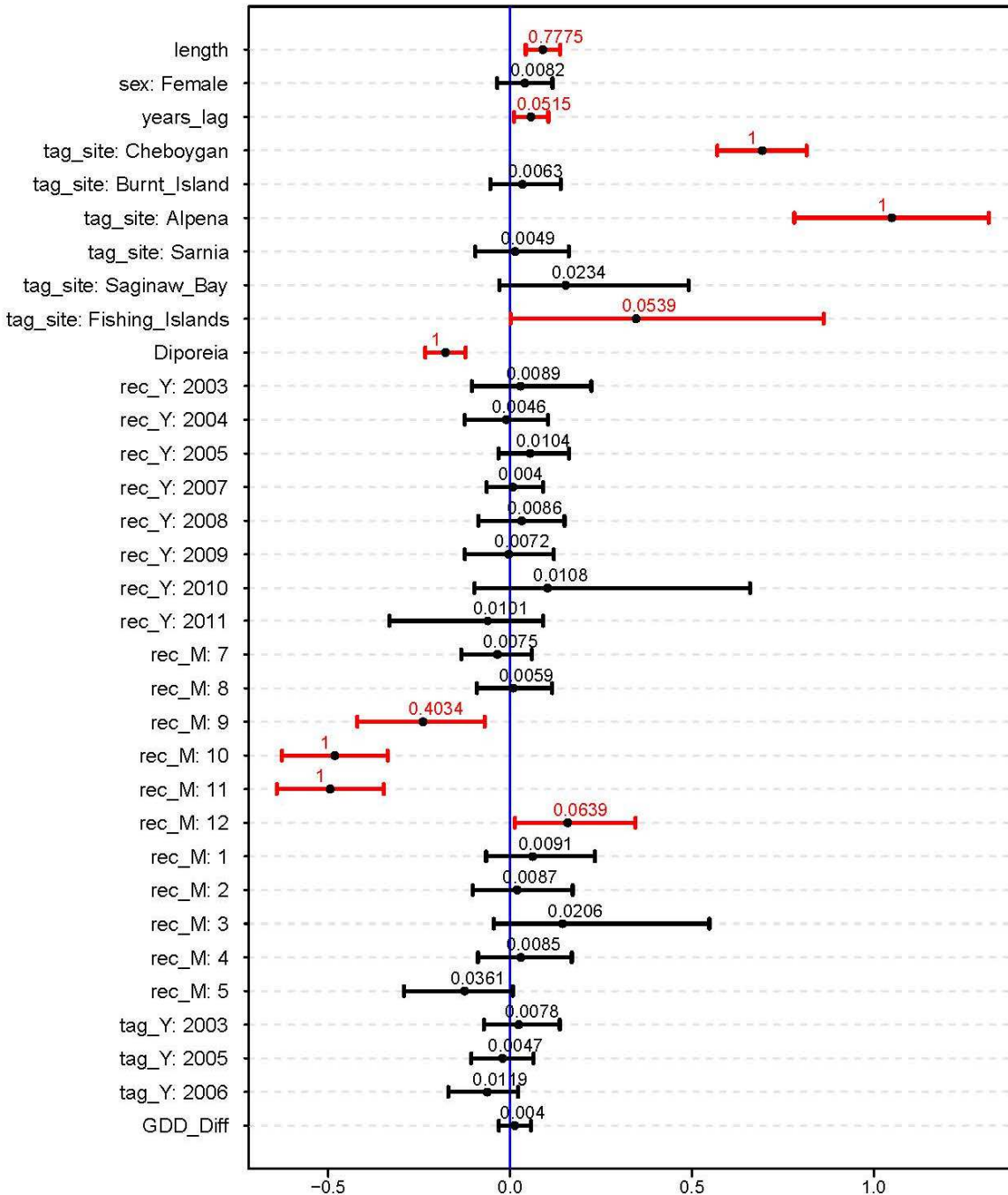


Figure 1.3. Variable-wise summary results (posterior mean with 95% credible intervals) of the effect of variables (the β_j), with variables named on y-axis for the case with GDD hypothesis 1. Bars are highlighted by red color when the 95% credible interval does not cover 0, which is defined as a consistent effect. The number above each bar is the marginal inclusion probability.

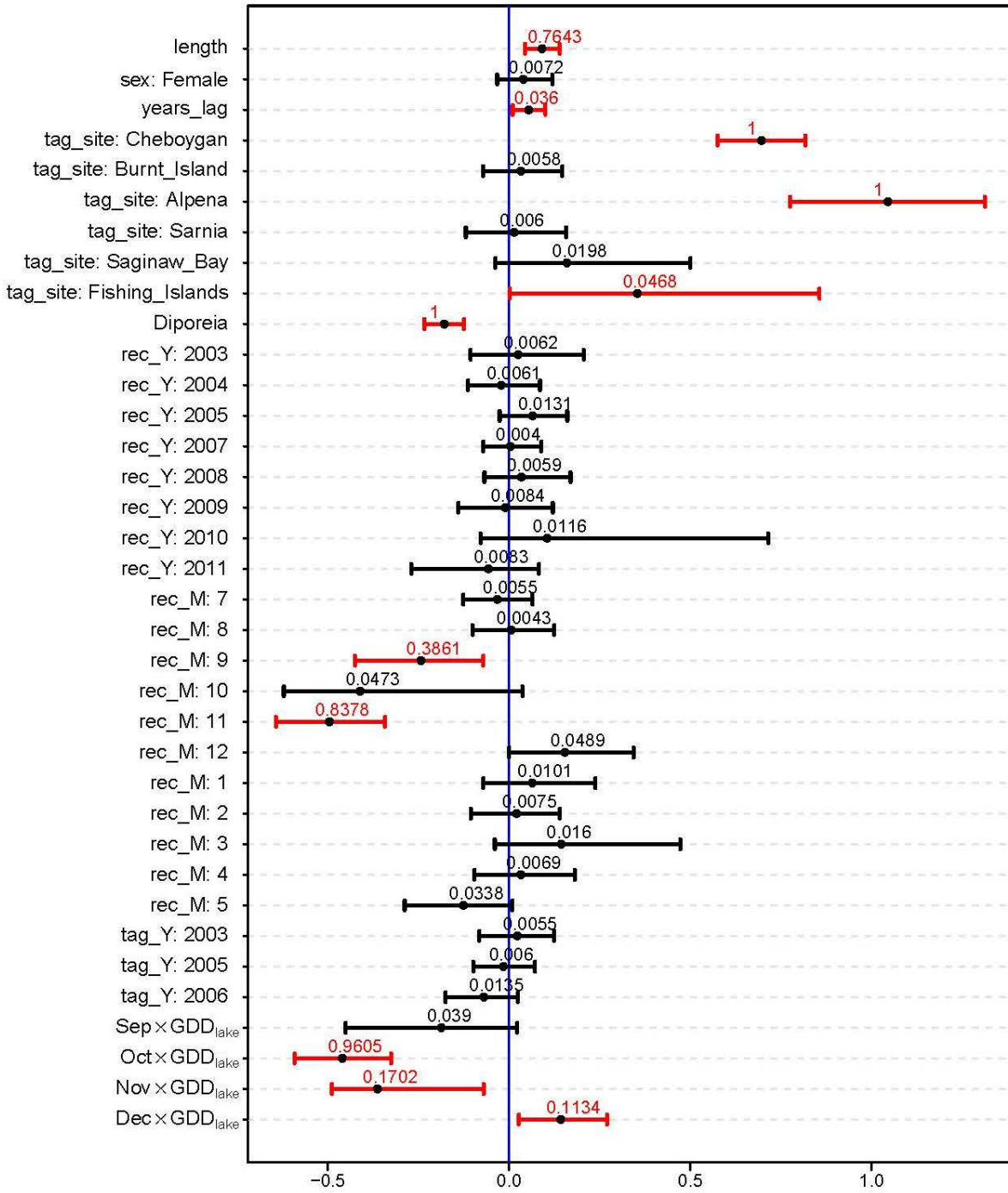


Figure 1.4. Variable-wise summary results (posterior mean with 95% credible interval of the effect β_j for the j th variable, as indicated in y-axis) for the case with GDD hypothesis 2. Bars are highlighted by red color when the 95% credible interval does not cover 0, which was defined as a consistent effect. The number above each bar is the marginal inclusion probability.

Model-wise summary

Given the variable selection result support our second GDD hypothesis, we only present model-wise summary for GDD hypothesis 2 case. One or more of the four interaction variables (recovery month \times GDD for the fish that were recovered from September, October, November, and December) were included in at least one out of the top 12 models. In addition to the three variables with marginal inclusion probability equals 1 in Figure 1.4 (*Diporeia*, tagging site: Cheboygan and Alpena), the interaction effect recovery month October \times GDD was also included in all 12 top models (Figure 1.5). Recovery month November was included in nine of the 12 top models (all but models 5, 7, and 10), while the interaction variable November \times GDD_{lake} was included in the other three top models. Total length of tagged fish was included in eight out of the 12 top models, recovery month September was included in four out of the top models, the interaction variable December \times GDD_{lake} was included in three out of the top models, and recovery month December and the interaction variable September \times GDD_{lake} were included in one out of the top models. The top two models both had a posterior probability greater than 0.14. These models were similar. The best model (i.e., the highest posterior probability model) included six variables and the second best model included all those variables, plus recovery month September. Most estimated β_j s are consistent across the top models, suggesting the effect of a variable was relatively uninfluenced by the presence of other variables in the models.

The best model (Model 1 in Figure 1.5) for the fit with GDD hypothesis 2 is summarized in Table 1.2. From the best model, fish that were tagged and released from tagging sites Cheboygan and Alpena had longer net distance than fish released at other tagging sites. Lake whitefish with greater total length also tended to have greater net distance. Fish that were

recovered in November consistently had shorter net distance than fish recovered in other months. In addition, shorter movement distance could be expected if the relative *Diporeia* density was higher near the spawning locations during the recovery year. The interaction term of month October and lake-average GDD resulted in shorter net distance when lake-average GDD was high.

Table 1.2. Posterior mean and 95% credible intervals for parameters of the highest posterior probability model.

Variable	Mean	Lower	Upper
Rec_M:11	-0.49	-0.63	-0.35
Oct × GDD _{lake}	-0.45	-0.58	-0.32
<i>Diporeia</i>	-0.17	-0.22	-0.12
length	0.09	0.05	0.14
tag_site: Cheboygan	0.69	0.57	0.80
tag_site: Alpena	1.04	0.78	1.30

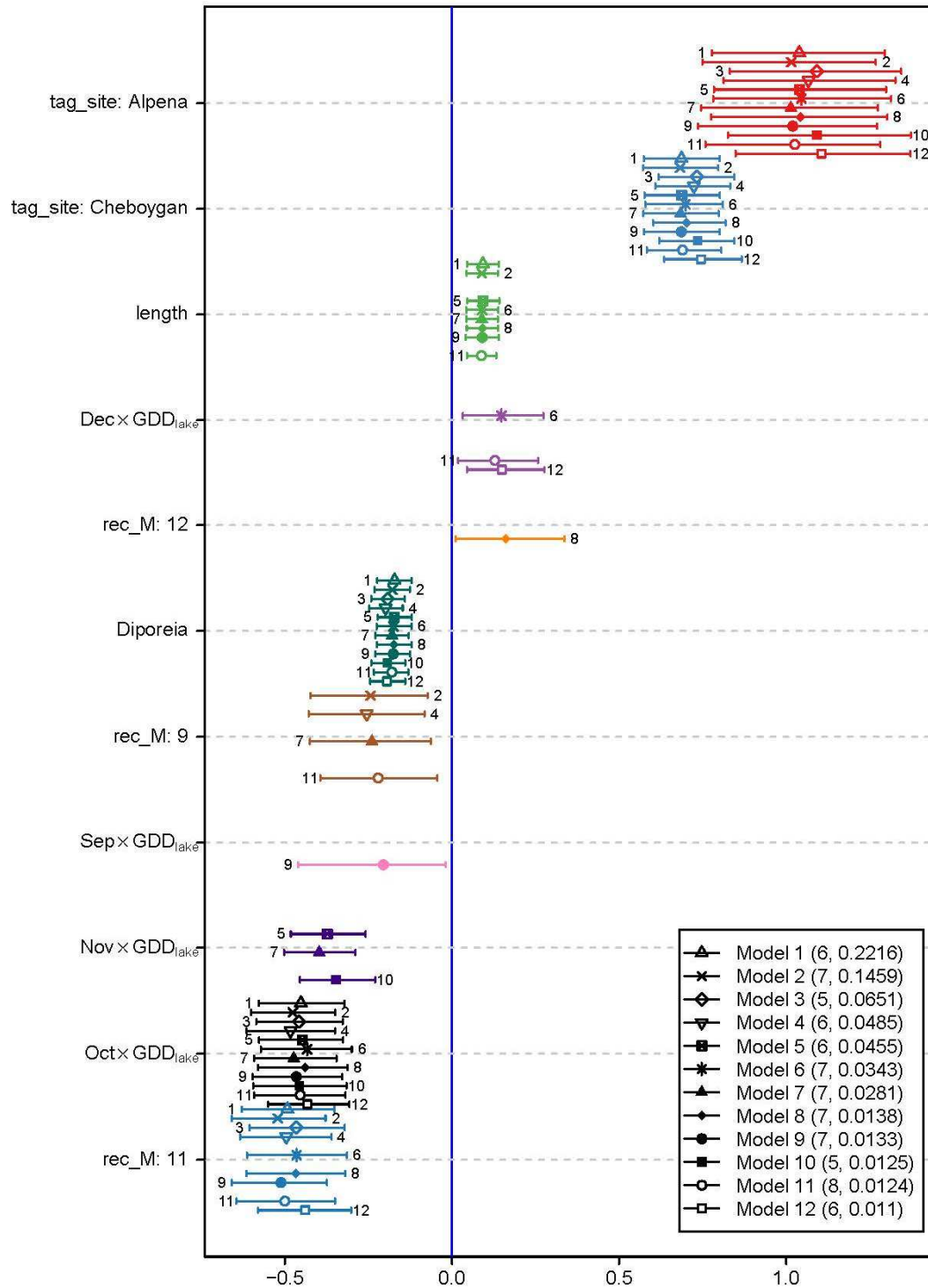


Figure 1.5. Model-wise summary for top 12 models ranked according to their posterior probability mass, for the case of GDD hypothesis 2. Variables that were included in the top 12 models are given on the y-axis. Horizontal bar represents posterior 95% credible intervals and symbols on each bar the posterior mean for each coefficient included in a model, with the associated model given to the left of the bar. Thus when more bars are given for a variable it was included in more models.

Model diagnosis, simulation study, and comparison with tree-based methods

Model diagnosis— There is no evidence for lack-of-fit of the top models under both GDD hypothesis cases. In particular the scatterplot of predicted and realized χ^2 appear consistent with a 1:1 relationship (Figure 1.A1 in online Appendix 1.C) and the Bayesian p-values are much larger than 0.05, indicating that the null hypothesis that the observed data follow the hypothesized model is not rejected. We also did a residual analysis for the top model of both GDD hypothesis cases, and plotted averaged standard residuals for the MCMC samples associated with those top models versus selected (including both continuous variables and two way combinations of categorical predictors). We did not observe any suspicious patterns from the plot given: 1) all residuals are nearly symmetric about zero, majority within (-3, 3), according to the 3-sigma rule, 2) there were no obvious trends in variation or mean across different values of the predictors.

Simulation study— In general, our BVS method had consistent performance at identifying important variables, and in identifying an appropriate model under scenarios with varying combinations of candidate variables (see online Appendix 1.D). Effects of interactions, and of continuous and categorical variables were all likely to be selected when they actually had effects, and not to be selected when they did not have effects on the response variable. Across all scenarios, the true model was very likely to be included in the top two models (i.e., probability ≥ 0.9), and most likely to be our top model (i.e., probability ≥ 0.74).

Comparison with tree-based methods— The top variables from both tree-based methods in Figure 1.A3 (online Appendix 1.E) are consistent with Bayesian variable selection (BVS) results, although there were several exceptions. The first exception was for the GDD hypothesis 1 case, where GDD_Diff was not selected as important variable by BVS, but was selected as top

variables by both boosted regression tree and random forest approaches. We believe that this is due to several high-leverage GDD_Diff observations (Figure 1.A4 in online Appendix 1.E), which the regression tree methods see as nonlinear effects. A second exception, also for the GDD hypothesis 1 case, was that fish length had a high inclusion probability (0.78 with BVS) and was also a top variable for boosted regression trees but was not included in the top list for the random forests approach. A third exception was that the rank of the variable September was lower for the tree-based approaches than for BVS, and this was true for both GDD hypotheses, albeit the three approaches rank variable importance in different ways (probability of inclusion for BVS, see X axis of Figure 1.A3 and Figure 1.A5 for tree-based methods).

Discussion

The goal of this study was to develop a model framework for analysis of how factors impact net fish movement distance in a larger water body, based on conventional tag-recovery results, and apply the framework to lake whitefish spawning stocks in Lake Huron of the Laurentian Great Lakes of North America. Our framework used a data-driven Bayesian variable selection (BVS) method, where the candidate variables represented hypothesis about drivers of net movement distance. The hypotheses we evaluated were that the net movement distance of adult lake whitefish in the main basin of Lake Huron was related to 1) fish total length, 2) sex, 3) tag and release year, 4) recovery year, 5) recovery month, 6) year(s) between tagging and recovery, 7) *Diporeia* spp. density near the spawning locations relative to the lake-wide *Diporeia* spp. density, 8) relative difference between the tagging site and lake-wide growing degree days, and 9) the interaction term between lake-wide growing degree days and recovery month. Some of the above hypotheses were well supported by the results presented.

There was a consistent positive relationship between lake whitefish net movement distance and fish total length at the time of tagging. This is consistent with conclusions from previous studies of stream-dwelling fish, in which longer movement and home range was observed for larger fish (Gunning and Shoop 1963, Gatz and Adams 1994). This greater movement may be due to the increasing mass-specific bioenergetic costs of mobility with decreasing body size (Roff 1991). Minns (1995) also found that the home range is related to body size in freshwater fisheries and is consistently larger in lakes than in rivers.

Because of the spawning site fidelity of lake whitefish, recovery months were expected to have effects on net movement distance. Ebener et al. (2010b), analyzing some of the same data but focused on different spatial and temporal scales with fewer predictor variables, also demonstrated that season of recapture played an important role in the distance moved by lake whitefish. Here, net movement distance was found to be negatively related to recovery months September, October and November, and positively related to December. This suggested that the spawning migration movement for lake whitefish generally occurred within months from September to November, and after that, fish tended to leave their spawning site and were actually further from the spawning location than in the baseline month of June.

Past research has documented that some life history events such as reproduction can be accelerated with warmer water temperature (Forseth et al. 1999). For example, the spawning of walleye has occurred earlier with earlier ice-out related to warmer temperature (Schneider et al. 2010). We found similar patterns in our study. When lake average GDD was higher, lake whitefish tended to move or stay closer to their spawning sites from September to November, and to be further away from their spawning sites in December. This suggests that fish may start their annual spawning migration runs earlier in warmer years after acquiring and processing

energy needed for spawning. The underlying mechanism could be that fish have to either achieve a critical condition before the cost of migration/spawning can be offset (Forseth et al. 1999), or to accumulate enough energy to survive a winter starvation period before spawning.

Although the decline of *Diporeia* spp. density in the Laurentian Great Lakes due to the establishment of dreissenid mussels has been argued as the main reason of lake whitefish expanding their movement range (Ebener et al. 2010b, Rennie et al. 2012), we know of no other direct evaluation of an effect of *Diporeia* density on movement. Our study evaluated this hypothesis by including relative *Diporeia* spp. density as a predictor for lake whitefish net movement distance, and we found that when relative *Diporeia* spp. density was high near the spawning grounds, lake whitefish tended to stay closer to their spawning site. This implied that fish might expand their foraging area when *Diporeia* density was low near their preferred habitat. Our analysis also found an effect of the relative density of *Diporeia* within a year, which suggests a pattern related to the density of this prey, not just a general change in movement over time throughout the lake as *Diporeia* declined.

Lake whitefish tagged and released from the tagging sites Cheboygan and Alpena had consistently greater net distance than those released from other areas. The underlying reasons may be related to the bathymetry and shoreline features of Lake Huron. Deep water (>80 m) near Cheboygan and Alpena may restrict the movement of Cheboygan and Alpena spawning stocks to north-south direction where there is a large area with relative shallow water. In contrast, the spawning stocks in Detour and Burnt Inlands may be constrained from moving south by the deep water in north of the main basin of Lake Huron, so that they tended to move in the east-west direction. Considering the shape of Lake Huron and the locations of those spawning stocks,

movement in the north-south direction allows longer movement distance than in the east-west direction.

There was similarity but also some differences in variable selection between our Bayesian variable selection and tree-based methods. One notable difference between tree-based methods and the Bayesian method is in the inclusion of GDD difference in GDD hypothesis 1 case for the tree-based methods but not by BVS. The overall neutral effect and low importance for the BVS was apparently because a few high-leverage points were treated as noise. By recursively partitioning the data according to different ranges of predictors, the tree-based methods are less sensitive to those points. However, such localized results based on small samples can hardly provide any general predictability. Rättsch et al. (2001) also found that overfitting can occur for regression tree-based methods using a boosting algorithm when there is a lot of noise.

Our BVS method can be used for various different species and any water system meeting our input requirements. For conventional tagging studies done in large lakes (e.g., Lake Huron as in our case) or oceans, shortest water distance can be used as response variable; while for a tagging study done in a river, a river network needs to be built /considered for calculating (net) movement distance. Given that our Bayesian variable selection method penalizes the number of selected variables, it has the potential to perform well for other cases with more candidate explanatory variables than we used in our application. In addition, the approach is adaptable to situations where residuals might be correlated. We assumed no such correlations given the spatial distribution of tagging sites and inclusion of spatial covariates (e.g., tagging sites), but in other situations there could be spatial structure that should be accounted for in random part of the model. In such cases correlations could be made a function of a measured quantity like distance

between tagging sites, and our code and detailed description of the model in the supplement outlines how this can be done. In addition, our Bayesian method also allows extra flexibility such as including: (1) random effects to cope with grouping variables with a large number of outcomes, which can greatly improve the prediction by better explaining the variability; (2) prior information for the effects of variables with flexible choices that can be leveraged from a broad catalog in the Bayesian variable selection literature. Thus we believe our work established a framework that could facilitate additional studies of animal movement based on conventional tagging data.

We made several simplifying assumptions and choices in our analysis. Firstly, we assumed 100% spawning site fidelity, so for the environmental factors *Diporeia* spp. density and GDDs, only data for the year of recovery were used. While fidelity is likely not 100%, available data suggest it is quite high for lake whitefish (Ebener et al. 2010b). Secondly, the T_0 used for the calculation of cumulative GDD is 5°C (Rennie et al. 2009), but it is possible that this is not the best threshold or that fish are responding to temperature in a different or more complex fashion than we assumed. We believe that violation of the 100% fidelity assumption and the GDD assumptions would act to obscure effects of *Diporeia* and tagging site rather than cause us to discover artefactual effects. Thirdly, we assumed similar tag reporting rates across all recovery basins, so data were not weighted across different recovery basins. Violation of this assumption could be influencing details of our results. However, we suspect the larger qualitative effects are real rather than artifacts of such a violation. If there were dominating differences in tag reporting rates among basin, we would have expected that to be reflected in consistent tagging site effects for sites within basins, which we did not see in our results.

Acknowledgements

Funding for this project was provided by Quantitative Fisheries Center supporting partners, Michigan DNR through the Partnership for Ecosystem Research and Management program, and grants from the Great Lakes Fishery Commission Fishery Research Program and the Great Lakes Fishery Trust (Project 2012.1250). We thank the commercial fisherman and the staff of the Inter-Tribal Fisheries and Assessment Program, the Hammond Bay Biological Station, the Great Lakes Fishery Commission, and Michigan DNR for their helps in tagging fish in very uncomfortable weather conditions. We thank Dr. Richard P. Barbiero for providing *Diporeia* abundances data, and Dr. Lacey Mason for providing surface water temperature data. We acknowledge the support of the Michigan State University High Performance Computing Center and the Institute for Cyber-Enabled Research. We wish to thank the Modeling Subcommittee of the Technical Fisheries Committee for 1836 Treaty-ceded waters for providing input on our research. This is publication 2017-XX of the Michigan State University Quantitative Fisheries Center.

APPENDICES

Table 1.A1. Recoveries that were excluded in this study.

Recoveries description	Recoveries number
Recoveries without recovery date recorded	107
Recoveries out of the main basin of Lake Huron	15
Recoveries without length recorded	6
Recoveries with unknown sex	340
Recoveries within two days of release	22
Recoveries without location information	234
Recoveries from year 2012	2
Recoveries from tagging site Douglas Point	2
Recoveries from tagging site South Bay Mouth	2

Appendix 1.A: Calculation of shortest water distance and GDD

Shortest water distance between release and recovery locations was calculated by using a Dijkstra type shortest path algorithm (Vincenty 1975), which constructs a path through a great circle distance weighted network so as to minimize the distance. To build the network (i.e., adjacency matrices), we first stored the connection (edge) from each grid to its four nearest neighbor grids of all 312 10-minute by 10-minute statistical grids of Lake Huron, with weights for each edge equal to the great circle distance between the grid centers by using `distVincentyEllipsoid` function in `geosphere` package (Hijmans et al. 2012). We then manually removed the cross-land connections in the network given fish cannot move overland. After that, we used `graph.adjacency` and `shortest.path` function from R package `igraph`, to build the adjacency matrix and calculate the shortest water distance (Csárdi and Nepusz 2006). The origin and destination grids for calculating the shortest path were determined by the release and recovery locations. We matched the coordinates for each release and recovery locations with their closest grid centers, and used such grid centers as the origin and destination. If exact coordinates were available for either release or recovery location, we added an additional great circle distance from the exact coordinates to the center of the matched grid.

GDD_{tag} and GDD_{lake} is given by: $GDD_i = \sum_{j=1}^D |SWT_{i,j} - T_0|$, where i is either tag or lake index, D is the day of the year a fish was recovered, SWT is the daily average surface water temperature of either the tagging location ($i = tag$) or the whole Lake Huron ($i = Lake$), and T_0 equals 5°C (Rennie et al. 2009). The daily average surface water temperature data were collected by GLSEA of Coast Watch Great Lakes node (<http://coastwatch.glerl.noaa.gov/>). The GLSEA produced a mean daily daytime surface temperature product covering the entire surface area of the Great Lakes at approximately 1800m resolutions, from 1995 to 2013, and we only

used data from 2003 to 2011. To use these data, we matched all release coordinates with their closest cell center of the new coordinate system.

Appendix 1.B: Model implementation

To make our model more general, I_N in equation 1 is replaced with a selected correlation matrix $R(\theta)$, which allows incorporation of the dependence among samples. For example, one would include the Conditional Autoregressive (CAR) to account for the spatial dependence via introducing a neighborhood structure (e.g., adjacency for areal data). In that case, the correlation matrix $R(\theta)$ can be written as:

$$R(\theta) = (M - \theta W)^{-1}$$

where W is the binary neighborhood matrix and M is the diagonal matrix storing the number of neighbors for each sample. θ is a parameter that can be updated through the MCMC iterations.

As an alternative example for $R(\theta)$, for spatial data with irregularly spaced locations, θ can be the range parameter in the frequently adopted Gaussian process models that capture the dependence via distance. Specifically, one popular choice for correlation structure under Gaussian process model is the exponential model, which can be written as

$$R(\theta) = e^{-D/\theta}$$

where D is the distance between samples. In this study, we used neither of these, and assumed homogenous, normal, and independent residual error, $R(\theta) = I_N$, because there is no natural way of defining the neighborhood system when these are not areal data, and the distance has been used as response variable in the model.

We proceed with the reversible jump Markov chain Monte Carlo (Green 1995) algorithm embedded in a Gibbs sampler for updating the parameters $\{J, \beta_j, \tau^2, \Lambda, \theta\}$. More specifically, we sequentially followed the steps described below at each iteration for updating each set of parameters from their full conditional distributions given the data and the remaining parameters:

Step 1: Update the index set \mathcal{J} given $\{\Lambda, \theta\}$

Conditional on the current \mathcal{J} , we proposed a new model (set of selected variables \mathcal{J}^*) from a proposal density function $g(\cdot)$. We specify $g(\cdot)$ to be either a forward-selection move (F-move) by including an extra variable that is not in \mathcal{J} , or a backward-elimination move (B-move) by excluding one existing variable in \mathcal{J} , with equal prior weight $P(\text{F-move} | \mathcal{J}) = P(\text{B-move} | \mathcal{J}) = 0.5$ for trans-dimensional exploration of \mathcal{J} with $1 < q < p$. Namely, we "tossed a coin" to determine if $q \rightarrow q + 1$ or $q \rightarrow q - 1$, and selected at random with equal probability a new variable to add from the remaining $p - q$ variables, or selected at random a variable to drop from the $q - 1$ variables (other than intercept) currently in the model. Note that for the boundary values $q = 1$ and $q = p$, we considered a forward-selection or backward-selection move, respectively, with probability 1.

We next considered the proposal densities for the parameters under the model, which plays a critical role for the posterior mixing. For a forward-selection move, we proposed the new Λ^* , τ^{2^*} and $\beta_{\mathcal{J}^*}$ under \mathcal{J}^* from a proposal density $h(\cdot)$. The probability of accepting the proposal $(\mathcal{J}^*, \Lambda^*, \tau^{2^*}, \beta_{\mathcal{J}^*})$ was calculated as:

$$\min \left\{ 1, \frac{g(\mathcal{J}|\mathcal{J}^*)}{g(\mathcal{J}^*|\mathcal{J})} \times \frac{\pi(\mathcal{J}^*, \Lambda^*, \tau^{2^*}, \beta_{\mathcal{J}^*} | Y)}{\pi(\mathcal{J}, \Lambda, \tau^2, \beta_{\mathcal{J}} | Y)} \times \frac{h(\Lambda, \tau^2, \beta_{\mathcal{J}} | \mathcal{J}^*, \tau^{2^*}, \beta_{\mathcal{J}^*}, \Lambda^*, \mathcal{J})}{h(\Lambda^*, \tau^{2^*}, \beta_{\mathcal{J}^*} | \mathcal{J}, \tau^2, \beta_{\mathcal{J}}, \Lambda, \mathcal{J}^*)} \right\}$$

under our choice of

$$h(\Lambda^*, \tau^{2^*}, \beta_{\mathcal{J}^*} | \mathcal{J}, \tau^2, \beta_{\mathcal{J}}, \Lambda, \mathcal{J}^*) = \pi(\Lambda^* | \mathcal{J}) \times \pi(\tau^{2^*} | \mathcal{J}^*, \Lambda^*, Y) \times \pi(\beta_{\mathcal{J}^*} | \mathcal{J}^*, \tau^{2^*}, \Lambda^*, Y)$$

More specifically, we first proposed λ_{q+1} for the newly added variable only, from its prior density, to obtain Λ^* from Λ . The proposal density for Λ^* in the denominator of the

acceptance ratio was then canceled out with its prior density in the numerator. We then proposed τ^{2*} from its posterior density, which marginalizes β_j out

$$\pi(\tau^{2*} | \mathcal{J}^*, \theta, \Lambda^*, Y) \propto \pi(\tau^{2*}) \int_{R^q} \pi(Y | \mathcal{J}^*, \tau^{2*}, \theta, \Lambda^*, \beta_j) \times \pi(\beta_j | \mathcal{J}^*, \tau^{2*}, \Lambda^*) d\beta_j$$

given this is an Inverse-Gamma density with shape $a_\tau + N/2$ and scale $b_\tau + Y'V(\theta, \Lambda^*)Y/2$, with $V(\theta, \Lambda^*) = R(\theta)^{-1} - R(\theta)^{-1}X_{j^*}(X'_{j^*}R(\theta)^{-1}X_{j^*} + \Lambda^{*-1})^{-1}X'_{j^*}R(\theta)^{-1}$ (Johnson et al. 2002). Next we generated β_{j^*} from its full conditional density $\pi(\beta_j | \mathcal{J}^*, \tau^{2*}, \theta, \Lambda^*, Y)$, which is a multivariate Normal distribution $N(\mu_\beta, \Sigma_\beta)$ with covariance matrix $\Sigma_\beta = \tau^{2*}(X'_{j^*}R(\theta)^{-1}X_{j^*} + \Lambda^{*-1})^{-1}$ and mean $\mu_\beta = (X'_{j^*}R(\theta)^{-1}X_{j^*} + \Lambda^{*-1})^{-1}X'_{j^*}R(\theta)^{-1}Y$. When drawing the parameters from their full conditional posterior density, they were marginalized from the acceptance ratio (Johnson et al. 2002). Consequently, the probability of acceptance reduces to a simpler form

$$\min \left\{ 1, \quad \frac{\pi(\mathcal{J}^*)}{\pi(\mathcal{J})} \times \frac{g(\mathcal{J} | \mathcal{J}^*)}{g(\mathcal{J}^* | \mathcal{J})} \times \frac{\pi(Y | \mathcal{J}^*, \Lambda^*)}{\pi(Y | \mathcal{J}, \Lambda)} \right\}$$

where the prior model ratio $\frac{\pi(\mathcal{J}^*)}{\pi(\mathcal{J})} = \frac{\exp\{-\kappa(q+1)\}/\binom{p-1}{q}}{\exp\{-\kappa q\}/\binom{p-1}{q-1}} = \exp\{-\kappa\} q/(p-q)$, and the

proposal model ratio $\frac{g(\mathcal{J} | \mathcal{J}^*)}{g(\mathcal{J}^* | \mathcal{J})} = \frac{P(B - move | \mathcal{J}^*)/q}{P(F - move | \mathcal{J})/(p-q)} = (p-q)/q$ when this search did not

involve the boundary values of q . The acceptance ratio then became a factor of $\exp\{-\kappa\}$

multiplied by the marginal likelihood ratio $\frac{\pi(Y | \mathcal{J}^*, \Lambda^*)}{\pi(Y | \mathcal{J}, \Lambda)}$. The marginal likelihood $\pi(Y | \mathcal{J}, \Lambda)$ is a

centered multivariate student's T-distribution with the logarithm of density (Johnson et al. 2002) determined by

$$-\left(\frac{1}{2}\right) \log |I_q + \Lambda X'_{j^*}R(\theta)^{-1}X_{j^*}| - (a_\tau + N/2) \log(1 + Y'V(\theta, \Lambda)Y/(2b_\tau))$$

Note at the boundary value $q = 1$, this ratio needed to be multiplied by an extra factor of $\frac{1}{2}$ given $P(\text{F-move} | \mathcal{J})=1$ rather than 0.5. For a backward-elimination move, the acceptance ratio is the reciprocal of Eq. A.1, and similar to the boundary condition for forward-elimination, we multiplied the ratio by 2.

Step 2: Update the noise level and fixed-effects (τ^2, β_j)

This pair of parameters was updated in step 1 when the proposed new \mathcal{J}^* was accepted. Since the newly proposed \mathcal{J}^* can also be rejected, in order to make sure this pair of parameters can always be updated, we further update them again in the Gibbs sampler to improve the mixing. Sampling β_j is again from the multivariate Normal distribution $N(\mu_\beta, \Sigma_\beta)$ in Step 1. Here, however sampling τ^2 will be conditional on β_j also, which is an Inverse-Gamma density with shape $a_\tau + N/2$ and scale $b_\tau + ((Y - X_j\beta_j)'R(\theta)^{-1}(Y - X_j\beta_j) + \beta_j'\Lambda^{-1}\beta_j)/2$.

Step 3: Update the signal-to-noise ratio and dependency (Λ, θ)

We draw a sample of λ_j from the Inverse-Gamma density with shape $a_\lambda + 1/2$ and scale $b_\lambda + \frac{\beta_j^2}{2\tau^2}$ for $j \in \mathcal{J}$. In our application, we assumed residual error, so $(\theta) = I_n$, and θ was an empty set. In cases where correlations are included and parameterized through θ , we recommend sampling θ from a density proportional to

$$|R(\theta)|^{-1/2} \times \exp\{(Y - X_j\beta_j)'R(\theta)^{-1}(Y - X_j\beta_j)/(2\tau^2)\} \times \pi(\theta)$$

using a Metropolis search.

Note that when the penalty parameter $v(\kappa)$ on q is assumed to be random, one can also update it in this step to allow data-driven penalty by reparametrizing κ with $v = 1 - \exp\{-\kappa\}$, assuming Beta prior $\pi(v)$ on v , and then updating v from $\pi(v|Y, q) \propto C(v)(1 - v)^{q-1}\pi(v)$.

Steps 1 to 3 were repeated through a large number of iterations until the convergence was considered committed. The convergence was monitored and we evaluated whether it was committed through examination of multiple MCMC runs with distinct starting values, by checking trace plots of full and marginal model likelihood, global parameters such as $\{\tau^2, \theta\}$, and model-dependent parameters $\{q, \lambda_j, \beta_j\}$, and the value of the corresponding potential scale reduction factors for model parameters (PSRF) close to 1, e.g., less than 1.2 for all model parameters (Brooks and Gelman 1998). Once convergence was well committed, posterior samples of $\{J, \beta_j, \tau^2, \Lambda, \theta\}$ well after the point of commitment were retained, and an overall posterior sample was retained by stacking all chains.

More specifically, we ran five independent chains for each of the GDD hypothesis cases, with each chain run for 100,000 iterations. Starting values were chosen to represent values that were believed to well dispersed (more variable than expected for the posterior distribution). We found that convergence was well committed after a burn-in period of the first 50,000 iterations. This was indicated by trajectories of the separate trace plots that converged in mean, variability and other visual patterns, and PSRFs for all model parameters close to 1. We then stacked the last 50,000 samples for each of the 5 chains. Given there was substantial correlation in the saved chains we then thinned the sample by retaining every 5th iteration in the stacked samples for analysis, to make management of the data easier without loss of substantial information. As a result, we obtained a final retained posterior sample of 50,000 sets of model parameters for each GDD case.

Appendix 1.C: Model diagnosis

We used the posterior predictive assessment of model fitness using the χ^2 -discrepancy as suggested by (Gelman et al. 1996) based upon which we calculate the Bayesian p-value. The predicted versus realized χ^2 -discrepancies with corresponding Bayesian p-value using our posterior samples for the top model of both GDD hypotheses are shown in Figure 1.A1.

Detailed description and calculation about χ^2 -discrepancy and Bayesian p-value can be found in Gelman et al. (1996) and we have included the calculation in the Bayesian variable selection code we provide online. Essentially each dot in Figure 1.A1 represents two values for each posterior sample that visits the top model: 1) the realized χ^2 -discrepancy is randomly drawn from the nominal distribution $\chi^2(n)$ with sample size n as the degree of freedom; 2) the predicted χ^2 from each posterior sample is the sum of the squared standardized residual, which is compared with the nominal distribution. The posterior predictive p-value is the average of upper-tail probability at each predicted discrepancy under the nominal distribution $\chi^2(n)$.

We also did residual analysis for the top model of both GDD hypothesis cases, and plotted averaged standard residual over MCMC samples with top model visited versus some example variables (including both continuous variables such as length and *Diporeia* density, and two way combinations of categorical predictors). We did not observe any suspicious patterns from the plot given: 1) all residuals are nearly symmetric about zero, majority within (-3, 3), according to the 3-sigma rule, 2) no obvious trend in variation across different ranges of the predictors.

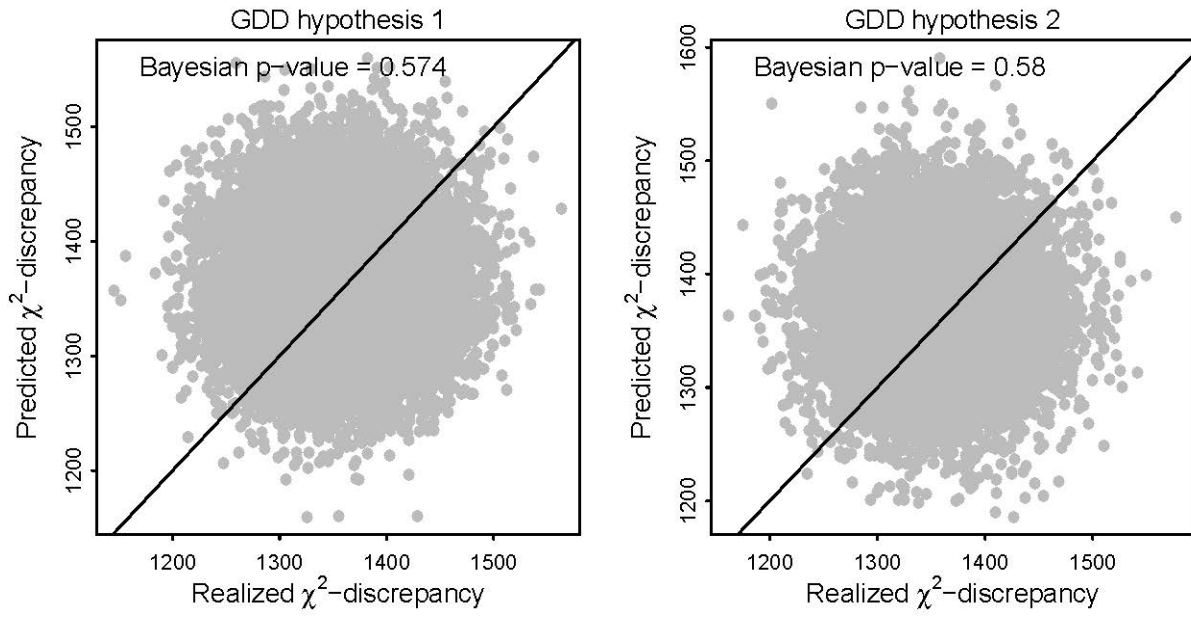


Figure 1.A1. χ^2 -discrepancy measures for both GDD hypothesis.

Appendix 1.D: Simulation study

There are five scenarios with alternative assumptions regarding the effects of month, length, and sex effects, and interactions between lake average growing degree-days and recovery month November. Based on the assumptions in each scenario and real data for the predictor variables, and an assumed standard deviation of the residuals (σ) of 0.1 which was set based on signal-to-noise ratio (we assume all month effects are detectable, i.e., the signal should exceed the noise level σ , hence we select σ with comparable scale to the smallest month effect, or weakest signal, 0.083 in Table 1.A3), we generated 100 sets of response variable outcomes (log transformed net movement distance), Then we conducted Bayesian variable selection (BVS) using the entire set of predictor variables (37 variables in the GDD hypothesis 2 case), to evaluate how well the procedure can discover the true set of important variables and estimate the corresponding effects. The five simulation scenarios are:

Scenario 1— only with month effects, no other effects.

Scenario 2— with month effects and interaction effect of Nov \times GDD_lake, no other effects.

Scenario 3—with month effects, interaction effect of Nov \times GDD_lake, and sex effect, no other effects.

Scenario 4— with month effects, interaction effect of Nov \times GDD_lake, and length effect, no other effects.

Scenario 5— with month effects, interaction effect of Nov \times GDD_lake, sex effect, and length effect, no other effects.

Each scenario represents different types of candidate variables combination. The month effects from November to next year October were simulated according to a quadratic function (i.e., to mimic natal homing), as shown in Figure 1.A2. The recapture month June was used as the reference category with zero effect. Note the effect for recapture month April was adjusted to be the same as May rather than using the value from the quadratic curve so that each moth (other than June) had a non-zero effect (see Figure 1.A2). The interaction effect of $\text{Nov} \times \text{GDD_lake}$, sex effect, and length effect were simulated treating the posterior means from the variable-wise summary of GDD hypothesis 2 case (rounding the estimates to 1 decimal place) as the true values (see Table 1.A3).

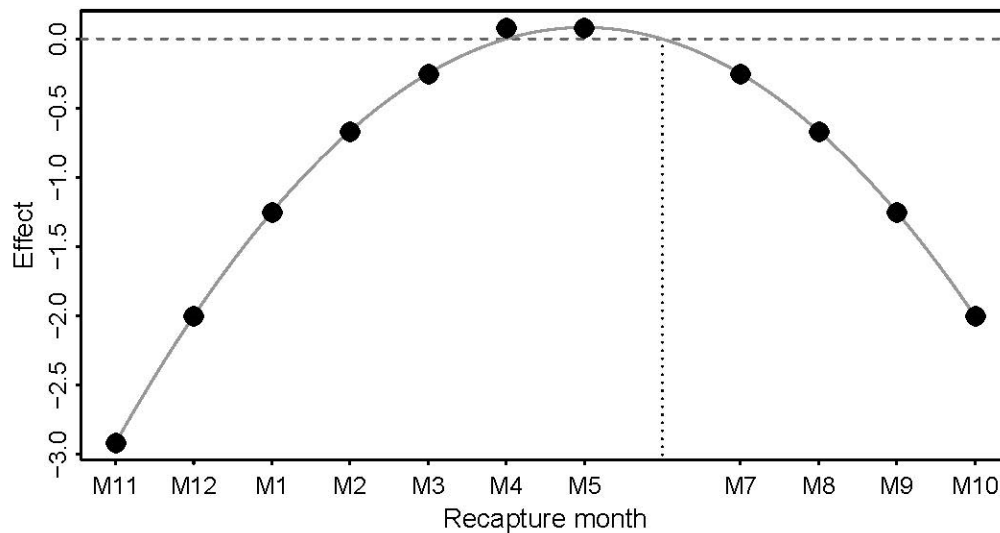


Figure 1.A2. Simulated quadratic function of month effects.

Let \mathcal{J}^* be the true index set of important variables, and \mathcal{J}_0 be the index set of estimators from variable-wise summary whose 95% credible interval did not span zero, i.e., BVS selected variables. We consider following measures based on the $B = 100$ simulations:

1. The agreement between the two sets defined as $\alpha(\mathcal{J}_0) = \text{card}(\mathcal{J}^* \cap \mathcal{J}_0) / \text{card}(\mathcal{J}^* \cup \mathcal{J}_0)$, where \cap and \cup are respectively the intersection and union of two sets, and card is the cardinality (i.e. the number of elements) of a set.

2. The probability that the selected variables are contained in the true set $P(\mathcal{J}^* \supseteq \mathcal{J}_0)$.

3. The probability that the selected variables contain the true set $P(\mathcal{J}^* \subseteq \mathcal{J}_0)$.

4. The probability that the selected variables match the true set $p_0 = P(\mathcal{J}^* = \mathcal{J}_0)$.

Those measures would evaluate how good our BVS method is at selecting important variables.

As for the top three models (\mathcal{J}_1 , \mathcal{J}_2 , and \mathcal{J}_3) from model-wise summary that ranked according to their posterior probability mass, we consider the same measures as above for the top model \mathcal{J}_1 , and additional probabilities that true model match at least one of the top two and three models. More specifically, $p_1 = P(\mathcal{J}^* = \mathcal{J}_1)$ is the probability that true model exactly matches the top model; $p_2 = P(\mathcal{J}^* = \mathcal{J}_1 \text{ or } \mathcal{J}_2)$ is the probability that the true model \mathcal{J}^* exactly matches with one of the top two models; and similarly, $p_3 = P(\mathcal{J}^* = \mathcal{J}_1 \text{ or } \mathcal{J}_2 \text{ or } \mathcal{J}_3)$ is the probability that the true model \mathcal{J}^* exactly matches with one of the top three models. Those measures would illustrate how good our BVS method is at identifying true model. Results are shown in Table 1.A2.

We further assess the model performance in terms of parameter estimation. Two measures are evaluated by averaging over the 100 simulated data sets: (1) average inclusion probability P_s ; and (2) average coverage probability P_c , i.e. percentage of the 100 simulations with the 95% credible interval constructed from variable-wise summary covering the true value of the simulation. The results are shown in Table 1.A3.

Table 1.A2. Summary of variable selection in the simulation study.

Scenario	1	2	3	4	5
$\alpha(\mathcal{J}_0)$	0.985	0.984	0.977	0.986	0.983
$P(\mathcal{J}^* \supseteq \mathcal{J}_0)$	0.82	0.84	0.8	0.87	0.84
$P(\mathcal{J}^* \subseteq \mathcal{J}_0)$	1	0.96	0.89	0.94	0.91
p_0	0.82	0.81	0.73	0.81	0.76
$\alpha(\mathcal{J}_0)$	0.985	0.975	0.979	0.977	0.981
$P(\mathcal{J}^* \supseteq \mathcal{J}_1)$	1	1	0.99	1	1
$P(\mathcal{J}^* \subseteq \mathcal{J}_1)$	0.84	0.74	0.77	0.74	0.75
p_1	0.84	0.74	0.77	0.74	0.75
p_2	0.99	0.95	0.9	0.93	0.92
p_3	1	0.97	0.93	0.95	0.97

Table 1.A3. Summary of parameter estimation in simulation study.

Scenario		1		2		3		4		5	
Variable	Effect	Ps	Pc	Ps	Pc	Ps	Pc	Ps	Pc	Ps	Pc
rec_M7	-0.25	0.97	1	0.94	1	0.98	1	0.93	1	0.95	1
rec_M8	-0.667	0.97	1	0.96	1	0.95	1	0.98	1	0.97	1
rec_M9	-1.25	0.9	1	0.97	1	0.94	1	0.93	1	0.92	1
rec_M10	-2	0.93	1	0.93	1	0.96	1	0.93	1	0.92	1
rec_M11	-2.917	0.97	1	0.86	1	0.85	1	0.8	1	0.85	1
rec_M12	-2	0.99	1	0.94	1	0.96	1	0.99	1	0.97	1
rec_M1	-1.25	0.98	1	0.91	1	0.92	1	0.93	1	0.98	1
rec_M2	-0.667	0.95	1	0.96	1	0.94	1	0.95	1	0.93	1
rec_M3	-0.25	0.92	1	0.97	1	0.97	1	0.94	1	0.95	1
rec_M4	0.083	0.8	0.828	0.77	0.81	0.78	0.851	0.84	0.878	0.8	0.876
rec_M5	0.083	0.83	0.992	0.83	0.984	0.86	0.99	0.88	0.991	0.95	0.999
GDD2_M11	-0.4			0.91	0.878	0.87	0.849	0.88	0.848	0.89	0.848
sexFemale	0.1					0.94	1			0.93	1
tag length	0.1							0.94	1	0.96	1

Appendix 1.E: Comparison with two tree-based methods

We consider two popular tree-based methods frequently used in ecological studies: the gradient boosting regression tree method (De'ath 2007) and the random forests approach (Breiman 2001). The variable ranking results for both methods for GDD hypothesis 1 are shown in Figure 1.A3. To investigate why GDD_Diff was ranked as the top variable with tree-based methods but not with BVS method, we show the centered fitted function for GDD_Diff from the boosted regression trees, partial dependence for GDD_Diff from the random forests, and the response variable versus GDD_Diff, as shown in Figure 1.A4. The first two plots similarly measure how the response reacts to GDD_Diff under the tree-based methods.

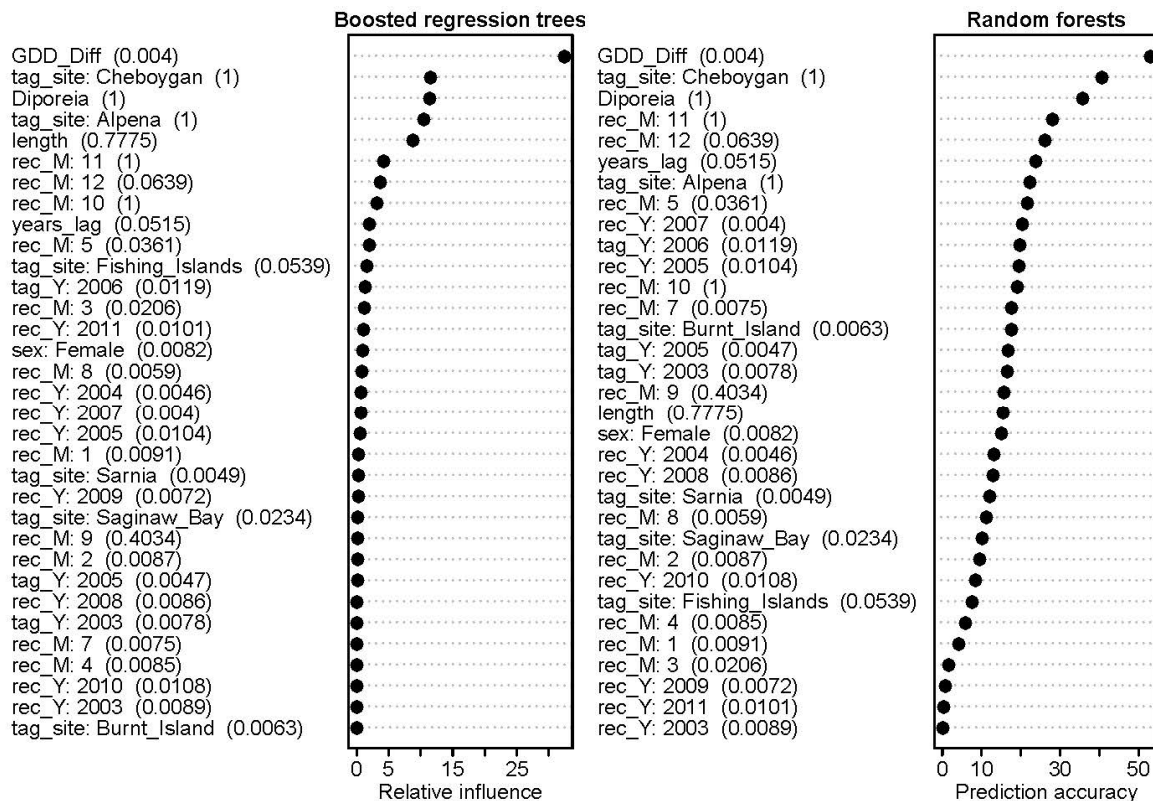


Figure 1.A3. Comparison of results from tree-based models for GDD hypothesis 1. Variables are ranked based on their importance according to the relative influence for boosted regression tree and out-of-bag prediction accuracy for random forests. The marginal inclusion probability from our Bayesian variable selection method is included in the parenthesis for comparison.

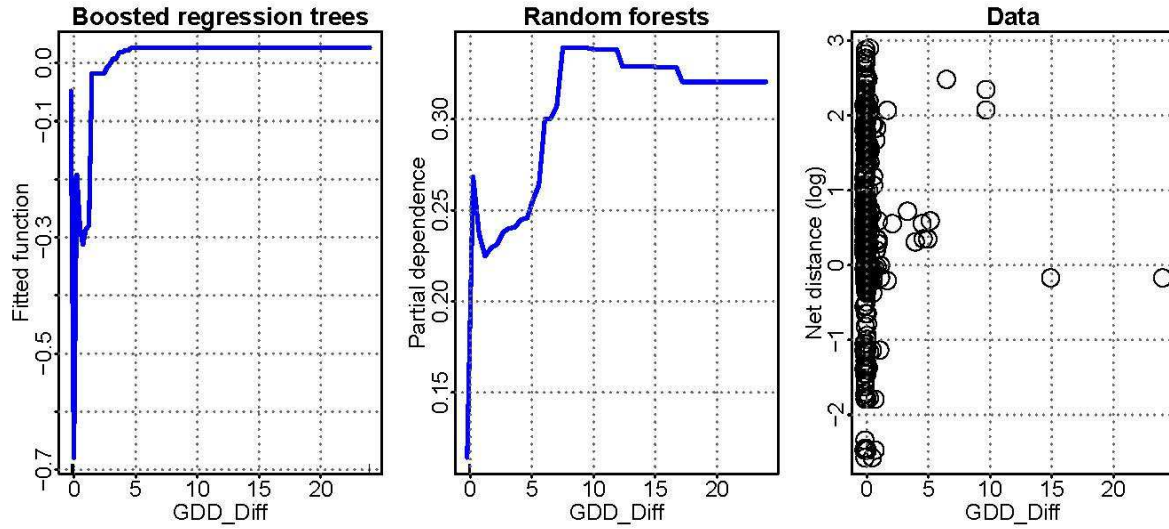


Figure 1.A4. Investigation of variable GDD_Diff in GDD hypothesis 1 case: a) centered fitted function for GDD_Diff from the boosted regression trees; b) partial dependence for GDD_Diff from the random forests; c) the response variable versus GDD_Diff.

