

Version	Description	Changed By	Date
1.0	First draft	B. Letcher and M. Morrissey	11/9/11

1 Index

1	Index.....	1
2	Problem Statement	2
3	Solution Statement.....	3
3.1	General solution statement	4
3.1.1	Specific research questions	7
4	Key Features	8
4.1	Available data	8
4.2	Model components	9
4.2.1	Module description: growth rate and body size.....	10
4.2.2	Module description: Module description: survival.....	10
4.2.3	Module description: movement	11
4.2.4	Module description: reproduction	12
4.3	Population projection	12
4.3.1	Further justification for model selection.....	12
4.3.2	Integral projection model construction	13
4.3.3	Critical parameters to be obtained by population projection.....	13
4.4	Data processing	15
5	Alternatives Considered and Rejected.....	15
6	Major Implementation Constraints	15
7	Major Risks and Dependencies.....	16
7.1	Major risks	16
7.2	Dependencies.....	16
8	Acknowledgments	16
9	Literature Cited	16
10	Appendix A – Diagrams/Figures	18
11	Appendix B – Model code	18

2 Problem Statement

Cold water fish species face multiple threats to persistence, including habitat loss and changes in key environmental conditions (stream temperature and flow). These will interact in complex ways with environmental changes driven by climate change. A key step in maximizing probabilities of local population persistence across native ranges of stream fishes is prioritizing streams for management actions and determining the best actions in the wide variety of stream types occupied by brook trout.

Population processes, such as survival, reproduction, and ultimately population growth and persistence are driven by many complex processes. Variation in vital rates (survival and reproduction) are largely driven by environmental variation, often mediated by traits such as growth and movement. We refer to vital rates and to processes such as growth and movement as 'population processes'. Meaningful prediction of 'population outcomes' of direct relevance to conservation, such as population growth rate and extinction risk thus require simultaneous consideration of multiple aspects of population biology, especially when simultaneous aspects of environmental variation (e.g., natural and anthropogenic variation flow, temperature, and connectivity) are thought to be relevant. Our goal is to develop models of multiple population processes (survival, reproduction, movement), and the dependence of these multiple processes on multiple aspects of environmental variation. These models will allow us to address problems such as: What aspects of the environment are most important in determining population outcomes (e.g., population growth rate and/or persistence)?, and, What conservation strategies are most likely to positively influence population outcomes? We focus on brook trout (*Salvelinus fontinalis*) populations within their native range. Brook trout are of particular conservation concern and are thought to be particularly sensitive to environmental changes, such as temperature, flow and connectivity. Importantly, long-term, individual-based studies of free-ranging brook trout have been conducted and indeed are ongoing, providing the large quantities of high quality data that are necessary to parameterize models of population processes and their dependence on various aspects of the environment.

The premise of the population persistence modelling component of the NALCC is that the best way to provide credible information to managers on population response to environmental change and management actions is to develop models of population dynamics that explicitly incorporate relationships between environmental drivers, local habitat, and demographic processes. Key features of the models are scalability across space and time, incorporation of uncertainty in relationships among model components, and generation of easily interpretable results.

The project consists of four primary components:

1. Models converting air temperature/precipitation in to stream temperature/stream flow (described in detail in NALCC 'Hydrology and stream temperature' document).
2. Incorporation of output from the terrestrial project (climate change forecasts, land use etc.) into demographic models flow (described in detail in NALCC 'Land/change land use modeling' document).
3. Demographic models for stream fishes.
4. Decision support tool to evaluate effects of alternate management actions on population persistence.

Here, we describe #3 in detail.

[↑ Index](#)

3 Solution Statement

To develop models that can effectively model population persistence of stream fishes, we need a flexible modeling framework to account for multiple data sources, models that account for uncertainty in parameter estimates at all stages of the modeling process, component models that can be linked easily, and good data sources. We are incorporating existing data from multiple study sites into new statistical modeling