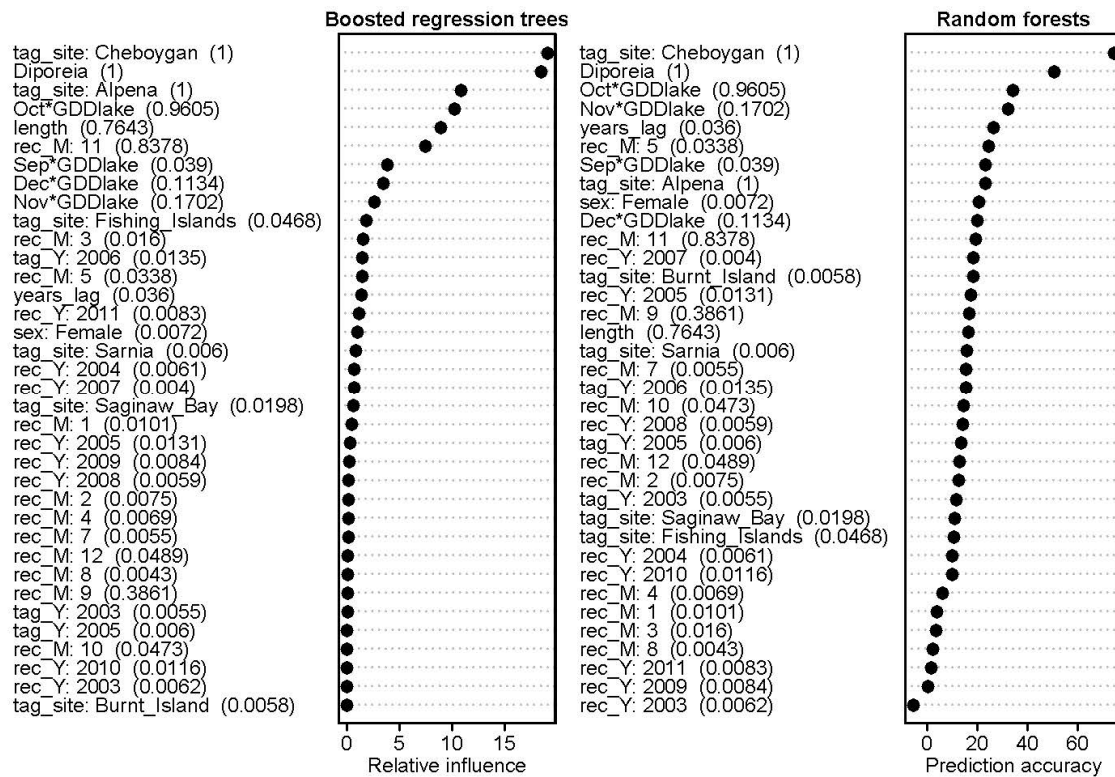


Figure 1.A5. Comparison of results from tree-based models for GDD hypothesis 2. Variables are ranked based on their importance according to the relative influence for boosted regression tree (Friedman 2001) and out-of-bag prediction accuracy for random forests (Breiman 2001). The marginal inclusion probability from our Bayesian variable selection method is included in the parenthesis for comparison.

Appendix 1.F

Codes for Bayesian variable selection, and tree based methods Code for Bayesian variable selection and tree-based methods can be found online, at <https://doi.org/10.6084/m9.figshare.5177206>.

BIBLIOGRAPHY

BIBLIOGRAPHY

- Albanese, B., Angermeier, P.L., and Dorai-Raj, S. 2004. Ecological correlates of fish movement in a network of Virginia streams. *Can. J. Fish. Aquat. Sci.* **61**(6): 857–869. doi:10.1139/f04-096.
- Barbieri, M.M., and Berger, J.O. 2004. Optimal predictive model selection. *Ann. Stat.* **32**(3): 870–897. doi:10.1214/009053604000000238.
- Barbiero, R.P., Schmude, K., Lesht, B.M., Riseng, C.M., Warren, G.J., and Tuchman, M.L. 2011. Trends in Diporeia populations across the Laurentian Great Lakes, 1997–2009. *J. Great Lakes Res.* **37**(1): 9–17. Elsevier B.V. doi:10.1016/j.jglr.2010.11.009.
- Breiman, L. 2001. Random forests. *Mach. Learn.* **45**(1): 5–32. doi:10.1023/A:1010933404324.
- Brooks, S.P., and Gelman, A. 1998. General Methods for Monitoring Convergence of Iterative Simulations. *J. Comput. Graph. Stat.* **7**(4): 434–455. doi:10.1080/10618600.1998.10474787.
- Chezik, K.A., Lester, N.P., Venturelli, P.A., and Tierney, K. 2014. Fish growth and degree-days II: selecting a base temperature for an among-population study. *Can. J. Fish. Aquat. Sci.* **71**(9): 1303–1311. doi:10.1139/cjfas-2013-0615.
- Claeskens, G., and Hjort, N.L. 2008. Model selection and model averaging. Cambridge University Press.
- Csárdi, G., and Nepusz, T. 2006. The igraph software package for complex network research. *InterJournal Complex Syst.* **1695**(5): 1–9. doi:10.3724/SP.J.1087.2009.02191.
- De'ath, G. 2007. Boosted trees for ecological modeling and prediction. *Ecology* **88**(1): 243–51. Available from <http://cran.r-project.org/web/packages/dismo/vignettes/brt.pdf>.
- Drouineau, H., Bau, F., Alric, A., Deligne, N., Gomes, P., and Sagnes, P. 2017. Silver eel downstream migration in fragmented rivers: use of a Bayesian model to track movements triggering and duration. *Aquat. Living Resour.* **30**. doi:10.1051/alr/2017003.
- Ebener, M.P., Brenden, T.O., and Jones, M.L. 2010a. Estimates of fishing and natural mortality rates for four lake whitefish stocks in northern Lakes Huron and Michigan. *J. Great Lakes Res.* **36**: 110–120. doi:10.1016/j.jglr.2009.06.003.
- Ebener, M.P., Brenden, T.O., Wright, G.M., Jones, M.L., and Faisal, M. 2010b. Spatial and temporal distributions of lake whitefish spawning stocks in Northern lakes Michigan and Huron, 2003–2008. *J. Great Lakes Res.* **36**: 38–51. doi:10.1016/j.jglr.2010.02.002.
- Ethier, D.M., Koper, N., and Nudds, T.D. 2017. Spatiotemporal variation in mechanisms driving regional-scale population dynamics of a Threatened grassland bird. *Ecol. Evol.* **7**(12): 4152–4162. doi:10.1002/ece3.3004.
- Forseth, T., Nesje, T.F., Jonsson, B., and Harsaker, K. 1999. Juvenile migration in brown trout: a consequence of energetic state. *J. Anim. Ecol.* **68**(4): 783–793. doi:10.1046/j.1365-2656.1999.00329.x.

- Friedman, J.H. 2001. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **29**(5): 1189–1232. Cambridge University Press, Cambridge. doi:10.1214/aos/1013203451.
- Fu, C., and Fanning, L.P. 2004. Spatial considerations in the management of Atlantic cod off Nova Scotia, Canada. *North Am. J. Fish. Manag.* **24**(3): 775–784. doi:10.1577/M03-134.1.
- Gatz, A.J., and Adams, S.M. 1994. Patterns of movement of centrarchids in two warmwater streams in eastern Tennessee. *Ecol. Freshw. Fish* **3**(1): 35–48. doi:10.1111/j.1600-0633.1994.tb00105.x.
- Gelman, A., Meng, X.-L., and Stern, H. 1996. Posterior predictive assessment of model fitness via realized discrepancies. *Stat. Sin.* **6**(4): 733–807. doi:10.1.1.142.9951.
- Gilliam, J.F., and Fraser, D.F. 2001. MOVEMENT IN CORRIDORS: ENHANCEMENT BY PREDATION THREAT, DISTURBANCE, AND HABITAT STRUCTURE. *Ecology* **82**(1): 258–273. doi:10.1890/0012-9658(2001)082[0258:MICEBP]2.0.CO;2.
- Green, P.J. 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* **82**(4): 711–732. doi:10.1093/biomet/82.4.711.
- Gunning, G.E., and Shoop, C.R. 1963. Occupancy of home range by longear sunfish, *Lepomis m. megalotis* (Rafinesque), and bluegill, *Lepomis m. macrochirus* Rafinesque. *Anim. Behav.* **11**(2–3): 325–330. doi:10.1016/S0003-3472(63)80119-0.
- Hijmans, R.J., Williams, E., and Vennes, C. 2012. Geosphere: spherical trigonometry.
- Hutchings, J.A. 1996. Spatial and temporal variation in the density of northern cod and a review of hypotheses for the stock's collapse. *Can. J. Fish. Aquat. Sci.* **53**(5): 943–962. doi:10.1139/f96-097.
- Li, Y., Bence, J.R., and Brenden, T.O. 2015. An evaluation of alternative assessment approaches for intermixing fish populations: a case study with Great Lakes lake whitefish. *ICES J. Mar. Sci.* **72**(1): 70–81. doi:10.1093/icesjms/fsu057.
- Lidicker, W.Z., and Stenseth, N.C. 1992. To disperse or not to disperse: who does it and why? *In* *Animal Dispersal*. Springer Netherlands, Dordrecht. pp. 21–36. doi:10.1007/978-94-011-2338-9_2.
- McNickle, G.G., Rennie, M.D., and Sprules, W.G. 2006. Changes in Benthic Invertebrate Communities of South Bay, Lake Huron Following Invasion by Zebra Mussels (*Dreissena polymorpha*), and Potential Effects on Lake Whitefish (*Coregonus clupeaformis*) Diet and Growth. *J. Great Lakes Res.* **32**(1): 180–193. doi:10.3394/0380-1330(2006)32[180:CIBICO]2.0.CO;2.
- Minns, C.K. 1995. Allometry of home range size in lake and river fishes. *Can. J. Fish. Aquat. Sci.* **52**(7): 1499–1508. doi:10.1139/f95-144.
- Mitchell, T.J., and Beauchamp, J.J. 1988. Bayesian Variable Selection in Linear Regression. *J. Am. Stat. Assoc.* **83**(404): 1023–1032. Available from <http://www.jstor.org/stable/2290129>.
- Mohr, L.C., and Nalepa, T.F. 2005. Proceedings of a workshop on the dynamics of lake whitefish (*Coregonus clupeaformis*) and the amphipod *Diporeia* spp. in the Great Lakes. *In*

- Great Lakes Fish. Comm. Tech. Rep. Available from Great Lakes Fish. Comm. Tech. Rep.
- Polacheck, T., Eveson, J.P., Laslett, G.M., Pollock, K.H., and Hearn, W.S. 2006. Integrating catch-at-age and multiyear tagging data: a combined Brownie and Petersen estimation approach in a fishery context. *Can. J. Fish. Aquat. Sci.* **63**(3): 534–548. doi:10.1139/f05-232.
- Price, H., Pothoven, S.A., McCormick, M.J., Jensen, P.C., and Fahnenstiel, G.L. 2003. Temperature Influence on Commercial Lake Whitefish Harvest in Eastern Lake Michigan. *J. Great Lakes Res.* **29**(2): 296–300. doi:10.1016/S0380-1330(03)70434-1.
- Radinger, J., and Wolter, C. 2014. Patterns and predictors of fish dispersal in rivers. *Fish Fish.* **15**(3): 456–473. doi:10.1111/faf.12028.
- Rätsch, G., Onoda, T., and Müller, K.-R. 2001. Soft Margins for AdaBoost. *Mach. Learn.* **42**(3): 287–320. doi:10.1023/A:1007618119488.
- Rennie, M.D., Ebener, M.P., and Wagner, T. 2012. Can migration mitigate the effects of ecosystem change? Patterns of dispersal, energy acquisition and allocation in Great Lakes lake whitefish (*Coregonus clupeaformis*). *Adv. Limnol.* **63**: 455–476. doi:10.1127/advlim/63/2012/455.
- Rennie, M.D., Sprules, W.G., and Johnson, T.B. 2009. Factors affecting the growth and condition of lake whitefish (*Coregonus clupeaformis*). *Can. J. Fish. Aquat. Sci.* **66**(12): 2096–2108. doi:10.1139/F09-139.
- Roff, D.A. 1991. Life History Consequences of Bioenergetic and Biomechanical Constraints on Migration. *Am. Zool.* **31**(1): 205–216. doi:10.1093/icb/31.1.205.
- Rothschild, B.J. 2007. Coherence of Atlantic cod stock dynamics in the Northwest Atlantic Ocean. *Trans. Am. Fish. Soc.* **136**(3): 858–874. doi:10.1577/T06-213.1.
- Runge, C.A., Martin, T.G., Possingham, H.P., Willis, S.G., and Fuller, R.A. 2014. Conserving mobile species. *Front. Ecol. Environ.* **12**(7): 395–402. doi:10.1890/130237.
- Schneider, K.N., Newman, R.M., Card, V., Weisberg, S., and Pereira, D.L. 2010. Timing of Walleye Spawning as an Indicator of Climate Change. *Trans. Am. Fish. Soc.* **139**(4): 1198–1210. doi:10.1577/T09-129.1.
- Schwarz, G. 1978. Estimating the Dimension of a Model. *Ann. Stat.* **6**(2): 461–464. doi:10.1214/aos/1176344136.
- Vandergoot, C.S., and Brenden, T.O. 2014. Spatially varying population demographics and fishery characteristics of Lake Erie walleyes inferred from a long-term tag recovery study. *Trans. Am. Fish. Soc.* **143**(1): 188–204. doi:10.1080/00028487.2013.837095.
- Vincenty, T. 1975. Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations. *Surv. Rev.* **23**(176): 88–93. doi:10.1179/sre.1975.23.176.88.
- Woznicki, S.A., Nejadhashemi, A.P., Abouali, M., Herman, M.R., Esfahanian, E., Hamaamin, Y.A., and Zhang, Z. 2016. Ecohydrological modeling for large-scale environmental impact assessment. *Sci. Total Environ.* **543**: 274–286. Elsevier B.V.

doi:10.1016/j.scitotenv.2015.11.044.

CHAPTER 2

CAN SPAWNING ORIGIN INFORMATION OF CATCH OR A RECRUITMENT PENALTY IMPROVE ASSESSMENT AND FISHERY MANAGEMENT PERFORMANCE FOR A SPATIALLY STRUCTURED STOCK ASSESSMENT MODEL?

Abstract

We used simulations based on Lake Whitefish (*Coregonus clupeaformis*) populations to explore the benefits of using spawning origin information for parsing catch to spawning populations in stock assessments for intermixed fisheries exhibiting an overlapping movement strategy. We compared this origin-informed assessment model with a standard assessment model that did not parse catch. We additionally evaluated the influence of including annual recruitment penalties. For standard assessment models, spawning stock biomass estimates could be unstable and biased (sometimes by more than 50%), depending upon population mixing and productivity, and in some cases estimated near average zero recruitment in the terminal year. Incorporating information on population-specific harvest age composition improved spawning stock biomass estimation (e.g., by sometimes essentially removing 50% biases, and improving accuracy). Assessments with recruitment penalties produced less biased terminal recruitment estimates (sometimes a 100% bias was removed). Under status quo target mortality rates improvements in assessments did not necessarily translate to improved fishery management performance (e.g., avoiding depletion of spawning biomass), but such improvements, and overall better performance, were seen at lower target mortality rates.

Introduction

Accurate estimation of spawning stock biomass and recruitment is important for the management of fishery stocks. Biased or imprecise estimates can influence measures of population productivity and year-class strength, stock-recruitment relationships, and management decisions (e.g., harvest regulations) that depend on these assessment results. When fish from distinct spawning populations intermix on fishing grounds during harvest periods (i.e., populations exhibit spatial structuring), estimating recruitment and spawning stock biomass dynamics for each spawning population from sampling programs that only target intermixed fisheries can be challenging. Statistical catch-at-age or catch-at-size models are commonly used for the assessment of commercial harvested fish populations for estimating biomass of spawning adults and recruitment dynamics. However, one known feature of such assessment models is that recruitment in the last several assessment years cannot be reliably estimated because there is little information about recruitment levels for those years. In addition, such assessment models typically ignore spatial structure and assume harvest is from a single population (i.e., the “unit stock” assumption).

When assessment data are collected from intermixed fisheries but a single population assumption is made in the stock assessment model, population abundance can be overestimated, which can further lead to inappropriate management advice especially for low productivity populations (Hutchings 1996; Fu and Fanning 2004; Ying et al. 2011; Hintzen et al. 2015; Li et al. 2015). For example, it has been argued that some Atlantic cod (*Gadus morhua*) and Pacific salmon (*Oncorhynchus* spp.) populations were overharvested due to intermixed fisheries that did not properly account for differences in population productivities (Hutchings 1996; Morishima and Henry 1999; Fu and Fanning 2004). To facilitate management of intermixed fisheries,

spatially-explicit stock assessment models can be used that either incorporate tagging data within the stock assessment framework (Eveson et al. 2009; Vincent et al. 2017), or incorporate mixing and migration rates in assessment models as fixed quantities (Guan et al. 2013; Li et al. 2015). Both approaches allow for spatially-explicit estimation of abundances, mortality components, and other dynamic rates within an integrated stock assessment model.

When accounting for spatial structure in stock assessments, two alternative movement strategies are commonly recognized: diffusion and overlap (Porch et al. 2001). The diffusion movement strategy, also known as meta-population mixing (Ying et al. 2011), assumes that the fraction of fish populations that move away from their original spawning areas become part of the spawning populations near to which they move (i.e., their spawning population identity changes according to their movement behavior). Conversely, the overlap movement strategy assumes 100% spawning site fidelity meaning that fish always move back to their original natal areas during the spawning season, and thus spawning population identity is maintained throughout a fish's lifetime. In this paper we focus on stock assessment models assuming an overlap movement strategy. While this is clearly a simplification for any given stock, it is a reasonable approximation of spatial structure for many stocks.

A known problem for assessment models, when applied to populations exhibiting spatial structuring with moderate to high levels of intermixing, is that population-specific estimates of recruitment are uncertain or not estimable, and estimates of spawning stock biomass are unstable or biased, even when mixing rates are assumed known (Ying et al. 2011; Molton et al. 2012; Li et al. 2015). Li et al. (2015) proposed an overlap stock assessment model in which an integrated statistical catch-at-age (SCAA) assessment model was fit to overlapping fish populations by incorporating actual mixing rates in the model. They found that mixing among areas caused

problems in estimating population-specific annual recruitments, and this led to substantial uncertainty and bias in estimation of recruitment and biomass. They hypothesized that this problem could be resolved if additional population-specific data were provided to the assessment model, such that harvest data could be allocated to source populations. Hintzen et al. (2015) evaluated the influence of fishery-independent survey data on the performance of an integrated catch-at-age method for intermixing fish populations, in which information on the classification of the catch to their spawning origin were used to inform survey indices (i.e., the proportions of survey sample to spawning populations). However, the catch data they used in the assessment model were not reallocated back to the spawning populations because their assessment model ignored spatial structure. Thus, mismatch between spatial structures in the assessment data and in the assessment model still existed. They found that spatially-explicit survey data marginally reduced bias in estimation of biomass, but when there were errors in classification rates inaccuracies could actually increase.

The goal of our research was to evaluate the benefits of including information on catch composition for the management of intermixing fish populations. Our research extended the overlap SCAA assessment model proposed by Li et al. (2015) by including information on population-specific harvest age composition, which could arise from having genetic or some other type of discriminatory characteristic (e.g., parasite community, meristic or morphometric feature) of the populations from a biological sample collected from the intermixed fisheries. Herein, we refer to the overlap assessment model proposed by Li et al. (2015) as the “standard assessment model”, and the extended one with additional data on population source as the “origin-informed assessment model”. In both assessment models, annual recruitments were estimated as free parameters, which is the same approach used by Li et al. (2015). We further

propose two alternative assessment models that are identical to these two models except that a penalty on annual recruitment residuals was incorporated in each model. Several studies conducted for single populations (no spatial structure) have shown that adding such penalties or other constraints can improve estimates of annual recruitment, particularly for terminal assessment years (Maunder and Deriso 2003; Methot et al. 2011; Korman et al. 2012). We tested how assessment and management performance of the standard and origin-informed assessment models were influenced by the magnitude of recruitment variation, assessment data quality, uncertainty regarding mixing rates, and target mortality rates.

The dynamics of our simulations were based on lake whitefish (*Coregonus clupeaformis*) populations and fisheries in the upper Laurentian Great Lakes of North America, although results should have general applicability to populations with similar life history and movement patterns given the stochastic modeling of uncertainty and the range of sensitivity analyses we report. An overlap movement strategy was assumed for the simulated lake whitefish populations, because evidence suggests that lake whitefish populations in the Laurentian Great Lakes region overlap during non-spawning seasons but move back to where they were born during the spawning season of each year (Ebener et al. 2010a; Stott et al. 2010; Li et al. 2017). Although tagging studies have suggested that considerable movement of lake whitefish in the Laurentian Great Lakes region from management units containing their spawning grounds to other management units during the non-spawning and harvest seasons (Ebener et al. 2010b; Li et al. 2017), they are still largely managed as unit stocks. To our best knowledge, our research is the first to evaluate the influence of including population-specific catch information on a spatial structured stock assessment model. Compared to Hintzen et al. (2015), we propose a different approach of using

such information for the management of intermixing stocks with a focus directed towards spatially structured stock assessments.

Methods

Simulation framework

Our simulation framework followed a management strategy evaluation approach (i.e., full closed-loop feedback simulation framework to evaluate alternative management procedures, Figure 2.1). These at simulations were designed to determine the long-term assessment and management performance for both standard and origin-informed assessment models with or without a lognormal penalty on annual recruitment residuals (Table 2.1). The operating model consisted of four hypothetical lake whitefish populations with age-structure and an overlap movement strategy (i.e., 100% natal fidelity was assumed) that intermixed across four areas of harvest. Observations from the four regions of harvest were then generated for input for the stock assessment models. Assessment models were fit to the observed data, and a harvest control rule was applied each year based on the assessment results so that target harvest levels (i.e., total allowable catch in our case) could be set. The management procedure then fed back to the operating model by implementing actual harvest based on the target with implementation error in the operating model of next year. Given we were considering alternative stock assessment models and the stock assessment results influenced dynamics, separate simulations were conducted for each assessment approach, albeit using the same random number seeds. To evaluate long-term performance of each assessment model, we ran each simulation for 100 years, and summarized results for the last 25 years. All symbols of index variables and accents used in the equations of this paper are identified in Table 2.2.

Figure 2.1. The full closed-loop feedback simulation framework, which followed a management strategy evaluation approach.

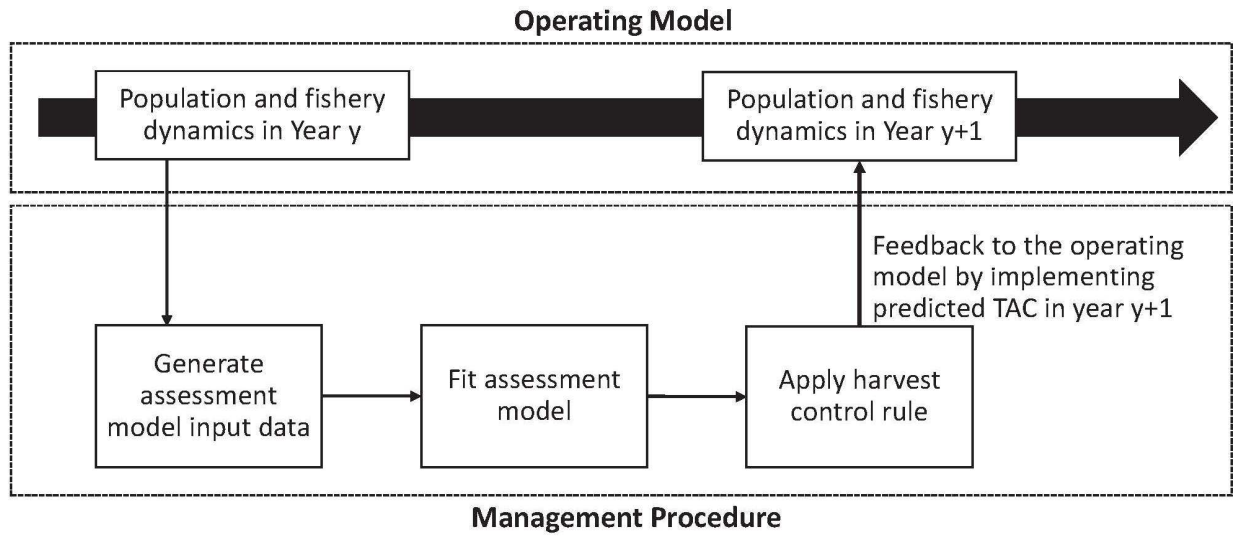


Table 2.1. Composition of the assessment input data and objective function for the four assessment models we evaluated.

Assessment model		Standard assessment model without a recruitment penalty (S)	Standard assessment model with a recruitment penalty (S W/Rec)	Origin-informed assessment without a recruitment penalty (O)	Origin-informed assessment with a recruitment penalty (O W/Rec)
Input data	Observed harvest	✓	✓	✓	✓
	Observed effort	✓	✓	✓	✓
	Aggregated harvest age composition	✓	✓		
	Population-specific harvest age composition			✓	✓
Objective function components (negative log likelihood or log-prior penalty for)	Area-specific fishery harvest	✓	✓	✓	✓
	Annual deviation from the general level of fishing mortality	✓	✓	✓	✓
	Aggregate harvest age composition	✓	✓		
	Population-specific harvest age composition			✓	✓
	Annual recruitment residuals		✓		✓

Table 2.2. Index variables and accents used in all equations.

Symbol	Definition
i	Population
j	Fishing ground
y	Year
a	Age
\sim	Observed variable
$\hat{\sim}$	Estimated variable
$'$	Derived variable

Operating model

The operating model was stochastic and age-structured (i.e., ages 3 to 12 with the last age class an aggregate group including age-12 and older fish), operated in annual time steps, and recognized four geographic fishing grounds that were presumed to surround the four spawning areas (i.e., each spawning area was associated and located within a fishing region). Yearly time steps were considered because evidence suggested that the movement of lake whitefish populations in the upper Great Lakes generally occurred soon after spawning (i.e., between late October and early December, Li et al. 2017). Thus, we assumed that fish moved away from their spawning areas on the first day of each year, and all surviving fish returned to their original spawning areas to spawn at the end of each year.

As described in detail below, many parameters of the operating model are taken from Li et al. (2015), which were based on a review of existing Lake Whitefish stock assessments. A single set of life history (growth, maturity) parameters was used, representative of those estimated from biological data used in those stock assessments. General levels of recruitment

stochasticity and productivity, and variations among populations were based on analysis of recruitment and spawning stock sizes from the existing assessments. The existing assessments are unit stock assessments, and the influence of this on perceived differences in recruitment productivity was taken into account when specifying varying productivity levels (Li et al. 2015). In real assessments, with spawning populations that differ in life history, it is likely that there would be additional advantages of biological data that is spawning population specific, which we have not evaluated here.

For each simulated population, we modeled recruitment (age-3 fish) from a Ricker stock-recruitment function with a first-order autoregressive process (AR1):

$$N_{i,y,a=3} = \alpha_i SSB_{i,y-3} e^{-\beta_i SSB_{i,y-3}} e^{\varepsilon_{R,i,y}}. \quad (2.1)$$

$$\alpha_i = \alpha_i' e^{-0.5\sigma^2}.$$

$$\varepsilon_{R,i,y} = \rho \times \varepsilon_{R,i,y-1} + \tau_{R,i,y}.$$

$$\tau_{R,i,y} \sim Normal(0, \sigma_R^2).$$

$$\sigma^2 = \frac{\sigma_R^2}{1-\rho^2}.$$

where $N_{i,y,a=3}$ is the abundance of age-3 fish from population i at the beginning of year y , $SSB_{i,y-3}$ is the spawning stock biomass of population i in year $y - 3$, and α_i and β_i are Ricker stock-recruitment function parameters for population i . The parameters ρ and σ_R defined the stochastic process for deviations of recruitment from the underlying Ricker stock-recruitment function, producing temporally autocorrelated recruitment. The level of process error presented in Table 2.3 was used for all simulated populations in the baseline scenario. Process error parameters were varied in the sensitivity analysis for evaluating the influence of recruitment

variation on modeling results. The stock-recruitment parameter α' , together with β , were chosen so that the deterministic stock recruitment would produce the desired average level of recruitment given stock size. For the simulations, α' was scaled by $e^{-0.5\sigma^2}$ so that the expectation of the stochastic form of the recruitment relationship would equal the deterministic value and not depend on the assumed level of recruitment variation.

Table 2.3. Coefficients for parameters used to generate different levels of productivity, data quality, recruitment variation, and annual-varying random generated rates in both operating and stock assessment models.

Coefficient name	Definition	Coefficient values		
		Low	Baseline	High
<i>Productivity levels</i>				
Steepness	S-R steepness	0.7	1.3668	1.9
α'	Ricker S-R parameter	0.0003169815	0.0007316319	0.001104342
β	Ricker S-R parameter	$1.511359e^{-10}$	$2.318631e^{-10}$	$2.716004e^{-10}$
<i>Data quality levels</i>				
$effN$	Effective sample size	25	50	100
Harvest CV	CV for observed harvest about actual harvest	0.4	0.15	0.1
Effort CV	CV for observed harvest about actual effort	0.8	0.3	0.2
<i>Annual-varying random generated rates</i>				
μ_ω	Mean of ω_y	2.313635	0.8472979	0.08004271
σ_ω^2	Variance of ω_y	0.3364	0.0625	0.21
<i>Recruitment variation levels</i>				
ρ	Autocorrelation coefficient	No autocorrelation 0	Baseline 0.45	High 0.45
σ_R	Innovative standard dev. in rec process error	0.8734	0.78	1.3395
σ	Stationary standard dev. in rec process error	0.8734	0.8734	1.5
<i>Target mortality levels</i>				
A	Annual total mortality rate	Low 0.55	Baseline (Status quo) 0.65	

Table 2.4. Biomass calculation in the operating model.

Model name	Model equation	Equation number
Age-specific SSB	$SSB_{i,y} = \sum_a Fem W_a m_a N_{i,y,a} Fec$	2.4.1
	where $Fem=0.5$ (from Li et al. 2015)	
Length at age	$L_a = L_\infty (1 - \exp(-\kappa(a - t_0)))$ where $L_\infty=60.9$ cm, $\kappa=0.1689$ year ⁻¹ , $t_0 = 0$ year (from Li et al. 2015)	2.4.2
Weight at age	$W_a = \gamma L_a^\psi$ where $\gamma = 8.06 \times 10^{-5}$, $\psi = 2.45$ (from Li et al. 2015)	2.4.3
Maturity at age	$m_a = \frac{m_\infty}{1 + \exp(-\vartheta(L_a - \delta))}$ where $\vartheta = 0.315$ cm ⁻¹ , $\delta = 37.86$ cm (from Li et al. 2015)	2.4.4

Total spawning stock biomass (SSB) for population i in year y was calculated as the product of female percentage in the population (50%), weight-, maturity-, and abundance-at-age, and weight-specific fecundity (19733/kg). All equations and parameter values used for calculating SSB are defined in Table 2.4, which are the same as used by Li et al. (2015).

For each population, post-recruitment (after age-3) abundances at age (a) at the beginning of each year were forward projected using an exponential mortality model with a constant natural mortality (M) of 0.25, and age-, year-, and region-specific (j) fishing mortality (F):

$$N_{i,y+1,a+1} = N_{i,y,a} \sum_j \theta_{ij} \exp(-M - F_{j,y,a}). \quad (2.2)$$

According to equation 2.2, fish from a spawning population either remained in the region surrounding their natal area during the non-spawning season or moved to one of the other harvest areas, depending on the assumed mixing rates θ_{ij} . Thus, the survival of fish in a population was a weighted average of the survival rates in each of the harvest regions, with weights equal to the proportions of fish from the population residing in the regions during the non-spawning season. In some scenarios, mixing rates varied among the populations in the operating model, but in all cases were temporally invariant for each population.

We used stay rate θ_{ii} (i.e., the proportion of fish from spawning population i that stay in the area surrounding that population's spawning area during the non-spawning season) to represent movement dynamics for population i , and assumed that a greater stay rate indicated higher-quality habitat, so that a greater proportion of fish from other population moved to that area (Table 2.5). Thus, mixing rates θ_{ij} (i.e., the proportion of fish from spawning population i that move to the area surrounding population j 's spawning area during the non-spawning season) were calculated as (Li et al. 2015):

$$\theta_{ij} = (1 - \theta_{ii}) \frac{\theta_{jj}}{\sum_{k \neq i} \theta_{kk}}. \quad (2.3)$$

where the summation is overall all areas k except the fishing grounds surrounding the spawning area of population i . Total allowable catch (TAC) for each harvest area was determined via the management procedure described below. Actual harvest (C) in each year was set equal to the TAC multiplied by a lognormal implementation error term with a coefficient of variation (CV) of 10%:

$$C_{j,y} = TAC_{j,y} \exp(\zeta_{j,y} - 0.5\sigma_{tac}^2). \quad (2.4)$$

$$\zeta_{j,y} \sim Normal(0, \sigma_{tac}^2).$$

where σ_{tac} is the standard deviation of TAC implementation error ζ . The fully selected fishing mortality rate f that produced the actual harvest level given age-specific abundances was solved for using a Newton-Raphson algorithm and Baranov's catch equation:

$$C_{j,y} = \frac{s_a f_{j,y}}{M + s_a f_{j,y}} (1 - e^{-M - s_a f_{j,y}}) \sum_i N_{i,y,a} \theta_{ij}. \quad (2.5)$$

Age-specific F s were set equal to the solved f multiplied by age-specific selectivities s_a :

$$F_{j,y,a} = S_a f_{j,y}. \quad (2.6)$$

Selectivities for age-3 and older ages were calculated from a gamma function that produced a dome-shape selectivity pattern with peak selectivity for age-10:

$$S_a = \frac{a^\eta \exp(-\tau a)}{10^\eta \exp(-\tau 10)}. \quad (2.7)$$

where selectivity parameters $\tau = 1.26 \text{ year}^{-1}$, $\eta = 13.074 \text{ cm}$ (from Li et al. 2015), were assumed to be the same for different populations.

Table 2.5. Simulation scenarios, including the baseline scenario and other combinations of productivity levels and stay rates, for four hypothetical populations used in the simulations.

Scenario index	Scenario	Population identifier	Productivity	Stay rate
Baseline (1)	Equal mixing with baseline productivity	Pop1	Baseline	70%
		Pop2	Baseline	70%
		Pop3	Baseline	70%
		Pop4	Baseline	70%
2	Equal mixing with different productivity	Pop1	Low	70%
		Pop2	Low	70%
		Pop3	High	70%
		Pop4	High	70%
3	Unequal mixing with baseline productivity	Pop1	Baseline	91%
		Pop2	Baseline	91%
		Pop3	Baseline	52%
		Pop4	Baseline	52%
4	Unequal mixing with different productivity (Positive correlation between productivity and stay rates)	Pop1	Low	52%
		Pop2	Low	52%
		Pop3	High	91%
		Pop4	High	91%
5	Unequal mixing with different productivity (Negative correlation between productivity and stay rates)	Pop1	Low	91%
		Pop2	Low	91%
		Pop3	High	52%
		Pop4	High	52%

We used the same approach as Li et al. (2015) to determine initialization abundances for each simulation. Specifically, initialization abundances for the populations were set to their equilibrium values based on the target mortality rate and a deterministic version of our model (equilibrium for populations at different productivity levels are shown as the intersections in Figure 2.2). As well, like Li et al. (2015), during the initial 20-year period of each simulation, the harvest control rule based on the target mortality rate was applied to the actual abundances at age (i.e., the assessment modeling was skipped). This was necessary as prior to year 20 the required

data time series for conducting assessments was not available. We were not interested in the transient dynamics during this initial period, and we set the starting conditions at the deterministic equilibrium solely to better ensure that the final 25 years of our 100-year simulations approximated steady-state conditions.

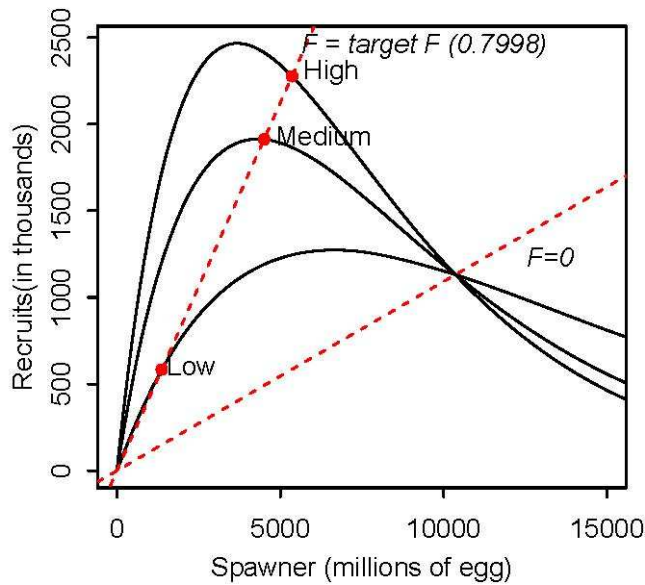


Figure 2.2. Ricker stock-recruitment relationships for populations with low, medium, and high level of productivity (Table 2.3). Two dashed lines represent the replacement lines for $F=0$ and target F and their intersections with stock-recruitment curves (dots) define equilibrium for low, baseline, and high productivity. Note that the target F is calculated based on the natural mortality rate and the status quo target total mortality ($A=0.65$).

Management Procedure

We attempted to emulate key aspects of the management procedures for lake whitefish in the 1836 Treaty-ceded waters, including data collection, stock assessment, and application of a constant total mortality harvest control rule (1836 Treaty Waters Modeling Subcommittee 2017). The underlying premises were that collected data were used to assess the populations (Figure 2.1), that the assessment results provided estimates of the abundance of fish present in each region, and that target harvests were set based on estimated abundances in an attempt to achieve the same target total mortality rate in each harvest region. All evaluated assessment models used

an integrated SCAA assessment model that correctly accounted for movements (i.e., stay and mixing rates were model inputs and were accurately known) among the regions, with the exception of the sensitivity analyses that evaluated the consequences of uncertain mixing rates. All assessment models fit the same population dynamic model to each of their observed data sets to estimate the parameters used to summarize population status and determine target harvest. When fitting the assessment models, only the most recent 20 years of data were used. We elected to use a fixed-length time series so that the amount of information available to an assessment remained stationary during the performance evaluation period (the last 25 years of each 100-year simulation). While relatively short by assessment standards, 20 years represents more than three times the expected period between birth and production of offspring, given the assumed life history, fishery selectivity, and target mortality rate in our operating model, based on Lake Whitefish. Simulations using a 40-year assessment period for a subset of scenarios produced nearly identical results to those with the 20-year assessment period. Age range of the assessment models was the same as that of the operating model. By minimizing the negative log-likelihood (see objective function subsection below), the assessment models were considered to have converged on a solution when the maximum gradient of the parameters was less than 0.001, and the Hessian matrix was positive definite. Convergence rate is defined as the fraction of simulations that met both of the above criteria.

For the standard assessment models with or without a recruitment penalty (i.e., S and S W/Rec in Table 2.1), observed harvest, effort, and aggregated (across populations) harvest age composition data were collected annually for each region. For the origin-informed assessment models (i.e., O and O W/Rec in Table 2.1), observed harvest, effort, and population-specific harvest age composition data were collected annually for each region. Observed harvest differed

from actual harvest as a result of observation error, which was modeled with a lognormal error term v :

$$\tilde{C}_{j,y} = C_{j,y} \exp(v_y - 0.5\sigma_c^2). \quad (2.8)$$

$$v_y \sim Normal(0, \sigma_c^2).$$

The observed fishing effort was a function of fishing mortality f , catchability q , and a lognormal observation error μ and we assumed $\sigma^2_F = 4 \sigma_c^2$.

$$E_{j,y} = \frac{f_{j,y}}{q} \exp(\mu_{j,y} - 0.5\sigma^2_F). \quad (2.9)$$

$$\mu_{j,y} \sim Normal(0, \sigma^2_F).$$

In the baseline scenario, baseline level of CVs for the error terms of observed harvest and effort were used (Table 2.3) while different levels of CVs were explored in the sensitivity analyses for data quality.

For the standard assessment models, aggregated observed age compositions for area-specific harvests were generated from multinomial distributions with probabilities equal to the actual age composition. For the origin-informed assessment models, observed population-specific age compositions for area-specific harvests were generated from multinomial distributions with probabilities equal to the actual population-specific age compositions in each region. The effective sample size (N_{eff}) for the multinomial distribution used to generate aggregated and population-specific age compositions was assumed at its baseline level (Table 2.3), except for the sensitivity analyses for data quality.

Recruitment ($\hat{N}_{i,y,a=3}$) of each assessment year, abundances at age (except age at recruitment) in the first assessment year ($\hat{N}_{i,y=1,a>3}$), gamma function selectivity parameters ($\hat{\tau}$,

$\hat{\eta}$), catchability (\hat{q}), the annual deviation from general level of fishing mortality ($\widehat{\varepsilon F}_{j,y}$, Fournier and Archibald 1982), and the standard deviation from observed harvest ($\hat{\sigma}_c$) were estimated during assessment model fitting. Recruitments in the standard and origin-informed assessment models without recruitment penalty were estimated as free parameters. For the assessment models that included a recruitment penalty, recruitment for each population was reparameterized as the product of average recruitment ($\widehat{R\mu}_i$) multiplied by an annual residual ($\varepsilon'_{i,y}$) that was exponentiated and bias corrected, so that the annual recruitment was assumed to come from a lognormal distribution:

$$N'_{i,y,a=3} = \widehat{R\mu}_i e^{\varepsilon'_{i,y} - 0.5\sigma'_R{}^2}. \quad (2.10)$$

$$\varepsilon'_{i,y} \sim \text{Normal}(0, \sigma'_R{}^2).$$

Post-recruit abundances at age in the first assessment year were estimated as free parameters. The fishing mortality in the assessment models was modeled in the same way as for the operating model, which was a product of selectivity at age and fully selected fishing mortality (same as in Equations 6 and 7, but here $\hat{\tau}$ and $\hat{\eta}$ were estimated parameters). The fully selected fishing mortality ($f'_{j,y}$) was modeled as a product of assessed catchability (\hat{q}), observed effort ($\tilde{E}_{j,y}$), and assessed annual deviation from general level of fishing mortality ($\widehat{\varepsilon F}_{j,y}$).

The natural mortality rates assumed in all assessment models were the same as those used for the operating model. The parameters of all assessment models were estimated in AD Model Builder (Fournier et al. 2012).

The population dynamics in all stock assessment models (i.e., S, S W/Rec, O, and O W/Rec) followed:

$$N'_{i,y+1,a+1} = N'_{i,y,a} \sum_j \theta_{ij} \exp(-M - F'_{j,y,a}). \quad (2.11)$$

$$C'_{j,y,i,a} = \frac{F'_{j,y,a}}{M+F'_{j,y,a}} (1 - e^{-M-F'_{j,y,a}}) N'_{i,y,a} \theta_{ij}. \quad (2.12)$$

$$C'_{j,y,a} = \sum_i C'_{j,y,i,a}. \quad (2.13)$$

For each harvest area, aggregated harvest age composition for the standard assessment models (Equation 2.14, Table 2.1), and population-specific harvest age composition for the origin-informed assessment models (Equation 2.15, Table 2.1) were:

$$p'_{j,y,a} = C'_{j,y,a} / \sum_a C'_{j,y,a}. \quad (2.14)$$

$$p'_{j,y,i,a} = C'_{j,y,i,a} / \sum_{i,a} C'_{j,y,i,a}. \quad (2.15)$$

Predicted SSB was calculated from estimated abundance at age $N'_{i,y,a}$ by using equation 2.1, and assuming weight, maturity at age and weight-specific fecundity were known (Table 2.4).

Objective function

The objective function for each assessment model was the summation of at least three negative log-likelihood and log-prior/penalty components (Table 2.1). All four assessment models assumed the same lognormal distributions for the log-likelihood component of total fishery annual harvest by harvest area and for the log-prior components associated with the fishing mortality-effort relationship for each harvest area.

The total negative log-likelihood component for the log of area-specific annual fishery harvest was based on a normal distribution

$$\ell_c = \sum_j (n \log_e(\hat{\sigma}_c) + \left(\frac{1}{2\hat{\sigma}_c^2}\right) \sum_y (\log_e(\frac{\tilde{c}_{j,y}}{\hat{c}_{j,y}}))^2), \quad (2.16)$$

where n was the number of assessment years (i.e., 20 years). A normal distribution was also assumed for the log-prior penalty associated with the log annual deviation from the general level of fishing mortality

$$\ell_{\varepsilon F} = \sum_j (n \log_e(\sigma_F') + \left(\frac{1}{2\sigma_F'^2}\right) \sum_y (\log_e(\widehat{\varepsilon F}_{j,y}))^2), \quad (2.17)$$

where $\sigma_F'^2$ was assumed to be four times greater than $\hat{\sigma}_c^2$, which matched what was assumed in the operating model. This penalty was equivalent to predicting effort as proportional to estimated fishing mortality and treating deviations between the log of observed and predicted fishing effort as normally distributed (Fournier and Archibald 1982).

The third likelihood component was associated with harvest age composition and was based on a multinomial distribution, but there were differences in this likelihood component for standard and origin-informed assessment models. For the standard assessment model (assessment models S and S W/Rec, Equation 2.18), the negative log likelihood component was for the aggregate harvest age composition for the harvest regions

$$\ell_a = -\sum_j \sum_y N_{eff} \sum_a (\tilde{p}_{j,y,a} \log_e p'_{j,y,a}). \quad (2.18)$$

where $\tilde{p}_{j,y,a}$ and $p'_{j,y,a}$ are the observed and estimated proportions of harvest in area j by age a in year y and N_{eff} is the assumed effective sample size. For the origin-informed assessment models (assessment models O and O W/Rec, Equation 2.19), the negative log likelihood component was for the population-specific harvest age composition for the harvest regions

$$\ell_{pa} = -\sum_j \sum_y N_{eff} \sum_{i,a} (\tilde{p}_{j,y,i,a} \log_e p'_{j,y,i,a}). \quad (2.19)$$

where $\tilde{p}_{j,y,i,a}$ and $p'_{j,y,i,a}$ are the observed and estimated proportions of harvest in area j by age a from population i in year y , respectively. For baseline scenarios, N_{eff} was set equal to 50 for both standard and origin-informed assessment models, but was varied in sensitivity analyses to evaluate the influence of data quality.

For standard and origin-informed assessment models that included a penalty on annual recruitment residuals (i.e., S W/Rec and O W/Rec in Table 2.1), the objective function included a log-penalty component that constrained the annual recruitment residuals $\varepsilon'_{i,y}$ of Equation 2.10 based on a normal distribution with standard deviation σ'_R equal to 2.0. In other words, the log-penalty on annual recruitment residuals was modeled as

$$\ell_R = \sum_j (\sum_y \log_e(\sigma'_R) + \frac{\varepsilon'_{i,y}{}^2}{2\sigma'^2_R}). \quad (2.20)$$

Application of the harvest control rule

To mimic the timing of implementing assessments and setting harvest targets of lake whitefish fisheries in 1836 Treaty-ceded waters, we included a one-year lag between data collection and incorporation in the four stock assessment models. More specifically, an annual assessment was conducted each year of a simulation based on data collected through the previous year, to set the harvest targets for the following year. In the lag year, abundances were projected based on an exponential population model where total mortality rates were assumed to be the mean of the last three years' value, and recruitments were assumed to be the mean of the most recent 10 years. During the year of setting harvest targets (after the lag year), we used the same approach as in the lag year to project abundance at the beginning of that year. We then used Baranov's catch equation (same as in Equation 2.12 and 13) to calculate harvest target, while the

fishing mortality rates were adjusted to the target fishing mortality rates, which can be calculated based on target mortality rates, estimated selectivity-at-age, and natural mortality rate.

Simulation Scenarios

We evaluated five productivity and movement scenarios (Table 2.5), and six sensitivity analysis scenarios (Table 2.3 and Table 2.6). We also evaluated all cross-combinations of productivity/movement scenarios and sensitivity analysis, and full results are available in the supplementary material. For each evaluated scenario, 200 simulations were conducted. In the baseline scenario (Table 2.5), we assumed the four simulated populations had equal stay rates and productivity levels to establish a baseline for comparison of assessment and management performance results. Then we explored alternative operating model settings with different productivity and movement assumptions, to evaluate the consequences of different combinations of productivity and movement dynamics of lake whitefish populations on stock assessments. We also evaluated outcome sensitivity to different quality of assessment data, uncertain mixing rates assumptions, and recruitment variability.

Table 2.6. Scenarios for sensitivity analyses. In each sensitivity scenario, except for the change described below all other parameters are at their baseline levels.

Scenario index	Description	Description of change from baseline scenario
Dat_L	Data quality levels (Table 2.3) all low.	Data quality
Dat_H	Data quality levels (Table 2.3) all high.	Data quality
MixV_Ass	Allowed mixing rates in the assessment model to vary annually about the true value assumed in the operating model.	Mixing rates in the assessment model
RecV_H	Recruitment variation levels (Table 2.3) all high.	Recruitment variation
RecV_0	Recruitment variation levels (Table 2.3) all no autocorrelation.	Recruitment variation
TarA=55%	Target mortality levels all low (Table 2.3).	Target mortality

Baseline scenario and alternative productivity and movement scenarios

We explored five scenarios of population-specific movement dynamics and productivity (scenario 1 is the baseline scenario) (Table 2.5). Overall, there were three different levels of productivity (i.e., low, baseline, and high), and three different stay rates during non-spawning season (low, medium, high). Each productivity level corresponded to a specific steepness parameter, and different productivity levels shared the same unfished equilibrium spawning stock size (Table 2.3). However, higher productivity levels would lead to greater fished equilibrium stock size and recruit levels (Figure 2.2). Target mortality rate (Target_A; annual

death rate=1.0-annual survival rate) was assumed to be 0.65 as a baseline level, which is the current rate used in 1836 Treaty-ceded management of lake whitefish, although as part of sensitivity scenarios explored the effects of a lower target mortality rate.

In the baseline scenario (scenario 1), the four populations had identical "baseline" productivity and stay rates set to "medium" levels. Scenario 2 explored a case in which the four populations still had equal medium levels of movement, but two of the populations had low productivity while the other populations had high productivity. In scenarios 3 to 5, the four populations had different stay rates and either had equal productivity levels (scenario 3) or unequal productivity levels (scenario 4: positive correlation between productivity level and stay rate; scenario 5: negative correlation between productivity level and stay rate).

Sensitivity Analyses

A total of six sensitivity scenarios (Table 2.6) were conducted to determine whether baseline results remained consistent after modifying specific conditions of the examined scenario (e.g., poor data quality). The purpose of the sensitivity analyses was to determine the general applicability of model results.

Data Quality—The first two sensitivity scenarios considered different levels of data quality available for assessment models: low and high (relative to the baseline level), by varying effective sample size (N_{eff}) and the CVs for harvest and effort (Table 2.3 and Table 2.6). The low and high levels of data quality were chosen to reflect the extreme data quality cases evaluated by Li et al. (2016) based on ranges seen in retrospective errors for actual lake whitefish stock assessments in the 1836 Treaty-ceded waters.

Uncertain Mixing Rates—In the baseline scenario, the mixing rates were consistent across populations and simulation years in the operating model, and assumed as correctly known parameters in the stock assessment model. In the third sensitivity scenario, we assumed that annual stay rates in the assessment models were still treated as known parameters, but did not match the true θ_{ii} in the operating model. The annually varying stay rates $\theta_{ii,y}$ used in the assessment model were parameterized by a ‘logistic’ function of re-parameterized rates (ω_y)

$$\theta'_{ii,y} = \exp(\omega'_y) / (\exp(\omega'_y) + 1). \quad (2.21)$$

The annual values for ω'_y were generated from a normal distribution (Table 2.3). Different sets of mean and variance values were assumed to ensure the annually varying stay rates used in the assessments were within 10% of the true θ_{ii} .

Recruitment Variation—For the next two sensitivity scenarios, we explored two recruitment variability levels (Table 2.3 and Table 2.6). In the high recruitment variability scenario, we kept the autocorrelation coefficient at 0.45 as in the baseline scenario but increased the stationary standard deviation in the recruitment process error to 1.5. For the second level, we removed the autocorrelation component of recruitment variation so that the recruitment variation was simply white noise, and kept the same stationary variance as for the baseline scenario.

Target mortality—For the last sensitivity scenario, a lower target mortality rate (Target_A) of 0.55 was implemented in the management procedure because this rate has been identified as sustainable for a wide range of lake whitefish populations with different productivities (Li et al. 2015).

Performance Statistics

Performance statistics for evaluating the different assessment models were average SSB, the proportion of years SSB was less than 20% of the unfished SSB level ($P(SSB < B_{20\%})$), average annual total yield and inter annual variation (IAV) in yield by area, relative error (RE) in the terminal assessment year SSB, and RE of estimating recruitment for all assessed years, over the last 25 years of the simulations. Relative error was calculated as $RE = (\bar{x} - x)/x$, where \bar{x} is the predicted value based on the assessment results and x is the true value generated from the operating model. We additionally estimated the autocorrelation in RE in the terminal assessment year SSB over the last 25 years for each simulation. This was intended to assess autocorrelation in assessment errors under stationary conditions. The autocorrelation was estimated by fitting an AR1 model to the time series of REs in terminal SSB resulting from each simulation by ordinary least squares. We used the `ar.ols` function from `stats` package in R 3.2.2 for the autocorrelation coefficient (ARC) calculation (R Core Team 2016). A large positive ARC would imply that the assessment errors tended to be similar for multiple years in a row. The distributions of the performance statistics calculated over all 200 simulations for an evaluated scenario, were summarized by the median and inter-quartile range. We choose to run 200 simulations because preliminary results of the baseline scenario suggested that results from 200 simulations were nearly identical from those based on 1000 simulations.

Results

In general, all four assessment models converged on solutions. Convergence rate of the assessments was >93% across all scenarios for the origin-informed model (O), the origin-informed model with recruitment penalty (O W/Rec), and the standard model with recruitment penalty (S W/Rec). Although the convergence rate of the standard assessment model (S) was

95% for the baseline scenario, it was less than 90% for other evaluated scenarios. Including a recruitment penalty increased the convergence rate for both standard and origin-informed models by 8.0% and 1.7% on average across all scenarios, with the largest improvement in convergence by 20.8% for the standard assessment approach under scenario 5 with low data quality (Table 2.5 and Table 2.6).

Baseline scenario

Under the baseline scenario, where the simulated populations had the same stay rates and productivity levels, the expected assessment and management performance was the same across all populations, and indeed the realized performance results were nearly identical (see full results in Supplementary material). Consequently, we summarize the results for only one of the four populations (i.e., Population 1 in Table 2.5). Compared to the standard assessment models (i.e., S and S W/Rec), adding population-specific harvest age composition in the origin-informed assessment models (i.e., O and O W/Rec) in general resulted in less bias and more weakly autocorrelated estimates of SSB in the terminal assessment year with smaller inter-quartile ranges (Figure 2.3a and Figure 2.3f), and less uncertainty in estimates of recruitment (based on smaller inter-quartile ranges of RE) over all assessment years except for the final two years (Figure 2.3b). However, the origin-informed assessment model performance did not translate into benefits in the management performance statistics, such as average true SSB and yield, with only slightly improvement in the IAV of yield (3c, 3d and 3e, and supplementary materials). When a recruitment penalty was added to both the standard and origin-informed assessment models (comparing S W/Rec and O W/Rec with S and O), this resulted in less IAV of yield (median IAV of yield decreased by 0.05 and 0.04 for standard and origin-informed models, Figure 2.3e), and lower bias in estimates of recruitment for the last two assessment years (Figure

2.3b), but slightly higher risk of SSB being lower than 20% of its unfished level (median $P(SSB < B_{20\%})$ increased by 3.8% and 7.7%, Figure 2.3c).

Both the standard and origin-informed assessment models without recruitment penalties had considerable difficulty in estimating recruitment levels in the terminal assessment year. In most simulations, the recruitment RE in the terminal assessment year was -100%, meaning that recruitment was being estimated at essentially 0 fish (Figure 2.3b and Figure 2.4). However, when a recruitment penalty was included in the assessments (comparing S W/Rec with O W/Rec), the origin-informed assessment model (i.e., O W/Rec) produced less biased estimates for the terminal assessment year recruitment (Figure 2.3b and Figure 2.4).

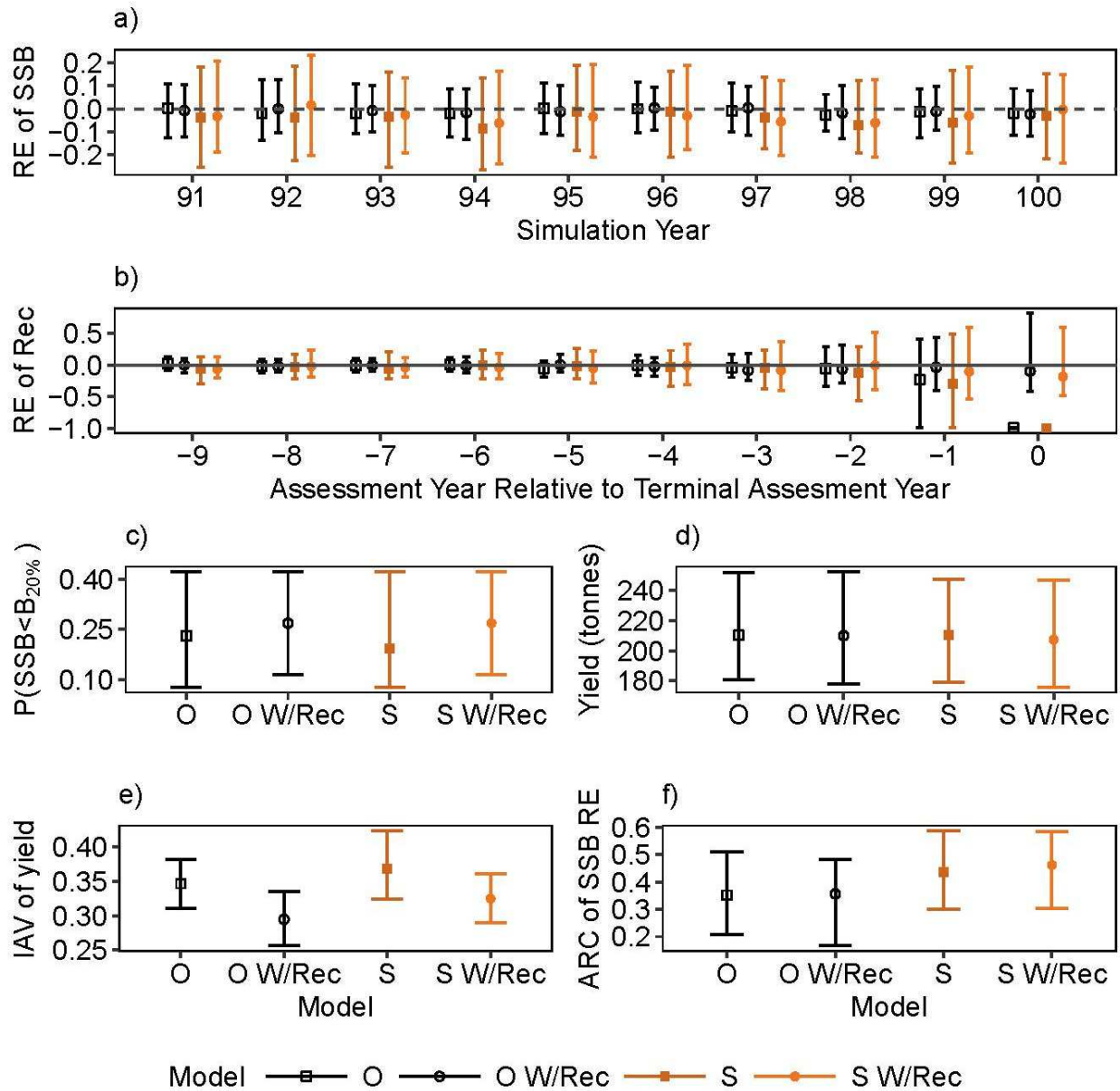


Figure 2.3. Simulation results (median \pm interquartile range) for population 1 (Table 2.5) in the baseline scenario. Full model names are in Table 2.1. (a) Relative error of estimating terminal assessment year SSB during simulation year 91 to 100. (b) In simulation year 100, relative error of estimating recruitment of the last ten assessment years. (c) Proportion of years SSB was lower than 20% of the unfished SSB level ($B_{20\%}$) over the last 25 years of simulations. (d) Mean annual yield for the fishing area surrounding spawning grounds of Pop1 over the last 25 years of simulations. (e) Mean interannual variation (IAV) in yield over the last 25 years of simulation. (f) Estimated autocorrelation for terminal year estimates of SSB during simulation years 75 to 100.

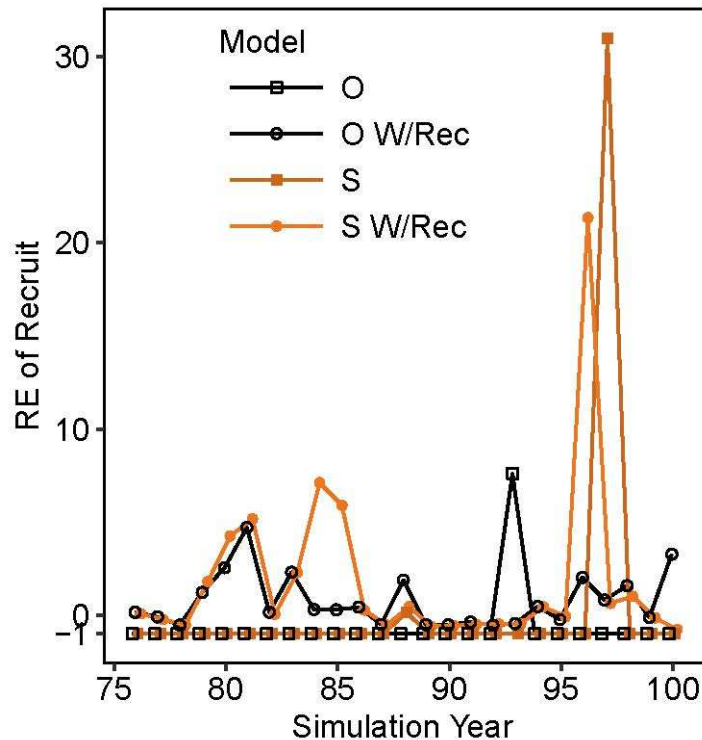


Figure 2.4. Relative error in estimates of recruitment for the terminal assessment year during the simulation year 76 to 100 for an example simulation. Full model names are in Table 2.1.

Alternative productivity and movement scenarios

For the alternative productivity and movement scenarios, we present results only for populations 1 and 3 because for these scenarios populations 1 and 2 and populations 3 and 4 had nearly identical results due to their same productivity and stay rates. When low and high productivity populations intermixed (Scenario 2, 4, and 5 in Figure 2.5), low productivity populations generally had high risk of being overfished (i.e., the interquartile ranges of average true SSB were below 20% of the unfished level) across all scenarios. Regardless of whether a penalty for annual recruitment residuals was included, the origin-informed assessment models (i.e., O and O W/Rec) substantially outperformed the standard assessment models (S and S W/Rec) in terms of estimation of SSB of the terminal assessment year for low productivity populations, but using population-specific harvest age composition data had only a slight

influence on estimation of SSB for high productivity populations. More specifically, for the low productivity populations, the RE of estimated terminal assessment year SSB in year 100 was less biased, and the autocorrelation for these estimates over the last 25 years was lower for assessment models O and O W/Rec than for S and S W/Rec. Such differences in assessment performance were greater for scenarios where there was a negative correlation between stay rates and productivity. For the scenario where populations had the same productivity but different stay rates (Scenario 3), assessment performance results were similar to those of the baseline scenario.

With respect to the estimation of terminal assessment year recruitment and for management performance statistics, results for all alternative productivity and movement scenarios were similar to those found in the baseline scenario. Neither the standard or origin-informed assessment models without recruitment penalties could produce reliable estimates of recruitment in the terminal assessment year. When low productivity populations intermixed with high productivity populations (Scenario 2, 4, and 5 in Figure 2.5), standard and origin-informed assessment models with recruitment penalties resulted in unbiased recruitment estimates in the terminal assessment year for high productivity population, but positive bias in recruitment estimates in the terminal assessment year for low productivity populations.

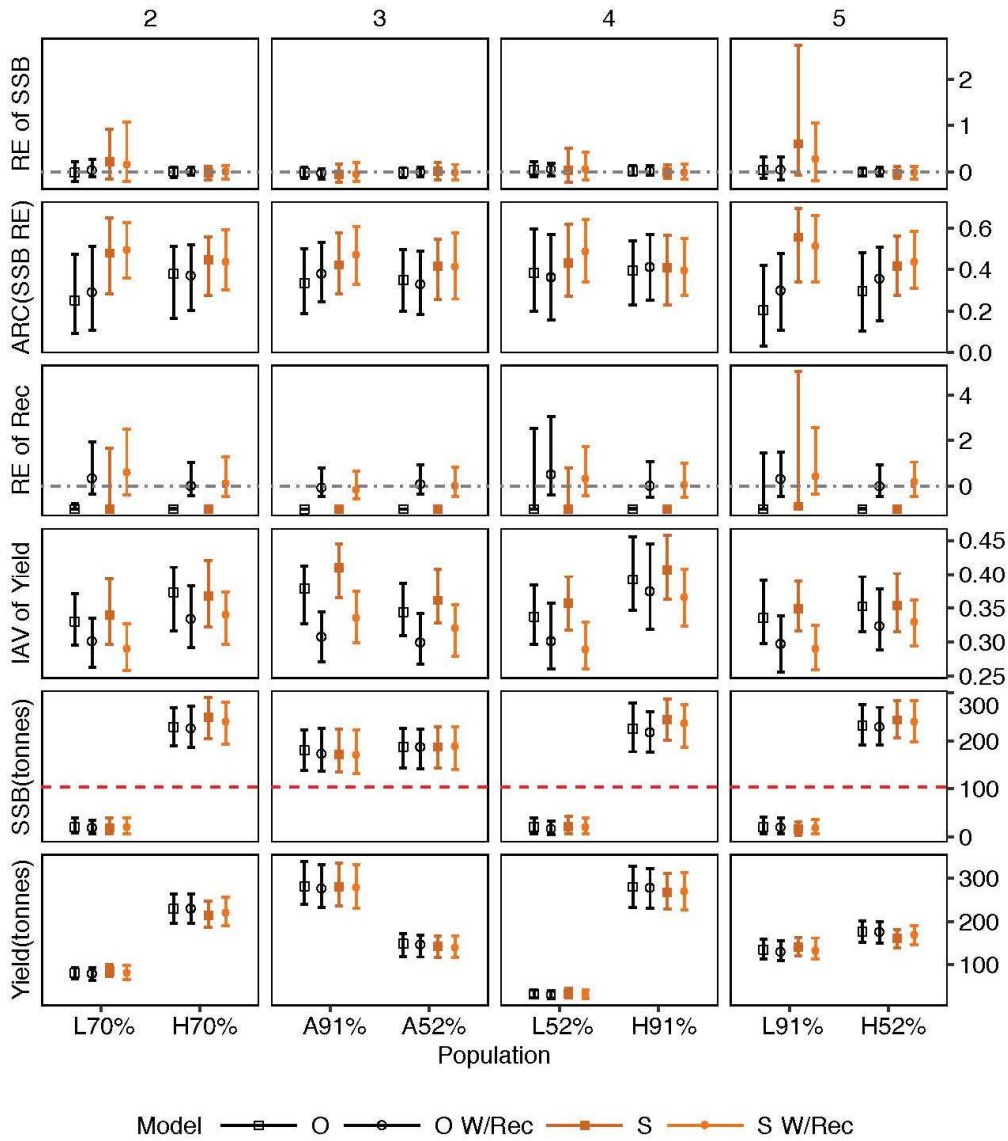


Figure 2.5. Simulation results (median \pm interquartile range) for populations 1 and 3 under scenarios 2 to 5 (Table 2.5). Full model names are in Table 2.1. Each column represents a different productivity and movement scenario, and each row presents a different performance statistic. The x-axis of each column indicates the productivity levels (L, A, H are low, average, and high productivity levels) and stay rates associated with the two populations results are presented for. For example, L70% means low productivity population with 70% stay rate. For each such productivity level and stay rate, results are given for the four different assessment methods, distinguished by different symbols. The second, fourth, and sixth rows represent the same performance statistics as for Figure 2.3c, 2.3e, and 2.3d. The first and third row are relative error of estimating terminal year SSB and recruitment in simulation year 100, respectively, with a 0 dashed line. The fifth row represents the average SSB over the last 25 years of simulation, and the dashed line is 20% of the unfished SSB.

Sensitivity Analyses

The assessment and management performances for all the assessment models were generally insensitive to changes in the magnitude of actual recruitment variation, target mortality, data quality, and to uncertain mixing rates assumptions (Figure 2.6), with patterns in performance statistics similar to those of the baseline scenario. There were only three exceptions. First, with a lower total mortality target (55%), the origin-informed assessment models both with and without recruitment penalties had better management and assessment performance than the standard assessment models, as evidenced by lower $P(SSB < B_{20\%})$ (median at 0.08 for O and at 0.12 for S), similar or even higher yield (median at 204.8 for O and at 204.2 for S), lower IAV of yield (median at 0.32 for O and at 0.35 for S), and less biased with smaller inter-quartile range (inter-quartile range [-0.13,0.10] for O and [-0.18,0.15] for S), and less autocorrelated estimates of SSB (median at 0.37 for O and at 0.44 for S) in the terminal assessment year. Second, when recruitment variation was high, $P(SSB < B_{20\%})$ was higher, and average yields were lower for all four assessment models. In addition, for this high recruitment scenario both assessment models with recruitment penalties tended to overestimate recruitment (RecV_H in Figure 2.6). Finally, when assessment data quality was low (RecV_L in Figure 2.6), all four assessment models tended to underestimate SSB, have greater IAV of yield, and greater inter-quartile range for the RE of estimating terminal year SSB.

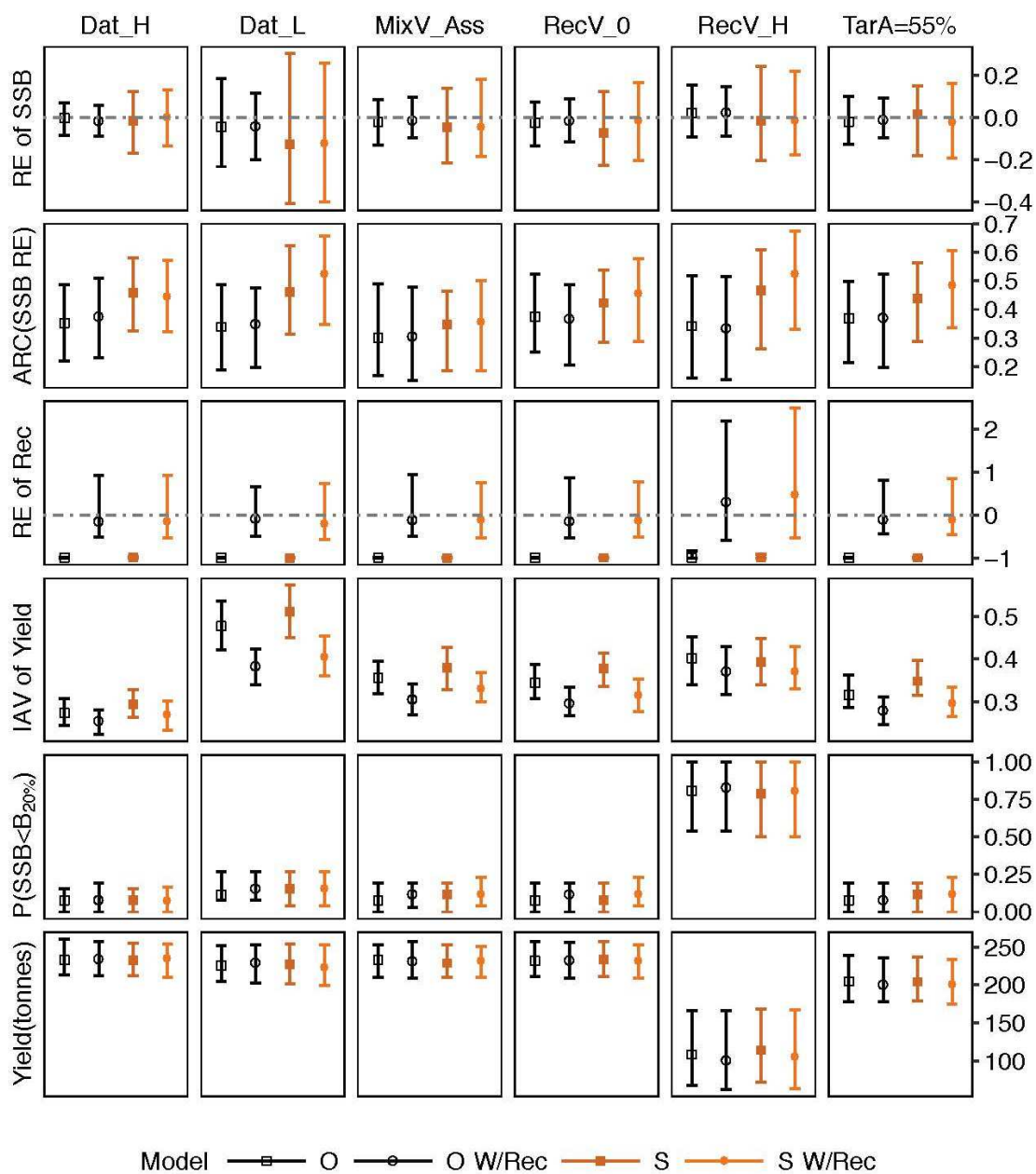


Figure 2.6. Simulation results (median \pm interquartile range) for Pop1 (Table 2.5) in sensitivity analyses. Full model names are in Table 2.1. Each column represents a sensitivity scenario, each row represents a performance metric (as described in Figure 2.5), and results in each panel are for the four assessment models.

Discussion

Attempting to account for movement in fish stock assessment models has become increasingly common for the management of intermixed fisheries (Cope and Punt 2011; Ying et al. 2011; Molton et al. 2012; Li et al. 2015; Vincent et al. 2017). In this study, we evaluated four spatially-structured SCAA models (standard assessment, standard assessment with recruitment penalty, origin-informed assessment, origin-informed assessment with recruitment penalty) for assessing lake whitefish populations that were assumed to exhibit an overlap movement strategy. We aimed to evaluate if considering additional assessment data about classification of catch to spawning origin, and adding a penalty for annual recruitment residuals, could improve the assessment and management performance of the overlap SCAA model proposed by Li et al. (2015). We found that data allowing parsing of catch from a management area to the specific spawning population the fish came from could result in less biased and less auto-correlated estimates of spawning stock biomass (SSB) in terminal assessment years, and less uncertainty in estimates of recruitment early in the time period assessed; while including a lognormal penalty on annual recruitment residuals in assessment models substantially improved the estimation of recruitment in the terminal assessment years. With the penalty, data on population source also led to improved terminal recruitment estimates.

When we used data on the classification of catch to spawning origin in our proposed overlap assessment models, we assumed a multinomial distribution of population-age composition for each year of harvest from an area. This is an extension of what we assumed in our standard SCAA model in which a multinomial distribution was assumed, as is often done, for age composition of harvest. Use of these additional data did provide better estimation of the spawning stock biomass (SSB) in the terminal assessment year. Hintzen et al. (2015) reached a

similar conclusion but with a small level of improvement when they used such data to inform survey indices for an integrated stock assessment model. This may be due to the mismatch between the spatial structures in their assessment data of catch and in the assessment model. Although spawning origin information allowed the assessment model to incorporate correct (or with uncertainty) survey indices, because their assessment model ignored spatial structure in the observed catch data such a mismatch can still lead to biased estimation of biomass and recruitment. Our results suggested that such improvements in assessment performance did not necessarily translate into improved management performance, except when we used a lower than status-quo mortality target. Under the status-quo mortality target, although the origin-informed assessment models provided better estimation of SSB than the standard overlap models, the calculated total allowable catch (TAC) based on the estimated SSB was still not sustainable. Coincidentally, because the standard assessment models tended to underestimate SSB, it resulted in a more “appropriate/conservative” TAC. This argument is evidenced by our sensitivity analysis with lower target mortality rate (Target_A=55%) in which origin-informed assessment models had better management and assessment performance than standard assessment models.

Past studies have found that when populations with different productivity levels intermix during harvest season, populations with lower productivity are generally more vulnerable to overharvest (Ricker 1958; Paulik et al. 1967; Hintzen et al. 2015; Li et al. 2015). The results from this study are consistent with those studies. We found that there was a high risk of being overfished for low productivity populations, especially when low productivity populations with high stay rates intermixed with high productivity populations with low stay rates. In such a case, for low productivity populations, standard assessment models tended to overestimate SSB, while the origin-informed assessment models provided nearly median unbiased estimation of SSB. We

suspect that the standard assessment model is challenged to identify the correct age composition for low productivity populations from the aggregate sample collected from each harvest area, because they consist of mixtures of age compositions from populations with different productivity, with contributions depending on population productivities and movement rates. Conversely, information on population-specific age compositions for area-specific harvests provides sufficient information to prevent inaccuracies in SSB estimates.

Our sensitivity analysis suggested that the improvement by including population-specific age compositions for area-specific harvests was limited to scenarios without high assessment data quality. In other words, when data quality is high, standard assessment models can provide sufficiently accurate estimates of population-specific SSB when supplied with accurate mixing rates. Thus, an origin-informed assessment model may not be necessary in conditions of high data quality and accurate information on mixing. We must emphasize that our consideration of data quality was focused on precision rather than potential biases in data. We also did not consider model misspecification except for the unmatched mixing rates assumed in the operating and assessment models in the sensitivity analyses, and our stochastic assumptions regarding recruitment for the models with recruitment penalties. A formal evaluation of how model misspecification affects the performance of spatially structured stock assessment model was outside the scope of our research but we would encourage investigations on this topic. We anticipate consequences of model misspecification to be case specific. Some cases of model misspecification may change the scale of biomass assessment, and this would not change the relative performance of the four assessment models we evaluated because target F in all assessment models would be adjusted to count for bias in similar manners. In other cases, however, model misspecification may lead to too high estimation errors. In such there may not

be a strong justification for collecting population-specific data because the advantages of origin-informed assessment models over the standard models may not be clear.

The other major finding from this research was that including a lognormal penalty on annual recruitment residuals in both standard and origin-informed assessment models markedly improved the estimation of recruitment at the end of the assessment period. This is consistent with what has been found in evaluations of stock assessments without spatial structure (Maunder and Deriso 2003; Methot et al. 2011; Korman et al. 2012). Although the inclusion of a recruitment penalty did not prevent recruitment from being overestimated when recruitment variation in the operating model was high, its performance was still better than when a recruitment penalty was not included. This overestimation may stem, in part, from the large standard deviation for the distribution governing the annual recruitment deviations in the assessment models with recruitment deviations. We also found that IAV of yield was lower when a recruitment penalty was incorporated. This may result from the more stable/reasonable estimation of recruitment at the end of the assessment year period. Such stabilization of recruitment estimates can lead to a more stable prediction of future abundance, and that is what the TAC calculation is based on. Also, because we included a 1-year lag between assessment data collection and assessment model implementation to mimic the real management procedure for lake whitefish in Laurentian Great Lakes region, the impact of recruitment estimation near the end of the time series is magnified, given we needed to project an additional year over what is assumed in some studies.

In summary, we found that for a spatially structured SCAA model that incorporated information on population-specific age composition of harvest resulted in less biased and less correlated estimates of spawning stock biomass (SSB) in terminal assessment years, and less

uncertainty in estimating recruitment in early assessment years. Including a lognormal penalty on annual recruitment residuals in the spatial structured SCAA model substantially improved the estimation of recruitment in the terminal assessment years, which we suggest as “best practice” for spatially-structured assessment models. Despite the improved assessment performance, preventing overharvest of low productivity populations when using such assessments will still require an appropriate harvest policy, such as lower target mortality rates or precautionary reference points. Different approaches for parsing catch to contributing populations are likely to have different levels of classification accuracy. For example, genetic classification methods may be more accurate than otolith microchemistry methods if there are not strong environmental differences among spawning locations. Further research into how assessment model performance is affected by classification accuracy would be beneficial. We also recommend additional investigation of factors such as the inclusion of more complex spatial structure (e.g., seasonal movement), alternative harvest policies, model misspecification, and alternative spatial structured stock assessment models (e.g., spatially structured virtual population analysis, tag integrated assessment model) to evaluate the benefits of parsing catch to spawning populations when it comes to the management of spatially-structured populations.

Acknowledgements

Funding for this project was provided by the Great Lakes Fishery Commission (Project 2013_BRE_44025) and Great Lakes Fishery Trust (Project 2012.1250). Additional funding support was provided by supporting partners of the Quantitative Fisheries Center and Michigan Department of Natural Resources through the Partnership for Ecosystem Research and Management program. We acknowledge the support of the Michigan State University High Performance Computing Center and the Institute for Cyber-Enabled Research. We wish to thank

the Modeling Subcommittee of the Technical Fisheries Committee for 1836 Treaty-ceded waters for providing input on our research. This is publication 201X-XX of the Michigan State University Quantitative Fisheries Center.

APPENDIX

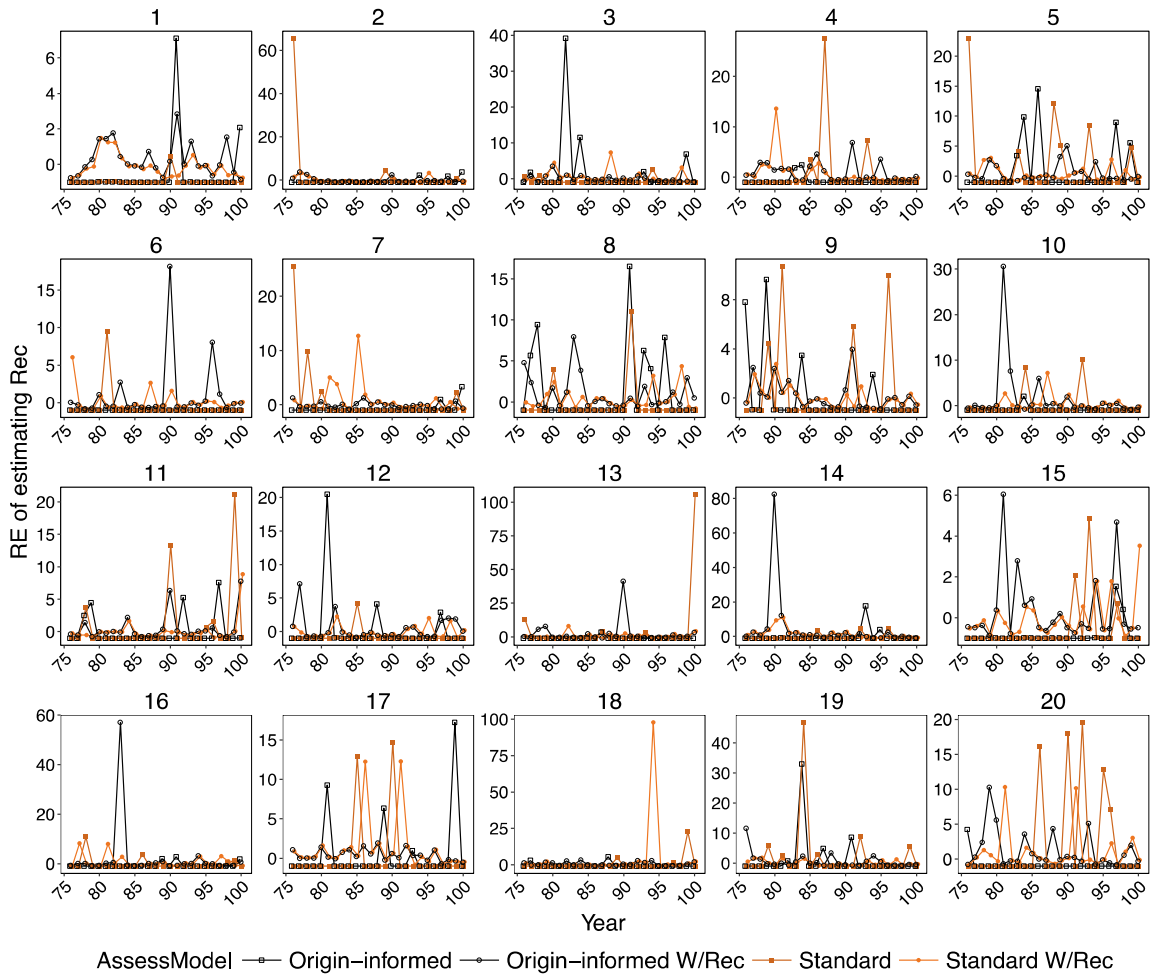


Figure 2.A1. Relative error of estimating recruitment for the terminal assessment year during the simulation year 75 to 100 in 20 example simulations.

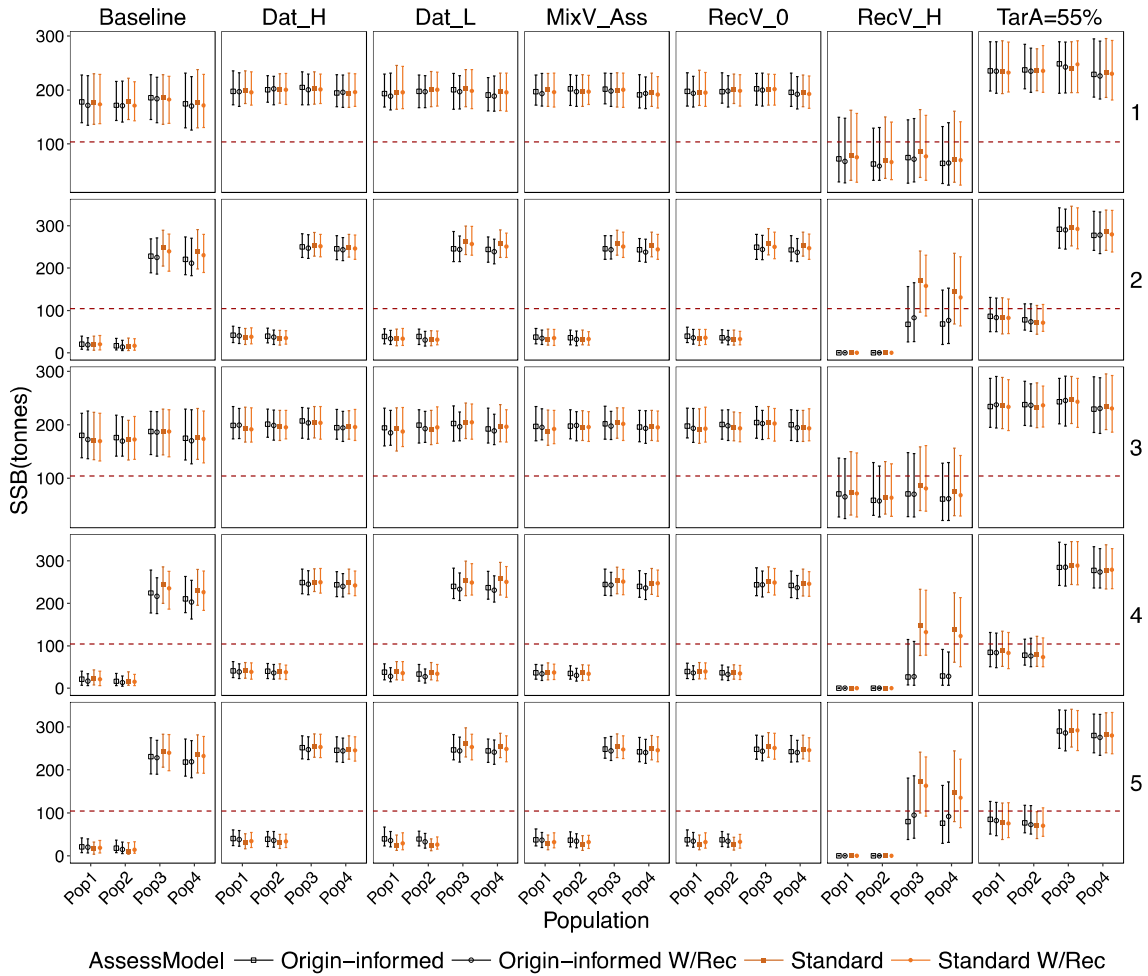
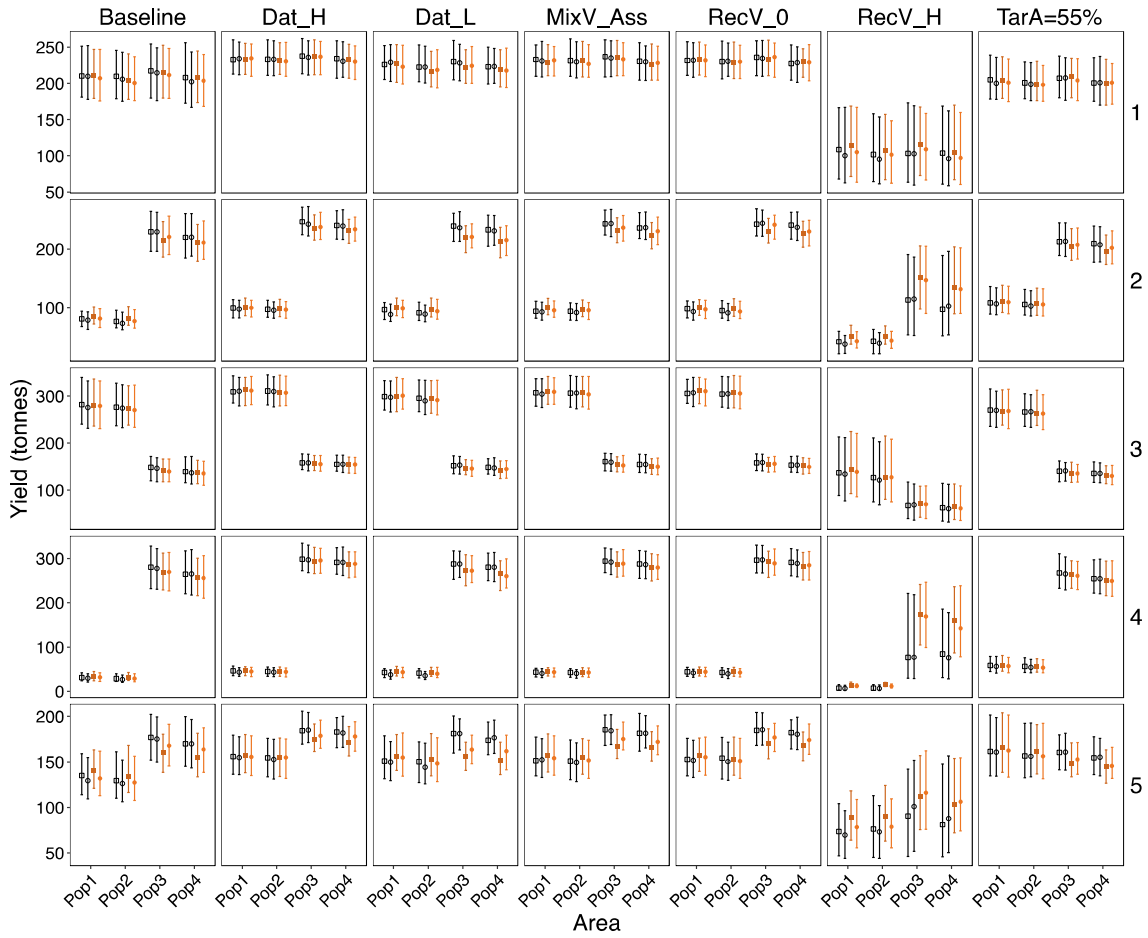


Figure 2.A2. Mean annual spawning stock biomass of each spawning population over the last 25 years of simulations. Each row of the panels represents different population productivity and stay rate scenarios (Table 2.5), and each column represents different scenarios of sensitivity analyses (Table 2.6). The x-axis of each panel represents the results of each spawning population. The dashed line is 20% of the unfished SSB. The symbols identify the median and the error bars identify the inter-quartile range of the simulations.



AssessModel —●— Origin-informed —●— Origin-informed W/Rec —■— Standard —■— Standard W/Rec

Figure 2.A3. As for Figure 2.A2, except y-axis is mean annual yield for the fishing area surrounding spawning grounds of each spawning population (populations are intermixed with each other in each fishing area during harvest season) over the last 25 years of simulations done for each scenario.

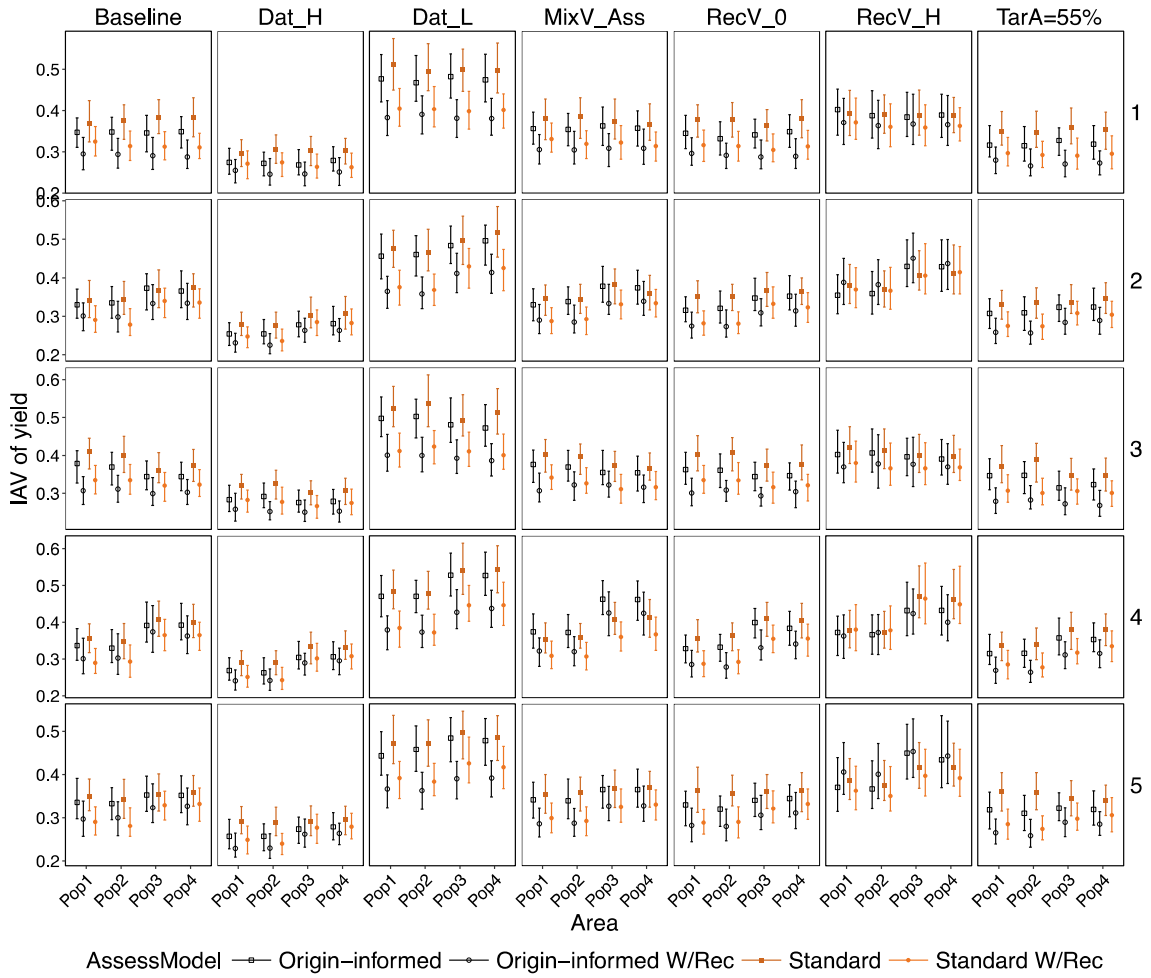


Figure 2.A4. As for Figure 2.A2, except y-axis is the mean interannual variation (IAV) in yield over the last 25 years of simulation.

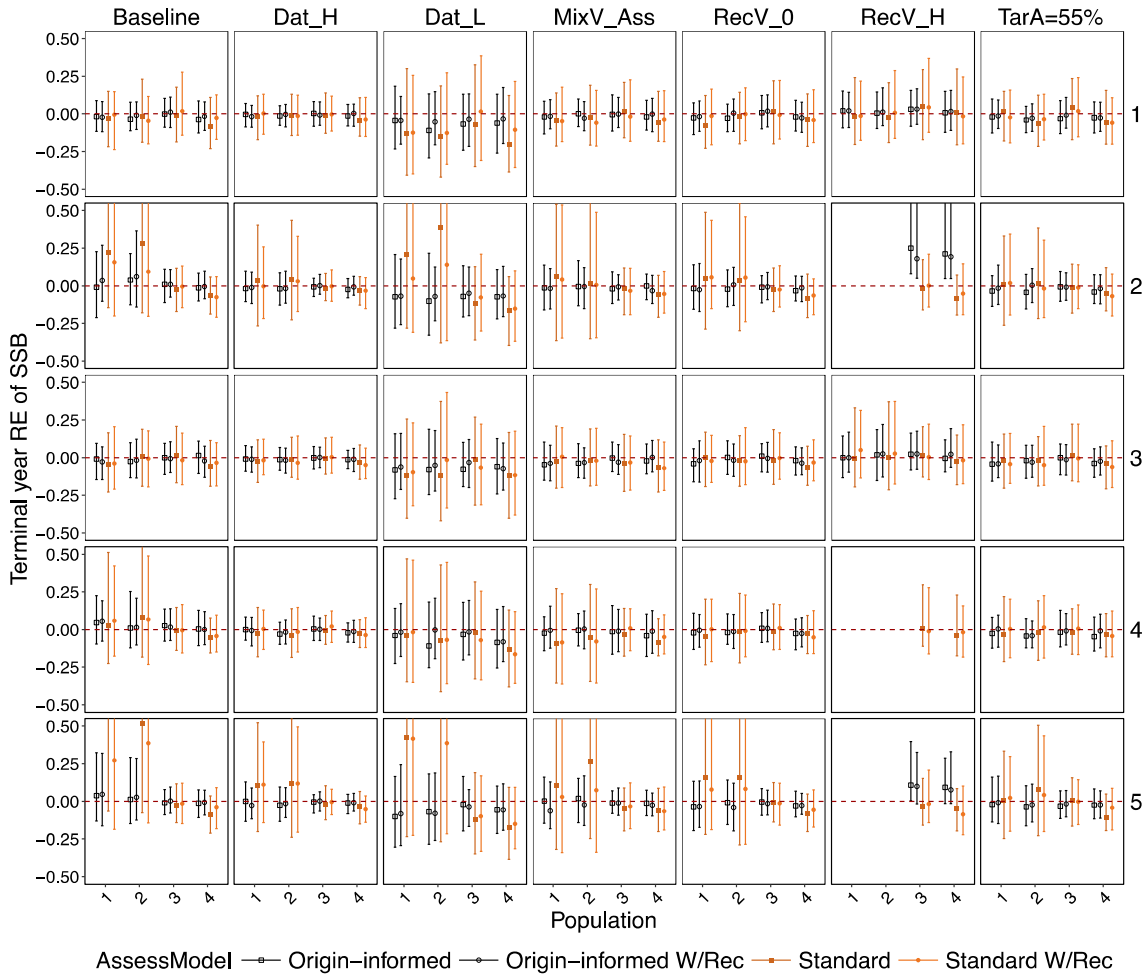


Figure 2.A5. As for Figure 2.A2, except y-axis is the relative error of estimating SSB of terminal assessment year.

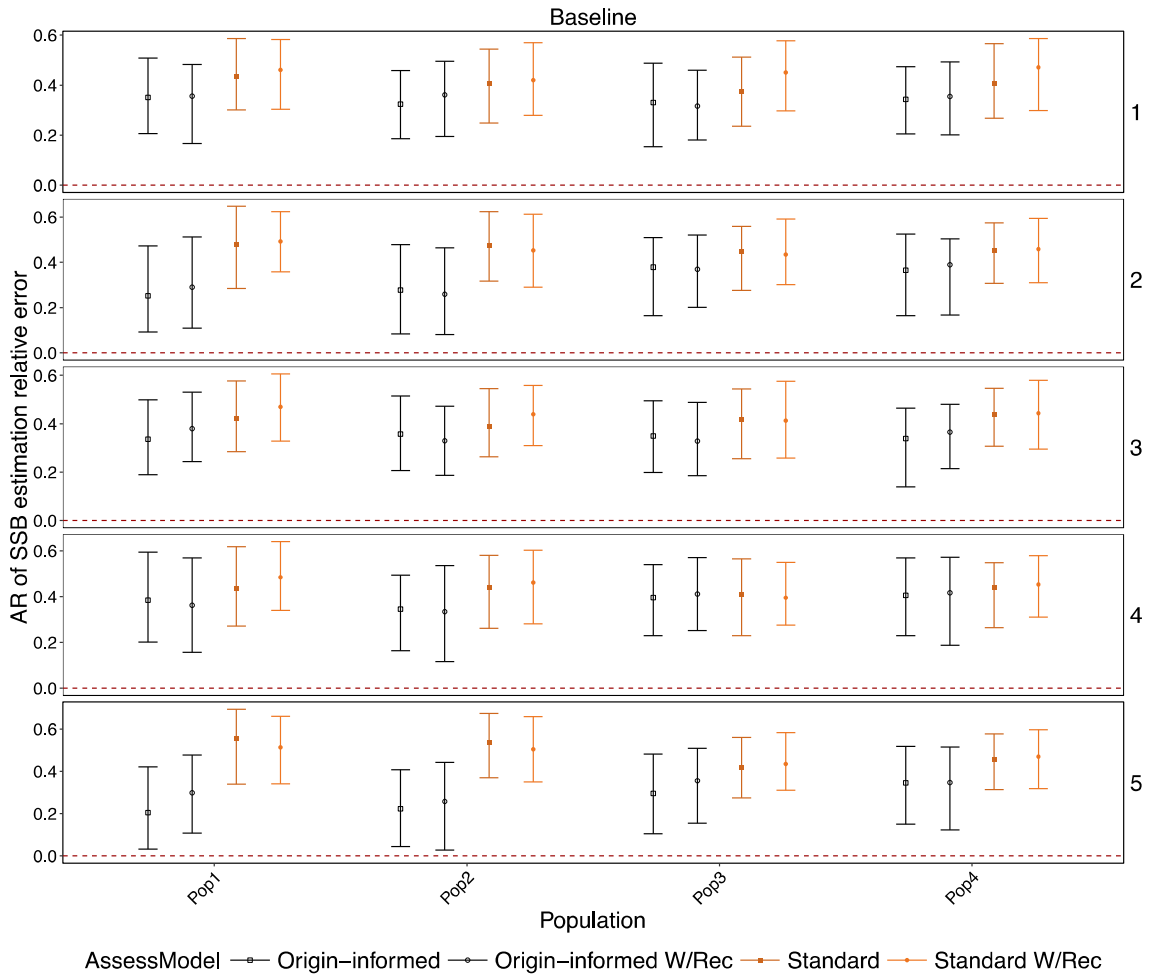


Figure 2.A6. Autocorrelation coefficient of estimating terminal year SSB during simulation year 75 to 100 (median \pm interquartile range) under all productivity and movement scenarios (include baseline scenario in the first row).

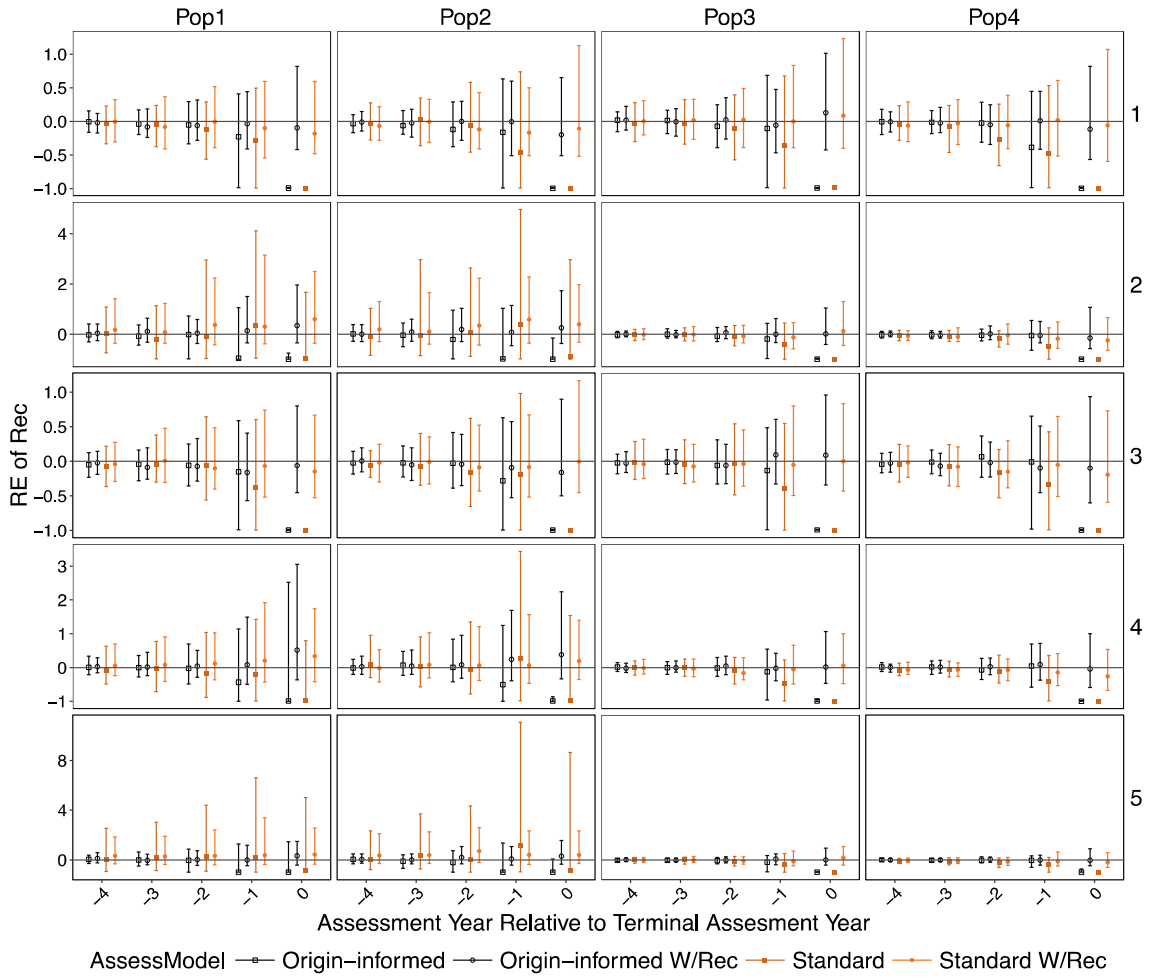


Figure 2.A7. Relative error (median \pm interquartile range) of estimating recruitment of assessment year 15 to 20 (i.e., year -4 to 0 relative to terminal assessment year) under all productivity and movement scenarios.

BIBLIOGRAPHY

BIBLIOGRAPHY

- 1836 Treaty Waters Modeling Subcommittee, Technical Fisheries Committee. 2017. Technical fisheries committee administrative report 2017: status of lake trout and lake whitefish populations in the 1836 treaty-ceded waters of Lakes Superior, Huron and Michigan, with recommended yield and effort levels for 2017. doi:10.13140/RG.2.2.19581.15844.
- Cope, J.M., and Punt, A.E. 2011. Reconciling stock assessment and management scales under conditions of spatially varying catch histories. *Fish. Res.* **107**(1–3): 22–38. doi:10.1016/j.fishres.2010.10.002.
- Ebener, M.P., Brenden, T.O., and Jones, M.L. 2010a. Estimates of fishing and natural mortality rates for four lake whitefish stocks in northern Lakes Huron and Michigan. *J. Great Lakes Res.* **36**: 110–120. doi:10.1016/j.jglr.2009.06.003.
- Ebener, M.P., Brenden, T.O., Wright, G.M., Jones, M.L., and Faisal, M. 2010b. Spatial and temporal distributions of lake whitefish spawning stocks in Northern lakes Michigan and Huron, 2003–2008. *J. Great Lakes Res.* **36**: 38–51. doi:10.1016/j.jglr.2010.02.002.
- Eveson, J.P., Laslett, G.M., and Polacheck, T. 2009. A spatial model for estimating mortality rates, abundance and movement probabilities from fishery tag-recovery data. *In* Modeling Demographic Processes In Marked Populations. doi:10.1007/978-0-387-78151-8.
- Fournier, D.A., Skaug, H.J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M.N., Nielsen, A., and Sibert, J. 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optim. Methods Softw.* **27**(2): 233–249. doi:10.1080/10556788.2011.597854.
- Fournier, D., and Archibald, C.P. 1982. A general theory for analyzing catch at age data. *Can. J. Fish. Aquat. Sci.* **39**(8): 1195–1207. doi:10.1139/f82-157.
- Fu, C., and Fanning, L.P. 2004. Spatial considerations in the management of Atlantic cod off Nova Scotia, Canada. *North Am. J. Fish. Manag.* **24**(3): 775–784. doi:10.1577/M03-134.1.
- Guan, W., Cao, J., Chen, Y., Cieri, M., and Quinn, T. 2013. Impacts of population and fishery spatial structures on fishery stock assessment. *Can. J. Fish. Aquat. Sci.* **70**(8): 1178–1189. doi:10.1139/cjfas-2012-0364.
- Hintzen, N.T., Roel, B., Benden, D., Clarke, M., Egan, A., Nash, R.D.M., Rohlf, N., and Hatfield, E.M.C. 2015. Managing a complex population structure: exploring the importance of information from fisheries-independent sources. *ICES J. Mar. Sci.* **72**(2): 528–542. doi:10.1093/icesjms/fsu102.

- Hutchings, J.A. 1996. Spatial and temporal variation in the density of northern cod and a review of hypotheses for the stock's collapse. *Can. J. Fish. Aquat. Sci.* **53**(5): 943–962. doi:10.1139/f96-097.
- Korman, J., Martell, S.J.D., Walters, C.J., Makinster, A.S., Coggins, L.G., Yard, M.D., Persons, W.R., and Quinn, T. 2012. Estimating recruitment dynamics and movement of rainbow trout (*Oncorhynchus mykiss*) in the Colorado River in Grand Canyon using an integrated assessment model. *Can. J. Fish. Aquat. Sci.* **69**(11): 1827–1849. doi:10.1139/f2012-097.
- Li, Y., Bence, J.R., and Brenden, T.O. 2015. An evaluation of alternative assessment approaches for intermixing fish populations: a case study with Great Lakes lake whitefish. *ICES J. Mar. Sci.* **72**(1): 70–81. doi:10.1093/icesjms/fsu057.
- Li, Y., Bence, J.R., and Brenden, T.O. 2016. The influence of stock assessment frequency on the achievement of fishery management objectives. *North Am. J. Fish. Manag.* **36**(4): 793–812. doi:10.1080/02755947.2016.1167145.
- Li, Y., Bence, J.R., Zhang, Z., and Ebener, M.P. 2017. Why do lake whitefish move long distances in Lake Huron? Bayesian variable selection of factors explaining fish movement distance. *Fish. Res.* **195**: 169–179. doi:10.1016/j.fishres.2017.07.014.
- Maunder, M.N., and Deriso, R.B. 2003. Estimation of recruitment in catch-at-age models. *Can. J. Fish. Aquat. Sci.* **60**(10): 1204–1216. doi:10.1139/f03-104.
- Methot, R.D., and Taylor, I.G. 2011. Adjusting for bias due to variability of estimated recruitments in fishery assessment models. *Can. J. Fish. Aquat. Sci.* **68**(10): 1744–1760. doi:10.1139/f2011-092.
- Modeling Subcommittee, T.F.C. 2017. Technical fisheries committee administrative report 2017: status of lake trout and lake whitefish populations in the 1836 treaty-ceded waters of Lakes Superior, Huron and Michigan, with recommended yield and effort levels for 2017. doi:10.13140/RG.2.2.19581.15844.
- Molton, K.J., Brenden, T.O., and Bence, J.R. 2012. Control rule performance for intermixing lake whitefish populations in the 1836 Treaty waters of the Great Lakes: A simulation-based evaluation. *J. Great Lakes Res.* **38**(4): 686–698. doi:10.1016/j.jglr.2012.09.014.
- Morishima, G.S., and Henry, K.A. 1999. The history and status of pacific northwest chinook and coho salmon ocean fisheries and prospects for sustainability. *In Sustainable Fisheries Management: Pacific Salmon. Edited by* E.E. Knudsen, C.R. Steward, D.D. MacDonald, J.E. Williams, and D.W. Reiser. CRC Press, Boca Raton, FL. pp. 219–236.
- Paulik, G.J., Hourston, A.S., and Larkin, P.A. 1967. Exploitation of Multiple Stocks by a Common Fishery. *J. Fish. Res. Board Canada* **24**(12): 2527–2537. doi:10.1139/f67-202.

- Porch, C.E., Turner, S.C., and Powers, J.E. 2001. Virtual population analyses of atlantic bluefin tuna with alternative models of transatlantic migration: 1970-1997. *Int. Comm. Conserv. Atl. Tunas Collect. Vol. Sci. Pap.* **52**(3): 1022–1045.
- R Core Team. 2016. R: A language and environment for statistical computing. R Found. Stat. Comput. doi:10.1007/978-3-540-74686-7.
- Ricker, W.E. 1958. Maximum Sustained Yields from Fluctuating Environments and Mixed Stocks. *J. Fish. Res. Board Canada* **15**(5): 991–1006. doi:10.1139/f58-054.
- Stott, W., VanDeHey, J.A., and Sloss, B.L. 2010. Genetic diversity of lake whitefish in lakes Michigan and Huron; sampling, standardization, and research priorities. *J. Great Lakes Res.* **36**(SUPPL. 1): 59–65. doi:10.1016/j.jglr.2010.01.004.
- Vandergoot, C.S., and Brenden, T.O. 2014. Spatially varying population demographics and fishery characteristics of Lake Erie walleyes inferred from a long-term tag recovery study. *Trans. Am. Fish. Soc.* **143**(1): 188–204. doi:10.1080/00028487.2013.837095.
- Vincent, M.T., Brenden, T.O., and Bence, J.R. 2017. Simulation testing the robustness of a multiregion, tag-integrated assessment model that exhibits natal homing and estimates natural mortality and reporting rate. *Can. J. Fish. Aquat. Sci.* **74**(11): 1930–1949. doi:10.1139/cjfas-2016-0297.
- Ying, Y., Chen, Y., Lin, L., and Gao, T. 2011. Risks of ignoring fish population spatial structure in fisheries management. *Can. J. Fish. Aquat. Sci.* **68**(12): 2101–2120. doi:10.1139/f2011-116.

CHAPTER 3

A BAYESIAN SPATIAL STRUCTURED TAGGING MODEL FOR ESTIMATING MOVEMENT PROBABILITIES, MORTALITY RATES, TAG REPORTING RATES AND ABUNDANCE FOR OVERLAPPING FISH POPULATIONS

Abstract

We propose a spatially structured tagging model framework for fish populations with overlapping spatial structure, and a Bayesian approach to fitting such models. The approach includes real-world applications for specification of prior probabilities for model parameters incorporating biological understanding, specific approaches to determining the posterior probability distribution to make inferences regarding parameters, and use of DIC to determine model parameterization. Our framework allows simultaneously estimations of movement, natural mortality and tag reporting rates, and can be used to evaluate the potential data conflict between catch-at-age and tag-recovery data sets. The results of our application to lake whitefish indicates that our model had robust estimates of movement rates, but natural mortality estimates were sensitive to assumptions about data limitation, such as combining adjacent basins without observer coverage, assuming same tag reporting rates for the combined basin, and assuming the movement rates equal to zeros for the regions with no observed recoveries. Our results emphasize the importance of having a robust tag recovery monitoring program across the entire area being considered for a spatial tagging model.

Introduction

A range of quantities can potentially be estimated by simultaneously analyzing multiyear tag-recovery and catch data, such as year-, and region-specific fishing and

natural mortality rates, and stock abundance (Polacheck et al. 2006, Goethel et al. 2015b). These quantities are useful for understanding the dynamics and productivity of populations. Compared to other assessment methods that also use catch-at-age data or catch-at-length data and indices of relative abundance, incorporating tag-recovery data can provide extra information on mortality rates, such as from the differential between expected returns at same age but from different tagging cohorts (Polacheck et al. 2006). With additional observer data, tag-reporting rate can also be estimated within tagging models by partitioning tag returns into those from observed, and unobserved components of catch, and assuming 100% tag reporting rates for the observed component (Pollock et al. 2002, Eveson et al. 2007).

For fisheries occurring over relatively large regions, where fish from a given harvest area originated from fish that spawned in several spatially distinct areas, parameters describing how mixing occurs are additional key parameters to be estimated from tag-recovery data (Eveson et al. 2009, Vandergoot and Brenden 2014). Models that use tagging data to estimate parameters describing this mixing, as well as population dynamics parameters, are called spatial tagging models. To use such models, fish are tagged and released from multiple locations, and their subsequent recovery locations reveal where the tagged fish moved to from the tagging sites. Eveson *et al.* (2009) proposed a spatial Brownie-Petersen model, which is a combination of two traditional but different approaches, a Brownie model (Brownie et al. 1985) and a Petersen-type model (Seber 1982), for analyzing such tag-recovery data. Vandergoot and Brenden (2014) applied a spatial Brownie model to walleye tag-recovery data in Lake Erie. However, existing applications of the spatial Brownie-Petersen model or spatial Brownie model

have implicitly or explicitly assumed that the probability of where a fish will be located in the next time period depends on the location of a fish in the current time period (i.e., metapopulation spatial structure, Guan *et al.*, 2013), and not the fish's natal site location (i.e., there is no accounting for homing, (Pollock *et al.* 2002, Eveson *et al.* 2007, 2009, Vandergoot and Brenden 2014). In contrast, some important marine and freshwater fish species demonstrate natal homing, such as Atlantic cod (*Gadus morhua*) and lake whitefish (*Coregonus clupeaformis*) (Thorrold 2001, Svedäng *et al.* 2007, Li *et al.* 2017). Hereafter we use “overlapping spatial structure” to refer to the situation where fish from several populations are mixed together on one or several common fishing grounds, but where the fish return to the natal location annually (usually to spawn), and where the probable locations during the harvest season depend on the natal location (Porch *et al.* 2001).

Misspecification of spatial structure in a stock assessment model that does not use tagging data can result in poor estimations of stock abundance and mortality rates (Rothschild 2007, Guan *et al.* 2013, Li *et al.* 2015), and similar results have been seen for tagging models (Eveson *et al.* 2009). Although the importance of specifying spatial structure correctly in assessment models is well established, there is little published research focusing on tagging models using an overlapping spatial structure assumption, and this is particularly true with respect to applications using real word tag-recovery data. Vincent *et al.* (2017) used simulations to evaluate the robustness of a multi-region tag-integrated assessment model for fish populations with an overlapping spatial structure. The tag-integrated assessment model required catch-at-age in addition to the tagging data required by traditional tagging models. The catch-at-age data is needed because the tag-

integrated assessment model estimates the number of recovered fish by assuming the tagged fish undergo the same dynamics as the whole population. Thus the tagging process works as a sub-model of the population dynamics model in the tag-integrated assessment model. Given that the tag integrated assessment model cannot be fit to tagging data alone, and that conflicts between data sources can lead to difficulties in actually fitting the integrated model (Eveson et al. 2007), data conflicts cannot be detected. That is, we typically do not have two sets of estimates to compare, one where all the data were used and another where only the tagging or the catch/index data were used.

In this study, we synthesized previous tagging models that either incorporates catch-at-age data (Polacheck et al. 2006), or observer data (Hearn et al. 1999, Eveson et al. 2007), or incorporated meta-population spatial structure assumption (Eveson et al. 2009), and proposed a tagging model framework for modeling multi-year tag-recovery data in a fishery context that incorporates catch-at-age, and observer data into spatial Brownie-Peterson models. Unlike the commonly used meta-population spatial assumption in the tagging models, our model requires fish to move back to their spawning site each year during the spawning period. As an example of application, we applied our model to the conventional tag-recovery, catch-at-age, and observer data of lake whitefish populations in Lake Huron collected from 2004 to 2007. We chose to use these data because a high degree of natal homing was observed for the Lake Huron lake whitefish populations in previous tagging studies (Ebener et al. 2010b). Lake whitefish have long supported the most robust and economically valuable commercial fishery in Lake Huron, and are of great ecological and economical importance for the Laurentian

Great Lakes. Although increasing evidence suggests widespread mixing among nominal management units, lake whitefish are still largely managed as unit stock in each management unit. In northern Lake Huron, four formerly distinct management units of lake whitefish had been aggregated into a larger one in 2011, in response to estimated population intermixing from a tagging study (Ebener et al. 2010b). However, there is still lack of quantitative assessment of mixing probabilities between those spawning populations from tagging models.

To our best knowledge, this study is among the first to introduce a spatial structured tagging model framework for fish populations with an overlap structure, and the most comprehensive to date in terms of the varying combinations of data sets one tagging model framework can be adapted to. Unlike the tag-integrated assessment models (Goethel et al. 2015a, Vincent et al. 2017), our model framework can work independently without catch-at-age data so that data conflict issues can be evaluated. In addition, our tagging model framework innovatively used Bayesian analysis to evaluate uncertainties in terms of parameters estimations, and we also explored model performance of a range of alternative model parameterizations in order to test which model fit the real data best, based on deviance information criteria (DIC), which is also a recent proposal (Vincent et al. 2017). The estimated movement probabilities for our application can also be of critical importance to fisheries managers and modelers for understanding the annual movement of Lake Huron lake whitefish populations, given those parameters have not been previously estimated.

Methods

Tagging model assumptions

First, we assumed fish moved from the fishing grounds to the spawning grounds immediately before each spawning season, and then back to the fishing grounds immediately after the spawning season (annual movement). Second, by definition fish become one year older right after each spawning season. We treated spawning as an instantaneous event, and defined a model year as a time period starting right after a spawning season and ending before the beginning of next spawning season. We assumed no fisheries occurred during the spawning season. Thus, all data were organized and all temporally varying fishery parameters were estimated by model year. We used the model year concept to avoid the need to separate the mortality rates into different proportions before and after spawning seasons. The basic assumptions Pollock et al. (1991) summarized for multi-year tagging models also apply here: 1) tagged and untagged fish are fully-mixed for each population; 2) fates of each fish are independent; 3) fish living in the same fishing ground at a given age have the same survival and exploitation probabilities; and 4) no tag-introduced mortality and tag shedding occurs, or if tag shedding occurs, tag-return probability need to be adjusted for tag loss, which requires estimation of additional parameters. If any of assumptions 1 to 3 are violated, overdispersion would occur and the variance of the tag counts would be underestimated (leading to overfitting of this data source).

Tagging model overview

In our framework, our full tagging model included both tag recovery dynamics and population dynamics. For our reduced model without catch-at-age data, only tag-recovery dynamics were included. We assumed an overlapping spatial structure in all cases, in which fish were allowed to move to other basins at the first day of each model year, but would move back to where they born and spawn at the end of each model year. As for the dimension of our model, we recorded the year (i), age (a), region (k) of tag-releasing and gear type (g), year (j) and region (k') of tag recovery. We also recorded either a recovered fish being observed on boat by observer(s) or not (l). We defined a group of tagged fish at the same age a and released from the same tag-releasing region k and year i as a “tagging group” iak . Our tagging model uses a full Bayesian framework, and we describe the likelihood components of data in the “tag recovery dynamics” and “population dynamics” sections, and the prior distribution in the “model implementation” section. All symbols used in model equations are defined in Table 3.1. The observed data, and estimated parameters are listed in the separate sections in Table 3.1. All equations except the most important five are listed in Table 3.2.

Tag recovery dynamics

The likelihood components for tag-recovery data (L_{tag}) were included in both the full model and reduced model (i.e., without catch-at-age data). We express the tag recovery likelihood (L_{tag}) as a product of Brownie likelihood ($P[R_{iakgjk'}]$) for the combined fishery (i.e., both monitored and unmonitored catch, here “·” represents

crossing all l s), and a conditional likelihood ($P[R_{iakgjk'l} | R_{iakgjk'}]$) with the probability equals to a recovered fish being monitored or not, as shown in Equation 3.2.1.

Table 3.1. Symbols used in the model equations.

Symbol	Description
Indicator variables	
i	Tagging year (I = total number of tagging years)
a	Tagging age (A = total number of age groups)
a'	Recovery age ($a' = \min(a + j - i, A)$ or $a'[y] = \min(a + y, A)$)
k	Tagging basin (K = total number of tagging basins)
g	Recovery gear type (G = total number of gear types)
j	Recovery year (J = total number of recovery years)
k'	Recovery basin (K' = total number of recovery basins)
l	The component of catch the recovery belongs to ($l=1$: monitored component of catch; $l=2$: unmonitored component of catch)
Observed (input) data	
T_{iak}	The number of total tags released in year i , at age a , from basin k (tagging group iak)
$\sum_{l=1}^L R_{iakgjk'l}$	The number of tags from tagging group iak recovered by gear type g , in year j , from basin k'
$C_{gjk'1}$	Monitored catch harvested by gear type g , in year j , from basin k'
$\sum_{l=1}^L C_{gjk'l}$	Catch harvested by gear type g , in year j , from basin k'
$\sum_{l=1}^L E_{gjk'l}$	Fishing effort by gear type g , in year j , from basin k'
$\sum_{i=1}^I \sum_{a=1}^A \sum_{k=1}^K R_{iakgjk'1}$	The number of tags recovered by gear type g , in year j and basin k' from the monitored component of catch
$\sum_{i=1}^I \sum_{a=1}^A \sum_{k=1}^K R_{iakgjk'2}$	The number of tags recovered by gear type g , in year j and basin k' from the unmonitored component of catch
$\rho_{gjk'a}$	The observed age composition (percentage) for catch harvested by gear type g , in year j , from basin k'

Table 3.1 (cont'd)

Estimated parameters	
Π_{ja}	Matrix whose (k^1, k^2) th element, $\pi_{jak^1k^2}$ is the probability that a fish at age a moves from region k^1 to region k^2 at the beginning of year j after spawning season of year $j - 1$.
$\lambda_{jk'}$	The probability that a tag would be reported from a tagged fish given it was caught by unmonitored component of catch during year j in recovery basin k' .
M_{jak}	The instantaneous rates of natural mortality for fish at age a , in year j and region k .
q	Catchability
η, τ	Selectivity parameters
R_{jk}	Recruitment (i.e., the first age class) for recovery year j in tagging site k (only included in the full model)
NO_{ak}	Initial abundance at age a for the first recovery year of tagging site k , $a \in$ [2nd age class, A] (only included in the full model)
Other symbols used in the model equations (equation #)	
$R_{iak\dots}$	total recovered tags from tagging group iak (Equation 3.2.1)
$R_{iakgjk'}$	Number of tags from tagging group iak recovered by gear type g , in year j , from basin k'
$p_{iakgjk'l}$	Tag return probability for a fish from tagging group iak being recovered by gear type g , in year j , basin k' , and the l th fishery component
$p_{iak\dots}$	Total probability a tag will be returned for a fish in tagging group iak
$U_{a'jgk'}$	probability that a fish at age a' is caught by gear type g in year j from basin k'
$\delta_{gjk'}$	conditional probability that a fish caught in year j by gear type g from basin k' was harvested by the monitored component of catch
$C_{gjk'}$	total catch harvested by gear type g , in year j , from basin k'
$C_{m_{gjk'}}$	the monitored component of $C_{gjk'}$
φ_{ij}	tag retention probability, for tagging and recovery year i and j
$\Delta_{iajkk'}$	multiannual survival-movement probability for tagging group iak to survive to the start of year j and be present that year in basin k' .
S_{jak}	survival probabilities for fish at age a , in year j and region k
F_{jakg}	instantaneous rate of fishing mortality for fish at age a , in year j and region k and being harvested by gear type g .
v_a	Selectivity at age a
$R_{\dots gjk'}$	the total number of tags recovered by gear type g , during year j from basin k' , from both monitored and unmonitored components of catch
$R_{\dots gjk'1}$	number of the tags recovered from monitored component of catch by gear type g during year j from basin k'
$R_{\dots gjk'2}$	number of tags recovered from unmonitored component of catch by gear type g , during year j from basin k'

Table 3.1 (cont'd)

$E[C_{gjk'a}]$	estimated catch at age a that harvested by gear type g , during year j from basin k'
$n_{gjk'}$	effective sample size, which treated as hypothetical observed fish count sampled by gear type g during year j from basin k'
$n_{gjk'a}$	hypothetical observed fish count at age a
$\rho_{gjk'a}$	Expected proportions at age in the harvest

Table 3.2. Equations for the tagging model framework.

Equation	#
$L_{tag} = P[R_{iakgjk'l}] = P[R_{iakgjk'l'}]P[R_{iakgjk'l} R_{iakgjk'l'}]$	3.2.1
$= \prod_{i=1}^I \prod_{a=1}^A \prod_{k=1}^K \prod_{g=1}^G \prod_{j=i}^J \prod_{k'=1}^{K'} (R_{iak\dots}, T_{iak} - R_{iak\dots}) \prod_{g=1}^G \prod_{j=i}^J \prod_{k'=1}^{K'} [p_{iakgjk'l'}^{R_{iakgjk'l'}} (1 - p_{iak\dots})^{T_{iak} - R_{iak\dots}}]$	3.2.2
<p style="text-align: center;">where $R_{iak\dots} = \sum_{g=1}^G \sum_{j=i}^J \sum_{k'=1}^{K'} \sum_{l=1}^L R_{iakgjk'l}$</p> $R_{iakgjk'l'} = \sum_{l=1}^L R_{iakgjk'l}$ $p_{iak\dots} = \sum_{g=1}^G \sum_{j=i}^J \sum_{k'=1}^{K'} \sum_{l=1}^L p_{iakgjk'l}$	
$S_{jak} = \exp[-(\sum_{g=1}^G F_{jakg} + M_{jak})]$	3.2.3
$U_{agjk} = \frac{F_{jakg}}{\sum_{g=1}^G F_{jakg} + M_{jak}} (1 - S_{jak})$	3.2.4
$F_{jakg} = v_a \cdot f_{jgk}$	3.2.5
$f_{jgk} = \widetilde{E}_{jgk} q_{jgk}$	3.2.6
$v_a = \frac{a^\eta \exp(-\tau a)}{10^\eta \exp(-\tau 10)}$	3.2.7
$P[R_{iakgjk'l} R_{iakgjk'l'}]$	3.2.8
$= \prod_{i=1}^I \prod_{a=1}^A \prod_{k=1}^K \prod_{g=1}^G \prod_{j=i}^J \prod_{k'=1}^{K'} \left(\begin{matrix} R_{iakgjk'l'} \\ R_{iakgjk'l_1} & R_{iakgjk'l_2} \end{matrix} \right) \left(\frac{\delta_{gjk'}}{(\delta_{gjk'} + (1 - \delta_{gjk'})\lambda_{jk'})} \right) \left(\frac{(1 - \delta_{gjk'})\lambda_{jk'}}{(\delta_{gjk'} + (1 - \delta_{gjk'})\lambda_{jk'})} \right)^{R_{iakgjk'l_2}}$	
$P[R_{iakgjk'l} R_{iakgjk'l'}]$	3.2.9
$\propto \prod_{g=1}^G \prod_{j=1}^J \prod_{k'=1}^{K'} \left(\begin{matrix} R_{\dots gjk'l'} \\ R_{\dots gjk'l_1} & R_{\dots gjk'l_2} \end{matrix} \right) \left(\frac{\delta_{gjk'}}{(\delta_{gjk'} + (1 - \delta_{gjk'})\lambda_{k'})} \right)^{R_{\dots gjk'l_1}} \left(\frac{(1 - \delta_{gjk'})\lambda_{k'}}{(\delta_{gjk'} + (1 - \delta_{gjk'})\lambda_{k'})} \right)^{R_{\dots gjk'l_2}}$	

Table 3.2 (cont'd)

$$P[R_{iakgjk'l} | R_{iakgjk' \cdot}] \propto \prod_{g=1}^G \prod_{j=1}^J \prod_{k'=1}^{K'} \frac{(\lambda_{k'})^{R_{\dots gjk'2}}}{(\delta_{gjk'} + (1 - \delta_{gjk'})\lambda_{k'})^{R_{\dots gjk'1} + R_{\dots gjk'2}}} \quad 3.2.10$$

$$L_{CT} = \prod_{g=1}^G \prod_{j=i}^J \prod_{k'=1}^K \frac{1}{\sqrt{2\pi\sigma^2_{gjk'}}} e^{-\frac{(\log(C_{gjk'}) - \log(\sum_{a=1}^A E[C_{gjk'a}]))^2}{2\sigma^2_{gjk'}}} \quad 3.2.11$$

$$L_{CP} = \prod_{g=1}^G \prod_{j=i}^J \prod_{k'=1}^K \left(n_{gjk'1} \dots n_{gjk'A} \right) \prod_{a=1}^A (\rho_{gjk'a})^{n_{gjk'a}} \quad 3.2.12$$

where $n_{gjk'a} = n_{gjk'} \times \rho_{gjk'a}$

$$\rho_{gjk'a} = \frac{E[C_{gjk'a}]}{\sum_{a=1}^A E[C_{gjk'a}]}$$

The Brownie likelihood $P[R_{iakgjk' \cdot}]$ for the combined components is the product of multinomial probabilities given by Equation 3.2.2. The idea is that the total tags released from each tagging group (T_{iak}) are partitioned into strata in terms of their recovery years (j), gear types (g), tag-recovery regions (k'), and either being observed on boat by observer(s) or not (l), with the probability of recoveries in each stratum modeled as a multinomial. Thus, for each tagging group iak , $\{R_{iak111 \cdot}, \dots, R_{iakGJK \cdot}, T_{iak} - R_{iak \dots}\} \sim MN(T_{iak}, p_{iak111 \cdot}, \dots, p_{iakgjk' \cdot}, 1 - p_{iak \dots})$. Note that although the Brownie likelihood is for the combined components of fisheries (all l s), we still need to consider whether the recovered fish being observed on boat by observer or not when we model the tag return probability $p_{iakgjk' \cdot}$, because we assumed that the tag reporting rate for the monitored component of catch ($l = 1$) is 100%, and for the unmonitored component ($l = 2$) is $\lambda_{k'}$ (Table 3.1).

Tag return probability $p_{iakgjk'l}$ is determined as a product of the survival and movement probability Δ , the exploitation rate U , the tag retention rate φ , either the probability the fish belonged to the monitored (δ) or unmonitored ($1 - \delta$) component, and the tag reporting rate (100% for $l = 1$ and $\lambda_{k'}$ for $l = 2$), given by:

$$p_{iakgjk'l} = \begin{cases} \Delta_{iajkk'} U_{a'jgk'} \delta_{gjk'} \varphi_{ij} & \text{if } l = 1 \\ \Delta_{iajkk'} U_{a'jgk'} (1 - \delta_{gjk'}) \varphi_{ij} \lambda_{k'} & \text{if } l = 2 \end{cases} \quad (3.1)$$

where $\delta_{gjk'} = C_{m_{gjk'}}/C_{gjk'}$ is the conditional probability that a fish caught in year j by gear type g from basin k' was harvested by the monitored component of catch. Thus the unconditional probability of a tag being returned (over the monitored and unmonitored components of catch) is:

$$p_{iakgjk'} = \sum_{l=1}^L p_{iakgjk'l} = \Delta_{iajkk'} U_{a'jgk'} \varphi_{ij} (\delta_{gjk'} + (1 - \delta_{gjk'}) \lambda_{k'}) \quad (3.2)$$

The survival-movement probability Δ_{iaj} matrix ($K \times K'$) with typical element $\Delta_{iajkk'}$ contains by spawning (rows, k) and fishing (columns, k') locations the probability that a fish from tagging group iak survives to the start of year j and is present in the fishing location k' at the start of year j :

$$\Delta_{iaj} = \begin{cases} \Pi_{j,a'} & j = i + 1 \\ D(H_{i,a,1}) \times D(H_{i,a,2}) \times \dots \times D(H_{i,a,j-i-1}) \times \Pi_{j,a'[j-i]} & j > i + 1 \end{cases} \quad (3.3)$$

where $a'[y] = \min(a + y, A)$ and $H_{i,a,m} = \Pi_{i+m,a'[m]} D(S_{i+m,a'[m]}) \mathbf{1}$ is a K' -by-1 column vector for $m = 1, \dots, j - i - 1$. The movement matrix Π_{ja} is defined in Table 3.1; and $S_{i,a}$ is a vector (length equal to number of regions K') whose typical element s_{iak} , is the probability that a fish at age a alive at the beginning of the year i in region k

survived the year. The operator $D(\theta)$ is an operator that transforms a K' vector θ into a $K' \times K'$ diagonal matrix containing the elements of θ on its diagonal. $\mathbf{1} = (1, \dots, 1)^T$. The terms $H_{i,a,m}$ are vectors of size K' , whose $(k)th$ element represents the probability of an individual from population k alive at the start of the year survives to the end of that year.

Because one goal of this tagging model framework is to estimate mortality rates, the survival and capture probabilities S_{jak} and U_{agjk} were written in terms of instantaneous natural mortality M_{jak} and fishing mortality F_{jak} , as shown in Equation 3.2.3 and 3.2.4 (Table 3.2).

The M_{jak} were not estimated freely for every year, age and area but were constrained to be constant spatially or temporally in different ways (see experimental design section). Note that for a population that is tagged in I consecutive release years, only $I - 1$ natural mortality rate parameters (per region) can be estimated, given natural mortality rates are estimated based on the difference between the expected returns at age $a + 1$ of fish released at age a and those released at age $a + 1$.

Fishing mortality is modeled as a product of selectivity v_a and fully selected fishing mortality f_{jgk} , as shown in Equation 3.2.5 (Table 3.2). The fully selected fishing mortality is modeled as the product of observed effort (\tilde{E}_{jgk}) and catchability q_{jgk} , as shown in Equation 3.2.6 (Table 3.2). When estimated, the q_{jgk} were not allowed to vary freely among regions and years, being constrained to be constant spatially or temporally in different ways (see experimental design section). In most cases, effort of different gear types fluctuate in synchrony, making it challenging to estimate separate catchabilities by gear type, when they are allowed to vary spatially and temporally. One approach, which

we adopt in our application, is to estimate catchability potentially varying for one gear type, and assume the catchability for the other gear type is a constant ratio of this (see application of Lake Huron Lake Whitefish). Selectivity at age v_a is modeled using a gamma function with parameters η and τ that produce a dome-shape selectivity pattern, as shown in Equation 3.2.7. In our application, we assume the function peaked at a value of 1.0 at age 10.

The likelihood $(P[R_{iakgjk'l}|R_{iakgjk'}])$ for the conditional probability of recovery in either the monitored or non-monitored catch is given by Equation 3.2.8. For each recovered fish from tagging group iak that were recovered in recovery years j and regions k' by gear types g , this conditional probability is a product of binomials:

$$R_{iakgjk'1} \sim Bin(R_{iakgjk'1}, \frac{\delta_{gjk'}}{\delta_{gjk'} + (1 - \delta_{gjk'})\lambda_{jk'}}). \text{ As } \lambda_{jk'} \text{ and } \delta_{gjk'} \text{ do not depend on } iak$$

we can simplify this likelihood to Equation 3.2.9. Note since $\delta_{gjk'}$ depends on the data but not the parameters and that the only parameter involved in Equation 3.2.9 is $\lambda_{k'}$.

Hence, the likelihood can be further simplify to Equation 3.2.10.

Population dynamics

The likelihood components for catch-at-age data (L_C) were only included in the full model with catch-at-age data, and this is also true of all additional parameters described in this section. We used catch at age likelihood equations that are widely applied in Statistical Catch at Age assessment models. The catch likelihood was broken into two subcomponents, one for the annual totals and one for the annual age compositions (proportions at age). By assuming that the catches for each region and year are statistically independent, we combined regions and years into these two

subcomponents. The overall catch component of likelihood $L_C = L_{CT}L_{CP}$. The subcomponent for the total, L_{CT} , is based on an assumed Gaussian distribution of errors for log-scale annual fishery total catch (Equation 2.11). The subcomponent for the harvest age-compositions, L_{CP} , is based on a multinomial distribution, with the effective sample size equal to 25 for each year, region, and gear type (Equation 2.12). We restricted attention to areas where tagged fish were released, and assumed catch of these fish in other areas was negligible.

The estimated catch $E[C_{gjk'a}]$ in equation 3.2.11 is based on estimated population abundance and where those fish are estimated to be within the fishing season. The population abundance at the beginning of year j , for years after the release year, before any movements, is modeled as:

$$N_{jak} = \begin{cases} \sum_{k'=1}^K N_{j-1,a-1,k} \pi_{j-1,a-1,k,k'} S_{j-1,a-1,k'} & (a < A) \\ \sum_{k'=1}^K N_{j-1,a-1,k} \pi_{j-1,a-1,k,k'} S_{j-1,a-1,k'} + \sum_{k'=1}^K N_{j-1,a,k} \pi_{j-1,a,k,k'} S_{j-1,a,k'} & (a = A) \end{cases} \quad (3.4)$$

where $\pi_{j-1,a-1,k,k'}$ and $S_{j-1,a-1,k'}$ represents the movement and survival rates in the year $j-1$, as described above for the tagging component of the likelihood. The recruitment for all tagging years R_{jk} (i.e., N_{j1k}) ($j \leq J$; $k \leq K$), and initial abundance-at-age in the first tagging year $N0_{ak}$ (i.e., N_{1ak}) ($a \geq 2$; $k \leq K$) are estimated as free parameters in the full model (Table 3.1).

The number of fish located in region k' at the start of the harvest season of year j , regardless of which of the K populations they are from, is modeled as a sum of products

of the population abundance at the beginning of year j before fishing intermixing and the movement rates for each population in year j :

$$N'_{jak'} = \sum_{k=1}^K N_{jak} \pi_{jakk'} \cdot \quad (3.5)$$

The estimated catch of these fish $E[C_{gjk'a}]$ for year j , region k' age a , and gear type g as:

$$E[C_{gjk'a}] = N'_{jak'} U_{jagk'} \cdot \quad (3.6)$$

where $U_{jagk'}$ is the catch probability from equation 3.2.4 in Table 3.2.

Example application to lake whitefish

We partitioned Lake Huron into nine distinct basins: MI 14, MI 23, MI central, MI south, ON north, ON central, ON south, North Channel, and Georgian Bay. Between 2003 and 2006 (i.e., $I = 4$, Table 3.1), lake whitefish were tagged and released in work coordinated by Mark Ebener then of the Chippewa-Ottawa Resource Authority from the main basin of Lake Huron during their spawning season (late October through early December), with each location selected as a known spawning site for a specific population. Tagging occurred in all basins of Lake Huron except the North Channel, and Georgian Bay (i.e., $K = 7$). All 35,285 tagged fish were measured before release, with spatial coordinates of tagging and release location, and date of release recorded. Fish were tagged on spawning grounds and then recovered (dead) during the fishing season. Most tagged fish were released in November, and previous research found that lake whitefish in Lake Huron tended to soon leave the spawning grounds and were on average as far or further from their tagging location by December as by the next summer (Li et al.

2017). We assumed the model year y for lake whitefish in Lake Huron spanned from December 1 of year $y-1$ through November 30 of year y , during which fish were vulnerable to harvest and were in mixed stocks. In order to meet the assumption of full mixing, we do not use the few recoveries from fish that are tagged and recovered within the same year and we assumed that all tagged fish survived to the next year, given that tagging of lake whitefish occurred near the end of the year during the spawning season.

Recoveries were made throughout the year during 2004–2011, with the majority recovered by commercial fishermen, and the rest recovered during fishery surveys. We excluded recoveries from Lake Michigan, North Channel, and Georgian Bay, thus restricting the analysis to fish caught in Lake Huron's main basin, and the number of recovery basins is the same as the number of the tagging basins (i.e., $K' = K = 7$, Figure 3.1). We restricted recovery (harvest) regions to regions where tagging was done because abundance could not be estimated for populations in regions that were not tagged because movement of unmarked populations could not be estimated. This restriction was reasonable because more than 97% of recovered fish were recovered from the main basin. We excluded recoveries from 2008 to 2011 (i.e., $J = 4$), thus the number of tag recovery years is the same as of the tag-release year. We did this because the major source of movement information is from tag recovery data and including years without tag releases can cause identifiability issues (Bailey et al. 2010, Cole and McCrea 2016). For example, survival rates in one year and exploitation rates in the next would be confounded for those years without tag releases for our reduced model with only tag-recovery and fishing effort data.

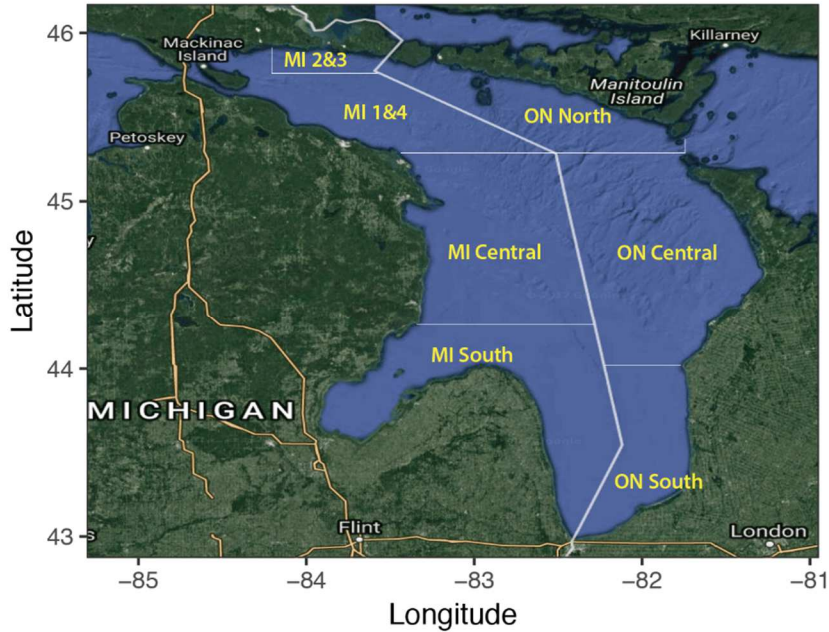


Figure 3.1. Map of Lake Huron with seven tag release and recovery basins.

We restricted our analysis to tagged fish with length information and recoveries with location and date information. We made these restrictions because age of tagged fish was based on length, and because year and basin of recovery was needed for use in the tagging model. In addition, we excluded recoveries that were made within two days of the tagging date as it was assumed that mixing would not have occurred this quickly. We only included recoveries from the commercial fishery (trapnet and gillnet), and excluded survey recoveries, as the recoveries were used to jointly estimate movement, survival, and exploitation (i.e., $G = 2$), and survey data were limited and differed geographically and were absent from some regions. The total number of reported recoveries was 2078, of which 1409 were used in this study (Table 3.3). Subsets of the data used here were previously reported by Ebener et al. (2010a, 2010b), and details of the tagging methodology are given by Ebener et al. (2010b).

Table 3.3. Exclusion criteria for fish recovered.

Criteria names	Number of recovered fish
Recovery without length data	5
Recoveries without location and recovery date information	225
Recoveries made within two days of the tagging date	131
Recoveries caught from fishery survey	138
Recoveries caught outside the main basin of Lake Huron	47
Recoveries caught from years 2008 to 2011	123

Ages of tagged fish were estimated by using tagging basin specific age-length keys. Age-length keys were developed based on age estimates determined from scale samples that were collected during annual population assessment surveys. Tagged fish were assigned to one of seven age groups (6, 7, 8, 9, 10, 11 and older [12+]) (i.e., $A = 7$). Age-6 and younger fish, and age-12 and older fish were combined into a single age group respectively.

Because the amount of the effort of different gear types tends to vary in the same way over time, for given j and k , q_{jgk} tends to highly correlated. Thus, we included an external calculation of the relative catchability ratio of different gear types ($g=1$ and 2 , for gill net and trap net) based on ratio of CPUEs: $r_{jk} = q_{j2k}/q_{j1k} = (C_{2jk}'/E_{j2k})/(C_{1jk}'/E_{j1k})$ for every year j and region k . Then for this lake whitefish application, fishing mortality in Equation 3.2.6 was replaced by:

$$\sum_{g=1}^2 f_{jgk} = \sum_{g=1}^2 \tilde{E}_{jgk} q_{jgk} = \tilde{E}_{j1k} q_{j1k} + q_{j1k} r_{jk} \tilde{E}_{j2k} \quad (3.7)$$

and we only estimated gill net catchability q_{j1k} . However, equation 7 can only work for a specific region and year (j and k) that had both gill net and trap net harvest, or had only

gill net harvest, but not when the only fishing gear was trap net. For a specific region and year with only trap net harvest, its catchability was estimated separately (see design of analysis).

As for the tag-retention rates in equation 1, we used previously published estimates for these tagging data by Ebener et al. (2010):

$$\varphi_{ij} = 1 - \frac{0.31}{1 + \exp(8.52 - t)} \quad (3.8)$$

where t is in months, we assumed $t = 1 + (j - i) \times 12 - 6$ because most tagged lake whitefish were released in November.

Design of Analysis

We evaluated alternative models with different parameterizations of tag reporting rate (three levels: varying by group of regions with three different grouping choices), and natural mortality (five levels: constant M , varying by group of years M_j , varying by group of ages M_a , or varying by group of regions with two different grouping choices), as shown in Table 3.4. We assumed movement rates varied by tagging basin (k), but was constant across different years and ages, due to the small sample size. For each tag basin k and recovery basin k' , we also assumed the movement rates $\pi_{k,k'}$ is 0 if there were no observed recoveries in recovery basin k' from tagging basin k

($\sum_{i=1}^I \sum_{a=1}^A \sum_{g=1}^G \sum_{j=1}^J \sum_{l=1}^L R_{iakgjk'l} = 0$). We also assumed constant catchability for all basins except for the MI_central and MI_south basins, in which only trap net harvest with different mesh sizes occurred during the research period. We thus estimated separate

catchability values for these two regions, and in total we estimated three catchability values in each model.

Table 3.4. Assumptions that were used to generate different parameterizations of tag reporting rate and natural mortality in the different alternative parameterizations of the assessment models. Parameter dependency refers to what leads to different values for parameter.

Parameter name	Parameter dependency	Index	If for group, grouping details
Tag reporting rate λ	Group of regions (option 1)	λ_{ks1}	(1) MI group: MI 14, MI 23, MI central, MI south; (2) ON group: ON north, ON central, ON south
	Group of regions (option 2)	λ_{ks2}	(1) MI north group: MI 14, MI 23 (2) MI central/south group: MI central, MI south (3) ON north/central group: ON north, ON central (4) ON south group: ON south
	Group of regions (option 3)	λ_{ks3}	(1) MI north group: MI 14, MI 23 (2) MI central/south group: MI central, MI south (3) ON north group: ON north (4) ON central/south group: ON central, ON south
Natural mortality M	Constant	M	/
	Group of years	M_{js}	(1) Early years group: 2004, 2005 (2) Later years group: 2006, 2007
	Group of ages	M_{as}	(1) Young group: 6, 7, 8 (2) Old group: 9, 10, 11, 12
	Group of regions (option 4)	M_{ks4}	(1) North group: MI 14, MI 23, ON north (2) Central group: MI central, ON central (3) South group: ON south, MI south
	Group of regions (option 1)	M_{ks1}	(1) MI group: MI 14, MI 23, MI central, MI south (2) ON group: ON north, ON central, ON south

Overall, we evaluated 15 models with different parameterizations, which crossed the parameterization levels for tag reporting rate, and natural mortality, as shown in Table 3.5. Each model is named based on its parameter dependency index of tag reporting rate, and natural mortality. For example, our baseline model assumed country-dependent tag reporting rates (group of regions based on option 1 in Table 3.4), and natural mortality. Thus, the name of the base model is $\lambda_{ks1} + M$. Note that when natural mortality rate is stated to be dependent, only I-1 parameters (3 parameters in our case) can be estimated, when I is the number of release years (see model description). Thus, we grouped years, age classes, and recovery basins to two or three groups, as shown in Table 3.2. If parameter dependency is for group of elements, say group of regions, we assumed the parameter value of all regions in that group were the same. The grouping of year and ages is simple: early and later years, and young and old ages. While the grouping of basin not only depends on the location and fisherman performance (grouping options 1 and 4 in Table 3.4), but also depends on the data (grouping options 2 and 3 in Table 3.4). Because there is no catch being observed for tag recoveries in MI south and ON central, we had to group this region with its adjacent regions for modeling tag-reporting rates.

Table 3.5. Model comparison results. The first four columns are four likelihood components, and the last four columns are D1, D2, p_D , and DIC for each model.

#	Model Name	D ₁	D ₂	p_D	DIC
1	$\lambda_{ks1} + M$	19988.8	19931.3	57.5	20046.2
2	$\lambda_{ks1} + M_{js}$	19964.9	19915.3	49.6	20014.6
3	$\lambda_{ks1} + M_{as}$	19979.9	19930.2	49.7	20029.6
4	$\lambda_{ks1} + M_{ks4}$	19930	19873.2	56.8	19986.8
5	$\lambda_{ks1} + M_{ks1}$	19989.7	19935.9	53.8	20043.5
6	$\lambda_{ks2} + M$	19998.1	19946.2	51.9	20050
7	$\lambda_{ks2} + M_{js}$	19973	19922.9	50.1	20023
8	$\lambda_{ks2} + M_{as}$	19995.2	19943.3	51.9	20047.1
9	$\lambda_{ks2} + M_{ks4}$	19913.6	19868.1	45.5	19959.1
10	$\lambda_{ks2} + M_{ks1}$	19976.9	19917.4	59.6	20036.5
11	$\lambda_{ks3} + M$	19918.8	19858.8	60	19978.7
12	$\lambda_{ks3} + M_{js}$	19907.8	19855.6	52.2	19960
13	$\lambda_{ks3} + M_{as}$	19913.4	19863.3	50.1	19963.5
14	$\lambda_{ks3} + M_{ks4}$	19891.2	19838.7	52.4	19943.6
15	$\lambda_{ks3} + M_{ks1}$	19901.4	19839.9	61.5	19962.9

Model implementation

We estimated the parameters using a two-step framework: In step 1, we computed maximum likelihood estimates for each parameters; In step 2, we used MLE as the initial value for each parameter and used a full Bayesian framework to incorporate prior knowledge, namely appropriate parameter ranges as reported in previous research. In general, for each parameter θ , we used Uniform prior distribution with lower bound θ_1 and upper bound θ_2 . To simulate the posterior distribution that combines both the prior distribution and the likelihood of the data, we used the random-walk Metropolis algorithm, a Markov Chain Monte Carlo (MCMC) method that widely used when conditional densities have an intractable form (Metropolis et al. 1953).

We used the adaptive MCMC algorithm (Roberts and Rosenthal 2009) to tune the proposal variance κ and to accelerate movement rates. We generated four MCMC chains with three distinct initial parameter values for convergence diagnostics (i.e., visual analysis via trace plots and Gelman and Rubin diagnostics). For each chain, we did a total of 6,000 iterations, repeating the above procedure for each parameter at each iteration. After the first 5,000 iterations as burn-in period, we sample the last 1000 iterations for the posterior inference per chain. Thus, a total of 3000 samples were sampled for the posterior inference.

Performance statistics

Deviance information criterion (DIC) was used to compare the full models with different parameterizations. The deviance statistics is calculated as twice the negative likelihood. We report: (1) the posterior mean deviance which averages over deviances of parameters at each posterior sample (D_1); (2) deviance of the posterior median parameter (D_2); (3) the effective number of parameter $p_D = D_1 - D_2$, and (4) $DIC = D_1 + p_D$. We chose to use deviance of the posterior median instead of posterior mean because it is less sensitive to large deviations (Hamada et al. 2008, Francois and Laval 2011). A smaller value of DIC indicates a better model fit. We report the posterior mean with an associated 95% credible interval for each estimated parameter.

For the top model, we summarized the posterior distributions of movement matrix, survival rates, selectivity at age, total abundance of all age group, and tag reporting rates, in terms of the posterior mean and 95% credible interval (CI). We also rerun our best model with no restriction in movement rates (i.e., no restriction model, fish can move to any recovery basins even the basins with no observed recovery) to evaluate

the consequences of our assumption about movement rates equaling zero for regions with no observed recoveries. We then fit a reduced version of our best model without catch-at-age data. We did this because differences in estimates between the best model and a reduced version provide evidence on whether there is a conflict between the tagging and fishery data. As a final step we summarized the posterior prediction distribution for number of recovered fish, catch, and age composition to assess the fit of our best model.

Results

Model comparison

When the same parameters were estimated for tag reporting rates (i.e., model comparisons within Model # 1-5, 6-10, or 11-15, Table 3.5), the best models (lowest DIC) had separate estimates of natural mortality rates for north, central and south groups (i.e., Model #4, 9, and 14) as well as the lowest posterior mean deviance (D1), and deviance of the posterior median (D2). Among models that parameterized natural mortality the same, the model which estimated separate tag reporting rates for MI north, MI central/south, ON north, and ON central/south groups (i.e., Model #11-15) always led to lowest DIC, D1, and D2. The effective number of parameters p_D varied from 45.5 to 61.5 among different models.

According to the DICs, our best model is Model 14, and the difference in DIC from all other models to our best model is greater than 10, which indicates that this model is much better at predicting the data than any alternative we considered. In the best model, we estimated separate tag reporting rates for MI north, MI central/south, ON north, and ON central/south groups; and different natural mortality rates for north, central, and south groups.

Summary of the best model: Model 14

Movement rates— In general, the stay rates (i.e., the probability of a spawning population remain in their original tag basin k , Π_{kk}) were considerably higher for the spawning populations in the ON basins than those in the MI basins (Figure 3.2). The posterior medians were greater than 97% for the spawning populations in the ON central and north basins, and 87% for the ON south basin. According to the posterior median, 10.5% of the spawning population in the ON south basin moved to the ON north basin, and 2.5% moved to the MI south basin. Lake whitefish that spawned in the MI central and south basins were more likely to move to ON basins than those spawned in the MI north basin. The probability that whitefish that spawned in the MI south basin remain in that region was 15%, while 62% of them moved to ON central (posterior median). The stay rate was 45% for the whitefish populations that spawned in MI central, while 33% of them moved to ON basins and 10% moved to MI south (posterior median). For the spawning populations in the MI 1_4, the probability that fish would remain there was 19%, while 64% of them moved to the MI central basin, and 11% of them moved to the MI 2_3 basin. The posterior median stay rate was 66% for whitefish in MI 2_3, and 24% of them moved to MI 1_4 and 4% of them moved to MI central. The highest uncertainty in estimated movement rates was observed for the MI central and south basins, with the broadest posterior distributions (density stays relatively high over wide range of values).

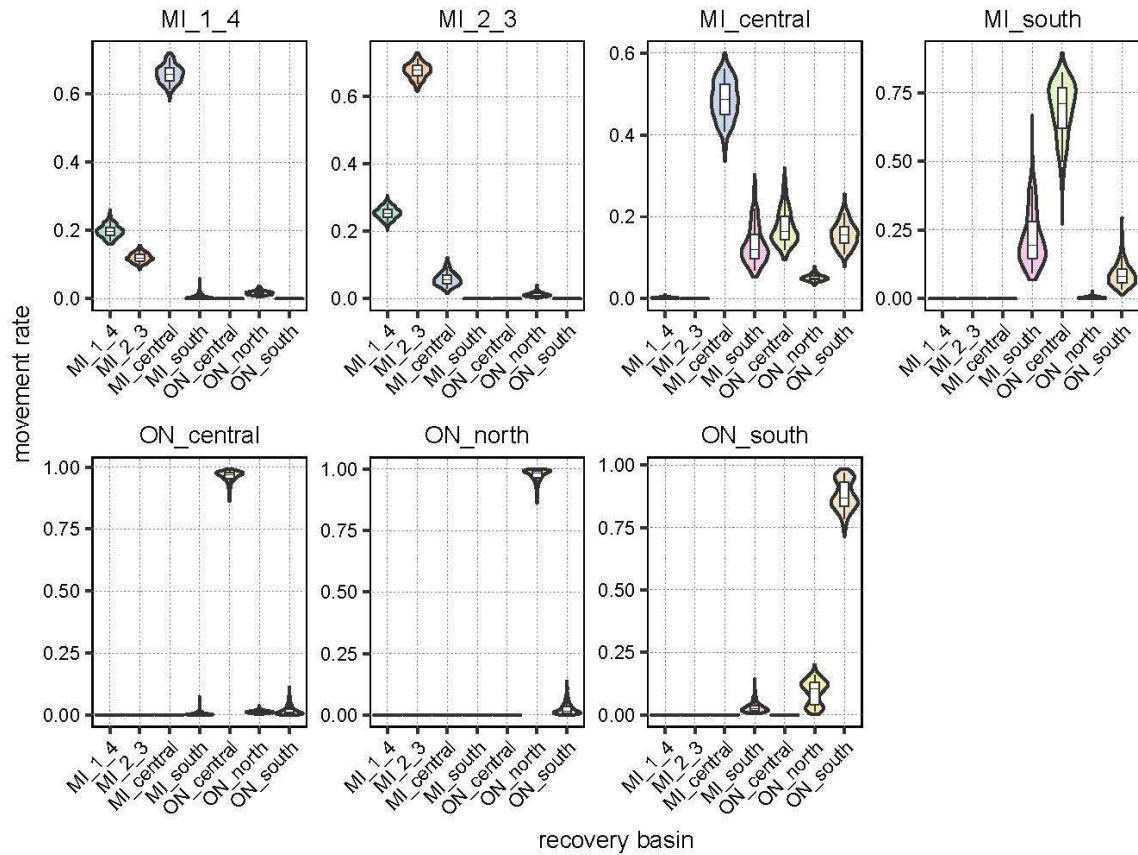


Figure 3.2. Posterior distribution of the movement probability that moving from each tag basin (panel name) to all recovery basins (the x-axis of each panel). Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Horizontal middle line: posterior median.

Tag reporting rates— Estimates of tag reporting rates were generally high in the north basins, and low in the central and south basins (Figure 3.3, Table 3.4). The posterior medians of tag reporting rates were greater than 0.95 y^{-1} for MI north and ON north groups separately, and were 0.17 and 0.10 y^{-1} for MI central/south and ON central/south groups.

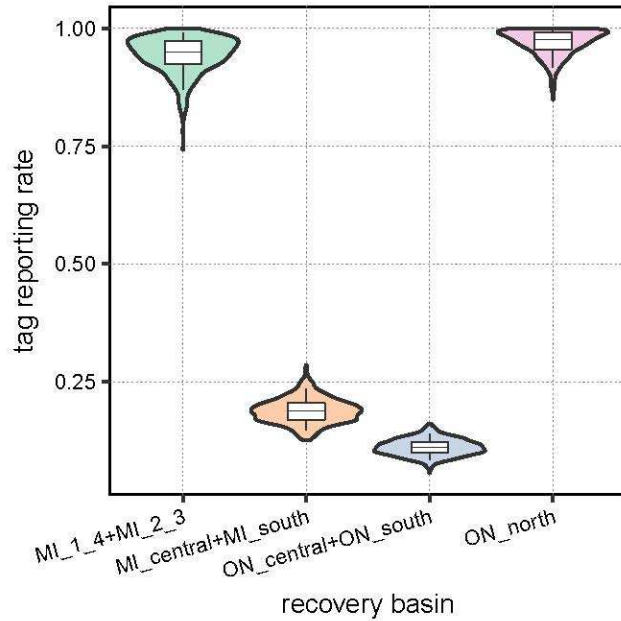


Figure 3.3. Posterior distribution of tag reporting rates by recovery basins. Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Middle dots: posterior median.

Instantaneous natural mortality rates—In general, the instantaneous natural mortality rates were high in the north and south basins, and low in the central basins (Table 3.6). The posterior standard deviation of estimated natural mortality rate was higher for the south basins (Table 3.6).

Table 3.6. Instantaneous rates of natural mortality (per year) for the best model.

	Posterior Mean	Posterior Standard Deviation	95% credible interval
North group:	0.79	0.06	(0.67, 0.92)
Central group:	0.11	0.06	(0.02, 0.23)
South group:	0.57	0.09	(0.40, 0.75)

Instantaneous fishing mortality rates—According to the posterior median, during the fishing season, lake whitefish in the central basins (both MI central and ON central basins) experienced much lower fishing mortality rates than those in the north and south basins (Figure 3.4). For all basins, on average, their estimated f_s increased slightly from

2006 to 2007, while the temporal trend of f_s changed differently in different recovery basins from 2004 to 2006. The estimated f_s increased from 2004 to 2005 in recovery basins MI 1_4, MI central and ON north, and in the later years f_s decreased in the MI 1_4 and ON north basins, but increased slightly in the MI central basin. Although the scales are different, on average, the temporal trends of f_s in the MI 2_3 and ON south are quite similar, in which lowest f_s were estimated in 2005. According to the posterior densities and CIs, the uncertainty in estimating f_s is greatest for fish in the MI south basin with CI varied from 0.05 to 0.68, and is lowest in the ON central and south basins. Selectivity at age was assumed to be time- and basin- invariant in our model, and was rescaled so that the selectivity at age 10 equals 1. Our estimated results show a dome-shaped selectivity, with a peak at age 9 and 10 (Figure 3.5).

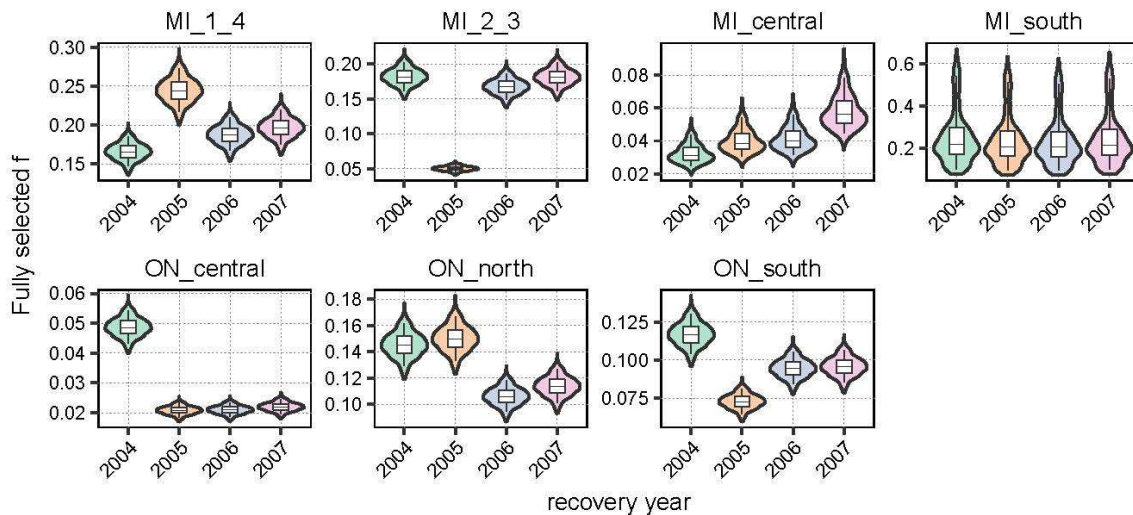


Figure 3.4. Posterior distribution of fully selected fishing mortality by recovery years for each basin. Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Middle dots: posterior median.

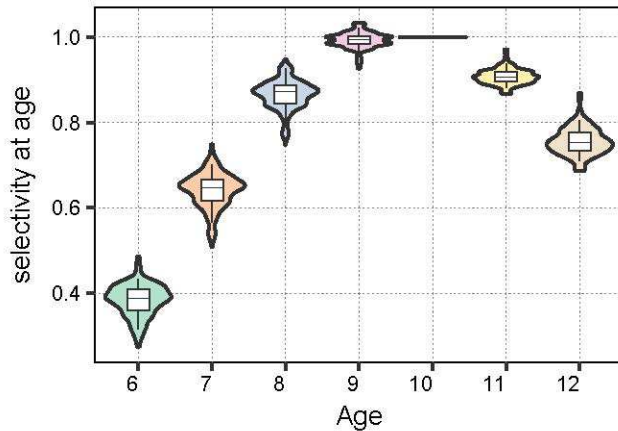


Figure 3.5. Posterior distribution of selectivity at age for all years, basins, and gear types. Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Middle dots: posterior median.

Abundance— Estimated abundance (based on posterior medians) at the beginning of each year were lowest for the spawning populations in the MI 2_3 and ON central basins, and was higher for those in the MI 1_4 and ON south basins (Figure 3.6). From 2004 to 2007, the estimated abundance decreased for the spawning populations in all MI basins and ON north basin, and slightly increased in ON central and ON south basins. Overall, uncertainties are high for the abundance estimates across all years and basins. However, the abundance for the mixed stock in each basin was more stable across different years (Figure 3.7). For the mixed stocks in all MI basins, the mixed abundance decreased over time for young fish (ages 6 to 9), and increased by over time for older fish (ages 10 to 12). There was no noticeable change in mixed abundance though years for the older group in all ON basins. While for younger fish, abundance of mixed stock fish in the ON central basin trended upward by years, whereas the abundance of mixed stock fish in the ON north basin trended downward. Mixed stock abundance at age was roughly constant over time in the ON south basin.

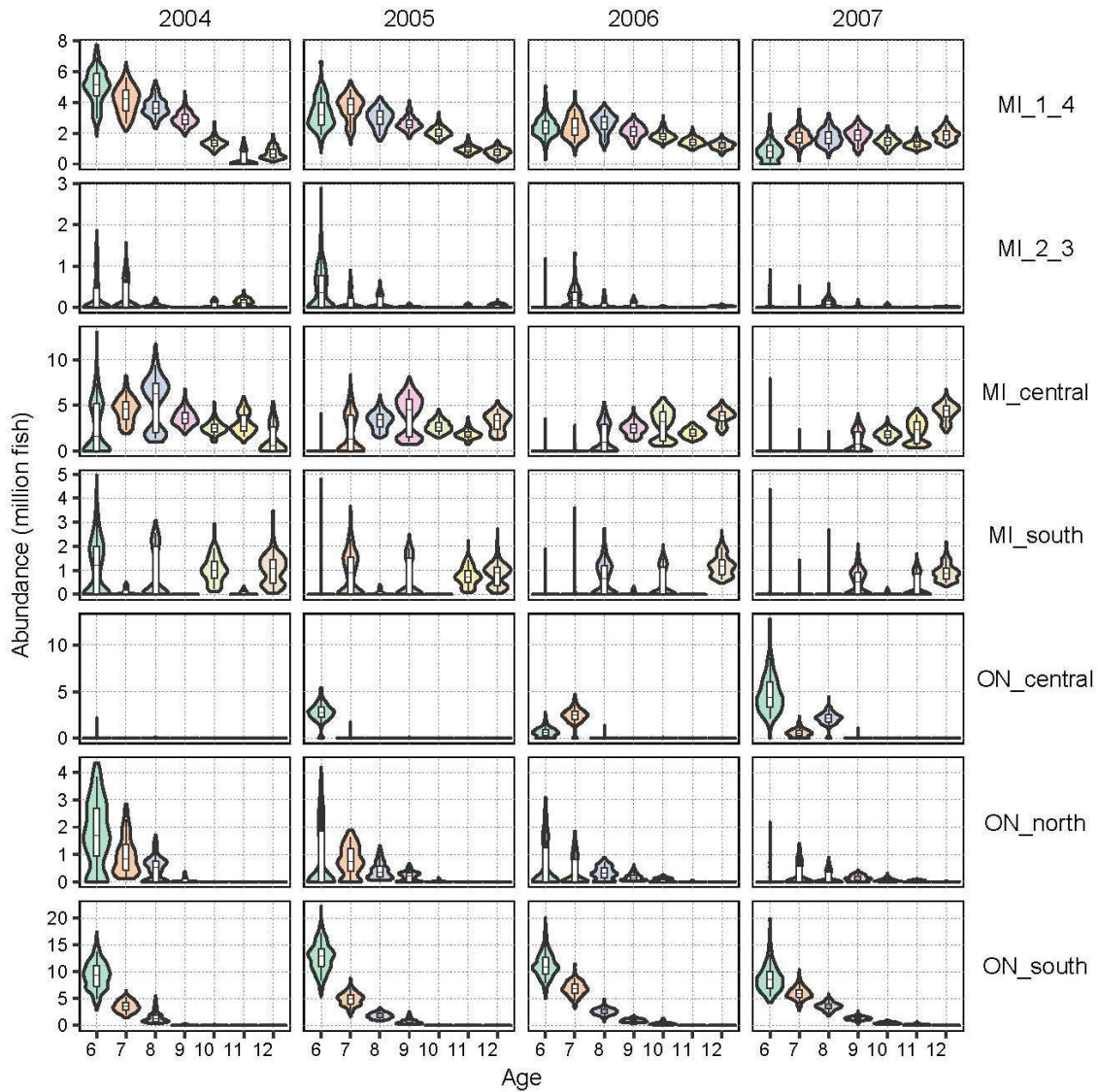


Figure 3.6. Posterior distribution of abundance at age for each spawning population (row names) and recovery years (column names). Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Horizontal middle line: posterior median.

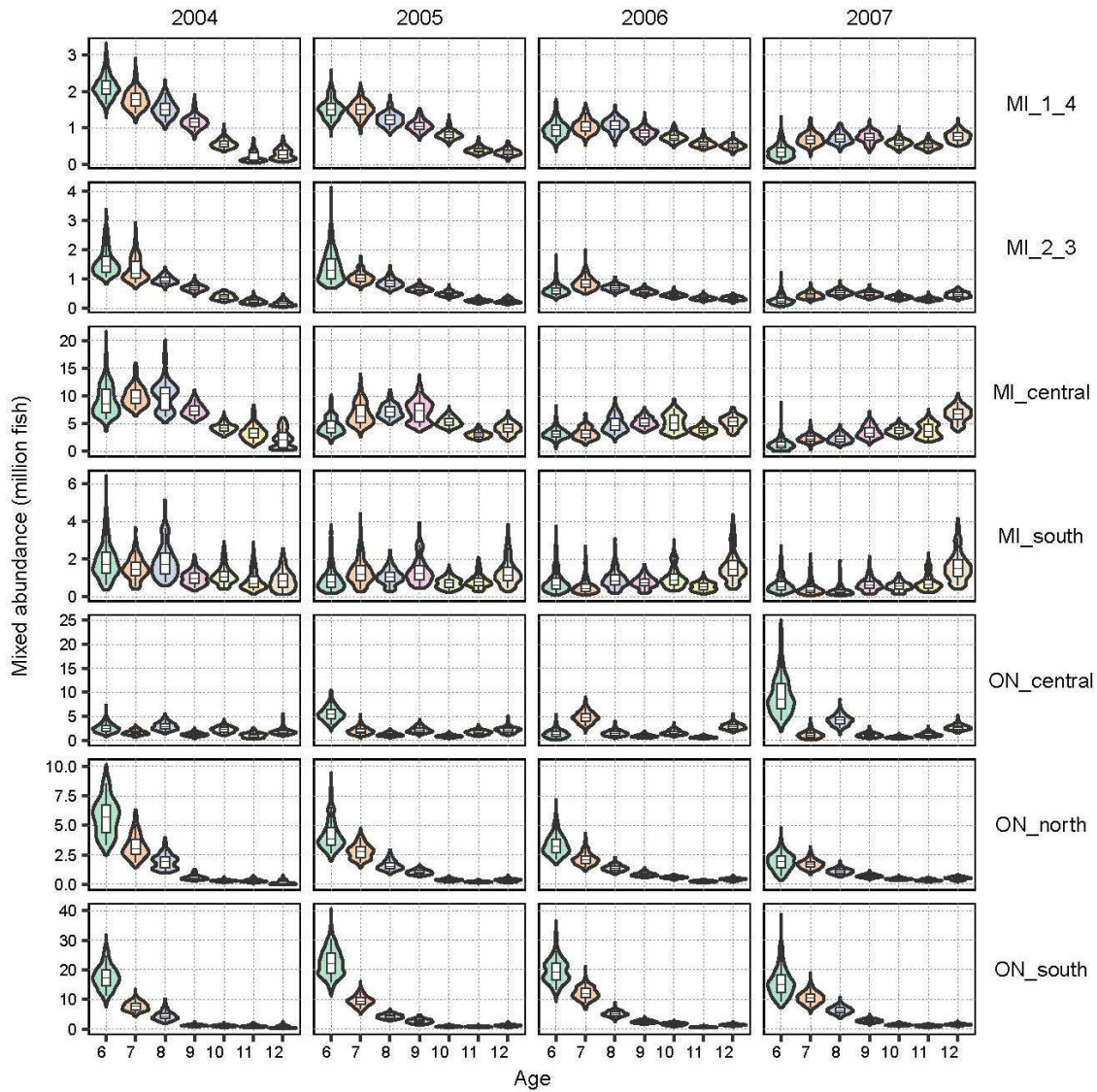


Figure 3.7. Posterior distribution of mixed abundance at age for each year and basin. Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Horizontal middle line: posterior median.

Assessing model fit

In most cases (by tagging basin and recovery basin), the observed number of recoveries were contained within the corresponding 95% highest posterior density interval (HPDI) for the predicted number of recovery (Figure 3.8). The only exceptions were the lake whitefish tagged in MI central, MI south, ON central and recovered in MI central, and those both tagged and recovered in ON south. Observed catch by basin and

year was generally contained within the corresponding 95% HPDI for predicted catch (Figure 3.9), with the exemptions of catch from MI 1_4 in 2004 and 2005, from MI south in 2004 and 2007, and from MI 2_3 in 2005, and from ON north in 2007. Most observed age composition proportions for catch by basin were contained in the corresponding 95% HPDI (Figure 3.10), with the exception of the first and last age groups fish in MI basins.

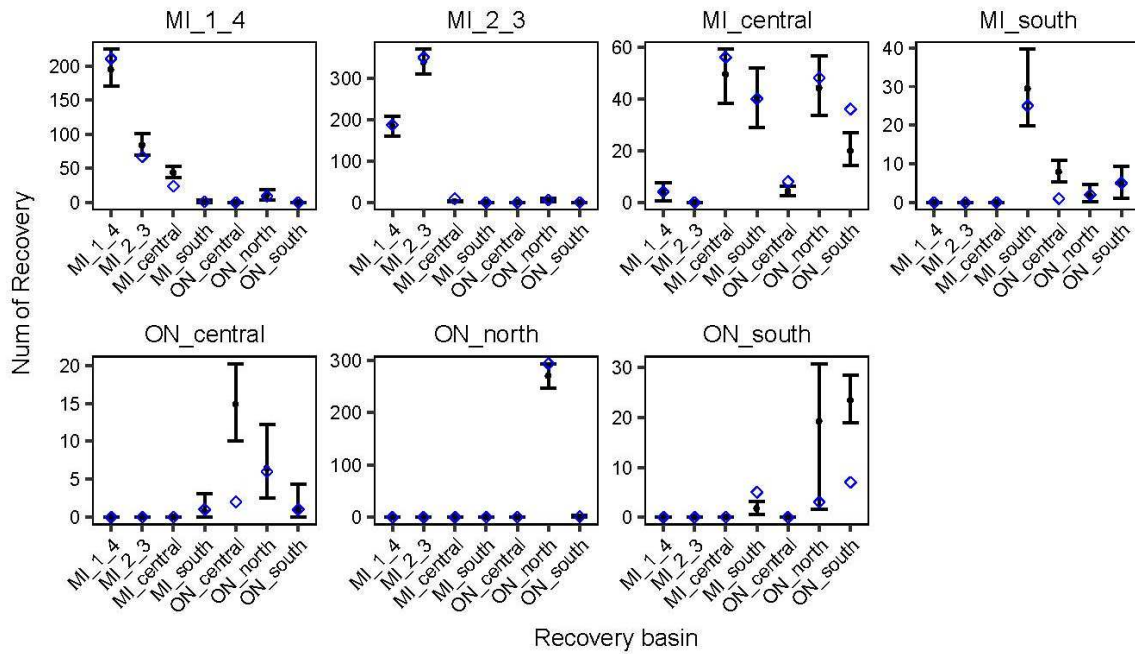


Figure 3.8. Comparison between the observed data and the posterior predictive distribution for the number of recovered fish, by recovery basin (x-axis) for each tag basin (panel). Error bar: 95% Bayesian highest posterior density interval (HPDI). Dot: Posterior median. Diamond: observed value.

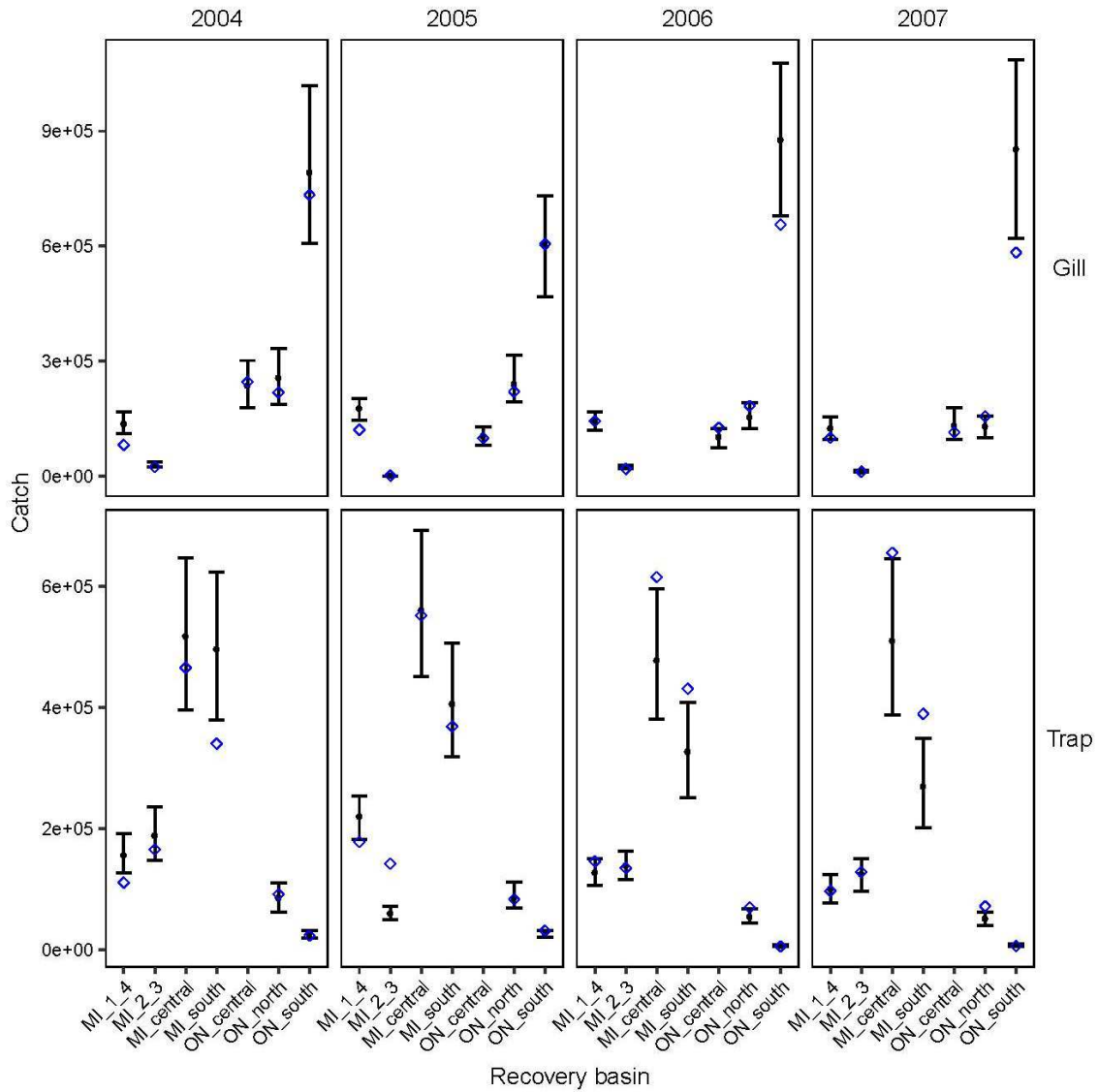


Figure 3.9. Comparison between the observed data and the posterior predictive distribution for the catch from each recovery basin (x-axis) in each year (column) by each gear type (row). Error bar: 95% Bayesian highest posterior density interval (HPDI). Dot: Posterior median. Diamond: observed value.

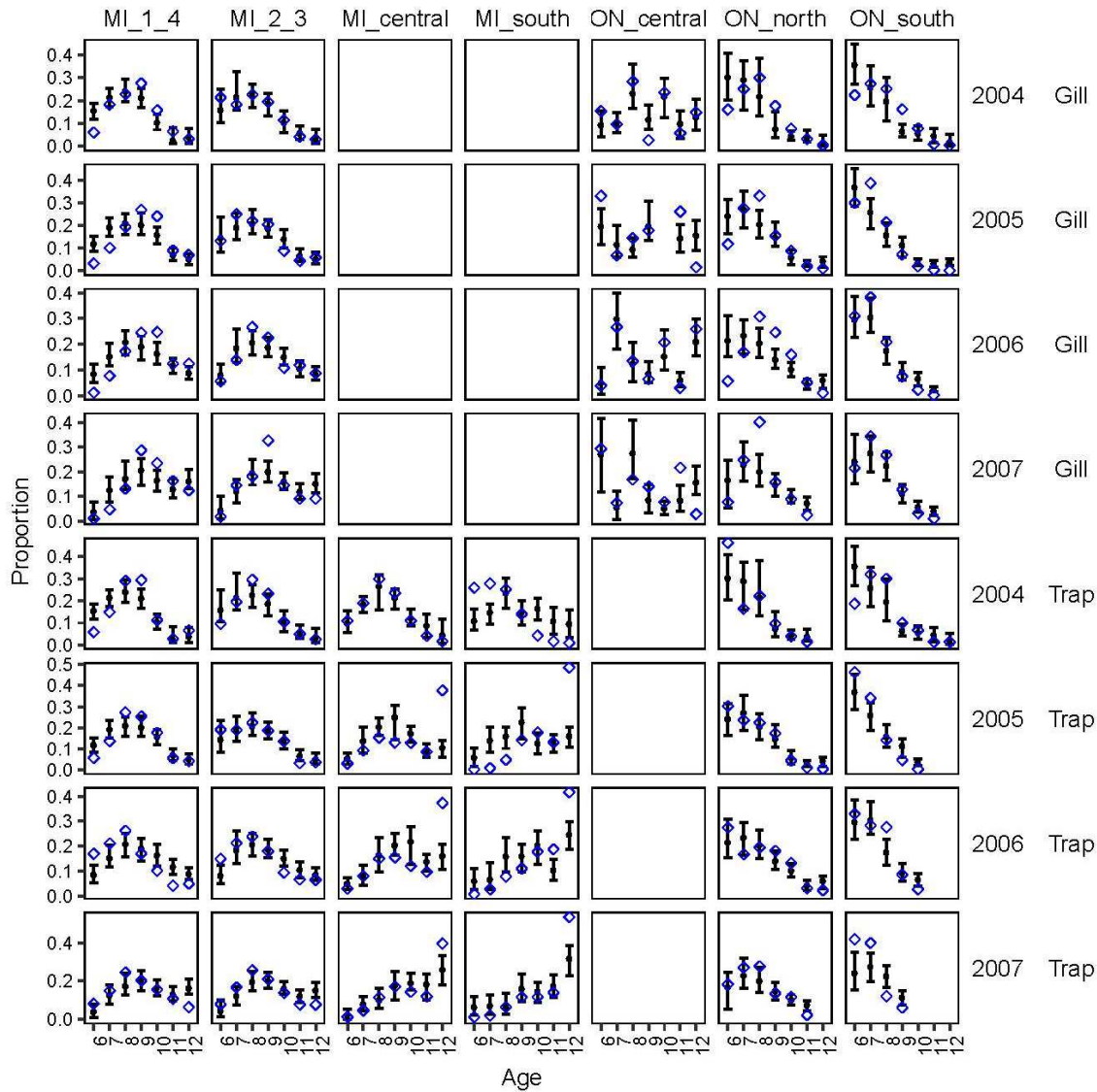


Figure 3.10. Comparison between the observed data and the posterior predictive distribution for the age composition proportion in each recovery basin (x- column) in each year and gear type (row). Error bar: 95% Bayesian highest posterior density interval (HPDI). Dot: Posterior median. Diamond: observed value

Sensitivity of parameter estimates to model parameterization

To test the sensitivity of our best model to how tag reporting rates and natural mortality rates were parameterized, we compared our best model with model #4 and 9, and with model #11, 12, 13, and 15, separately. Results of the full model were included in the supplementary materials. Based on the comparison with the results of model #4 and 9, estimated tag reporting rates, movement rates, and fishing mortality rates are

sensitive to how tag reporting rates were modeled. Of particular note, for lake whitefish tagged or recovered in MI south, MI central, and ON central basins, there were substantial shifts in posterior medians and much wider CI ranges, in comparison with the best model. Both models also had quite different estimates of natural mortality rates for the central basins (MI central and ON central basin): for model #4, the 95% CI range of central basins tag reporting is from 0 to 7.3 with posterior median equal 3.9; and from 0 to 5.8 with a posterior median equals 3.2 for model #9. Based on comparing results from our best model with those from models #11, 12, 13, and 15, we found that estimates (other than natural mortality) were not sensitive to how natural mortality was parameterized. The only exception was modest shifts in posterior medians and larger CI ranges for the fully selected fishing mortality rates and abundance estimates, in comparison with the best model.

Overall, results are sensitive to our assumption about the restriction in movement rates, especially for the dynamics of spawning population in ON north. Compared to our best model, the full model without movement restriction led to lower stay rate for the spawning population in the ON north (Figure 3.A1). Around 35% (posterior median) of the ON north population moved to ON central, while in our best model, we forced that probability to equal 0. The movement rate from ON south to ON north also decreased in the full model compared to our best model. Although the movement rates estimates of other spawning populations were not sensitive to our assumption of movement rates restriction, such different estimates of movement rates for the ON north spawning population also influenced the overall estimates of other parameters. Compared to our best model, the estimated natural mortality rates in the full model were higher for the

north and south groups, and were lower for the central group. The fully selected fishing mortality rates were higher across all years and basins, while the selectivity for age 6 to age 8 were slightly lower in the full model (Figure 3.A3 and Figure 3.A4). The estimates of tag reporting rates, abundance before mixing, and mixed abundance were lower for all other basins except ON north.

Comparison between best model and reduced model

In general, our reduced best model without catch-at-age data had similar estimates of fully selected fishing mortality rates, but different estimates of natural mortality rates, movement rates (moving to), and tag reporting rates for the central and south regions (see Table 3.A2 and Figure 3.A10-3.A14 in the supplementary material). Natural mortality rates in the reduced best model were high for the central region and low for the south region, which was opposite from our best model. For most spawning populations, estimated movement rates to MI south regions substantially increased in the reduced model compared to the best model. The tag reporting rates in the MI central and south region and in the ON central and south region were both greater, and uncertainty in estimated selectivity at age also increased for the reduced model.

Discussion

Although the Brownie-type tagging models have been widely used in the past few decades, spatially structured tagging model framework is relatively new (Eveson et al. 2009, Vandergoot and Brenden 2014), especially for fish populations with natal homing. In this study, we proposed a spatially structured tagging model framework for fish populations with overlapping spatial structure, and 100% natal homing rate were assumed, and provided a series of approaches to real-world applications about prior

incorporation, Bayesian analysis for parameter estimation and uncertainty quantification using DIC to determine model parameterization. Although in the other branch of tagging models, integrated tagging and catch-at-age analysis, natal homing of fish populations has recently been considered (Vincent et al. 2017), their focus was to use tagging data as additional information for estimating movement within a spatial structured catch-at-age model, and their framework cannot work independently without catch-at-age data. Our framework focused more on fitting models to the tagging data, and has the flexibility to either incorporate catch-at-age, scientific observer, and tag-recovery data (full model), or work independently without catch-at-age data. Thus our framework can be used to evaluate the potential data conflict between the catch-at-age and the tag-recovery data sets. Data conflict is a common issue when tag-recovery data were analyzed together with other data sources. It can be caused by different reasons such as observation errors, ageing error, low tag return rates, or ignored tag shedding and tag mortality. Although in the real-world applications it is always difficult to detect the specific reasons of data conflict, the awareness of data conflict is still of great importance for interpreting the parameter estimations from the tagging models. Our framework used Bayesian analysis for parameter estimations, which is widely recognized as a robust approach of evaluating the uncertainties of parameter estimations. All those efforts were made to address the most common issues that real-world tagging studies can encounter: low accuracy and precision in estimating parameters (Goethel et al. 2015a, Kerr et al. 2016). To the best of our knowledge, only few studies of spatial structured tagging models or integrated models have been applied to the real-world tagging data (Goethel et al. 2015a), and most of them encountered difficulties in reliable estimating parameters, especially without

fixing reporting rates or natural mortality rates. Our framework allows simultaneously estimations of movement, natural mortality and tag reporting rates, and used DIC to select best model that was supported by the real-world data from different model parameterizations.

Although we aimed at developing a general model that makes use of all available data to estimate parameters, the setting of a Brownie-type tagging model largely limited the use of the catch-at-age data. In many real world applications, we have longer time series and broader geographic coverage for catch-at-age data than for the tag-recovery data. For our framework, however, in order to allow estimations of movement and mortality rates from tagging data along (our reduced model) and avoid making arbitrary assumption about how movement and mortality changed in time and space, we have to make sure the spatial and temporal scales of tag-recovery and catch-at-age data are the same. Such sacrifice of catch-at-age data may not guarantee better estimates of parameters, especially for the estimates of recruitment, abundance, and age composition. In our application to lake whitefish, we can only use four years of catch-at-age data to estimate recruitment and initial abundance, which obviously not enough for a fish species like lake whitefish that survives past age 12 (the aggregated oldest age we used in age-structured assessments). For the real world stock assessment of lake whitefish, without considering movement and tag-recovery data, 40 years or even longer time series of catch-at-age data were considered in statistical catch-at-age models. For cases like this, to cut the length of years for catch-at-age data, or to ignore the movement between populations, becomes a question that needs to be considered before analysis. In our application of lake whitefish, we chose to cut the length of years for catch-at-age data

because our goal was to estimate movement and natural mortality rates, but we did find that the estimations of recruitment and abundance were very low, likely unreasonably so, for some basins.

The results of our application to lake whitefish indicates that when there are particular limitations to the data, the results can quite uncertain and sensitive to reasonable assumptions. For example, we combined some basins without observer coverage and assumed the tag reporting rates in those basins were the same as their adjacent basins. We did that not because we have prior hypothesis about tag reporting rates in those regions, but because those regions without information of tag reporting rates were black boxes of our model framework, and ignoring it would make the whole model crash. The same issues apply for the parameterization of natural mortality rates. As for those arbitrary assumptions we made, we did provide different hypotheses and used DIC to select the “best” model that was supported by data. We believe this is one of the best practices for real-world cases with limited data. Due to the low number of tag returns, we also made “ad hoc” assumption about the movement rates by forcing them to be zeros for the regions with no observed recoveries. Although our sensitivity results indicated that the estimates of some parameters were sensitive to this assumption, the results of our full model without restriction on movement rates tended to be unrealistic. Based on the results of the full model, more than 60% of the spawning populations from all other basins except for MI_23 and ON_north and central moved to MI south according to the posterior median, with a high natural mortality rates and tag reporting rates, and extreme low fishing mortality rates estimated for MI south (posterior median around 0.0075). We suspect there were positive biases in estimating natural mortality

rates, and negative bias in estimating fishing mortality rates, which suggests that fewer returns were seen than would be predicted in the full model.

Given the robustness of the movement estimates, we believe these provide useful inferences regarding lake whitefish in Lake Huron. The estimates of fishing mortality and natural mortality are questionable, given their sensitivity and the limitations of the data. The sensitivity of the natural mortality estimates to restrictions on movement rates and assumptions about reporting rates emphasizes the importance of having a robust tag recovery monitoring program across the entire area being considered for a spatial model. The short length of the time-series provides little ground truth on whether the magnitudes of natural mortality that were estimated are consistent with long-term fishery dynamics. This is important because catch-at-age stock assessments of lake whitefish in Lake Huron, using longer times series have found that it is necessary to assume much lower natural mortality in order to match observed catch-at-age data (MSC 2017). One avenue to move forward would be to use the tagging data with longer-term catch at age data, and with more restrictive assumptions on natural mortality (e.g., assume constant rates over time or constant at least for the period with tagging data) and movement rates (e.g., assume constant or known rates over time for the period without tagging data). Such an integrated tagging model, although not allowing a full evaluation of how parameters vary spatially and temporally, could potentially provide estimates suitable for fishery management, and address the substantial movement we know is occurring among areas. A significant issue is that past work has suggested that spatial catch-at-age models without tagging data can have identifiability issues, even when movements are known (Li

et al. 2015). It is an open question on whether short tagging series would ameliorate this issue.

Although we increased model complexity compared to previous tagging studies, by including additional dimensions such as age groups, gear types, and being observed on boat or not, we still made some simplifications for estimated parameters such as time-invariant catchabilities and same natural mortality rates for some age, year, or region groups. Given our study was short-term, with only four years of data, assuming constancy of some parameters may not be a serious issue. However, it may lead to biased estimations for long-term tagging studies.

APPENDIX

Table 3.A1. Instantaneous rates of natural mortality (per year) for the best model without restriction on movement.

	Posterior Mean	Posterior Standard Deviation	95% credible interval
North group:	1.04	0.09	(0.85, 1.20)
Central group:	0.04	0.02	(0.01, 0.09)
South group:	0.72	0.13	(0.52, 1.05)

Table 3.A2. Instantaneous rates of natural mortality (per year) for the reduced best model without catch and age composition data.

	Posterior Mean	Posterior Standard Deviation	95% credible interval
North group:	0.88	0.06	(0.76, 0.98)
Central group:	1.03	0.08	(0.86, 1.10)
South group:	0.09	0.08	(0.01, 0.28)

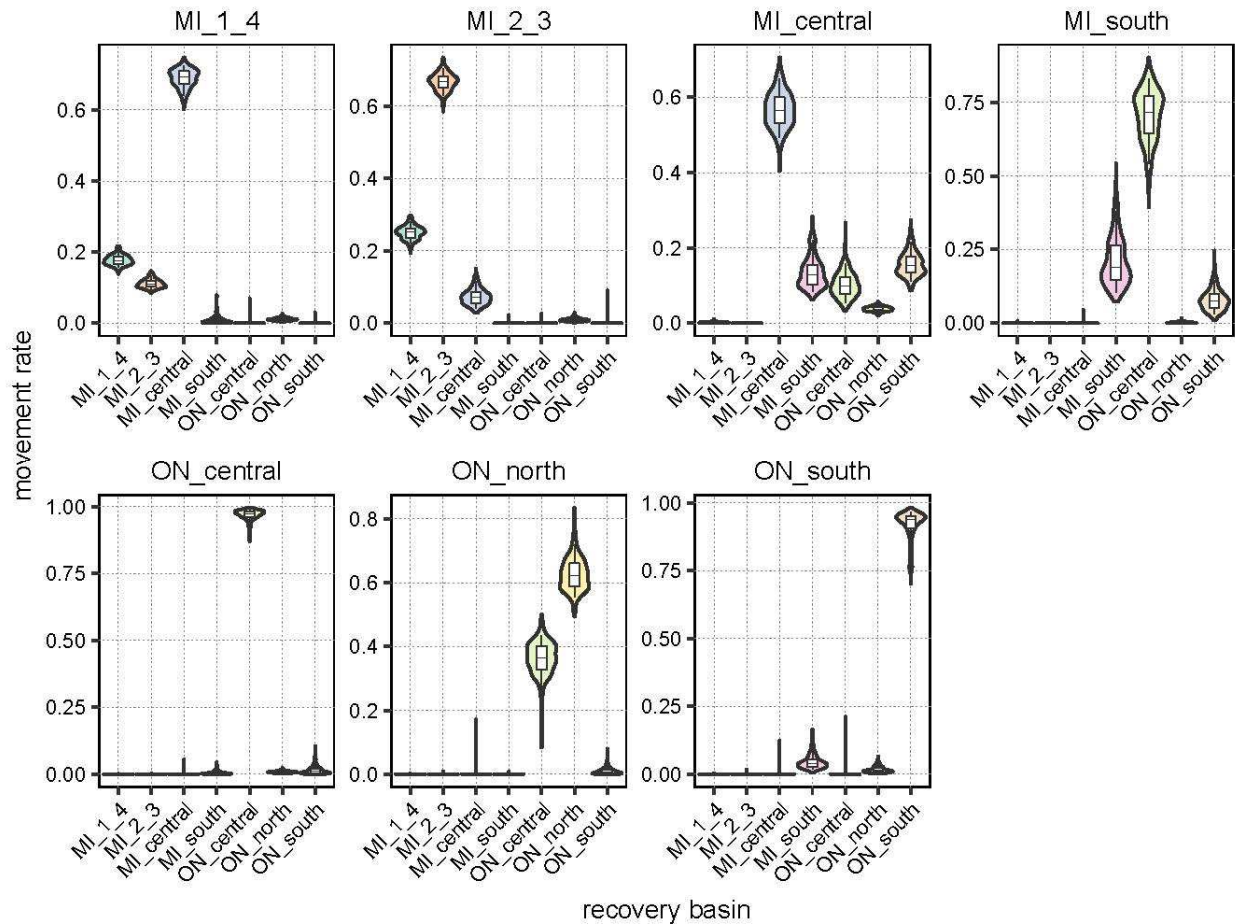


Figure 3.A1. Posterior distribution of the movement probability for moving from each tag basin (panel name) to all recovery basins (the x-axis of each panel). Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Horizontal middle line: posterior median.

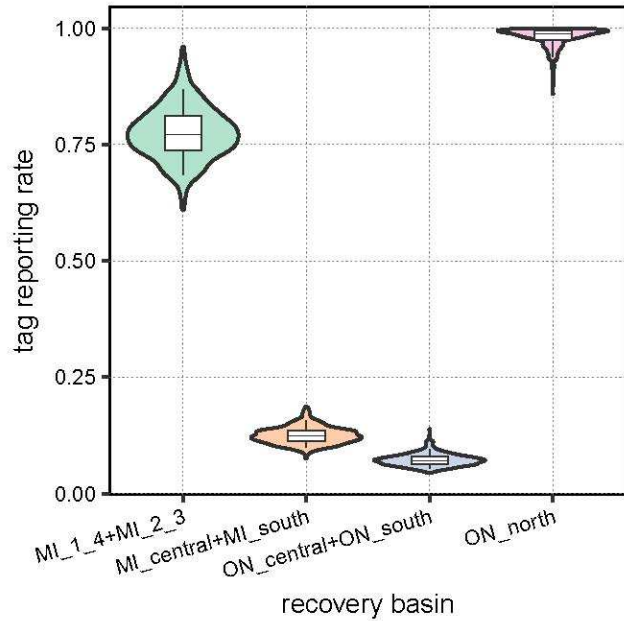


Figure 3.A2. Posterior distribution of tag reporting rates by recovery basins. Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Middle dots: posterior median.

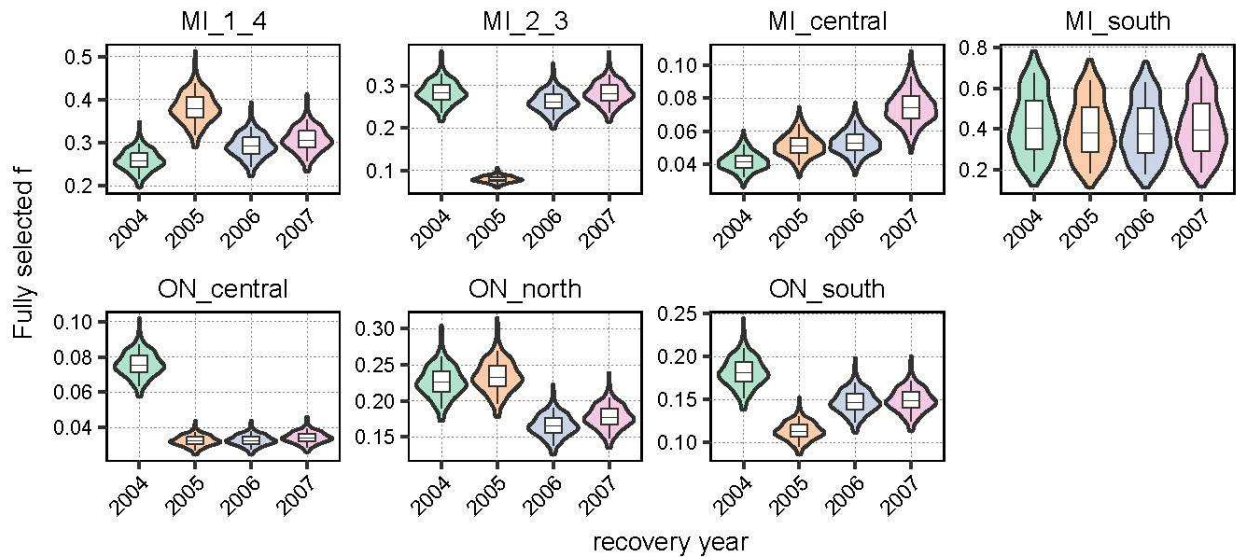


Figure 3.A3. Posterior distribution of fully selected fishing mortality by recovery years for each basin. Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Middle dots: posterior median.

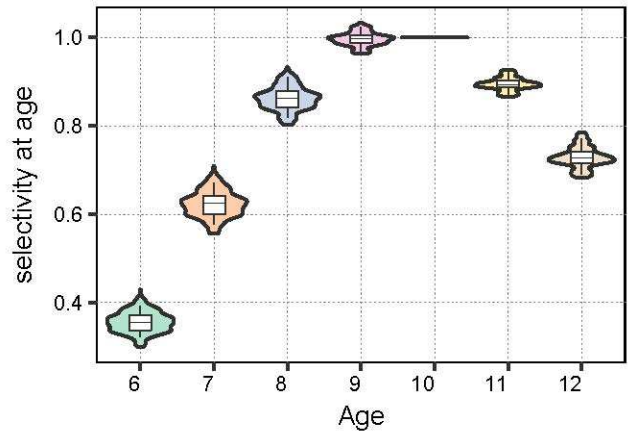


Figure 3.A4. Posterior distribution of selectivity at age for all years, basins, and gear types. Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Middle dots: posterior median.

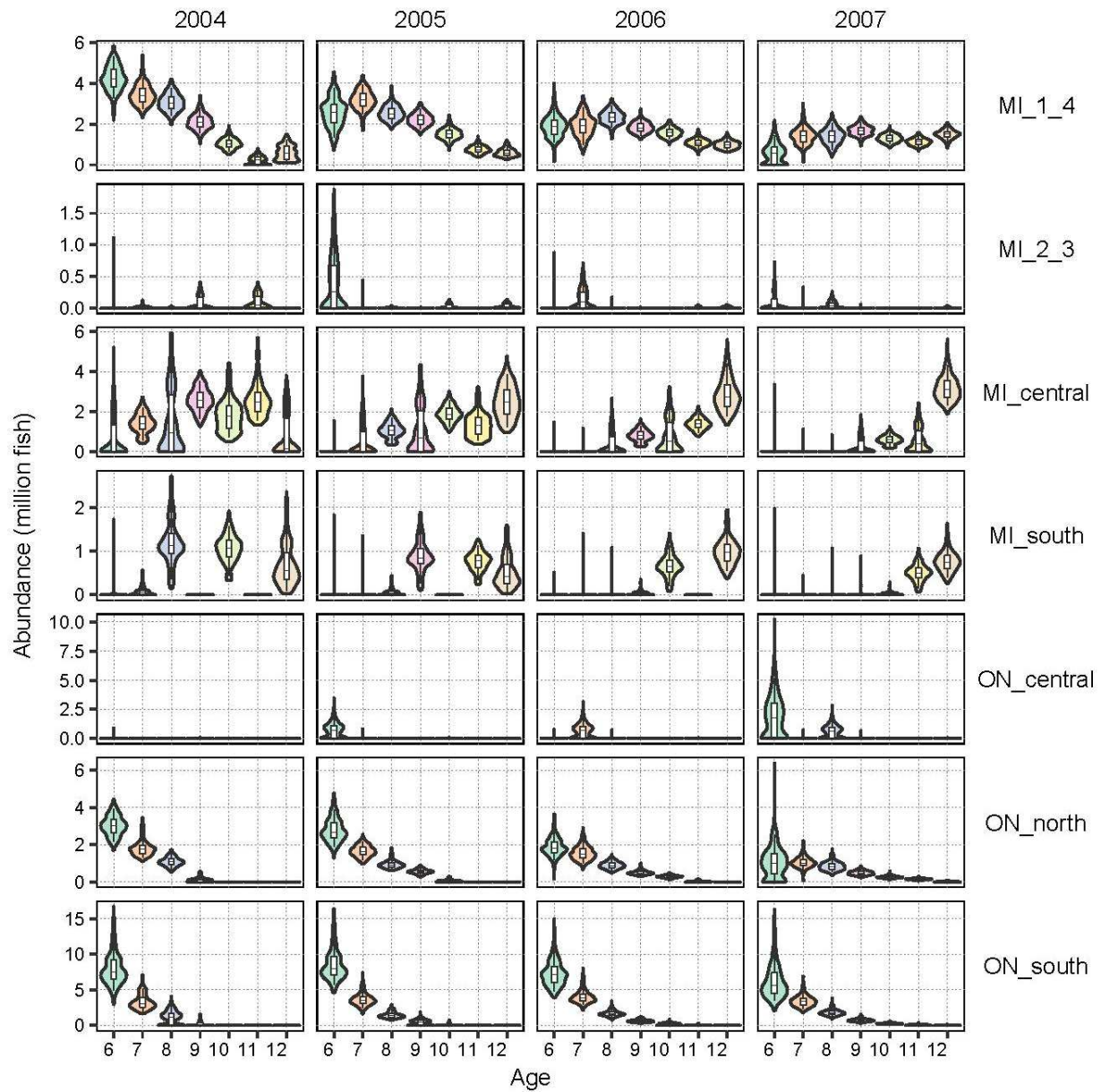


Figure 3.A5. Posterior distribution of abundance at age for each spawning population (row names) and recovery years (column names). Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Horizontal middle line: posterior median.

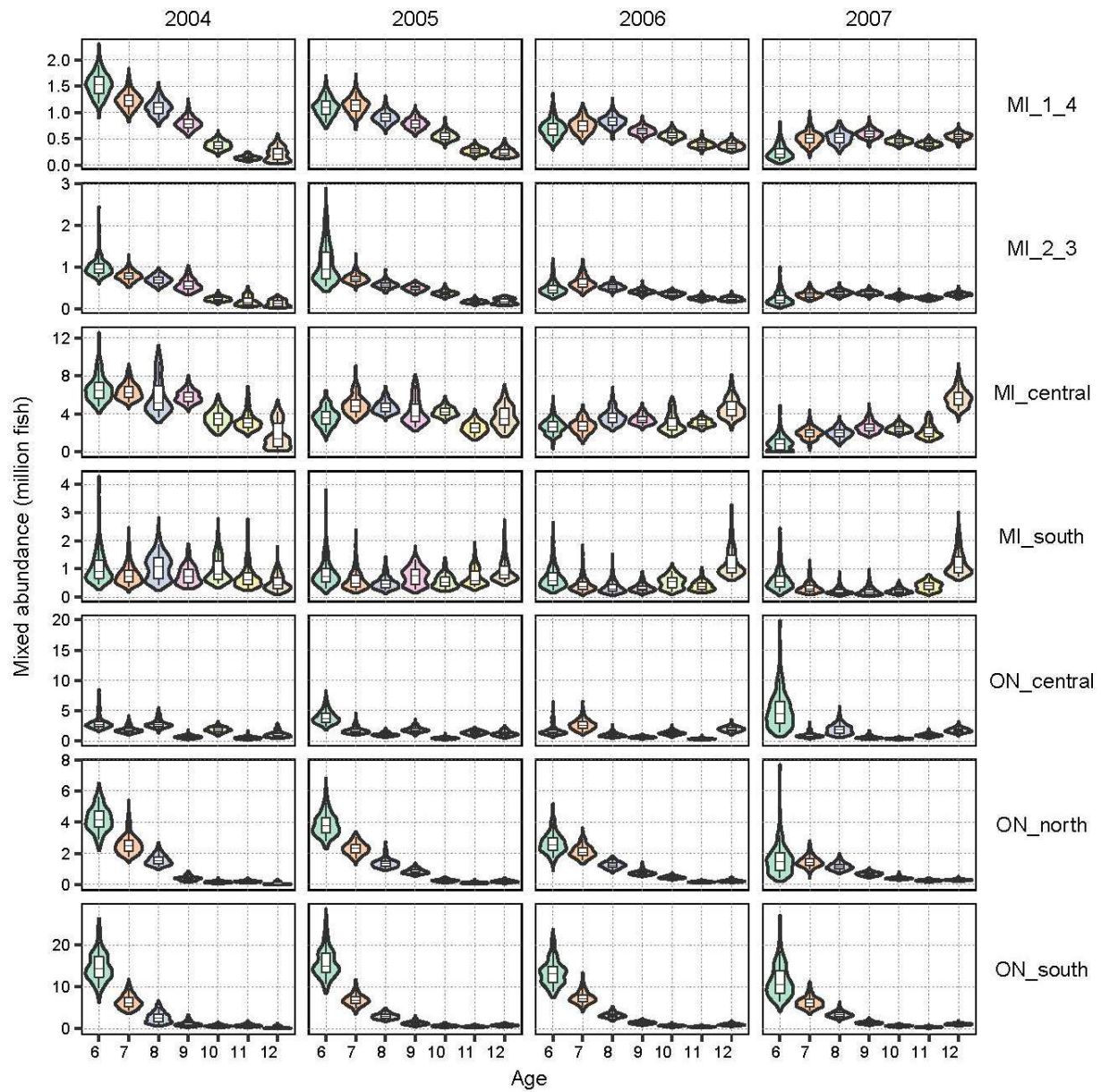


Figure 3.A6. Posterior distribution of mixed abundance at age for each year and basin. Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Horizontal middle line: posterior median.

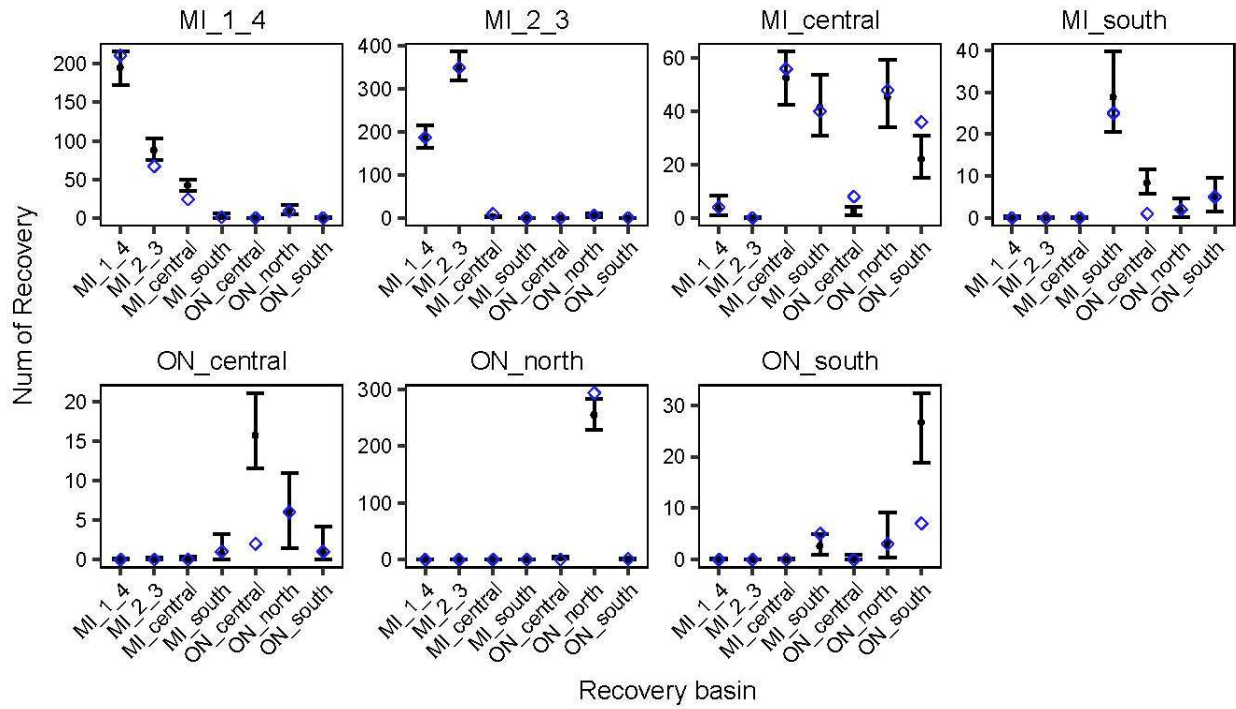


Figure 3.A7. Comparison between the observed data and the posterior predictive distribution for the number of recovered fish, by recovery basin (x-axis) for each tag basin (panel). Error bar: 95% Bayesian highest posterior density interval (HPDI). Dot: Posterior median. Diamond: observed value.

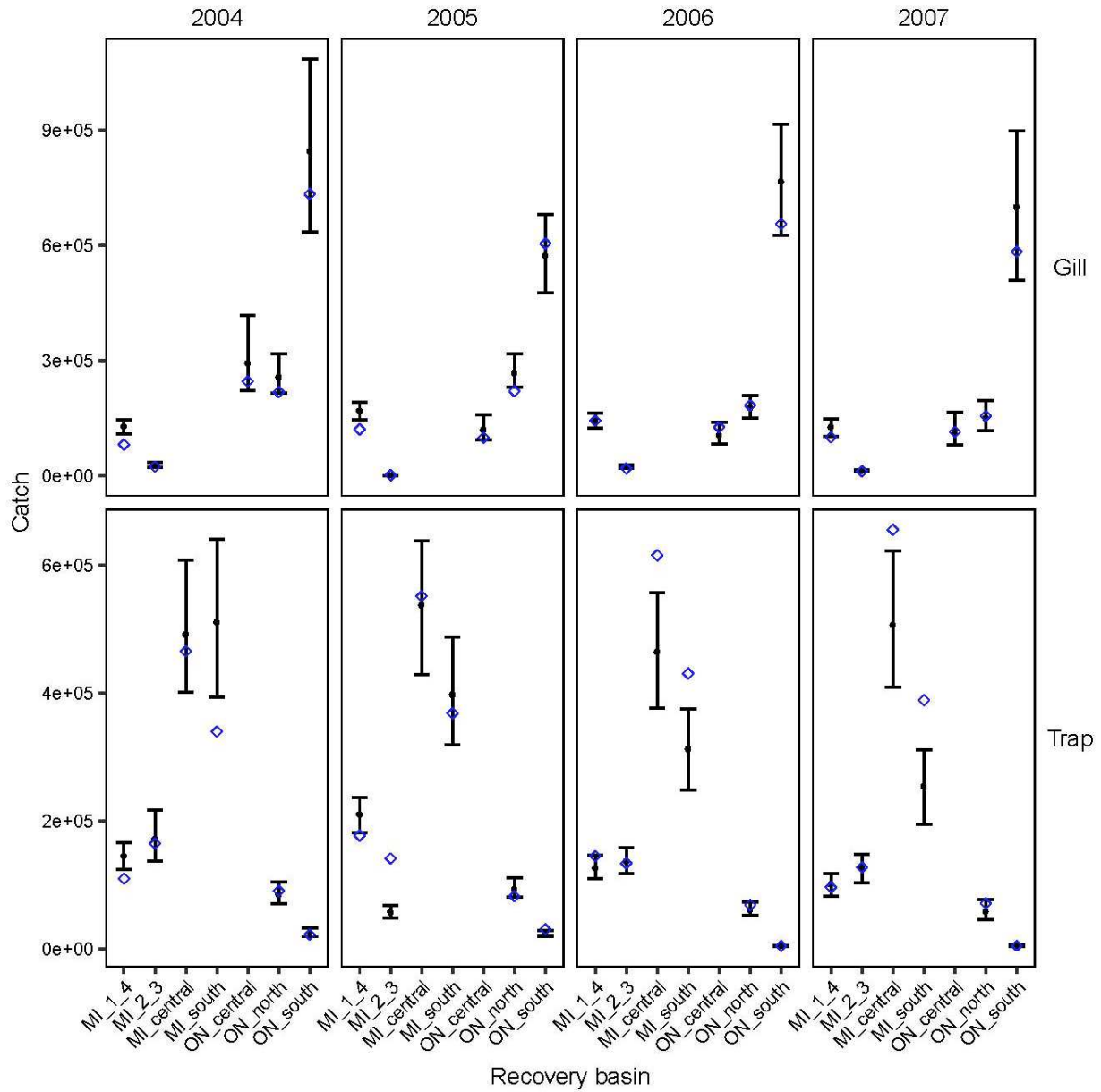


Figure 3.A8. Comparison between the observed data and the posterior predictive distribution for the catch from each recovery basin (x-axis) in each year (column) by each gear type (row). Error bar: 95% Bayesian highest posterior density interval (HPDI). Dot: Posterior median. Diamond: observed value.

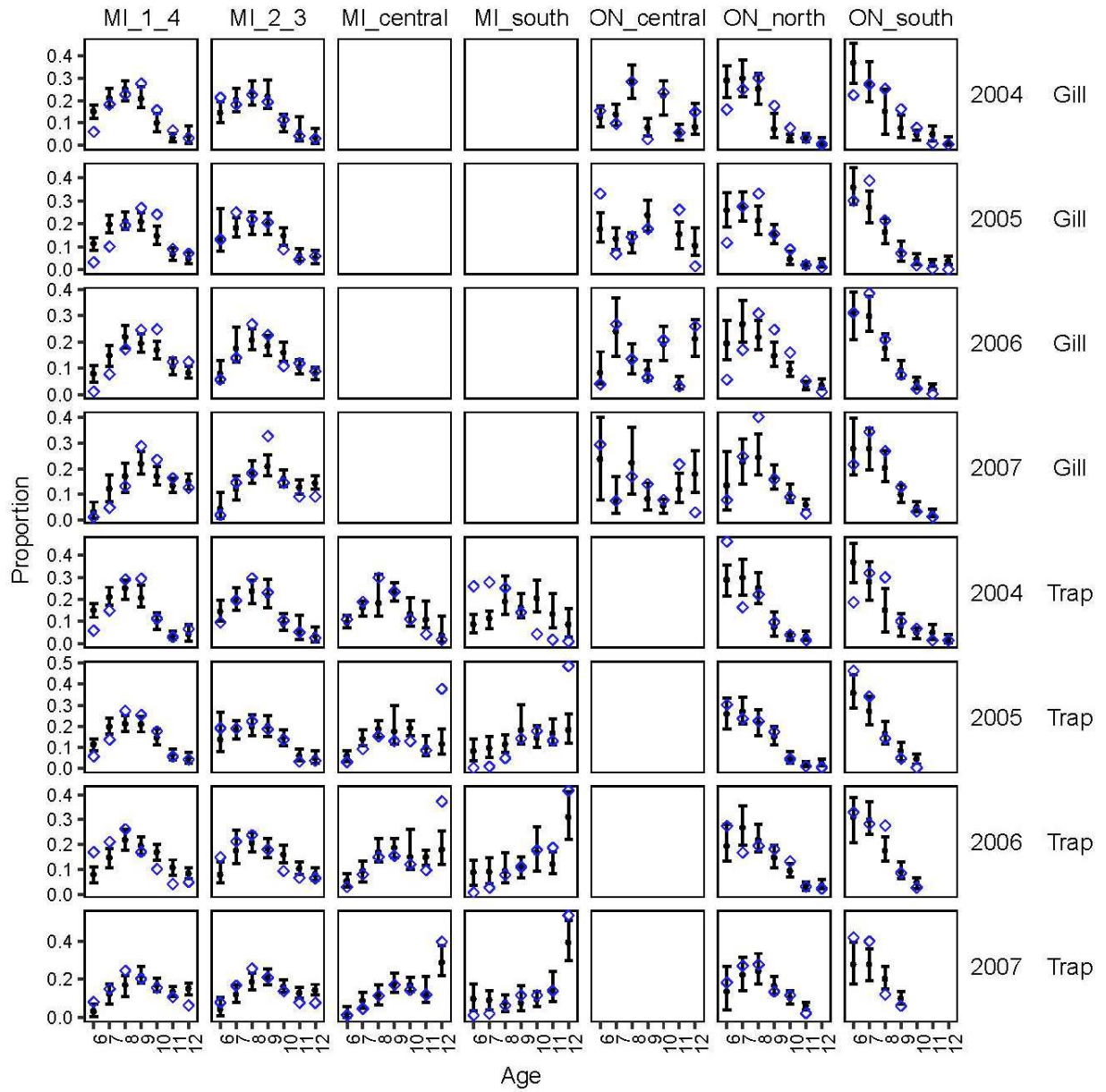


Figure 3.A9. Comparison between the observed data and the posterior predictive distribution for the age composition proportion in each recovery basin (x- column) in each year and gear type (row). Error bar: 95% Bayesian highest posterior density interval (HPDI). Dot: Posterior median. Diamond: observed value.

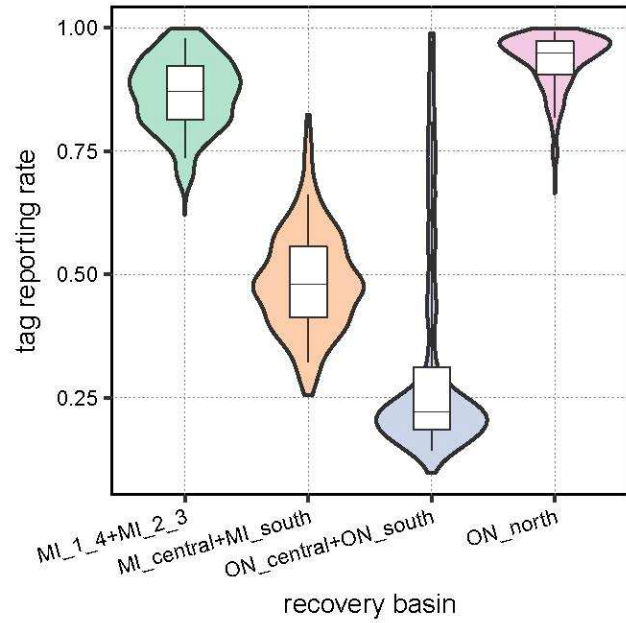


Figure 3.A10. Posterior distribution of tag reporting rates by recovery basins for the reduced best model. Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Middle dots: posterior median.

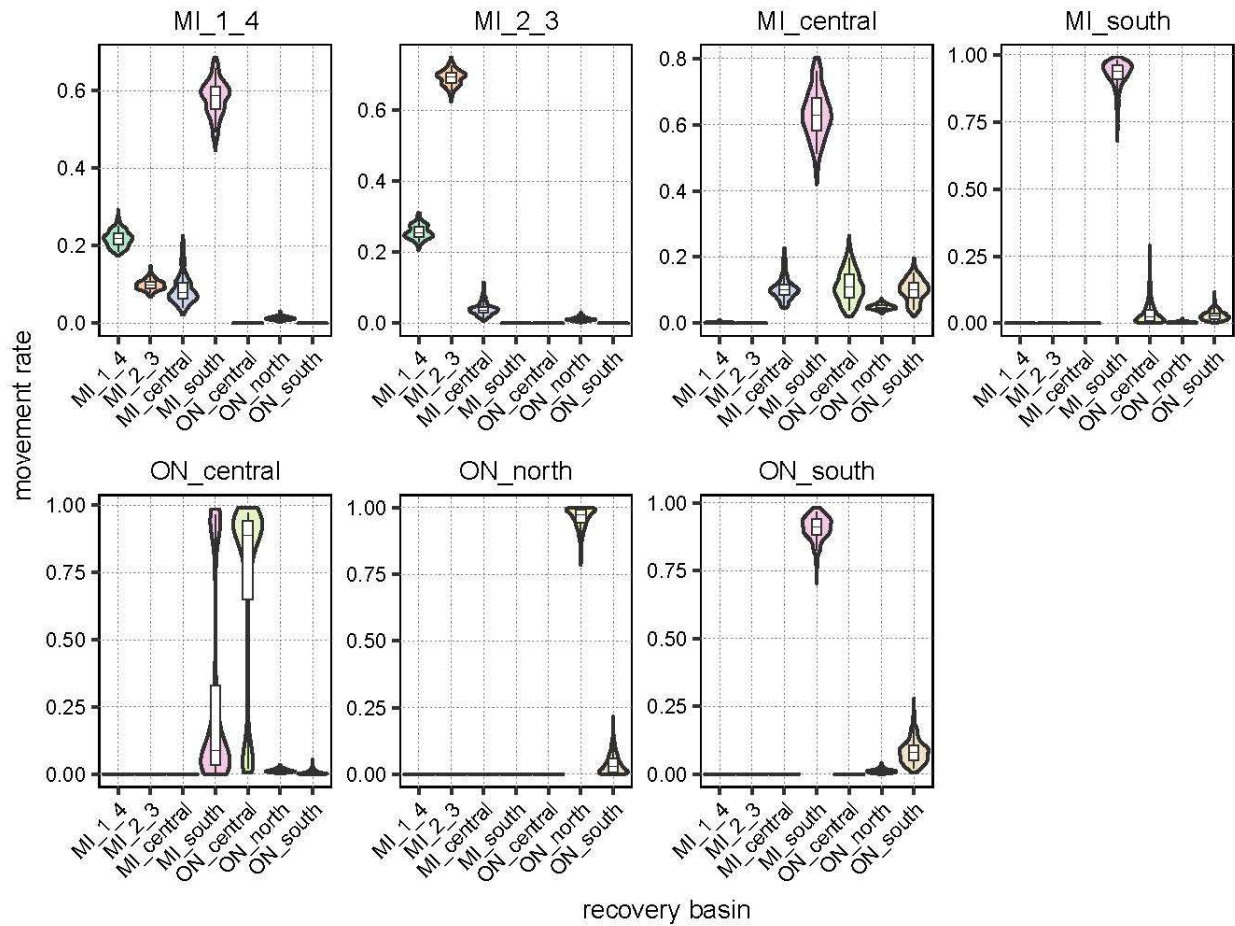


Figure 3.A11. Posterior distribution of the movement probability that moving from each tag basin (panel name) to all recovery basins (the x-axis of each panel) for the reduced best model. Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Horizontal middle line: posterior median.

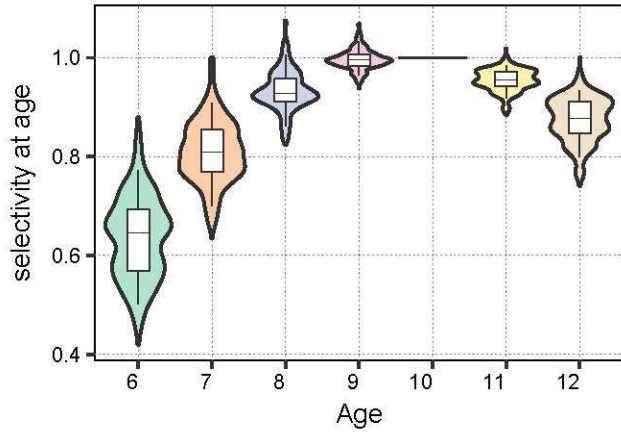


Figure 3.A12. Posterior distribution of selectivity at age for all years, basins, and gear types for the reduced best model. Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Middle dots: posterior median.

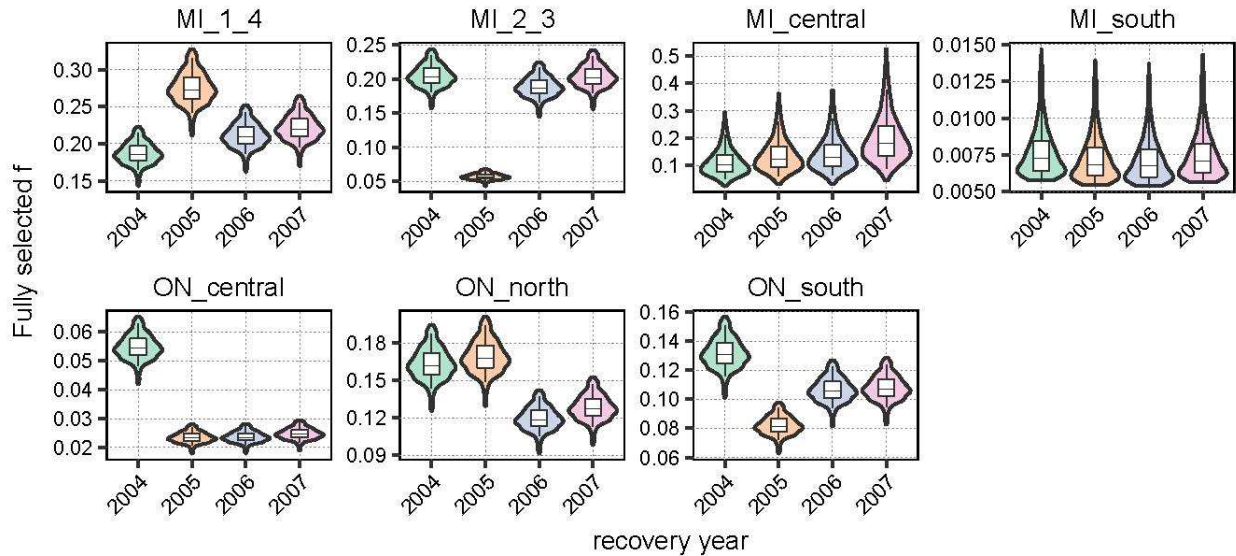


Figure 3.A13. Posterior distribution of fully selected fishing mortality by recovery years for each basin for the reduced best model. Violin: mirrored posterior densities. Boxes: 50% Bayesian credibility interval. Whiskers: 95% Bayesian credibility interval. Middle dots: posterior median.

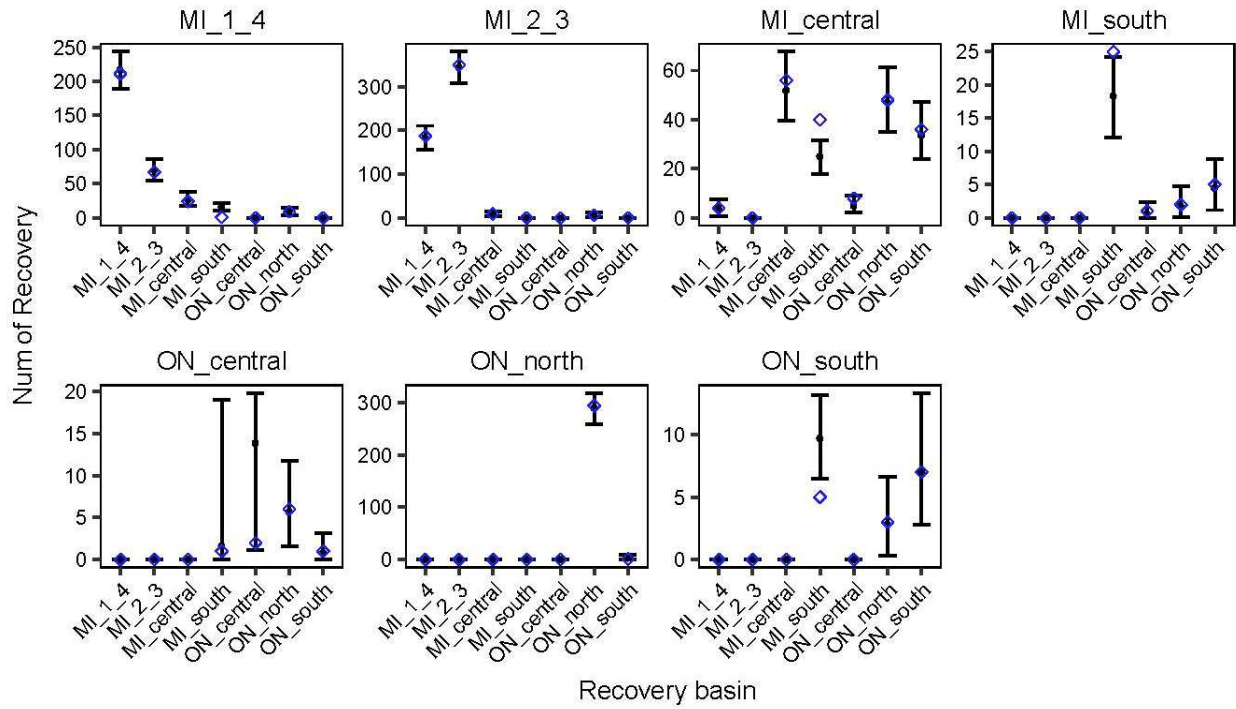


Figure 3.A14. Comparison between the observed data and the posterior predictive distribution for the number of recovered fish, by recovery basin (x-axis) for each tag basin (panel) for the reduced best model. Error bar: 95% Bayesian highest posterior density interval (HPDI). Dot: Posterior median. Diamond: observed value.

BIBLIOGRAPHY

BIBLIOGRAPHY

- Bailey, L.L., Converse, S.J., and Kendall, W.L. 2010. Bias, precision, and parameter redundancy in complex multistate models with unobservable states. *Ecology* **91**(6): 1598–1604. doi:10.1890/09-1633.1.
- Brownie, C., Anderson, D.R., Burnham, K.P., and Robson, D.S. 1985. *Statistical inference from band recovery data - a handbook*.
- Cole, D.J., and McCrea, R.S. 2016. Parameter redundancy in discrete state-space and integrated models. *Biometrical J.* **58**(5): 1071–1090. doi:10.1002/bimj.201400239.
- Ebener, M.P., Brenden, T.O., and Jones, M.L. 2010a. Estimates of fishing and natural mortality rates for four lake whitefish stocks in northern Lakes Huron and Michigan. *J. Great Lakes Res.* **36**: 110–120. doi:10.1016/j.jglr.2009.06.003.
- Ebener, M.P., Brenden, T.O., Wright, G.M., Jones, M.L., and Faisal, M. 2010b. Spatial and temporal distributions of lake whitefish spawning stocks in Northern lakes Michigan and Huron, 2003–2008. *J. Great Lakes Res.* **36**: 38–51. doi:10.1016/j.jglr.2010.02.002.
- Eveson, J.P., Laslett, G.M., and Polacheck, T. 2009. A spatial model for estimating mortality rates, abundance and movement probabilities from fishery tag-recovery data. *In Modeling Demographic Processes In Marked Populations*. doi:10.1007/978-0-387-78151-8.
- Eveson, J.P., Polacheck, T., and Laslett, G.M. 2007. Incorporating fishery observer data into an integrated catch-at-age and multiyear tagging model for estimating mortality rates and abundance. *Fish. Bull.* **105**(4): 493–508. Available from <http://fishbull.noaa.gov/1054/eveson.pdf>.
- Francois, O., and Laval, G. 2011. Deviance Information Criteria for Model Selection in Approximate Bayesian Computation. arXiv **stat.CO**: 22. doi:10.2202/1544-6115.1678.
- Goethel, D.R., Legault, C.M., and Cadrin, S.X. 2015a. Demonstration of a spatially explicit, tag-integrated stock assessment model with application to three interconnected stocks of yellowtail flounder off of New England. *ICES J. Mar. Sci.* **72**(1): 164–177. doi:10.1093/icesjms/fsu014.
- Goethel, D.R., Legault, C.M., and Cadrin, S.X. 2015b. Testing the performance of a spatially explicit tag-integrated stock assessment model of yellowtail flounder (*Limanda ferruginea*) through simulation analysis. *Can. J. Fish. Aquat. Sci.* **72**(4): 582–601. doi:10.1139/cjfas-2014-0244.
- Guan, W., Cao, J., Chen, Y., Cieri, M., and Quinn, T. 2013. Impacts of population and fishery spatial structures on fishery stock assessment. *Can. J. Fish. Aquat. Sci.* **70**(8): 1178–1189. doi:10.1139/cjfas-2012-0364.
- Hamada, M.S., Wilson, A.G., Reese, C.S., and Martz, H.F. 2008. Bayesian Reliability. *In*

October. doi:10.1007/978-0-387-77950-8.

- Hearn, W.S., Polacheck, T., Pollock, K.H., and Whitelaw, W. 1999. Estimation of tag reporting rates in age-structured multicomponent fisheries where one component has observers. *Can. J. Fish. Aquat. Sci.* **56**(7): 1255–1265. doi:10.1139/cjfas-56-7-1255.
- Kerr, L.A., Hintzen, N.T., Cadrin, S.X., Clausen, L.W., Dickey-Collas, M., Goethel, D.R., Hatfield, E.M.C., Kritzer, J.P., and Nash, R.D.M. 2016. Lessons learned from practical approaches to reconcile mismatches between biological population structure and stock units of marine fish. *ICES J. Mar. Sci. J. du Cons.* **74**: fsw188. doi:10.1093/icesjms/fsw188.
- Li, Y., Bence, J.R., and Brenden, T.O. 2015. An evaluation of alternative assessment approaches for intermixing fish populations: a case study with Great Lakes lake whitefish. *ICES J. Mar. Sci.* **72**(1): 70–81. doi:10.1093/icesjms/fsu057.
- Li, Y., Bence, J.R., Zhang, Z., and Ebener, M.P. 2017. Why do lake whitefish move long distances in Lake Huron? Bayesian variable selection of factors explaining fish movement distance. *Fish. Res.* **195**: 169–179. doi:10.1016/j.fishres.2017.07.014.
- Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., and Teller, E. 1953. Equation of state calculations by fast computing machines. *J. Chem. Phys.* **21**(6): 1087–1092. doi:10.1063/1.1699114.
- Modeling Subcommittee, Technical Fisheries Committee (MSC). 2017. Technical Fisheries Committee Administrative Report 2017: Status of Lake Trout and Lake Whitefish Populations in the 1836 Treaty-Ceded Waters of Lakes Superior, Huron and Michigan, with Recommended Yield and Effort Levels for 2017.
- Polacheck, T., Eveson, J.P., Laslett, G.M., Pollock, K.H., and Hearn, W.S. 2006. Integrating catch-at-age and multiyear tagging data: a combined Brownie and Petersen estimation approach in a fishery context. *Can. J. Fish. Aquat. Sci.* **63**(3): 534–548. doi:10.1139/f05-232.
- Pollock, K.H., Hearn, W.S., and Polacheck, T. 2002. A general model for tagging on multiple component fisheries: An integration of age-dependent reporting rates and mortality estimation. *Environ. Ecol. Stat.* **9**(1): 57–69. doi:10.1023/A:1013715008683.
- Pollock, K.H., Hoenig, J.M., and Jones, C.M. 1991. Estimation of fishing and natural mortality when a tagging study is combined with a creel survey or port sampling. *Am. Fish. Soc. Symp.* **12**: 423–434.
- Porch, C.E., Turner, S.C., and Powers, J.E. 2001. Virtual population analyses of atlantic bluefin tuna with alternative models of transatlantic migration: 1970-1997. *Int. Comm. Conserv. Atl. Tunas Collect. Vol. Sci. Pap.* **52**(3): 1022–1045.
- Roberts, G.O., and Rosenthal, J.S. 2009. Examples of Adaptive MCMC. *J. Comput. Graph. Stat.* **18**(2): 349–367. doi:10.1198/jcgs.2009.06134.
- Rothschild, B.J. 2007. Coherence of Atlantic cod stock dynamics in the Northwest Atlantic

- Ocean. Trans. Am. Fish. Soc. **136**(3): 858–874. doi:10.1577/T06-213.1.
- Seber, G. a F. 1982. The estimation of animal abundance and related parameters. New York **2**: 654. doi:10.1002/iroh.19740590517.
- Svedäng, H., Righton, D., and Jonsson, P. 2007. Migratory behaviour of Atlantic cod *Gadus morhua*: natal homing is the prime stock-separating mechanism. Mar. Ecol. Prog. Ser. **345**: 1–12. doi:10.3354/meps07140.
- Thorrold, S.R. 2001. Natal homing in a marine fish metapopulation. Science (80-.). **291**(5502): 297–299. doi:10.1126/science.291.5502.297.
- Vandergoot, C.S., and Brenden, T.O. 2014. Spatially varying population demographics and fishery characteristics of Lake Erie walleyes inferred from a long-term tag recovery study. Trans. Am. Fish. Soc. **143**(1): 188–204. doi:10.1080/00028487.2013.837095.
- Vincent, M.T., Brenden, T.O., and Bence, J.R. 2017. Simulation testing the robustness of a multiregion, tag-integrated assessment model that exhibits natal homing and estimates natural mortality and reporting rate. Can. J. Fish. Aquat. Sci. **74**(11): 1930–1949. doi:10.1139/cjfas-2016-0297.

CHAPTER 4

CAN A TAGGING MODEL FRAMEWORK BE ADAPTED TO DIFFERENT SPATIAL STRUCTURE ASSUMPTIONS? A BAYESIAN SPATIALLY STRUCTURED ASSESSMENT MODEL FRAMEWORK INCORPORATING HOMING PROBABILITY

Abstract

We propose an integrated assessment model framework for modeling tag-recovery, tag monitoring, fishing effort, and catch-at-age data that allows for a continuum of spatial structures through modeling homing probability in the assessment model. Based on simulations of six hypothetical lake whitefish populations, we explored how the degree of homing, the extent of spatial movements, and which data used (only tag-recovery and effort data, or with additions of tag monitoring or tag monitoring and catch-at-age) influenced estimability of parameters of interest in our model framework. We also conducted sensitivity analyses to evaluate how data quality and recruitment variation influenced results. Our results suggested that our model framework with only tag-recovery and fishing effort data had robust assessment performance in estimating movement rates, homing probability, natural mortality, and fully selectivity fishing mortality rates. When tag monitoring data were also available, our framework can provide reasonable estimate of tag reporting rate simultaneously. Including additional catch-at-age data did substantially improve the estimates of selectivity at age, slightly improved estimates of tag reporting and natural mortality rates, but the bias in estimating recruitment and spawning stock biomass can be high, especially for low productivity populations. Our results in general is sensitive to data quality, but not sensitive to high recruitment variation.

Introduction

Spatial structure and movement patterns of fish populations have increasingly been investigated in the last few decades, largely due to the advanced tagging methodologies and development of spatial stock assessment models (Cadrin and Secor, 2009; Goethel *et al.*, 2011; Ying *et al.*, 2011). For fish populations, in addition to the “unit stock” type, which assumes fish spawning populations are isolated throughout the fishing and spawning periods, there are two alternative box transfer types of population spatial structures that have been widely recognized: diffusion and overlap (Vandergoot and Brenden, 2014; Goethel *et al.*, 2015a, 2015b). The diffusion (also called metapopulation) structure assumes that fish are allowed to spawn in any basin they are in at the time of spawning, and thus their spawning population identity changes when they move to a new basin. The overlap structure assumes all intermixing fish move back to their natal site before the spawning period of each year, and thus their natal spawning population identity is maintained. The diffusion structure has been assumed in most tagging or tag-integrated studies, either explicitly, or implicitly by the equations used (Vandergoot and Brenden, 2014; Goethel *et al.*, 2015a, 2015b), whereas the overlap structure has only sparingly been used in stock assessments or tagging models (Goethel *et al.*, 2011; Vincent *et al.*, 2017). Many fish species, however, exhibit natal homing, such as Atlantic Bluefin tuna (*Thunnus thynnus*), weakfish (*Cynoscion regalis*), and lake whitefish (*Coregonus clupeaformis*).

The diffusion and overlap structures assume either 0% or 100% homing, but actual fish populations will not generally perfectly match these categories. Tagging, genetic, or hard structure chemical signature data can be used to estimate the degree of

homing. When evaluated, even fish considered to have a high degree of homing have had some straying (Thorrold, 2001; Rooker et al., 2008; Ebener et al., 2010). Hereafter we refer to populations with some degree of homing that is less than 100% as having an “incomplete overlap structure”. Previous studies have revealed that ignorance or misspecification of spatial structure in stock assessment models can lead to biased estimates of population parameters and stock status, inappropriate harvest targets, and depletion of local populations, especially for low productivity populations (Ying *et al.*, 2011; Molton *et al.*, 2012; Li *et al.*, 2015). Thus, it is potentially important to incorporate information on the degree of homing for populations with an incomplete overlap structure, rather than treating them as having either an overlap or diffusion structure in spatially structured tagging models or stock assessment models (Stewart *et al.*, 2003).

Herein we present an integrated assessment model framework that allows for an incomplete overlap structure. The model has the potential to use a variety of data sources. In the most data sparse case, it uses only tag-recovery and fishery effort data, whereas it has the potential to also use catch-at age data, and tag monitoring data on the fraction of observed harvested tags that are reported. The model can be viewed as modification of the “spatial Brownie-Peterson” (SBP model) (Eveson *et al.*, 2009), allowing for observer data and incomplete overlap structure. We have, however, also integrated some aspects developments from the “integrated tagging and catch-at-age analysis”(ITCAAN) model (Goethel *et al.*, 2011, 2015b). The SBP and ITCAAN models are two well-known methods being widely used for estimating fish movement from tag-recovery and catch-at-age data. The underlying assumption for both types of models is that tagged fish and their spawning population at large experienced the same movement

and survival dynamics, and thus the subsequent recovery locations of tagged fish and harvest collected from different basins reveal where the cohort of the spawning population moved to from their spawning sites (i.e., tagging sites, because tagged fish are usually released during the spawning period). The SBP model evolved as a tagging model, and can be implemented without catch data (Vandergoot and Brenden, 2014), but simulations for populations without intermixing suggested that catch-at-age data improved estimates of both mortality rates and abundance (Polacheck *et al.*, 2006). The ITCAAN model treats tagging data as an additional data sources within a statistical catch-at-age assessment model (Goethel *et al.*, 2011, 2015b). The major apparent difference between SBP and ITCAAN models is how they treat the tag return process. While formulated in different ways, we show that the ITCAAN tagging submodel is equivalent to the SBP without catch-at-age data (Appendix 4.A). This is important because it allows us to take advantage of what has been learned about ITCAAN models in the context of an SBP-based approach.

We incorporated the idea of homing probability into a SBP-based approach. The basic idea here is that for fish that live outside the spawning region during the fishing season there is some probability the fish will return to their last spawning ground before the spawning season. This homing probability can take values from zero to 1, representing a continuum of spatial structures ranging from diffusion to overlap. Note that the homing probability for the incomplete overlap spatial structure is not the probability of fish moving back to where they born. Only when homing probability equals 1 do all fish will move back to their natal ground (where they born) during each year spawning season, and this special case is the spatial overlap structure. The homing

probability, instead of the probability of fish moving back to where they born, was considered for the incomplete overlap spatial structure because it seems unrealistic that fish can remember where they born after spawn in different spawning ground(s) for one or more years. We included the capability of penalizing annual recruitment deviations, which is a common practice of catch-at-age models, including ITCAAN models (Goethel *et al.*, 2015b; Vincent *et al.*, 2017), but has rarely been used in SBP models. Chapter 2 suggested that for spatially structured statistical catch at age stock assessment model such a penalty is of critical important for estimating recruitment of terminal assessment years. Our model is built under Bayesian framework so that expert opinions and some empirical or historical used-parameter ranges can be incorporated through prior distributions. The overarching objective for this study was to evaluate for an incomplete overlap structure, how different assumed dynamics and availability of different types of data influenced the estimability of movement, abundance, mortality, and other model parameters. Specifically, we explored how the degree of homing, the extent of spatial movements, which data used (only tag-recovery and effort data, or with additions of tag monitoring or tag monitoring and catch-at-age) influenced estimability of fisheries interested parameters. We also conducted sensitivity analyses to evaluate how data quality and recruitment variation influenced results.

We based our simulation study on six hypothetical populations (Figure 4.1). We based the life history and parameter values on lake whitefish populations in the upper Laurentian Great Lakes. As an ecologically and economically importance native fish species in the Great Lakes, increasing evidence from tagging studies had suggested widespread mixing. However, lake whitefish are still largely managed as unit stock in

each management unit due to the lack of effective ways of understanding movement dynamics. There is increasing interest to estimate the movement dynamics of lake whitefish by analyzing multiple sources of data. The intent of our simulation study is to provide guidance to those charged with managing intermixing populations in terms of how different sources of data can be synthesized to estimate movement dynamics and other parameters relevant to fishery management. Although our study is based on simulations of lake whitefish, we believed that our comprehensive spatial tagging model framework has general applicability to other intermixing fish species. To the best of our knowledge, our model is the first to allow for a continuum of spatial structures through modeling a homing probability, and one of the most general tagging model framework that can be adapted to alternative movement assumptions, and different data availabilities.

Methods

Simulation framework Overview

An operating model was used to simulate the true population and tagging dynamics, and by adding observation errors, to generate observed data that were used as the input data for the tagging models. The operating model used in this study simulated six hypothetical spawning populations from six adjacent management/harvest basins (Figure 4.1): north-east (NE), north-west (NW), central-east (CE), central-west (CW), south-east (SE), and south-west (SW). For each spawning population, the spawning site was inside of one of the harvest basins (Figure 4.1). We define a spawning population as a group of fish that spawned at the same spawning site during the most recent spawning

period (or if they had not previously spawned in their natal site), and a mixed stock as a group of fish that may be composed of fish from multiple spawning populations, but were at risk of harvest in the same harvest basin. In our operating model fish remained in the same harvest basin throughout the fishing season for a given year. Three terms frequently used in this study are: movement rate, stay rate, and homing probability. Movement rate, $\pi_{kk'}$, is the proportion of spawning population k that moves to harvest basin k' right after their spawning period. We assumed such movement only occurs once a year, and is an instantaneous process. Stay rate, π_{kk} , is a special case of movement rate, which corresponded to the proportion of a spawning population k that stays close to their most recent spawning site k , and did not move across the boundaries of the harvest basin which contains that spawning site (Figure 4.1). Homing probability, ξ , is the probability a fish that did not stay close to its spawning population's spawning site during the harvest (non-spawning season) will move back to that spawning site (i.e., return to its beginning of year spawning population, if not already there) before the next year's spawning period. We defined the spawning "period" as the instantaneous spawning instant that occurred at the end of each year.

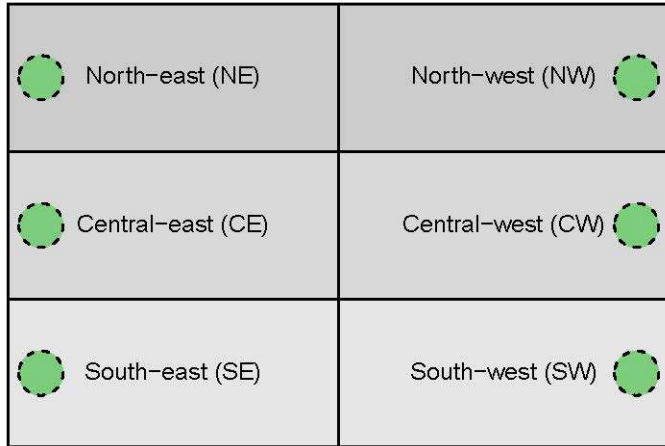


Figure 4.1. Map of the six hypothetical management/harvest basins (i.e., solid lines show the boundaries of each basin). The dotted circle inside of each basin is its spawning site (i.e., where tagged fish being released).

The process of simulation for a single year is as follows. At the start of the year, fish from each spawning population either stayed in the basin that contained their spawning site (i.e., did not move across their management boundaries in Figure 4.1) or moved to other basins depending on the movement rates for that spawning population. The mixed stocks were then exploited in different harvest basins throughout the year. At the end of the year, right before the spawning period, fish that had moved to a harvest basin that did not surround their spawning site at the beginning of the year either joined the spawning population associated with the spawning site contained within that harvest basin, or remained in their original spawning population, depending on the homing probability. Fish then moved to the spawning site associated with their spawning population. For example, if the homing probability was zero, all fish would go to the spawning site within the harvest basin they resided in during the fishing season, illustrating what is known as a diffusion type of spatial structure (Porch *et al.*, 2001). If the homing probability were one, all fish would move back to their original spawning population's spawning site, which would also be where they were born, as individuals

would never change their spawning population. This special case is known as the overlap type of spatial structure (Porch *et al.*, 2001). If the homing probability was between zero and one, some of the fish that had moved away from the harvest basin surrounding their spawning populations spawning site at the beginning of the year would return to the beginning of year (last) spawning site and the rest would join new spawning populations. The probability a fish would remain in its beginning of the year spawning population (i.e., that it will move back to the associated spawning site) is ξ , and the probability of joining a new population (associated with the spawning site within its current harvest basin) and moving to the nearby associated spawning site is $1 - \xi$. Tagging of fish occurs during the spawning season at the spawning sites. At the beginning of next year, all fish became one year older, and the process repeats.

Both the operating and tagging models followed the dynamics of fish from recruitment at age- R , through to an aggregated age class A , consisting of age- A and older fish. Symbols are defined in Table 4.1. All equations of the operating model (except two core equations listed below) are given in Table 4.2, and all equations of the tagging model (except five core equations listed below) are given in Table 4.3.

Table 4.1. Symbols for all equations, except for estimated parameters that are defined in Table 4.4 and data sources that were described in Table 4.5.

Symbol	Description
i	Index variable for tagging year (I = total number of tagging years)
y	Index variable for year varying from tag year and recovery year (Y = total number of years)
a	Index variable for tagging age (A = total number of age groups)
k	Index variable for tagging basin (K = total number of tagging basins)
j	Index variable for recovery year (J = total number of recovery years)
k'	Index variable for recovery basin (K' = total number of recovery basins)
l	Index variable for the component of harvest the recovery belongs to ($l=1$: monitored component of harvest; $l=2$: unmonitored component of harvest)
\sim	Observed values from the operating model
$\hat{}$	Predicted values in the assessment model
N	Abundance at beginning of the year
α	Ricker S-R parameter
β	Ricker S-R parameter
SSB	Spawning stock biomass
σ_R	Standard deviation in recruitment process error
ρ	Autocorrelation coefficient in recruitment process error
τ_R	Innovative standard deviation in recruitment process error
S	Survival rate
F	Instantaneous fishing mortality
Sel	Selectivity
f	Fully selected fishing mortality
η	Selectivity-at-age relationship parameter
τ	Selectivity-at-age relationship parameter
E	Fishing effort
q	Catchability
N'	Intermixed stock in certain basin during the harvest season
ξ	Homing probability
N''	Abundance at the end of the year during the spawning season
w	Weight
mat	Maturity
Fem	Female ratio
σ_F	Observed fishing effort standard deviation
σ_c	Observed harvest standard deviation
p	Tag return probability
ω	survival-movement probability
U	Exploitation rate
δ	conditional probability that a fish caught was harvested by the monitored component of harvest
Const.	constant value in likelihood equation that does not depend on the parameters

Table 4.2. Equations for the population dynamics in the operating model.

Equation Number	Equation name	Equation
4.2.1	Ricker stock-recruit relationship	$N_{k,y,a=R}$ $= \alpha_k SSB_{k,y-R} e^{-\beta_k SSB_{k,y-R}} e^{\varepsilon_{Rk,y}}$ $\alpha_k = \alpha_k' e^{-0.5\sigma_R^2}$ $\varepsilon_{Rk,y} = \rho \times \varepsilon_{Rk,y-1} + \tau_{Rk,y}$ $\tau_{Rk,y} \sim Normal(0, \sigma_R^2)$ $\sigma^2 = \frac{\sigma_R^2}{1 - \rho^2}$
4.2.2	Survival rate	$S_{k',y,a} = \exp[-(F_{k',y,a} + M)]$
4.2.3	Fishing mortality at age	$F_{k',y,a} = Sel_a f_{k',y}$
4.2.4	Selectivity at age	$Sel_a = \frac{a^\eta \exp(-\tau a)}{10^\eta \exp(-\tau 10)}$
4.2.5	Fishing effort and catchability	$E_{k',y} = f_{k',y}/q$
4.2.6	Intermixed abundance	$N'_{k',y,a} = \sum_{k=1}^K N_{k,y,a} \pi_{kk'}$
4.2.7	Catch at age	$C_{k',y,a} = N'_{k',y,a} U_{k',y,a}$
4.2.8	Exploitation rate	$U_{k',y,a} = \frac{F_{k',y,a}}{M + F_{k',y,a}} (1 - S_{k',y,a})$
4.2.9	Spawning stock biomass	$SSB_{k,y} = \sum_{a=R}^A N''_{k,y,a} w_a mat_a Fem$
4.2.10	Observed fishing effort	$\tilde{E}_{k',y} = E_{k',y} \exp(\mu_{k',y} - 0.5\sigma_F^2)$ $\mu_{k',y} \sim Normal(0, \sigma_F^2)$
4.2.11	Observed harvest	$C_{k',y} = \sum_{a=3}^A C_{k',y,a}$ $\tilde{C}_{k',y} = C_{k',y} \exp(v_y - 0.5\tilde{\sigma}_c^2)$ $v_y \sim Normal(0, \tilde{\sigma}_c^2)$

Table 4.3. Equations for the tagging model.

Equation Number	Equation name	Equation
4.3.1	Tag-recovery likelihood	$L_{tag} = P[R_{i,a,k,j,k',l}]$ $= P[R_{i,a,k,j,k'}]P[R_{i,a,k,j,k',l} R_{i,a,k,j,k'}] = L_1L_2$
4.3.2	Brownie likelihood	$L_1 = P[R_{i,a,k,j,k'}]$ $= \prod_{i=1}^I \prod_{a=1}^A \prod_{k=1}^K \binom{T_{i,a,k}}{R_{i,a,k}, T_{i,a,k} - R_{i,a,k}}$ $\prod_{j=i}^J \prod_{k'=1}^{K'} [p_{i,a,k,j,k'}^{R_{i,a,k,j,k'}} (1 - p_{i,a,k})^{T_{i,a,k} - R_{i,a,k}}]$ $-\ln(L_1) = \text{Const.} - \sum_{i=1}^I \sum_{a=1}^A \sum_{k=1}^K \left((T_{i,a,k} - R_{i,a,k}) \log(1 - p_{i,a,k}) + \sum_{j=i}^J \sum_{k'=1}^{K'} R_{i,a,k,j,k'} \cdot \log p_{i,a,k,j,k'} \right)$
4.3.3	conditional probability for fish being monitored	$\delta_{j,k'} = \frac{C_{\underline{m}}}{C_{j,k'}}$
4.3.4	fully selected fishing mortality	$f_{y,k'} = \tilde{E}_{k',y} q_y$
4.3.5	Catchability	$q_{y=1} = \Phi \text{ for } y = 1$ $q_y = q_{y-1} e^{\hat{\varepsilon} F_y} \text{ for } 1 < y \leq Y$
4.3.6	Catchability penalty	$-\ln(\ell_{\varepsilon F}) = \text{Const.} + \sum_{y>1} \ln(\hat{\sigma}_F) + \frac{(\ln \hat{\varepsilon} F_y)^2}{2\hat{\sigma}_F^2}$
4.3.7	Recruitment	$N_{y,a=R,k} = \widehat{R}\mu_k e^{\varepsilon'_{k,y} - 0.5\sigma'_R{}^2}$
4.3.8	Recruitment penalty	$-\ln(L_R) = \text{Const.} + \sum_k \sum_y \ln(\sigma'_R) + \frac{\varepsilon'_{k,y}{}^2}{2\sigma'_R{}^2}$

Operating model

Age-R fish were generated for each spawning population according to a Ricker stock-recruit relationship based on the spawning stock biomass for that spawning population R years prior to the year of recruitment (Equation 4.2.1). Process error for the Ricker stock-recruitment function was temporally autocorrelated, and process error parameters were varied depending on the simulation scenario (see simulation scenarios).

The abundance at age for different populations at the beginning of each simulation were set by assuming that all populations were at equilibrium without considering movement. Abundance at age, for ages $a > R$, for population k at beginning of year y (before annual movement) were forward projected based on an exponential population model that is composed of two components. The first component inside the summation (for $1 < a < A$) represents fish that were in spawning population at the start of the previous year, and the second represent fish that join population k from other spawning populations (i.e., straying):

$$N_{k,y,a} = \begin{cases} \sum_{k'=1}^K \xi N_{k,y-1,a-1} \pi_{kk'} S_{k',y-1,a-1} + (1 - \xi) N_{k',y-1,a-1} \pi_{k'k} S_{k,y-1,a-1} & \text{if } 1 < a < A \\ \sum_{k'=1}^K \xi N_{k,y-1,a-1} \pi_{kk'} S_{k',y-1,a-1} + (1 - \xi) N_{k',y-1,a-1} \pi_{k'k} S_{k,y-1,a-1} + \xi N_{k,y-1,a} \pi_{kk'} S_{x,y-1,a} + (1 - \xi) N_{k',y-1,a} \pi_{k'k} S_{k,y-1,a} & \text{if } a = A \end{cases}$$

(4.1)

In the first component, $N_{k,y-1,a-1} \pi_{kk'} S_{k',y-1,a-1}$ is the age $a - 1$ and year $y - 1$ specific fish from spawning population k ($N_{k,y-1,a-1}$) that moved from their spawning site k to harvest basin k' ($\pi_{kk'}$) at beginning of year $y - 1$, and survived there

throughout the year ($S_{k',y-1,a-1}$). Then the survival components are multiplied by the homing probability (ξ , fish remaining/moving back in its beginning of year spawning site k), and the summation calculates the total survives of spawning population k at the beginning of year $y - 1$ that remained in spawning basin k at the end of year $y - 1$ due to homing. In the second component $N_{k',y-1,a-1}\pi_{k'k}$ is the abundance of each spawning population k' that resides in harvest basin k . The summation calculates the total abundance of fish that resided in harvest basin k , and then multiplies this by the proportion that survive and join the k th spawning population, $(1 - \xi)S_{k,y-1,a-1}$. The equation for age-A fish separately tracks the survival and movements of age-(A-1) fish and age-A fish from the year before and adds them together.

Survival rate $S_{k',y,a}$ was calculated as a function of fishing mortality rate $F_{k',y,a}$ and natural mortality rate M (Equation 4.2.2), M was assumed to be time- and age-invariant, and $F_{k',y,a}$ was modeled as a product of fully selected fishing mortality rate (value is specified in the selection of parameter section) and selectivity at age (Equation 4.2.3). Selectivity at age was modeled by a gamma function (Equation 4.2.4). We assumed all populations had the same selectivity at age. The fishing effort $E_{k,y}$ were calculated as fully selected fishing mortality divided by catchability (Equation 4.2.5), and the catchability q was assumed to be constant across all years and basins (details is given in Appendix 4.B).

The abundance of the mixed stock for each harvest basin at beginning of year y was calculated by Equation 4.2.6. Catch at age was calculated by a Baranov's catch equation (Equation 4.2.7). The exploitation rate was calculated as Equation 4.2.8. Recall

that we assumed spawning occurred at the end of each year right after homing, thus abundance at the time of spawning is:

$$N''_{k,y,a} = \sum_{k'=1}^K \xi N_{k,y,a} \pi_{kk'} S_{k',y,a} + (1 - \xi) N_{k',y,a} \pi_{k'k} S_{k,y,a} \quad (4.2)$$

The spawning stock biomass was calculated based on $N''_{k,y,a}$ and other life history parameters (Equation 4.2.9), which were assumed to be known, as described in the Appendix 4.B.

We assumed the tagged fish follow the same dynamic rules as for the entire population. During the spawning period of each year, we assumed a given number of fish were tagged and released from their spawning sites. We defined a tagging cohort, $T_{i,a,k}$, as a group of tagged fish that were released at the same age (a) from the same tagging year (i) and spawning basin (k). For each tagging cohort, we generated the number of recovered tags during subsequent fishing seasons. The basic tagging recoveries are assumed to be by reporting of tags by fishermen, so not all tags are reported. For each recovery year and basin, we either assumed there was no tag monitoring program and tag reporting rate for all tagged fish were 50%, or modeled a tag monitoring program in which a given number of harvested fish were observed. With a tag monitoring program, we assumed the tag reporting rates for the harvested fish that were monitored as part of the monitoring program ($l = 1$) were 100%. The tag reporting rate for harvested fish that were not monitored as part of the monitoring program ($l = 2$) were 50%, and were time-, basin-, and age- invariant. Thus we assumed that fishers returned 50% of the tags when the harvest was not monitored and that all tags were detected by the monitoring program for the harvest that was monitored.

Simulated data sources included: 1) tag recovery, 2) tag monitoring, 3) fishing effort, and 4) catch-at-age data, and details of each data sources are described below and summarized in Table 4.5.

- 1) Tag-recovery data source include total number of tagged fish (i.e., tagging cohort, $T_{i,a,k}$), and number of recovered fish ($R_{i,a',k,j,k',.}$) from each tagging cohort that were recovered from different years and basins. The total number of tagged fish released in each year and basin ($T_{i,.k}$) was assumed to be time- and basin- invariant, and was varied in different data quality scenarios (see simulation scenarios). The age composition of the tagged fish were generated from multinomial distributions with probability equal to the actual age composition of abundance at time of spawning, $T_{i,a,k} \sim MN(T_{i,.k}, N''_{k,y=i,a} / \sum_a N''_{k,y=i,a})$. The number of recovered fish for each tagging cohort were generated from multinomial distributions with probability equal to the actual tag return probability of each tagging cohort in each recovery year and basin, $R_{i,a',k,j,k',.} \sim MN(T_{i,a,k}, \bar{p}_{i,a,k,j,k',.})$. The actual tag return probability $\bar{p}_{i,a,k,j,k',.}$ is a function of homing probability, movement, survival, and exploitation rates, and unconditional tag reporting rate for both monitored and unmonitored tagged fish with tag monitoring program (or tag reporting rate without tag monitoring program). It was calculated based on the same equations that we used for the tagging model (same as the calculations of $p_{i,a,k,j,k',.}$ in equation 4 in the tagging model section below), but with all parameters at their true values.

- 2) Tag monitoring data were available only when there was a tag monitoring program. It includes number of harvested fish being monitored (C_m), total number in the observed harvest ($\tilde{C}_{k',j}$), and number of recovered fish being monitored in each year and basin ($R_{,,,,j,k',l=1}$). C_m was assumed to be time- and basin- invariant, but was varied in different data quality scenarios (see simulated scenarios). Lognormal observation error was included in generating observed harvest (Equation 4.2.11), and the CV of error term was varied in different data quality scenarios (see simulated scenarios). Number of recovered fish being monitored ($R_{i,a,k,j,k',l=1}$) in each recovery year and basin from each tagging cohort were generated from binomial distributions with probability equal to a recovered fish being monitored in each year and basin, $R_{i,a,k,j,k',l=1} \sim Bin \left(R_{i,a,k,j,k'} \cdot \frac{C_m/C_{k',y}}{(C_m/C_{k',y} + (1 - C_m/C_{k',y})\lambda)} \right)$, where $C_{k',y}$ is the real harvest, and λ is the tag reporting rate for the unmonitored tags (50%). We assumed tagging cohort information was not available from the tag monitoring program and only total number of recoveries that were monitored from each recovery year and basin was recorded as assessment data sources (i.e., $R_{,,,,j,k',l=1} = \sum_{i,a,k} R_{i,a,k,j,k',l=1}$).
- 3) Observed fishing effort data differed from the actual effort as a result of lognormal observation error (Equation 4.2.11).
- 4) Catch-at-age data includes the effective sample size for age compositions (here constant over years), and observed age proportion by year, basin, and age. The observed age composition (proportions at age) of harvest arose from a multinomial distribution with the probabilities equal to the real age

composition of the harvest. The effective sample size for these observed age compositions was varied in different scenarios (see simulated scenarios).

One or more of the data sources were then used as input data for different tagging models with different data availability (Table 4.6). Given that the existence of the tag monitoring program influenced the tag reporting rate assumption, separate simulations were conducted to produce the simulated input data with and without tag monitoring program, albeit using the same random number seeds. We assumed no aging error in all data sets, and tagging-introduced mortality, tag loss, and tag shedding were assumed to not have occurred.

Tagging model framework and its applications to different data availabilities

Our Bayesian tagging model framework has three steps: 1) incorporating prior probability distributions, 2) modeling dynamics and structuring the likelihood, and 3) using posterior simulations to quantifying uncertainty. For the estimated parameters in each tagging model, we assumed experts can only provide boundaries for each estimated parameters based on either previous research, or biological sense. Thus uniform prior probability distributions were incorporated (Step 1, Appendix 4.C). Then we modeled the probabilities of each tagged fish being recovered (Figure 4.2), and/or population dynamics depending on the tagging model. Different tagging models also had different likelihood structures (Step 2, Table 4.6). We then simulated posterior probability distributions based on the prior probability distribution and the likelihood of the data, and summarized MCMC output (posterior distributions) for all estimated parameters and quantities of interest, and their uncertainty.

Table 4.4. Description of estimated parameters (first and second column) and their inclusion in different tagging models (last three columns). The names of the last three columns “STR”, “STRO”, and “Full” indicate three assessment models: spatial tag recovery (STR) model, spatial tag recovery and observer (STRO) model, and the Full model. For each row, “X” indicates that the parameter was estimated in the assessment model; “O” indicates that parameter was assumed to be known without error in the assessment model; and “Flexible” indicates that the parameterization of the parameter was flexible and its estimability was evaluated in the assessment model.

Symbol	Description	STR	STRO	Full
Π	Movement matrix whose (k, k') th element, $\pi_{kk'}$ is the probability that a fish at age a moves from basin k to basin k'	X	X	X
$\Phi; \hat{\sigma}_F; \widehat{\varepsilon F}_y$	Catchability in the first year; standard deviation for catchability random walk; annual catchability deviation for $2 \leq y \leq Y$	X	X	X
η, τ	Selectivity parameters	X	X	X
M	The time-, basin-, and age- invariant instantaneous rates of natural mortality for fish	Flexible	Flexible	Flexible
ξ	Homing probability	Flexible	Flexible	Flexible
$\lambda_{k'}$	The probability that a tag will be reported from a tagged fish caught in recovery basin k' .	O	X	X
$R\mu_k; \varepsilon'_{k,y}$	Average recruitment in tagging site k ; annual recruitment deviation in year y and tagging site k			X
$\Gamma_k; \Delta_{k,a}$	Log scaled average initial abundance for population k ; log scaled age-specific abundance at age $4 \leq a \leq A - 1$ for population k			X

Table 4.5. Description of data sources.

Data source name	Data included	Description
Tag-recovery	T_{iak}	Number of tagged fish released by tagging year, age, and basin
	$R_{ia'kjk'}$	Number of recovered fish by tagging year, age, and basin and recovery year, and basin
Tag monitoring	$R_{\dots,j,k',l=1}$	Number of recovered fish being monitored from tag monitoring program by recovery year and basin
	C_m	Number of harvest being observed by observer on boat by year, and basin
	$C_{jk'}$	Total harvest harvested by year, and basin
Fishing effort	$E_{k',y}$	Fishing effort by year and basin
Catch-at-age	N_{eff}	Effective sample size of age composition
	$\rho_{jk'a'}$	Observed age proportion by year, basin, and age

Table 4.6. Inclusion of data set(s) for different tagging models.

Tagging models	STR	STRO	Full	
Data set(s)	Tag-recovery	X	X	X
	Fishery effort	X	X	X
	Tag monitoring		X	X
	Catch-at-age			X
Likelihood components	Tag-recovery proportion by cohort	X	X	X
	Catchability penalty	X	X	X
	Tag monitoring		X	X
	Total harvest			X
	Initial abundance			X
	Age composition			X
	Recruitment penalty			X

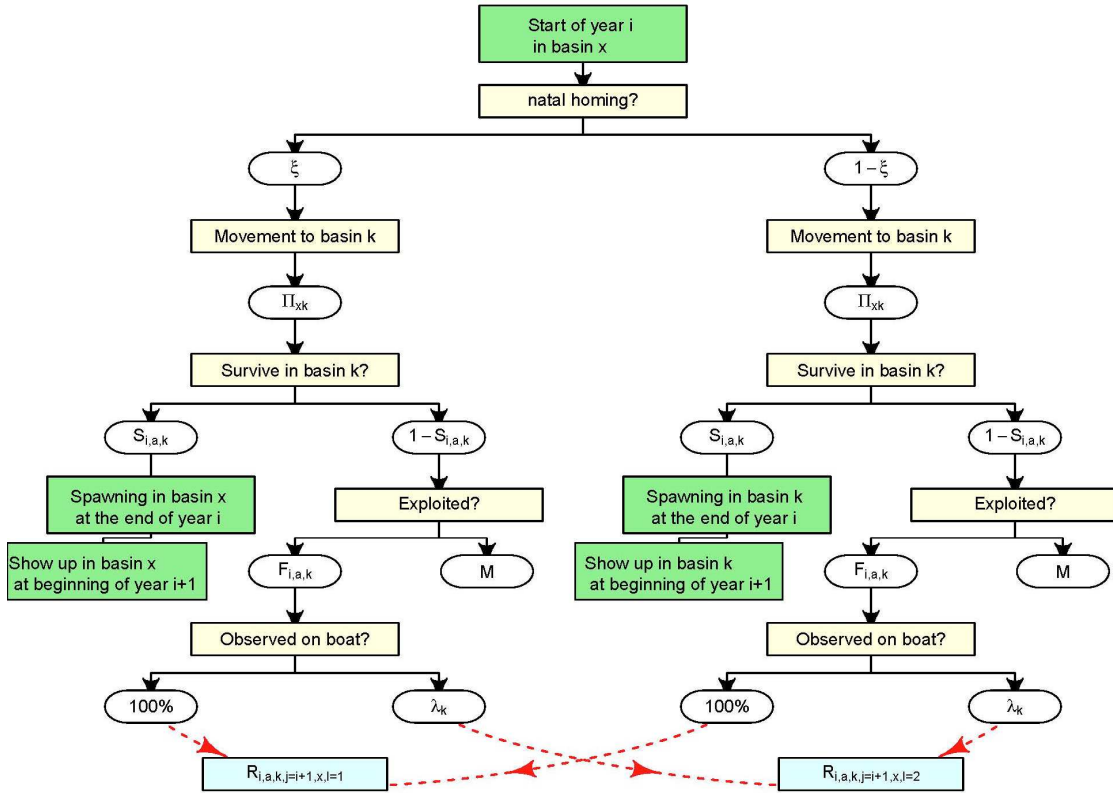


Figure 4.2. Flow chart for of the multi-state tag return probability for the tagged fish from cohort iak (tagged and released in year i at age a from basin k). Each rectangular box represents a state.

Tagging models

In this study we compared three tagging models with different data availability: spatial tag recovery (STR), spatial tag recovery and observer (STRO), and Full models. We assumed that a tag monitoring program was not implemented and tag monitoring data were not available as input data for the STR model, while they were available in both STRO and Full models. The input data for the STR model included tag-recovery and fishing effort data. In the STRO model, tag-recovery, tag monitoring, and fishing effort data were available, and tag-reporting rates were estimated within the tagging model. In the Full model, in addition to same data that were included in the STRO model, catch-at-

age data were also available (Table 4.6). Recall that the existence of a tag monitoring program influenced the tag reporting rate assumption in the operating model, and that the same random number seed was used for each simulation. Thus in each individual simulation, the tag-recovery data for the STR model were the different from those that were generated for the STRO and Full models, while the tag-recovery and tag monitoring data were the same for the STRO and Full models.

For each tagging cohort $T_{i,a,k}$, in the STR model, the tag-recovery likelihood was modeled as a Brownie likelihood (Equation 4.3.2); while in the STRO and Full model, the tag-recovery likelihood was modeled as a product of Brownie likelihood for the combined components fishery (i.e., both monitored and unmonitored harvest) and a conditional likelihood with the probability equal to the a recovered fish being monitored or not, as shown in Equation 4.3.1. The Brownie likelihood component was a product of multinomial probabilities, which was similar to that of Eveson *et al.* (2009), and the conditional likelihood for the number of recovered fish from each recovery year and basin being either monitored or unmonitored was a product of binomials, which was similar to that of Hearn *et al.* (1990).

Tag return probability $p_{i,a,k,j,k'}$ in the Brownie likelihood (Equation 4.3.2) was composed of a survival-movement probability ω , an exploitation rate U , and a tag reporting rate λ (Figure 4.2). Recall that we assumed fish spawned at the end of each year, and movement occurred at the beginning of each year. For the STR model without tag monitoring data, we assumed the tag reporting rates λ were known without error (0.5), and tag return probability is the product of a survival-movement probability ω , an exploitation rate U , and a known tag reporting rate λ . For the other tagging models, the λ

for the unmonitored tagged fish were estimated. For tagged fish being harvested from monitored component of harvest ($l = 1$) in year j from basin k' , an additional conditional probability of whether fish belonging to the monitored component ($\delta_{j,k'}$) were added (Equation 4.3.3), and we assumed their tag reporting rate were 100%; while for those fish being harvested from unmonitored component of harvest ($l = 2$), the conditional probability of whether fish belonging to the unmonitored component were added ($1 - \delta_{j,k'}$), with the tag reporting rate equals $\lambda_{k'}$, given by:

$$p_{i,a,k,j,k',l} = \begin{cases} \omega_{i,a,j,kk'} U_{a',j,k'} \delta_{j,k'} & \text{if } l = 1 \\ \omega_{i,a,j,kk'} U_{a',j,k'} (1 - \delta_{j,k'}) \lambda_{k'} & \text{if } l = 2 \end{cases} \quad (4.3)$$

Thus the unconditional probability of a tag being returned over the monitored and unmonitored components of harvest is:

$$p_{i,a,k,j,k'} = \sum_{l=1}^L p_{i,a,k,j,k',l} = \omega_{i,a,j,kk'} U_{a',j,k'} [\delta_{j,k'} + (1 - \delta_{j,k'}) \lambda_{k'}] \quad (4.4)$$

Although equation 4.4 was initially designed for STRO and Full models, it was also applicable to the STR model without tag monitoring data, because when no harvest was monitored, $\delta_{j,k'} = 0$.

Exploitation rate U , fishing mortality, and selectivity at age in the tagging models were modeled in the same way as in the operating model (Equations 2.8, 2.3 and 2.4), but the selectivity at age parameters ($\hat{\tau}$ and $\hat{\eta}$) were estimated parameters in the tagging models (see estimated parameters section below and Table 4.4).

The fully selected fishing mortality was modeled as the product of observed effort ($\tilde{E}_{y,k'}$), and catchability q_y (Equation 4.3.4). The q_y modeled using a random walk

process (Equation 4.3.5). Note that we incorporated a model misspecification here: we assumed constant catchability and lognormal observation error in observing fishing effort data in the operating model, while we used a random walk process to model catchability, and ignored observation error in fishing effort data in the assessment model.

We used matrix calculations to model survival-movement probability $\omega_{i,a,j,kk'}$.

The survival-movement probability matrix $[\Omega_{i,a,j}]$ is a K -by- K matrix, which was modeled as:

$$[\Omega_{i,a,j}] = \begin{cases} [\Pi] & j = i + 1 \\ \prod_{y=i+1}^{j-1} [H_{y,a'(y)}][\Pi] & j > i + 1 \end{cases} \quad (4.5)$$

where $[\Pi]$ was the movement rates matrix (Table 4.4). The subscript $a'(y) = \min(a + y - i, A)$, was the fish of age in year y . The annual survival matrix $H_{y,a'(y)}$ was a K -by- K matrix for $y = i + 1, \dots, j - 1$. The annual survival rates in year i were not considered here because tagged fish were released at the end of each year in our simulation.

The annual survival matrix $[H_{y,a}]$ modeled the tagging dynamics of the tagged fish, and the population dynamics of the whole population (States 2 and 3 in Figure 4.2), whose (kk') th element, $H_{y,a,kk'}$ is the probability of an individual at age a that alive and show up in the spawning site k at the start of the year y survives to the end of that year and be present in the spawning site k' by the end of that year during the spawning period.

To better explain $[H_{y,a}]$, we introduce an annual movement-survival split matrix $[O]$, to model fish annual movement and survival before homing. For simplification, we

drop subscripts for year and age in this in this explanation. We assumed fish movement occurred at beginning of year y , and they can move to any basin (varying from 1 to K) depending on their movement rates matrix $[\Pi]$ (State 1 in Figure 4.2). Only if they can survive in their destination basin can they enter into a spawning population in the next year (Figure 4.2). The annual movement-survival split matrix $[O]$ was presented by:

$$[O] = \Pi D([S]) = \begin{bmatrix} \pi_{11}S_1 & \pi_{12}S_2 & \cdots & \pi_{1K}S_K \\ \pi_{21}S_1 & \pi_{22}S_2 & \cdots & \pi_{2K}S_K \\ \cdots & \cdots & \cdots & \cdots \\ \pi_{K1}S_1 & \pi_{K2}S_2 & \cdots & \pi_{KK}S_K \end{bmatrix} \quad (4.6)$$

The operator $D([S])$ was an operator that transformed a vector $[S]$ (i.e., survival rates of all basins) with length K into a $K \times K$ diagonal matrix containing the elements of $[S]$ (i.e., S_1, S_2, \dots, S_K). Then we used homing probability ξ as a scalar for that matrix so that the matrix $[O]$ was split into a homing component $\xi[O]$ and a diffusion component $(1 - \xi)[O]$. For the homing component, the annual survival rates before the spawning period were the row summation of the scaled split matrix (i.e., $\xi[\Pi D([S])\mathbf{1}]$, a K by 1 column vector whose k th element was the annual survival rate for spawning population k . The K by 1 column vector $\mathbf{1} = (1, \dots, 1)^T$. Recall that regardless of which basin fish moved to, their spawning population identity can always be maintained with homing. For the diffusion component of annual movement-survival matrix $(1 - \xi)[O]$, fish join the spawning population of their destination basin right before the spawning period and their spawning population identity changed correspondingly. Thus their scaled split matrix can be maintained, and be used directly for the population dynamics of the next year (i.e., $(1 - \xi)[\Pi D([S])]$, a K by K matrix).

By adding year (y) and age (a) index, the annual survival matrix $H_{y,a}$ can be presented by: $[H_{y,a}] = D\{\xi[\Pi D([S_{y,a}])\mathbf{1}]\} + (1 - \xi)[\Pi D([S_{y,a}])]$. The k th element of $[S_{y,a}]$, $S_{k,y,a}$, is the survival rate for age- a fish in basin k and year y . The survival and capture probabilities $S_{k,y,a}$ in the tagging model were modeled in the same way as in the operating model (Equation 4.2.2).

The conditional likelihood for a recovered fish being monitored or not, L_2 (Equation 4.3.1), was only included in STRO and Full models (Table 4.6). This follows a similar approach to that used by Hearn et al. (1990), and additional modeling details on L_2 are in the Appendix 4.D.

Lognormal distributions were assumed for the the annual catchability deviations, leading to a penalty that constrained those deviations (Equation 4.3.6). This penalty was included in all our tagging models (Table 4.6).

The Full tagging model is distinguished from the STR and STRO models in that it attempts to model the full population abundance at age, rather than just mortality and movement. This is based on additional data and requires estimation of additional parameters. There are four additional log-likelihood components for the Full tagging model (Table 4.6), for total harvest, initial abundance penalty, age composition, and recruitment penalty. Recruitment and the initial abundance at age in the population need to be estimated (see estimated parameters). Recruitment was parameterized as the product of average recruitment ($\widehat{R\mu}_k$) multiplied by an annual residual ($\varepsilon'_{k,y}$) that was exponentiated and bias corrected. The annual residuals contributed to the additional

likelihood component for recruitment (Equation 4.3.8), which effectively assumes that annual recruitment varied about an average according to a lognormal distribution.

Post-recruit abundances at age in the first assessment year were modeled as:

$$N_{j=1,a,k} = e^{\Gamma_k + \Delta_{k,a}} \text{ for } a > 3 \quad (4.7)$$

where for each k , $\sum_{a=4}^A \Delta_{k,a} = 0$. Thus, for each population k , only $A - 1$ parameters for Δ were estimated, and we assumed $\Delta_{k,A} = -\sum_{a=4}^{A-1} \Delta_{k,a}$ (see estimated parameters). The population dynamics in the tagging models followed the same dynamics as in the operating model.

Catch at age was calculated by Baranov's catch equation (same as Equation 4.2.7). The negative log-likelihood component for the log of area-specific annual fishery harvest, the log-likelihood component for the harvest age composition, and log-likelihood penalty for the initial abundance-at-age white noise deviations followed the general approaches that were widely used in statistical catch-at-age models. Modeling details are described in Appendix 4.E.

The last likelihood component, the log-penalty that constrained the annual recruitment deviations ε'_y , treats those deviations as following a normal distribution with standard deviation σ'_R assumed to be 2.0 (Equation 4.3.8), which was intended to provide only a weak restriction on parameter estimates.

Estimated parameters

Estimated parameters in each assessment model are described in Table 4.4. In summary, movement rates, catchability in the first year, standard deviation for catchability random walk, annual catchability deviation, and selectivity parameters were

estimated in all three models. The movement rates for each population were modeled using a logit transformation, which constrained movement rates to be between 0 and 1, and the sum equals 1. Tag reporting rates for the unmonitored component of harvest were only estimated in the STRO and Full models, and average recruitment for each tagging site, annual recruitment deviation, initial abundance for each spawning population, and initial abundance at post-recruit ages for each spawning population were only estimated in the Full model. Natural mortality and homing probability were either assumed to be known or estimated in three assessment models under different parameterization scenarios (see simulation scenarios). Note that for a population that is tagged in I consecutive years, only $I-1$ natural mortality rate parameters (per basin) can be estimated, given natural mortality rates is estimated based on the difference between the expected returns at age $a+I$ of fish released at age a and those released at age $a+1$. Thus for I consecutive release years, estimates can only be obtained for $I-1$ of the natural mortality parameters.

Selection of parameters for the operating model

In general, the parameters of the operating model were based on estimated parameters for lake whitefish populations in the Upper Great Lakes regions. To evaluate the performance of the framework more generally we varied some parameters from what are likely for lake whitefish. For example, based on previous tagging studies, lake whitefish populations in Lake Michigan and Lake Huron demonstrate a high degree of homing (Ebener et al. 2010, Stott et al. 2010), but we evaluated different levels of homing in our simulation scenarios. We assumed high productivity for the spawning populations of the north basins, medium productivity for those of the central basins, and

low productivity for those of the south basins (Figure 4.2). Populations at different productivity levels had the same unfished spawning stock size ($SSB_0 = 519.70$ t), but different steepness values (h). Plausible steepness levels were chosen to represent low, medium, and high levels of population productivity for lake whitefish populations in Upper Great Lakes regions (Table 4.7). We converted the steepness parameterization of Ricker stock-recruit function to the standard form (Appendix F, Table 4.A2). We assumed that the recruitment occurred at Age-3 fish, and the last age group was Age-12 and older fish. We used the Ricker stock-recruit function because of evidence of overcompensation in lake whitefish recruitment (Healey, 1978; Henderson *et al.*, 1983). The values of life history parameters (i.e., weight- and maturity-at-age, female ratio, selectivity parameters and fully selected age), and steepness values for different productivity levels are the same as used by Li *et al.* (2017) (Appendix 4.B, Table 4.7). Stochastic modeling parameters (i.e., process errors, observation errors, and effective sample sizes) are defined in Table 4.7 and Table 4.8. In the operating model, we assumed a natural mortality rate of 0.25, and the fully selected fishing mortality rate was the same as its target level based on the actual harvest policy of lake whitefish populations in much of the Upper Great Lakes, that strives to limit total annual target mortality rate to 0.65 (as a maximum over ages).

Table 4.7. Three levels of population productivity we assumed for six hypothetical management/harvest basins (Figure 4.1, high for NE, NW; medium for CE, CW, low for SE, SW), and two levels of recruitment variation we assumed in the baseline and sensitivity scenarios.

Productivity levels	
low	Steepness $h=0.7$
Medium	Steepness $h = 1.3668$
High	Steepness $h = 1.9$
Recruitment variation	
baseline	Autocorrelation coefficient $\rho=0.45$ Innovative standard dev in rec process error $\sigma_R=0.78$ Stationary standard dev. in rec process error $\sigma^2=0.8734$
High	Autocorrelation coefficient $\rho = 0.45$ Innovative standard dev in rec process error $\sigma_R=0.3395$ Stationary standard dev. in rec process error $\sigma^2=1.5$

Table 4.8. Different levels of data quality.

Data quality levels	Low	Baseline	High
Number of years of tagging and releasing (I and J)	5	10	15
Number of fish being released in each year and basin	1000	1500	2000
Number of fish being observed in each year and basin	5000	10000	20000
Effective sample size of sampling age composition	100	200	400
Observation errors of observing harvest	CV=0.4	CV=0.15	CV=0.1
Observation errors of observing effort	CV=0.8	CV=0.3	CV=0.2

Simulation scenarios

Baseline scenarios

The baseline scenarios were used to compare the relative performance of the three tagging models that made use of different data (i.e., STR, STRO, and Full models, Table 4.6). Baseline scenarios are composed by four scenarios with alternative parameterization of natural mortality rate and homing probability: both assumed to be known ($fix_{Xi} + fix_M$); each estimated with the other known ($est_{Xi} + fix_M$; $fix_{Xi} + est_M$); and both estimated ($est_{Xi} + est_M$). In all baseline scenarios, we assumed the spatial structure for all populations was incomplete overlap, and the homing probability in the operating model was 50% (Table 4.9). We assumed the natural mortality rate and homing probability in all tagging models were time-, basin-, and age- invariant, as they were in the operating model. The purpose of the baseline scenarios was to evaluate how additional data influenced the ability to estimate parameters common to all the models. In addition we were specifically interested if adding reporting data allowed for estimation of reporting rate and the degree to which data availability influence the ability to estimate natural mortality rate and homing probability. We assumed movement rate, recruitment variation, and data quality were all at their baseline levels (Tables 4.7 and 4.8).

Table 4.9. Different levels of homing probability and movement assumptions we evaluated in a four (homing probability) by five (movement assumption) factorial experiment.

Homing probability	
No (diffusion)	0
Baseline (incomplete overlap)	50%
Complete (overlap)	100%
Movement assumptions	
Low	Stay rate = 90% for all six basins
Baseline	Stay rate = 60% for all six basins
High	Stay rate = 30% for all six basins
Unequal mixing with positive correlation between stay rate and productivity levels	Stay rate = 90% for high productivity populations (NE, NW)
	Stay rate = 60% for medium productivity populations (ME, MW)
	Stay rate = 30% for low productivity populations (SE, SW)
Unequal mixing with negative correlation between stay rate and productivity levels	Stay rate = 30% for high productivity populations (NE, NW)
	Stay rate = 60% for medium productivity populations (ME, MW)
	Stay rate = 90% for low productivity populations (SE, SW)

Alternative homing and movement scenarios

In these scenarios, we evaluated the influence of the homing probability, and movement intensity on the performance of the Full model. We evaluated the performance of the Full estimation model when the operating model parameters were changed, considering two additional levels for homing probability with the baseline movement rates, and four additional levels of movement rates with the baseline homing probability (Total of six scenarios, Table 4.9). For these scenarios, the natural mortality rate was estimated and the homing probability was assumed be to known in the estimation model. Other than the specific aspects mentioned here other attributes of the operating and estimation model were the same as for the baseline scenarios.

Recruitment variation scenario

In this scenario, we evaluated the model performance of the full model when recruitment variation was higher in the operating model than was assumed for the baseline scenarios (Table 4.9). Natural mortality rate and homing probability were estimated, and all other details of the operating and estimation model were the same as the baseline scenarios.

Data quality scenarios

We evaluated the performance of the full model with different levels of data quality (low, high), as shown in Table 4.7. The levels of data quality were determined by the number of years of tag and recovery, total number of tagged fish released each year, number of observed fish each year, effective sample size for sampling age compositions, and variance for observation errors for observed harvest and effort (Table 4.7). Natural mortality rate and homing probability were estimated and all other details were the same as in the baseline scenarios.

Performance metrics

Under each scenario, tagging model(s) were fit to 200 simulated data sets from the operating model. We investigated two statistical criteria, bias and coverage probability, of estimating movement, natural mortality, fully selected fishing mortality rates, selectivity at age, reporting rates, annual recruitment, SSB, and homing probability. Relative error (RE) was used to represent the bias and uncertainty in estimating parameters, which was calculated as the relative difference between the parameter posterior median \bar{X} from its true value X : $RE = (\bar{X} - X)/X$. Box plots were used to

visualize the distribution of relative error for all simulations. The coverage probability was calculated as the number of simulation runs, in which the true value from the operating model was located inside the two-sided 95% credible interval divided by the total number of simulation runs. The convergence of the MCMC chain was evaluated based on the potential scale reduction factor with a value close to 1 indicating convergence for multiple chains with distinct initial parameters (Brooks and Gelman, 1998).

Results

In all scenarios, NE, CE, and SE populations had the same productivity, movement, and homing probabilities as for NW, CW, and SW populations, their expected assessment performance were the same, and indeed the realized performance results were nearly identical. We thus only summarized the results for the NE, CE, and SE populations.

Baseline scenarios

The parameterizations of homing probability and natural mortality rate for all three assessment models influenced estimation of the movement rates (Figure 4.3), tag reporting rate, homing probability, and natural mortality rates (Figure 4.4), but did not have noticeable impact on the estimates of selectivity at age, fully selectivity fishing mortality rates, SSB, and recruitment. Populations with different productivity levels did have different assessment performance, in terms of estimating SSB and recruitment for the Full model, but the estimation of movement rates and fishing mortality rates was not markedly influenced by productivity for all three models. Thus, we only summarized

results of estimating movement rates of the medium productivity population (population in CE basin) moving from their spawning site to all harvest basins, fully selected fishing mortality in CE basin, and results of estimating selectivity at age, SSB, and recruitment under one baseline parameterization scenario ($fix_{Xi} + est_M$, fixed homing probability and estimated natural mortality) for each assessment model.

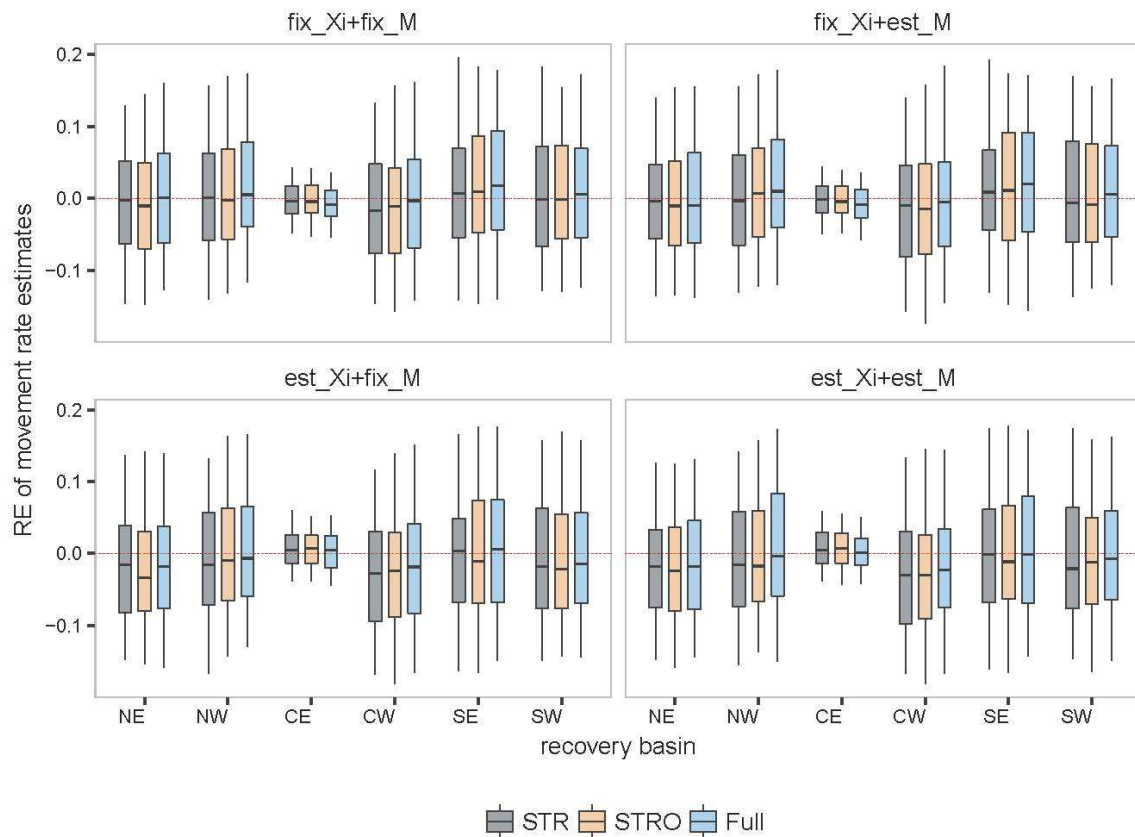


Figure 4.3. Relative error of movement probability estimates for CE spawning population (medium productivity population) moving from their spawning site to all six recovery basins (the x-axis) under the baseline scenarios with alternative parameterizations of the natural mortality rate and homing probability in the assessment models (panel name). The triplet in each panel are the results for each assessment model. Whiskers, boxes, and horizontal middle line are 95% and 50% inter quantile ranges, and median of 200 simulations.

In general, the three assessment models led to similar assessment performance in estimating movement rates (Figure 4.3). For all three models, the uncertainty (i.e.,

interquartile ranges, IQRs) in estimating movement rates to other harvest basins was higher than for using the harvest basin surrounding their last spawning site (stay rates). For all tagging models, when homing probability was estimated, there was a slight positive bias in estimating stay rates and negative bias in estimating movement rates to other basins, while such biases were largely absent when homing probability was assumed to be known.

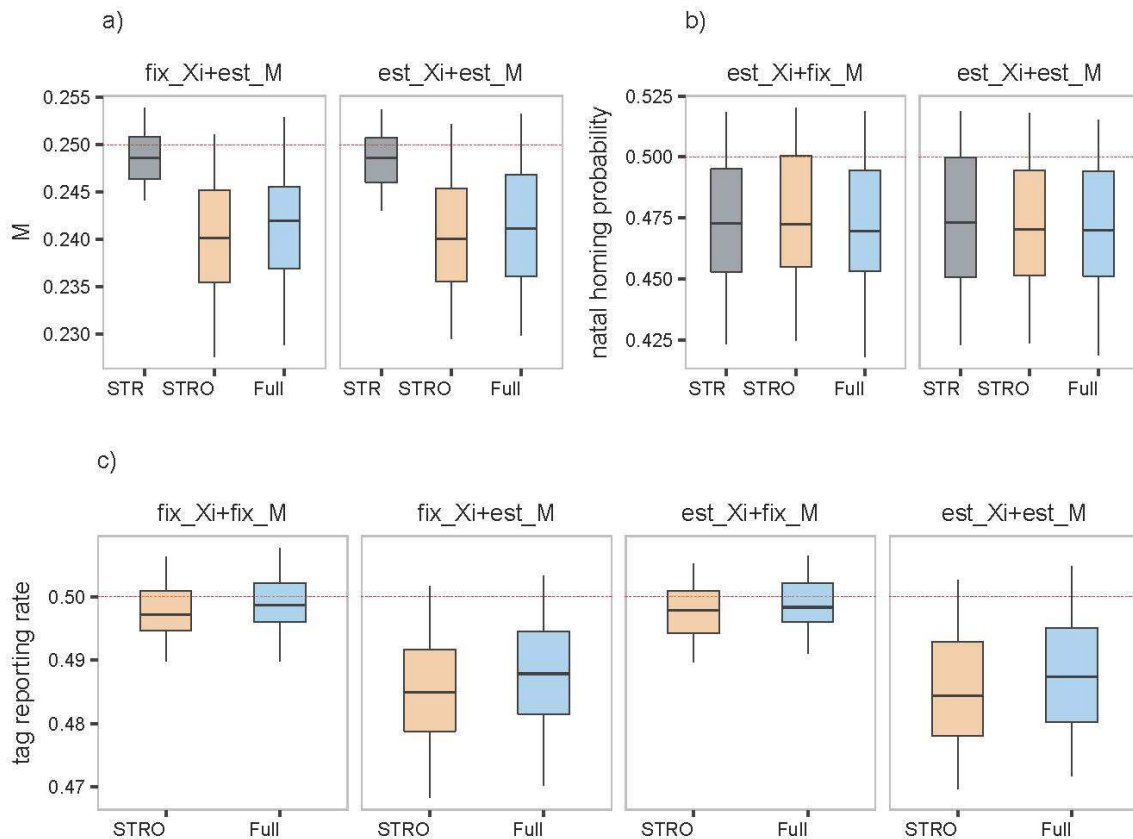


Figure 4.4. The natural mortality rate (a), homing probability (b), and tag reporting rate estimates under the baseline scenarios with alternative parameterizations of natural mortality rate and homing probability (panel name). The dashed line in (a) is the true value that was assumed in the operating model. The x-axis in each panel are the assessment models. Whiskers, boxes, and horizontal middle line are 95% and 50% interquartile ranges, and median of 200 simulations.

Compared to the STRO and Full models in which tag reporting rate was estimated, the STR model, which assumed tag reporting was known without error had the

best assessment performance in estimating natural mortality in terms of smaller bias and less uncertainty (Figure 4.4a). All three models had similar negative bias in estimating homing probability, but such bias is quite small (Figure 4.4b, true value at 0.5 versus predicted median at 0.475). The Full model resulted in slightly less median biased median estimate of tag reporting rate than the STRO model (Figure 4.4c). When natural mortality was estimated, the magnitude of negative bias in estimating tag reporting rates was greater than when natural mortality was known, for both models for which this was estimated (Figure 4.4c). All three assessment models had slightly positively biased estimates of f , with the greatest bias for the STRO model (Figure 4.5a). The Full model had the best assessment performance in estimating selectivity at age, i.e., the smallest bias and less uncertainty (Figure 4.5a). Recruitment and spawning stock biomass (SSB) were only estimated in the Full model. For high productivity population (i.e., pop of NE), the estimates of recruitment were nearly median-unbiased during the early recovery years, and were negatively biased with great uncertainty during the last two years (Figure 4.6a). For populations with medium and low productivity (pops of CE and SE), the estimates of recruitment were negatively biased across all recovery years, and uncertainty were increased for later recovery years. The bias of estimating recruitment were greater for low productivity population than for medium productivity population. For all populations, less biased estimates of SSB with lower uncertainty were obtained in the later years (Figure 4.6b). During the first five recovery years, the Full model tended to underestimate SSB for populations with low productivity, to overestimate SSB for populations with high productivity, and provided nearly unbiased estimates of SSB for populations with medium productivity.

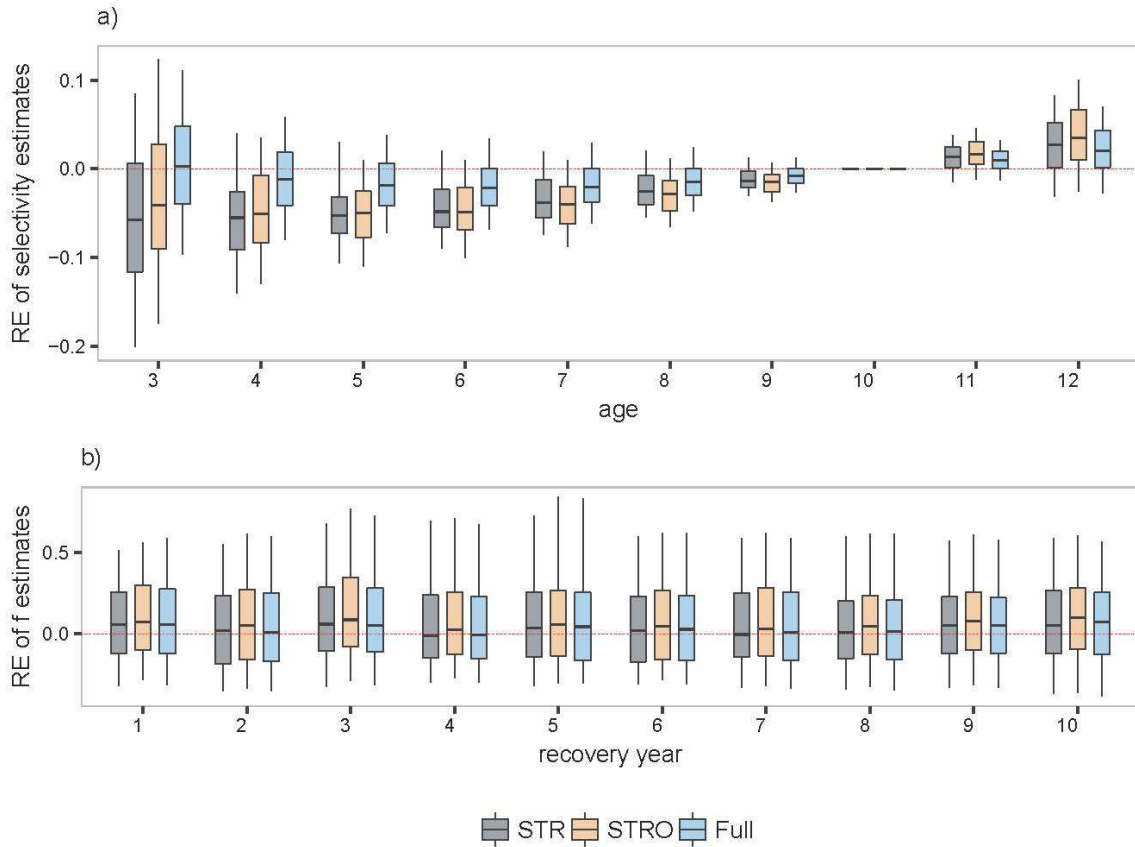


Figure 4.5. Relative error of selectivity at age, and fully selected fishing mortality estimates in CE basin under the baseline scenario in which natural mortality rate was estimated and homing probability was assumed to be known in the assessment models. The triplet in each panel are the results for each assessment model. Boxes, and horizontal middle line are 50% inter quantile ranges, and median of 200 simulations.

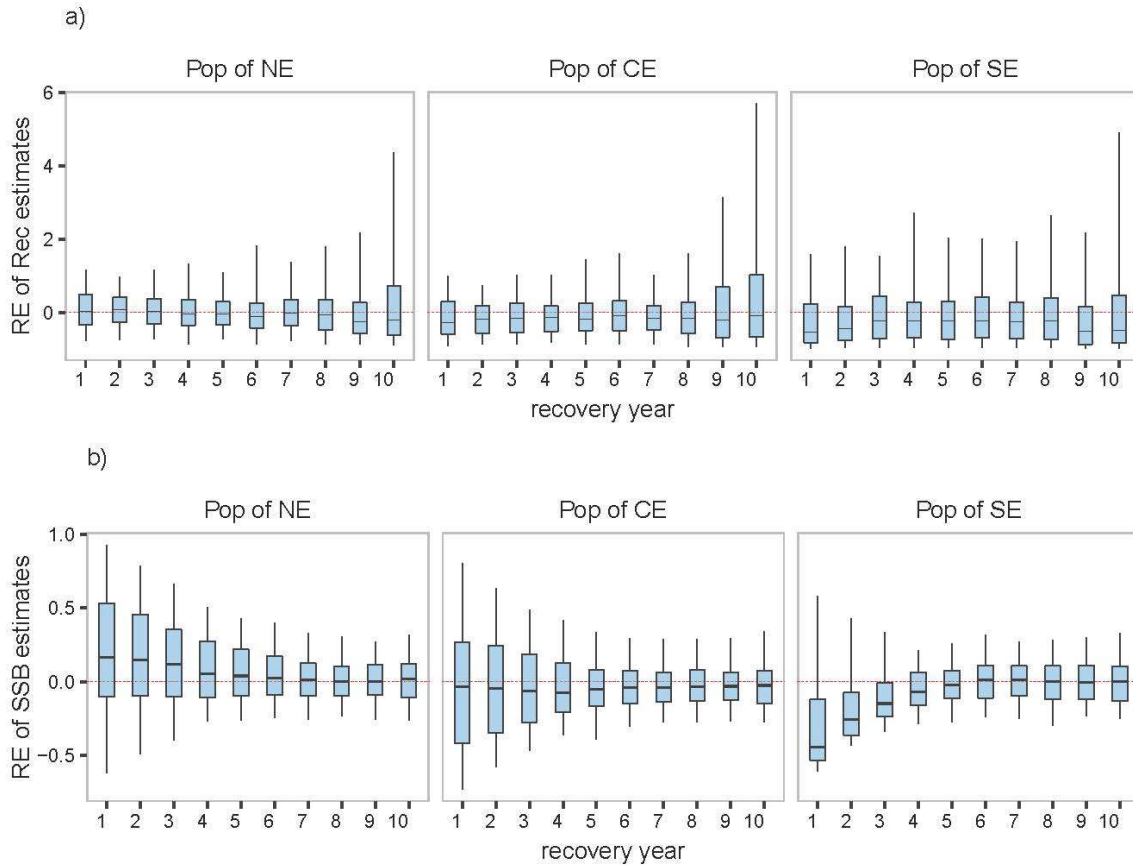


Figure 4.6. Relative error of estimating recruitment (a) and SSB (b) in the Full model under the baseline scenario in which natural mortality rate was estimated and homing probability was assumed to be known in the assessment models. Whiskers, boxes, and horizontal middle line are 95% and 50% inter quantile ranges, and median of 200 simulations.

For all three models, the coverage probabilities (covering true value in 95% credible intervals) in general were consistently high for movement rates (median ~ 0.8), recruitment and SSB (median ~ 0.66 , and ~ 0.53), and were consistently low for fishing mortality (median ~ 0.15), and selectivity at age (median ~ 0.3) (Figure 4.7). The STR model had the highest coverage probability for M (~ 0.59), while such coverage probabilities of STRO and Full were both around 0.2. The Full model had lowest coverage probability for the homing probability (~ 0.43), while this coverage probability in the STR and STRO were above 0.47. The Full model had the greatest coverage

probability for the selectivity at age among all three models (~0.37). The coverage probability of all parameters were consistent across alternative parameterization scenarios, except for the tag reporting rate estimates in the STRO and Full models, the homing probability estimate in the STR model, and the movement rate estimates in all three models. Regardless of whether homing probability (ξ) was estimated in the assessment model, the coverage probability of tag reporting rate were around 0.78 when M was fixed at its true value, and were only around 0.32 when M was estimated in the STRO and Full models. For the STR model, the coverage probability of homing probability increased from 0.49 to 0.54 when M was fixed at its true value. For all three models, the coverage probability for movement rate was increased by 2% when ξ was fixed at its true value.

Alternative homing and movement assumptions

Compared to the baseline scenario in which natal the homing probability was assumed to be 50%, the overlap spatial structure assumed 100% homing, and the diffusion spatial structured assumed no homing, in both operating and assessment models. As in the baseline scenario, higher uncertainty in estimating recruitment was observed during the later recovery years, in terms of the IQRs of RE (Figure 4.8). The Full model with an overlap spatial structure resulted in less biased estimates of recruitment with lower uncertainty for the populations with medium and high productivity than with low productivity (Figure 4.8). Bias in recruitment estimates was not obviously influenced by productivity, however, for the diffusion spatial structure.

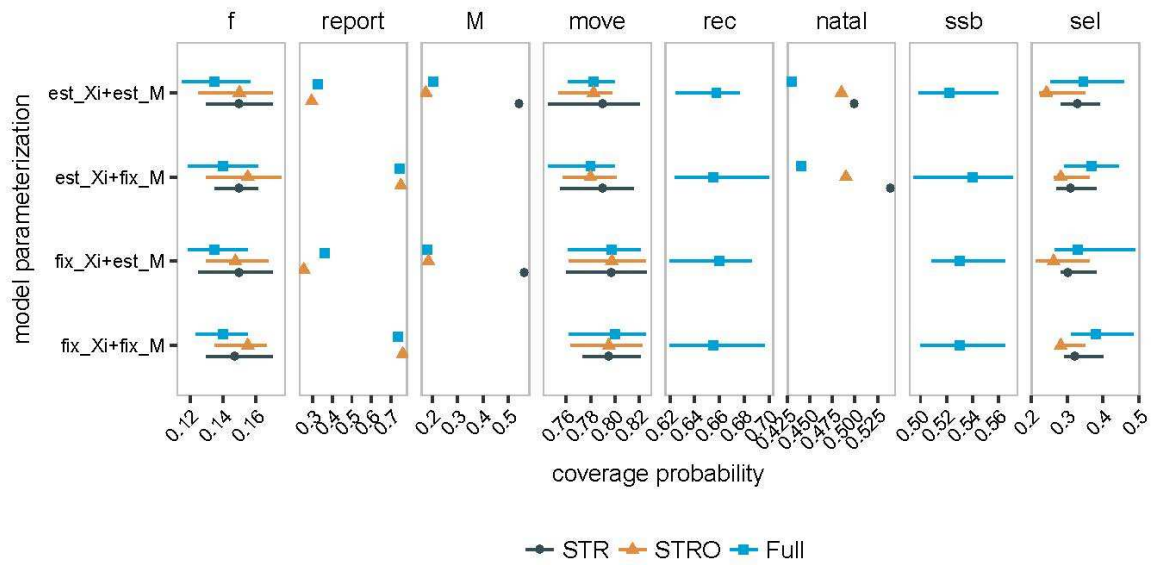


Figure 4.7. The coverage probability of covering true value from the operating model in the 95% Bayesian credibility interval under the baseline scenarios with alternative parameterization of natural mortality rate and homing rate (row name). Each combination of shape and color represent results of an assessment model. Each panel represents an estimated parameter (group): f, fully selected fishing mortality rate; report: tag reporting rate; M, natural mortality rate; move: movement rate; rec, recruitment; natal, homing rate; ssb, spawning stock biomass; sel, selectivity at age. For some parameters that are varying by year, basin, and/or age, the coverage probability of each dimension/element was calculated. The middle point represents the median value, and the error bar represents the 50% interquartile ranges across all elements/dimensions.

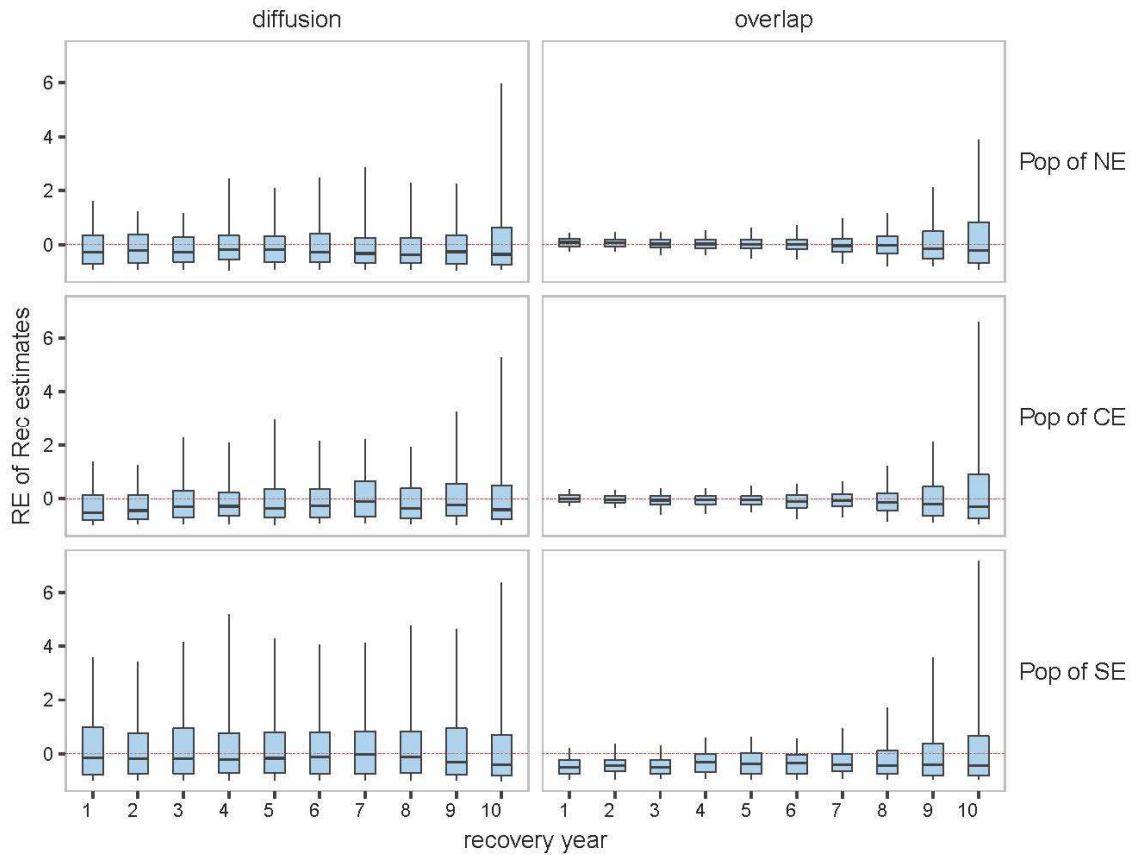


Figure 4.8. Relative error of recruitment estimates for the Full model under the scenarios with alternative homing assumptions. Each column represents a homing scenario, and each row represents a population. Whiskers, boxes, and horizontal middle line are 95% and 50% inter quantile ranges, and median of 200 simulations.

For the diffusion spatial structure, the Full model tended to provide negatively biased estimates of SSB with high uncertainty for early assessment years, and such bias was greater for population with lower productivity level (Figure 4.9). For the overlap spatial structure, the Full model tended to overestimate SSB for populations with high productivity, and to underestimate SSB for populations with low productivity, and the magnitude of these biases were greater during the earlier recovery years. The bias in estimating movement rates were slightly higher when spatial structure was overlap than when it was diffusion (Figure 4.10). With the overlap spatial structure, the Full model tended to overestimate movement rates into regions where low productivity populations

were located from the high and medium productivity populations, and to underestimate the movement rates into regions where high and medium productivity populations were.

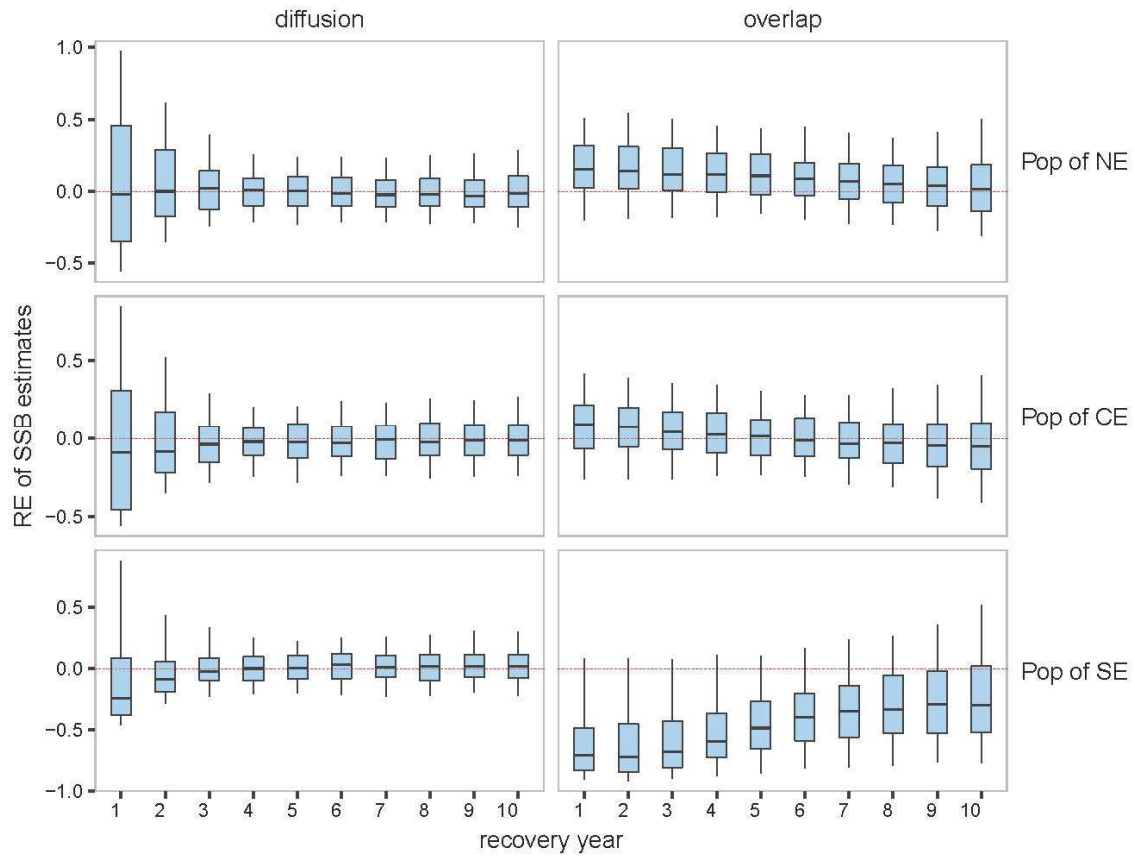


Figure 4.9. Relative error of SSB estimates for the Full model under the scenarios with alternative homing assumptions. Each column represents a homing scenario, and each row represents a population. Whiskers, boxes, and horizontal middle line are 95% and 50% inter quantile ranges, and median of 200 simulations.

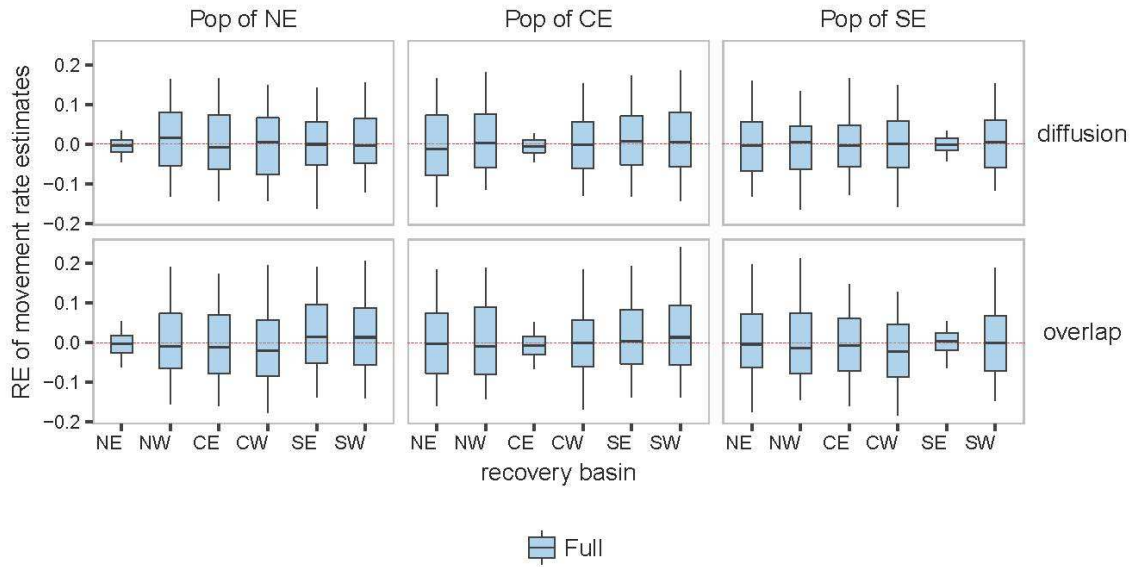


Figure 4.10. Relative error of movement probability estimates for spawning populations of NE, CE, and SE basins (panel name) that moving from their spawning site to all six recovery basins (the x-axis) from the Full model under the scenarios with homing assumptions (row name). Whiskers, boxes, and horizontal middle line are 95% and 50% inter quartile ranges, and median of 200 simulations.

We found that the bias in estimating selectivity at age, recruitment, and SSB were high under the high movement scenario, and were low under the low movement scenario (Figure 4.11, Figure 4.12, and Figure 4.13). Under the high movement scenario, the Full model tended to underestimate SSB in the early recovery years for all populations with great uncertainty (Figure 4.13).

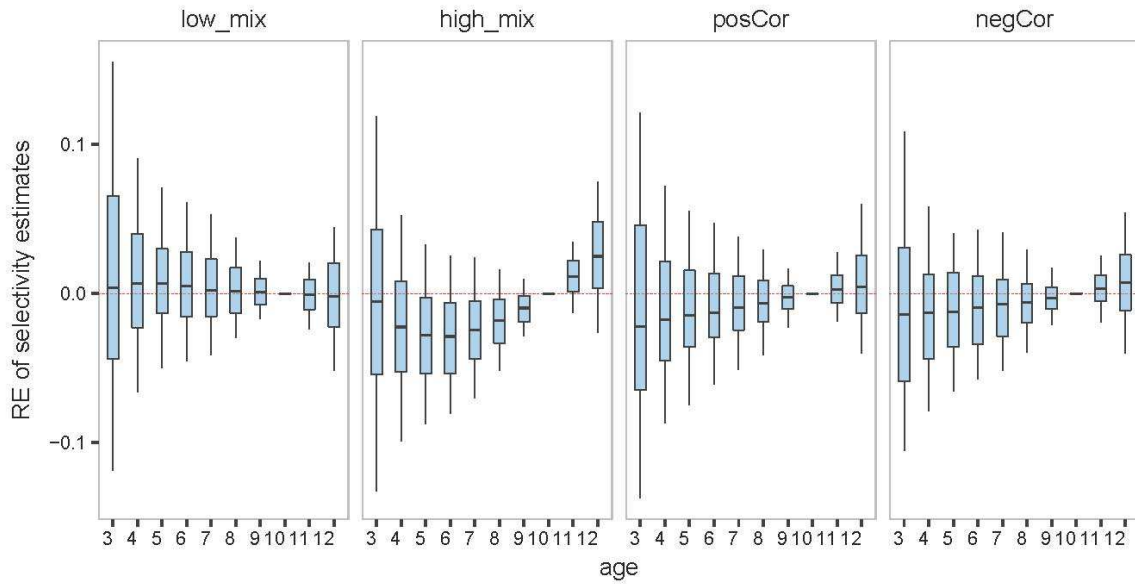


Figure 4.11. Relative error of selectivity at age estimates for the Full model under the scenarios with alternative movement assumptions. Each panel represents a movement scenario. Whiskers, boxes, and horizontal middle line are 95% and 50% inter quantile ranges, and median of 200 simulations.

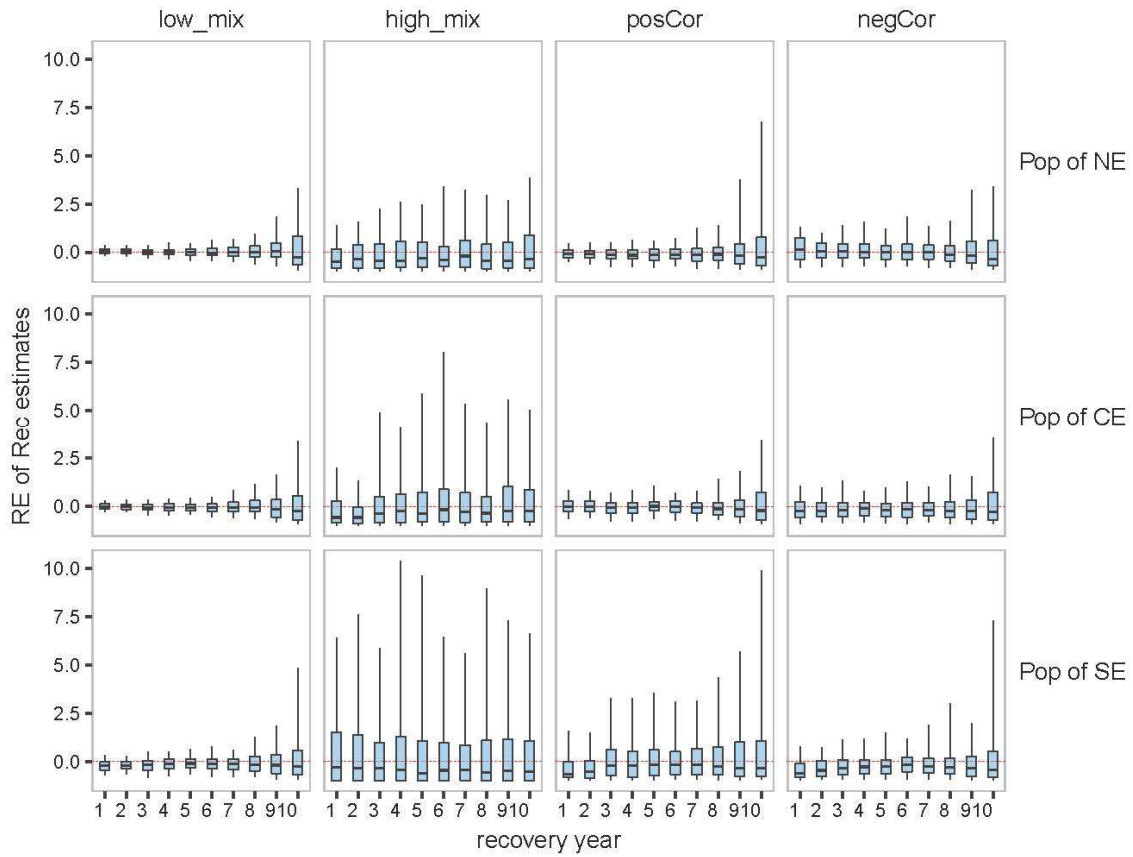


Figure 4.12. Relative error of recruitment estimates for the Full model under the scenarios with alternative movement assumptions. Each column represents a movement scenario, and each row represents a population. Whiskers, boxes, and horizontal middle line are 95% and 50% inter quantile ranges, and median of 200 simulations.

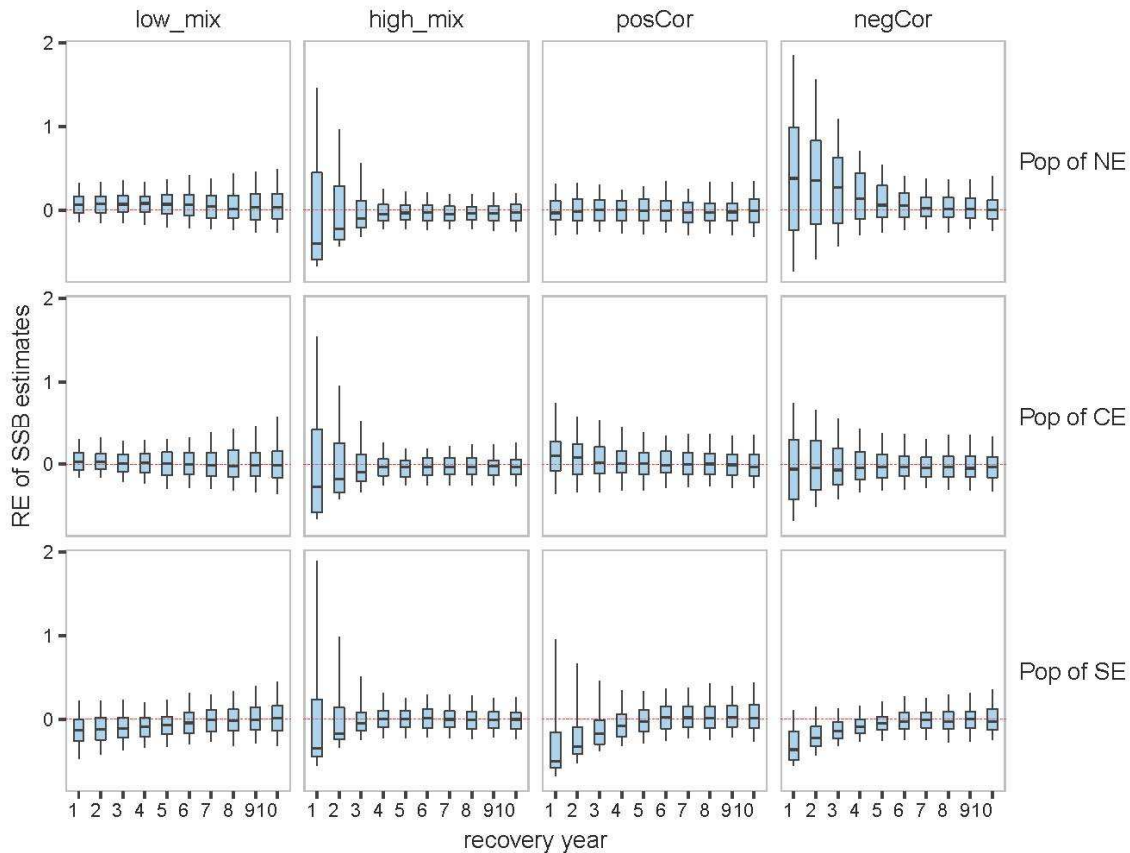


Figure 4.13. Relative error of SSB estimates for the Full model under the scenarios with alternative movement assumptions. Each column represents a movement scenario, and each row represents a population. Whiskers, boxes, and horizontal middle line are 95% and 50% inter quantile ranges, and median of 200 simulations.

Sensitivity scenarios

In general, the results of the baseline scenario ($fix_{X_i} + est_M$) are sensitive to data quality, but not sensitive to high recruitment variation (Figure 4.14). With low data quality, both tag reporting rate and natural mortality rate were underestimated, and f was greatly overestimated by the Full model. The selectivity at age estimates were also more biased with low data quality. In contrast, the Full model with high data quality resulted in nearly median-unbiased estimates of selectivity at age, tag reporting rate, and M .

Although both the bias and uncertainty in estimating recruitment and SSB were decreased for the high data quality scenario compared to the baseline scenario (Figure 4.15 versus

Figure 4.6), similar tendencies, such as biased estimates of recruitment in terminal assessment years and SSB in early assessment years persisted.

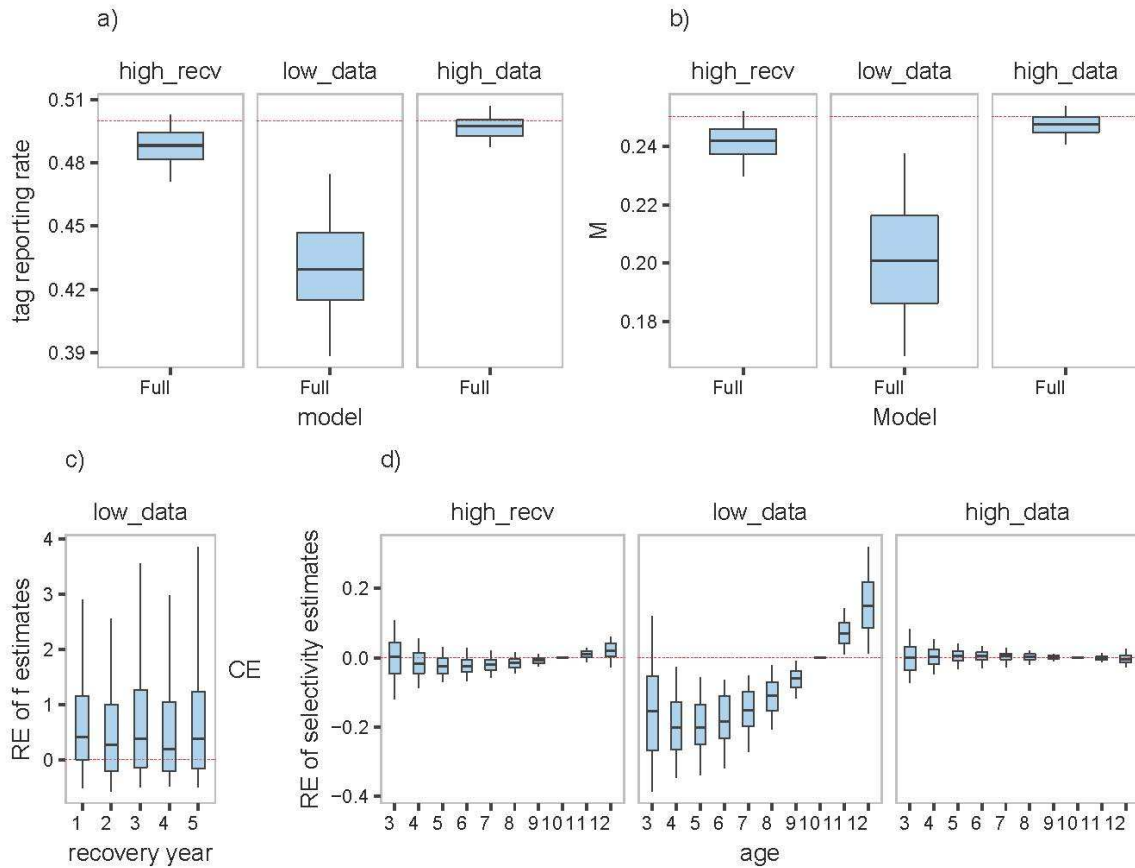


Figure 4.14. Assessment results for the Full model in sensitivity analysis. Shown are tag reporting rate estimates (a), natural mortality rate estimates (b), relative error of full selected fishing mortality estimates in all recovery years (c), and relative error of selectivity at age estimates (d). The panel name in all sections represents a sensitivity scenario. Whiskers, boxes, and horizontal middle line are 95% and 50% inter quantile ranges, and median of 200 simulations.

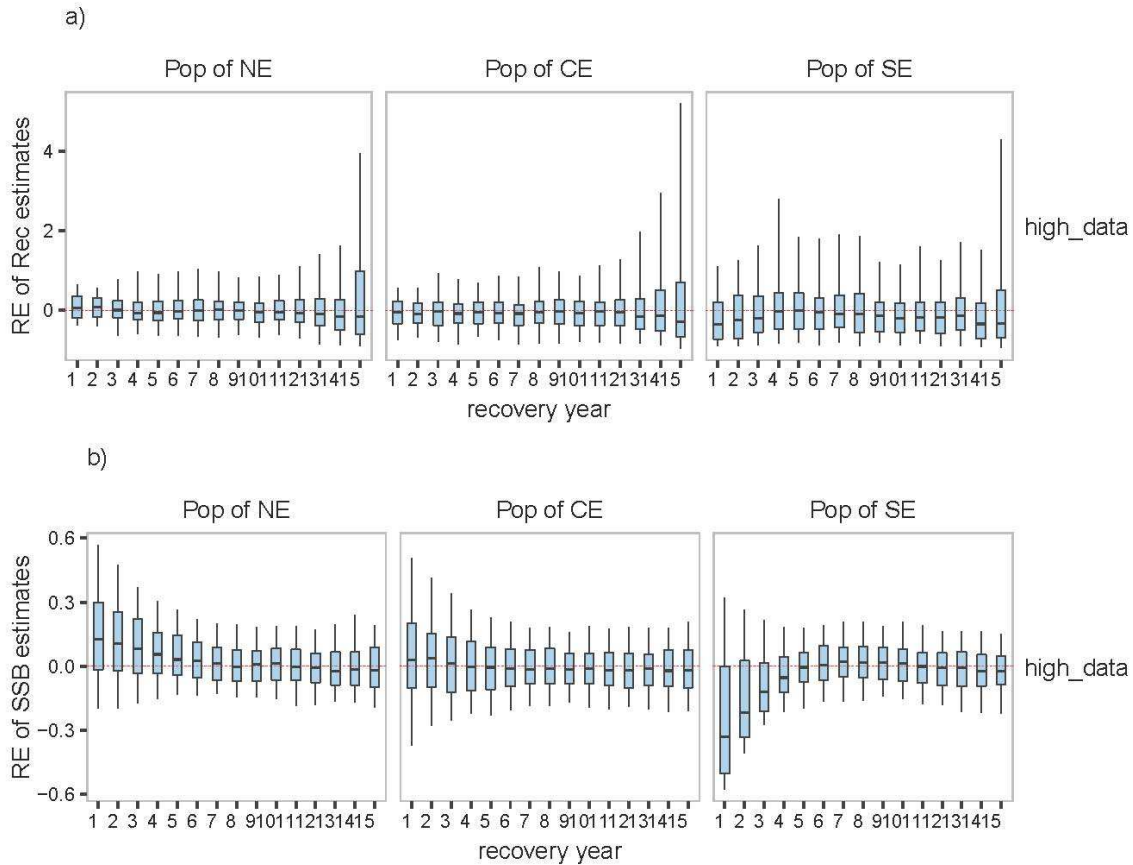


Figure 4.15. Assessment results for the Full model in sensitivity analysis with high data quality. Shown are relative error of recruitment estimates (a) and relative error of SSB estimates (b) for all spawning populations (panel name) in all recovery years (x-axis). Whiskers, boxes, and horizontal middle line are 95% and 50% inter quantile ranges, and median of 200 simulations.

Except for the scenarios with low data quality, high movement, and with alternative homing assumptions, the coverage probability of all performance metrics was consistent across different movement, homing, and sensitivity scenarios (Figure 4.16). Under the scenario with low data quality, the coverage probability for tag reporting, M , f , and selectivity at age (median of all dimensions) were all below 0.1. Under the high movement scenario, the coverage probability of recruitment and SSB was obviously lower than those under other scenarios (both 20% lower). The coverage probabilities for

all performance metrics were substantially higher when the spatial structure was overlap than diffusion, except for the movement rates.

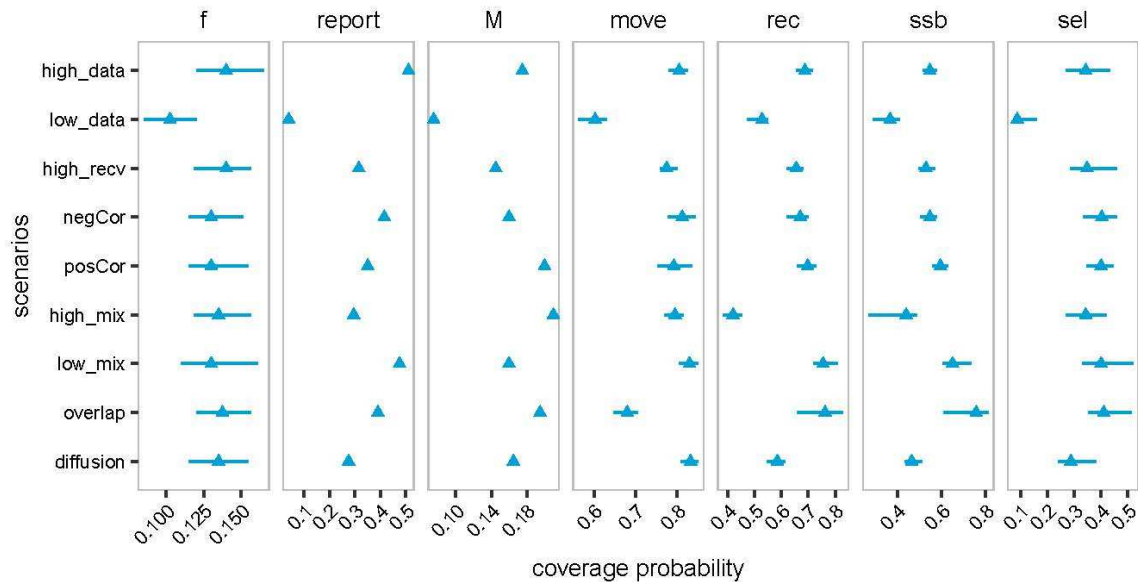


Figure 4.16. Same as Figure 4.9, except that each row represents a scenario with alternative homing assumption or movement assumption, or a sensitivity scenario.

Discussion

Traditional Brownie type of tagging models or integrated tagging and catch-at-age analysis (ITCAAN) commonly assume either diffusion or overlap spatial structure for intermixed fisheries, but actual fish populations will not always perfectly match these categories. We proposed an integrated assessment model framework for modeling tag-recovery, tag monitoring, fishing effort, and catch-at-age data that allowed for a continuum of spatial structures through modeling homing probability in the assessment model. This innovation provides the useful ability to adapt the model to alternative movement assumptions (overlap, incomplete overlap, and diffusion) by assuming a known homing probability. It also provides an opportunity to evaluate whether the data are informative regarding the spatial structure by estimating homing probability. Another

useful aspect is that our framework can be adapted to different data availabilities so that not only it can make use of all available data sources in addition to tag-recovery and fishing effort data (such as tag monitoring data and catch-at-age data), but also can be used to evaluate the data conflicts between tag-recovery and catch-at-age data. Our results suggested that our model framework with only tag-recovery and fishing effort data had robust assessment performance in estimating movement rates, homing probability, natural mortality, and fully selectivity fishing mortality rates. When tag monitoring data were also available, our framework can provide a reasonable estimate of the tag reporting rate. Including additional catch-at-age data did substantially improve the estimates of selectivity at age, slightly improved estimates of tag reporting and natural mortality rates, but the bias in estimating recruitment and spawning stock biomass can be high, especially for low productivity populations.

Our simulations also provided a detailed understanding of a continuum of spatial structures, and how homing probability influence the population dynamics of spatially structured populations. Homing probability has been widely used as an indicator of spatial structure for intermixed populations, and can be estimated from some types of discriminatory studies such as geochemical signature in otolith and molecular genetics (Stewart *et al.*, 2003; Rooker *et al.*, 2008, 2014). Other promising source of information for evaluating homing probability is from telemetry and geolocation (Goethel *et al.*, 2011). Our results suggest that the homing probability can also be accurately estimated as part of fitting an assessment model, even with just tag-recovery and fishing effort data. Tag monitoring and catch-at-age do not provide a strong signal of fish movement and

homing, and thus as we expected, did not improve the estimates of homing probability in our simulations.

The robustness of estimating movement rates for our integrated assessment model framework is consistent with previous SBP and ITCAAN simulation studies (Eveson *et al.*, 2009; Goethel *et al.*, 2015b; Vincent *et al.*, 2017). This could be a factor of the simplified movement dynamics assumed in the operating and assessment models. We assumed time-invariant annual movement that occurred at the first day of each year, but in the real world fish movement rate can vary seasonally, occur at different timing of each year, be age-specific, or as a function of environmental variables or population density. However, increased number of parameters associated with more complex movement assumptions would no doubt increase the difficulty in model convergence, and sometimes results in even worse model performance (Goethel *et al.*, 2015a, 2015b; Vincent *et al.*, 2017). The trade-off between parameterization complexity and model realism has long been an issue for spatial structured models (Lauretta and Goethel, 2017; Vincent *et al.*, 2017). We thus urge caution in making complex movement assumptions in our integrated model without preliminary simulations to evaluate parameter estimability.

In our operating model, we simulated data sets with and without a tag monitoring program so that our model framework could be broadly applicable to real world situations with different data sources. With data from tag monitoring program, the STRO and Full models both provided relatively accurate estimations of tag reporting rates with only slight negative bias, and such bias can be removed if natural mortality rate (M) were assumed to be known in the assessment models. When M was estimated, small negative bias in estimating M were observed in both models. Together with the small positive bias in

estimating fully selected fishing mortality and selectivity for older age fish, it seems that both models were forced to estimate a lower tag reporting rate to provide reasonable fit to tag-recovery data, and to estimate a lower natural mortality rate so as to obtain a correct overall survival rate, due to the overestimated fishing mortality rate. The slightly smaller bias in estimating M and tag reporting rate in the Full model compared to the estimates in the STRO model, is consistent with this interpretation, given that the estimates of selectivity at age and fully selected fishing mortality in the Full model were less biased than in the STRO model. The greater bias in estimating selectivity and fishing mortality in the STRO model could be due to the lack of age information in the tag-recovery data, short length of the tagging time series, and the model misspecification we assumed for the catchability and fishing effort in the operating and assessment models. We assumed constant catchability, and lognormal observation error in fishing effort data in the operating model, while we used random walk process to model catchability, and ignored observation error in fishing effort data in the assessment model.

In the STR model, the tag reporting rate was assumed to be constant and known without error and as expected the estimates of M were more unbiased than the other two models in which the tag reporting rate was estimated. However, since the tag reporting rate depends on fishermen's behavior and outreach effort, it is unlikely to be consistent across years, or to have such accurate external estimates of tag reporting rates. For future studies, we recommend sensitivity tests to evaluate how model misspecification of tag reporting rate could influence the model performance of the STR model. For future real world tagging study with tag monitoring program, we recommend simultaneously run STR model with external estimates of tag reporting rates, and STRO model with internal estimates of

tag reporting rates, and compare the results to see if the two models provide similar conclusions.

Including additional catch-at-age data does not always guarantee more accurate estimates of parameters, and it appears that the usefulness of catch-at-age data within our integrated assessment model framework can be highly case-specific. For our Full model with catch-at-age data, although the bias in estimating selectivity at age, tag reporting rates and natural mortality rates all decreased compared to the STRO model, there were considerable bias in estimating recruitment and SSB, especially for low productivity populations. One possible reason could be that the assessment model was challenged at identifying correct age composition data for low productivity population, because the age composition data collected from each harvest basin were dominated by fish from populations with higher productivity. Another reason, could be because we assumed recruitment was stationary when it was not. In particular the recruitment penalty we incorporated in the assessment model assumed annual recruitment of each spawning population were lognormal distributed around a constant mean. But due to the incomplete overlap spatial structure, there were fish that originally came from populations with high and medium productivity that joined the low productivity population annually, which could have caused the recruitment to increase over time in the low productivity populations. This is also suggested by results from our diffusion and overlap scenarios. In those, less biased and more certain estimates of recruitment were obtained for the overlap spatial structure than for the diffusion spatial structure for medium and high productivity populations. Nevertheless a recruitment penalty of some type appears necessary: in our preliminary simulations that did not include a recruitment penalty, and in other studies for spatially

structured stock assessment models without recruitment penalties, recruitment in terminal assessment years could not be estimated (Li *et al.*, in press; Vincent *et al.*, 2017). The problem might also be associated with the short time series of data we used in our baseline scenario, as suggested by sensitivity scenarios with high data quality in which 15 years of data were incorporated. In that case, the bias and uncertainty in estimating all parameters was much reduced. A normal statistical catch-at-age model for a species like lake whitefish usually incorporates more than 30 years of catch-at-age data. However, a 30-year tagging study in most cases is unrealistic. For example, the actual tagging program done for Lake Huron lake whitefish population is only four years long (Li *et al.* 2017), which is even shorter than our evaluated low data quality scenario. Under the low data quality scenario, except for movement rates, many parameters could not be reliably estimated. Thus, for future tagging studies that intend to collect data for an integrated assessment model framework like ours, we suggest to first run simulations to provide guidance on the needed length of tagging time series and other aspects of study design.

We found some systemic errors in estimating SSB for populations with the overlap spatial structure in our integrated assessment models. We believe this may due to the collapse of low productivity populations. Our results, in agreement with Vincent *et al.* (2017) suggested that a spatial structured statistical catch-at-age model has difficulty in estimating abundance for a population that is close to collapse, given harvest, age composition, and tag-recovery data from all harvest basins are dominated by other populations. However, instead of overestimating recruitment for low productivity population as found in many other studies of spatially structured stock assessments, in which movement rates were assumed to be known (e.g., Li *et al.*, 2015), our model tended

to over-estimate movement from high productivity population to low productivity population and under-estimate movement from low productivity population to high productivity populations. This suggests that a spatially structure estimation model can explain mixed stock abundance by either moving fish among regions or “creating” them through recruitment and initial abundance (Goethel *et al.*, 2015b). Although tag-recovery data can introduce additional information on movement, such signals can be weak if the size of the low productivity population is low relative to other populations.

We believe our integrated assessment model results are generally applicable to tagging studies with different data availability and spatial structure assumptions, but there are limitations to our simulation study that need to be considered. For example, the simplification of movement, tag reporting, and tag monitoring process in our simulations may need to be reevaluated in the context of the reality of a specific case. In the real world case where often time series of tagging data are shorter than in our simulations while longer-term catch-at-age data are available, alternative adaptations of our model framework should be considered for unbalanced length of time series of tagging and catch-at-age data, although more restrictive assumptions on natural mortality and movement rate will be necessary in that case.

APPENDICES

Appendix 4.A: Reparameterization of the tagging dynamics of the ITCAAN model

The major difference between SBP and ITCAAN models is in how they treat the tag return process. Although both models treat recovered tags as multinomial, the SBP model treats the tag return probabilities as a direct product of several conditional probabilities such as movement, survival, exploitation, tag retention, and tag reporting probabilities. In contrast, the ITCAAN model estimates the number of recovered fish by assuming the tagged fish undergo the same dynamics as the whole population, and then back-calculates the tag return proportion by dividing the estimated number returned by the total number of tagged fish released. Here we reparameterize the tagging dynamics in the ITCAAN model, and illustrate how the tag recaptures proportions in ITCAAN model can be calculated directly as a function of survival, movement, tag reporting, and mortality rates.

In Goethel et al.(2014 and 2015), tag recaptures (r) was modeled using Baranov's catch equation, given by:

$$r_{j,y}^l = n_{j,y}^l \beta_j \frac{F_{j,y}}{M+F_{j,y}} (1 - e^{-(M+F_{j,y})}) \quad (4.A1)$$

where l, j, y are indexes for cohort (a combination of released area and year), stock and year separately. Parameter β is the tag reporting rates, and F and M are instantaneous fishing and natural mortality rates. n is tag cohort abundance, which was modeled as:

$$n_{j,y}^l = \begin{cases} n_{j,y}^l & \text{for } y = 1 \\ \sum_{k=1}^K n_{k,y-1}^l e^{-(M+F_{k,y})} T_{k,j} & \text{for } y > 1 \end{cases} \quad (4.A2)$$

where $T_{k,j}$ is the movement proportion moving from area k to area j . the $n_{j,1}^l$ is the total number of released tags at the first year, and abundance for the same tag cohort l in later years y , $n_{j,y}^l$, were modeled as total fish in all areas that survive from the previous year that moved to area j . Here for simplification, we ignored the separate time subscript t in Goethel et al.(2014).

The tag recaptures were converted into proportions by “state”. For each cohort l , a recovery state was defined as the combination of recapture region and year, and the final “not recaptured” (NR) state which subsumed all possible states for which a tagged fish may not be recaptured. The tag proportion P_{tag} in each state for a given cohort was the number of tags recaptured (or not captured for the final state) in that state divided by the total number of releases for the cohort:

$$p_{j,y}^l = r_{j,y}^l / n_{j,1}^l \quad (4.A3)$$

$$p_{j,y,NR}^l = 1 - \sum_{y=1}^Y \sum_{j=1}^J p_{j,y}^l \quad (4.A4)$$

Then the tag recaptures were assumed to follow a multinomial distribution, with probability the proportions by “state” we calculated before.

To reparameterize the above dynamics, we treated the total number of released tag $n_{j,1}^l$ as scaler in equation 4.A2, and $n_{j',y}^l$ was modeled as a function of $n_{j,1}^l$, given by:

$$n_{j',y}^l = n_{j,1}^l \omega_{j,j',y}^l \quad (4.A5)$$

where $\omega_{j,j',y}^l$ is the survival-movement probability for cohort l that showed up at the beginning of year 1 in area j that survived and moved to area j' at the beginning of year

y. We used matrix calculation to model $\omega_{j,j',y}^l$ as given in the main text (same as $\omega_{i,a,y,kk'}$ by assuming homing rate equals to 1, and a specific combination of $i = 1, a, k = j$ correspond to cohort l). Then we treated $n_{k,1}^l$ as a divisor in equation 4.A3, and canceled. This shows that the tag recaptures proportions by “state” can be calculated directly as a function of survival, movement, tag reporting, and mortality rates.

Appendix 4.B: Life history parameters used in our simulations

We used the same values of life history parameters (i.e., weight- and maturity-at-age, female ratio, selectivity parameters and fully selected age), and steepness values for different productivity levels as used by Li *et al.* (In press). Equations for weight- and maturity-at-age calculation, and parameter values for selectivity at age are in Table 4.A1. The fully selected age is assumed to be age-10 for lake whitefish. The catchability in the operating model were assumed to be consistent for all years, and we used the same value as used by Li *et al.* (In press). Ricker stock-recruitment parameterization in the operating model are in Table 4.A2.

Table 4.A1. Weight- and maturity-at-age calculation and selectivity at age parameter values in the operating model.

Model name	Model equation
Length at age	$L_a = L_\infty (1 - \exp(-\kappa(a - t_0)))$ <p>where $L_\infty = 60.9$ cm, $\kappa = 0.1689$ year⁻¹, $t_0 = 0$ year (from Li et al. 2015)</p>
Weight at age	$W_a = \gamma L_a^\psi$ <p>where $\gamma = 8.06 \times 10^{-5}$, $\psi = 2.45$ (from Li et al. 2015)</p>
Maturity at age	$m_a = \frac{m_\infty}{1 + \exp(-\vartheta(L_a - \delta))}$ <p>where $\vartheta = 0.315$ cm⁻¹, $\delta = 37.86$ cm (from Li et al. 2015)</p>
Selectivity at age	selectivity parameters in Equation 4.2.4: $\tau = 1.26$ year ⁻¹ , $\eta = 13.074$ cm (from Li et al. 2015)

Table 4.A2. Equations for Ricker stock-recruitment parameterization in the operating model.

Model name	Model equation
Steepness	$h = \frac{R_{0.2S_0}}{R_{S_0}} = \frac{\alpha 0.2S_0 e^{-\beta 0.2S_0}}{\alpha S_0 e^{-\beta S_0}} = 0.2e^{\beta 0.8S_0}$ $\beta = \frac{\ln(h/0.2)}{0.8S_0}$
Unfished abundance at age during spawning period	$N_{spawn_a} = N_{a+1} = N_a e^{-(M)}$ $N_{a=1} = R_{S_0}$
Unfished spawning stock	$S_0 = \sum_a N_{spawn_a} Fem m_a W_a$ <p>where $Fem=0.5$ (from Li et al. 2015)</p>
Unfished stock per recruit	$SPR_0 = \frac{S_0}{R_{S_0}} = \frac{S_0}{\alpha' S_0 e^{-\beta S_0}} = \frac{e^{\beta S_0}}{\alpha'}$ $\frac{S_0}{R_{S_0}} = \frac{e^{\beta S_0}}{\alpha'}$
Solving Ricker parameter α'	$\alpha' = \frac{e^{\beta S_0}}{S_0} = \frac{e^{\beta S_0}}{\sum_a N_{spawn_a} Fem m_a W_a}$ <p>when $R_{S_0}=1$</p>

Appendix 4.C: Prior Probability Distributions

The prior probability distributions for all estimated parameters (Table 4.4) in different tagging models are in Table 4.A3.

Table 4.A3. Prior distributions for all estimated parameters.

Parameter	Uniform prior:
logit transformation of $\pi_{kk'}$	$(-\infty, +\infty)$
$\lambda_{jk'}$	(0,1)
M	(0,3)
Φ	Log scale (-15, -10)
η	Log scale (-3,5)
τ	Log scale (-3,3)
$R\mu_k$	Log scale (9,19)
$\varepsilon'_{k,y}$	Log scale (-5,5)
Γ_k	Log scale (9,19)
$\Delta_{k,a}$	Log scale (-7,7)
$\hat{\sigma}_F$	Log scale (-5,5)
$\widehat{\varepsilon F}_y$	Log scale (-5,5)
ξ	[0,1]

Appendix 4.D: Likelihood component for models including tag monitoring data

For each $iakjk'$, because recovered fish from the observed component

$R_{i,a,k,j,k',1} \sim \text{Bin} \left(R_{i,a,k,j,k',\cdot} \frac{\delta_{j,k'}}{(\delta_{j,k'} + (1 - \delta_{j,k'})\lambda_{k'})} \right)$, L_2 is the product of binomials, given by:

$$L_2 = P \left[R_{i,a,k,j,k',l} \mid R_{i,a,k,j,k',\cdot} \right]$$

$$= \prod_{i=1}^I \prod_{a=1}^A \prod_{k=1}^K \prod_{j=i}^J \prod_{k'=1}^{K'} \left(\begin{matrix} R_{i,a,k,j,k',\cdot} \\ R_{i,a,k,j,k',1} & R_{i,a,k,j,k',2} \end{matrix} \right) \left(\frac{\delta_{j,k'}}{(\delta_{j,k'} + (1 - \delta_{j,k'})\lambda_{k'})} \right)^{R_{i,a,k,j,k',1}}$$

$$\left(\frac{(1 - \delta_{j,k'})\lambda_{k'}}{(\delta_{j,k'} + (1 - \delta_{j,k'})\lambda_{k'})} \right)^{R_{i,a,k,j,k',2}}$$

As $\lambda_{k'}$ and $\delta_{j,k'}$ do not depend on iak we can further simplify this likelihood to

$$L_2 \propto \prod_{j=1}^J \prod_{k'=1}^{K'} \left(\begin{matrix} R_{\dots,j,k',\cdot} \\ R_{\dots,j,k',1} & R_{\dots,j,k',2} \end{matrix} \right) \left(\frac{\delta_{j,k'}}{(\delta_{j,k'} + (1 - \delta_{j,k'})\lambda_{k'})} \right)^{R_{\dots,j,k',1}}$$

$$\left(\frac{(1 - \delta_{j,k'})\lambda_{k'}}{(\delta_{j,k'} + (1 - \delta_{j,k'})\lambda_{k'})} \right)^{R_{\dots,j,k',2}}$$

Note since $\delta_{j,k'}$ is from the data, the only parameter involved here is $\lambda_{k'}$, hence the likelihood is proportional to

$$L_2 \propto \prod_{j=1}^J \prod_{k'=1}^{K'} \frac{(\lambda_{k'})^{R_{\dots,j,k',2}}}{(\delta_{j,k'} + (1 - \delta_{j,k'})\lambda_{k'})^{R_{\dots,j,k',1} + R_{\dots,j,k',2}}}$$

where $R_{\dots,j,k',\cdot} = \sum_{i=1}^I \sum_{a=1}^A \sum_{k=1}^K \sum_{l=1}^L R_{i,a,k,j,k',l}$ is the total number of tags recovered during year j from basin k' , from both monitored and unmonitored components of harvest.

$R_{\dots jk'1} = \sum_{i=1}^I \sum_{a=1}^A \sum_{k=1}^K R_{i,a,k,j,k',1}$ is number of the tags recovered from monitored component of harvest during year j from basin k' . $R_{\dots jk'2} = \sum_{i=1}^I \sum_{a=1}^A \sum_{k=1}^K R_{i,a,k,j,k',2}$ is the number of tags recovered from unmonitored component of harvest during year j from basin k' . Thus,

$$\begin{aligned}
 -\ln(L_2) = \text{Const.} & - \sum_{j=1}^J \sum_{k'=1}^K \left(R_{\dots j,k',2} \log \lambda_{k'} \right. \\
 & \left. - \left(R_{\dots j,k',1} + R_{\dots j,k',2} \right) \log \left(\delta_{jk'} + (1 - \delta_{jk'}) \lambda_{k'} \right) \right)
 \end{aligned}$$

Appendix 4.E: Likelihood components for models including catch at age data

Predicted harvest in the assessment model were modeled as:

$$\hat{C}_{k',y,a} = \left(\sum_{k=1}^K N_{k,y,a} \pi_{kk'} \right) \frac{F_{k',y,a}}{M_{k',y,a} + F_{k',y,a}} (1 - S_{k',y,a})$$

$$\hat{C}_{k',y} = \sum_{a=3}^A \hat{C}_{k',y,a}$$

$$\hat{p}_{k',y,a} = \frac{\hat{C}_{k',y,a}}{\hat{C}_{k',y}}$$

Thus, the negative log-likelihood component for the log of area-specific annual fishery harvest L_c was based on a normal distribution, and we assumed $\hat{\sigma}_F^2$ was four times greater than σ_c^2 , which matched what was assumed in the operating model:

$$-\ln(L_c) = \text{Const.} + \sum_{y=1}^Y \sum_{k=1}^K \ln(\sigma_c^2) + \frac{(\ln(C_{k',y}) - \ln(\hat{C}_{k',y}))^2}{2\sigma_c^2}$$

$$\text{We assumed } \sigma_c^2 = 0.25 \hat{\sigma}_F^2$$

A multinomial distribution was assumed for the log-likelihood component for the harvest age composition, given by:

$$-\ln(\ell_{ac}) = \text{Const.} - \sum_{k'} \sum_y N_{eff} \sum_a p_{k',y,a} \ln \hat{p}_{k',y,a}$$

where N_{eff} is the effective sample size defined in Table 4.8.

The log normal penalty for initial abundance at age were modeled as:

$$-\ln(L_I) = \text{Const.} + \sum_{k=1}^K \sum_{a=4}^{A-1} \left(\ln(\sigma'_I) + \frac{\Delta_{k,a}^2}{2\sigma'^2_I} \right)$$

The log-standard deviation σ'_I for the initial abundance deviations were set equals to 4.0, to provide only a weak restruictiin on parameter estimations.

BIBLIOGRAPHY

BIBLIOGRAPHY

- Brooks, S. P., and Gelman, A. 1998. General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics*, 7: 434–455. <http://www.tandfonline.com/doi/abs/10.1080/10618600.1998.10474787>.
- Cadrin, S. X., and Secor, D. H. 2009. Accounting for Spatial Population Structure in Stock Assessment: Past, Present, and Future. *In The Future of Fisheries Science in North America*, pp. 405–426. Springer Netherlands, Dordrecht. http://link.springer.com/10.1007/978-1-4020-9210-7_22.
- Eveson, J. P., Laslett, G. M., and Polacheck, T. 2009. A spatial model for estimating mortality rates, abundance and movement probabilities from fishery tag-recovery data. *In Modeling Demographic Processes In Marked Populations*.
- Goethel, D. R., Quinn, T. J., and Cadrin, S. X. 2011. Incorporating spatial structure in stock assessment: Movement modeling in marine fish population dynamics. *Reviews in Fisheries Science*, 19: 119–136.
- Goethel, D. R., Legault, C. M., and Cadrin, S. X. 2015a. Demonstration of a spatially explicit, tag-integrated stock assessment model with application to three interconnected stocks of yellowtail flounder off of New England. *ICES Journal of Marine Science*, 72: 164–177. <https://academic.oup.com/icesjms/article-lookup/doi/10.1093/icesjms/fss041>.
- Goethel, D. R., Legault, C. M., and Cadrin, S. X. 2015b. Testing the performance of a spatially explicit tag-integrated stock assessment model of yellowtail flounder (*Limanda ferruginea*) through simulation analysis. *Canadian Journal of Fisheries and Aquatic Sciences*, 72: 582–601. <http://www.nrcresearchpress.com/doi/10.1139/cjfas-2014-0244>.
- Healey, M. C. 1978. Fecundity Changes in Exploited Populations of Lake Whitefish (*Coregonus clupeaformis*) and Lake Trout (*Salvelinus namaycush*). *Journal of the Fisheries Research Board of Canada*, 35: 945–950. <http://www.nrcresearchpress.com/doi/abs/10.1139/f78-155>.
- Henderson, B. A., Collins, J. J., and Reckahn, J. A. 1983. Dynamics of An Exploited Population of Lake Whitefish (*Coregonus clupeaformis*) In Lake Huron. *Canadian Journal of Fisheries and Aquatic Sciences*, 40: 1556–1567. <http://www.nrcresearchpress.com/doi/abs/10.1139/f83-180>.
- Kerr, L. A., Hintzen, N. T., Cadrin, S. X., Clausen, L. W., Dickey-Collas, M., Goethel, D. R., Hatfield, E. M. C., *et al.* 2016. Lessons learned from practical approaches to reconcile mismatches between biological population structure and stock units of marine fish. *ICES Journal of Marine Science: Journal du Conseil*, 74: fsw188. http://fdslive.oup.com/www.oup.com/pdf/production_in_progress.pdf%5Cnhttp://w

www.ncbi.nlm.nih.gov/pubmed/27899565%5Cn<https://academic.oup.com/icesjms/article-lookup/doi/10.1093/icesjms/fsw188>.

- Lauretta, M. V., and Goethel, D. R. 2017. The robustness of Brownie tag return models to complex spatiotemporal dynamics evaluated through simulation analysis. *Canadian Journal of Fisheries and Aquatic Sciences*, 74: 1845–1861. <http://www.nrcresearchpress.com/doi/10.1139/cjfas-2016-0291>.
- Li, Y., Bence, J. R., and Brenden, T. O. 2015. An evaluation of alternative assessment approaches for intermixing fish populations: a case study with Great Lakes lake whitefish. *ICES Journal of Marine Science*, 72: 70–81. <https://academic.oup.com/icesjms/article-lookup/doi/10.1093/icesjms/fss023>.
- Molton, K. J., Brenden, T. O., and Bence, J. R. 2012. Control rule performance for intermixing lake whitefish populations in the 1836 Treaty waters of the Great Lakes: A simulation-based evaluation. *Journal of Great Lakes Research*, 38: 686–698. <http://dx.doi.org/10.1016/j.jglr.2012.09.014>.
- Polacheck, T., Eveson, J. P., Laslett, G. M., Pollock, K. H., and Hearn, W. S. 2006. Integrating catch-at-age and multiyear tagging data: a combined Brownie and Petersen estimation approach in a fishery context. *Canadian Journal of Fisheries and Aquatic Sciences*, 63: 534–548. <http://www.nrcresearchpress.com/doi/abs/10.1139/f05-232>.
- Porch, C. E., Turner, S. C., and Powers, J. E. 2001. Virtual population analyses of atlantic bluefin tuna with alternative models of transatlantic migration: 1970-1997. *International Commission for the Conservation of Atlantic Tunas (ICCAT) Collective Volume of Scientific Papers*, 52: 1022–1045.
- Rooker, J., Secor, D., DeMetrio, G., Kaufman, A., Belmonte Ríos, A., and Ticina, V. 2008. Evidence of trans-Atlantic movement and natal homing of bluefin tuna from stable isotopes in otoliths. *Marine Ecology Progress Series*, 368: 231–239. <http://www.int-res.com/abstracts/meps/v368/p231-239/>.
- Rooker, J., Arrizabalaga, H., Fraile, I., Secor, D., Dettman, D., Abid, N., Addis, P., *et al.* 2014. Crossing the line: migratory and homing behaviors of Atlantic bluefin tuna. *Marine Ecology Progress Series*, 504: 265–276. <http://www.int-res.com/abstracts/meps/v504/p265-276/>.
- Stewart, I. J., Quinn, T. P., and Bentzen, P. 2003. Evidence for fine-scale natal homing among island beach spawning sockeye salmon, *Oncorhynchus nerka*. *Environmental Biology of Fishes*, 67: 77–85. <http://link.springer.com/10.1023/A:1024436632183>.
- Vandergoot, C. S., and Brenden, T. O. 2014. Spatially varying population demographics and fishery characteristics of Lake Erie walleyes inferred from a long-term tag recovery study. *Transactions of the American Fisheries Society*, 143: 188–204. <http://www.tandfonline.com/doi/abs/10.1080/00028487.2013.837095>.

Vincent, M. T., Brenden, T. O., and Bence, J. R. 2017. Simulation testing the robustness of a multiregion, tag-integrated assessment model that exhibits natal homing and estimates natural mortality and reporting rate. *Canadian Journal of Fisheries and Aquatic Sciences*, 74: 1930–1949. <http://www.nrcresearchpress.com/doi/10.1139/cjfas-2016-0297>.

Ying, Y., Chen, Y., Lin, L., and Gao, T. 2011. Risks of ignoring fish population spatial structure in fisheries management. *Canadian Journal of Fisheries and Aquatic Sciences*, 68: 2101–2120. <http://www.nrcresearchpress.com/doi/abs/10.1139/f2011-116>.

CONCLUSIONS AND RECOMMENDATIONS

Chapter 1 Conclusions and Recommendations

- Consistent with conclusions from previous studies of stream-dwelling fish, there was a consistent positive relationship between lake whitefish net movement distance and fish total length at the time of tagging. Therefore, I recommend that for future research the movement rates or intensity of lake whitefish could be modeled as a function of length or age, when possible.
- The spawning migration for lake whitefish generally occurred within months from September to November, and after that in December, fish tended to leave their spawning site and were actually further from the spawning location than in the baseline month of June. When additional information is not available, I recommend using the first day of December as the timing of instantaneous annual movement for modeling lake whitefish movement in Upper Great Lakes regions.
- My results suggest that fish may start their annual spawning migration runs earlier in warmer years after acquiring and processing energy needed for spawning. Therefore, I recommend taking into account growing degree days when defining movement and spawning times.
- When relative *Diporeia* spp. density was high near the spawning grounds, lake whitefish tended to stay closer to their spawning site. This implies that fish might expand their foraging area when *Diporeia* density was low near their preferred habitat. To further evaluate such hypotheses, future research could explore how changes in the food web influence lake whitefish movement within a broader context. For example, density distributions of zebra mussels (*Dreissena*

polymorpha), sea lamprey (*Petromyzon marinus*), and round goby (*Neogobius melanostomus*) could all be incorporated as predictors.

- Lake whitefish tagged and released from the tagging sites Cheboygan and Alpena had consistently greater net distance than those released from other areas.

Together with previous genetic and food web marker studies of Lake Huron lake whitefish (Stott et al. 2010, Eberts et al. 2017), which both indicate that current management units are too small in general, I recommend that it is worthwhile to carefully reevaluate the management boundaries in Lake Huron to see if they are biologically accurate. It is a complex issue involving tradeoffs between biology, politics, and management, and the management strategy evaluate approach might be useful to evaluate the long-term performance of alternative management boundaries.

- The Bayesian framework for analysis of conventional tagging data developed in Chapter 1 has potential for wide applicability, and detailed model details and the code are provided to facilitate this.

Chapter 2 Conclusions and Recommendations

- Incorporating information on population-specific age composition of harvest in a spatially structured statistical catch-at-age (SCAA) model resulted in better estimates of spawning stock biomass (SSB, less biased and less correlated estimates of SSB in terminal assessment years), and less uncertainty in estimating recruitment in early assessment years.
- The improvements in assessment performance of the assessment models with spawning origin information of catch did not necessarily translate into improved

management performance, except when I used a lower than status-quo mortality target or when low productivity populations intermixed with high productivity populations. While I do recommend considering spawning origin data of catch and models that can use these data, it is clear that benefits from improved assessments rely on applying an appropriate harvest policy. I recommend considering a revision of current harvest policies to lower mortality rates, as has been recommended in past studies.

- Including a lognormal penalty on annual recruitment residuals in the spatial structured SCAA model substantially improved the estimation of recruitment in the terminal assessment years. I therefore recommend that including recruitment penalty should be considered as one of the “best practice” for spatially-structured assessment models.

Chapter 3 Conclusions and Recommendations

- The movement rate estimates were relatively robust in the application of extended spatial Brownie-Petersen (SBP) tagging model to Lake Huron lake whitefish data (i.e., chapter 3). My results suggest that spawning populations in the U.S. main basin had higher probability to overlap with other populations during the fishing season compared to populations in Canadian waters.
- The estimates of fishing mortality, natural mortality, and tag reporting rates for lake whitefish in Chapter 3 are questionable, given their sensitivity and limitations of data. My application highlights the importance of data quality for my spatially structure stock assessment models.

- The setting of the SBP model largely limits the use of catch-at-age data. The SBP model required the same spatial and temporal scales of tag-recovery and catch-at-age data. I recommend that future research explore how to model unbalanced years of tagging and catch-at-age data for spatially structured stock assessment models, given in most real world cases the time series of catch-at-age data is longer than the time-series of tagging data.

Chapter 4 Conclusions and Recommendations

- My simulation study in Chapter 4 provides a detailed understanding of a continuum of spatial structures, and of how homing probability influence the population dynamics of spatially structured populations. Future research or management could consider using my framework to identify and/or verify the fish population spatial structure, when their conventional tagging and fishing effort data are available.
- In general, with only tag-recovery and fishing effort data, my framework in Chapter 4 had robust assessment performance in estimating movement rates, homing probability, natural mortality, and fully selectivity fishing mortality rates. When tag monitoring data were also available, my framework can provide reasonable estimate of tag reporting rate simultaneously. Including additional catch-at-age data did substantially improve the estimates of selectivity at age, slightly improved estimates of tag reporting and natural mortality rates, but bias in estimating recruitment and spawning stock biomass was sometimes high, especially for low productivity populations.

- The assessment performance of my integrated model was found to be sensitive to data quality. For future tagging studies that intend to collect data for an integrated assessment model framework like ours, I suggest to first run simulations to provide guidance on the needed length of tagging time series and other aspects of study design.
- My integrated assessment model approach is generally applicable to tagging studies with different data availability (either with or without tagging monitoring program and catch-at-age data) and spatial structure assumptions. Given the model can be applied with or without tagging data, it can be used to evaluate the data conflicts between tag-recovery and catch-at-age data.

Final thoughts

Although each chapter represents a stand-alone model framework with different emphasis on understanding and modeling fish movement, they could in fact be used synergistically and provide extended benefits. The variable selection procedure in Chapter 1 could be used for the preliminary study of understanding fish movement mechanisms, such as identifying the timing of the spawning migration and when fish began to leave their spawning sites, whether movement intensity is influenced by environmental factors, and life history traits such as length and sex. This information could be critically important for developing spatially structured stock assessment models (e.g., models introduced in the other chapters), by influencing choices of how movement and population dynamics is modeled. The integrated assessment model in the Chapter 4 could be used to identify the correct spatial structure assumption (whether it is overlap or diffusion, or incomplete overlap with a degree of homing), and estimate movement rates

between regions and homing together with other parameters by simultaneously analyzing conventional tagging, tag monitoring, and fishing effort data. Based on my simulations, including additional catch-at-age data in my integrated assessment model seems not that necessary, given it can only improve the estimates of selectivity at age and slightly improved estimates of tag reporting and natural mortality rates, but still resulted in high bias in estimating recruitment and spawning stock biomass, especially for low productivity populations. With prior information of movement rates between regions and spatial structure of populations (those values can also be estimated from models introduced in chapter 4), a spatially structured statistical catch-at-age model like that of Chapter 2 can be used to further estimate the stock size, recruitment, other relevant parameters, and to predict the total allowable catch. The MSE-based simulation approach in Chapter 2 could also be used to evaluate the long-term stock assessment and fisheries management performance of the spatially structured assessment models.

In Chapter 3, I applied extended spatial Brownie-Petersen (SBP) tagging models to Lake Huron lake whitefish data and selected the “best” model supported by data, but unfortunately the estimation results were highly sensitive to assumptions and thus are questionable. This highlights the importance of experimental design of conventional tagging program and tag monitoring program. For example, extreme short time-series of tagging data, incomplete coverage of the tag monitoring program, and basins with tag releases but with no recoveries forced us to make “ad hoc” assumptions about the parameterization in my application. I therefore recommend that simulation testing as a first step for the experimental design of tagging program and tag monitoring program, in

order to ensure the investment in tagging can produce the desired results. The sensitivity analyses in Chapter 3 highlight some of the design features that should be considered.

Lake whitefish in Lake Huron have undergone significant changes in their abundance, body conditions, diets, and distribution over recent decades. Given the results of Chapter 1 it would not be surprising if their movement patterns have changed substantially, and spatial assumptions for stock assessments and management of lake whitefish are being actively reconsidered on that lake, in part due to the research reported in this dissertation. I believe that spatial movement patterns should not be considered fixed, in general. Assessment scientists and managers should periodically synthesize the best available movement and stock identification information from tagging, genetics, otolith chemistry, life history, morphometric, parasites, and other sources. The extent of changes should be considered to determine if historical assessment and management structure still fits the biology of the fish populations.

BIBLIOGRAPHY

BIBLIOGRAPHY

- Eberts, R.L., Wissel, B., Simpson, G.L., Crawford, S.S., Stott, W., Hanner, R.H., Manzon, R.G., Wilson, J.Y., Boreham, D.R., and Somers, C.M. 2017. Isotopic Structure of Lake Whitefish in Lake Huron: Evidence for Regional and Local Populations Based on Resource Use. *North Am. J. Fish. Manag.* **37**(1): 133–148. doi:10.1080/02755947.2016.1245225.
- Stott, W., VanDeHey, J.A., and Sloss, B.L. 2010. Genetic diversity of lake whitefish in lakes Michigan and Huron; sampling, standardization, and research priorities. *J. Great Lakes Res.* **36**(SUPPL. 1): 59–65. doi:10.1016/j.jglr.2010.01.004.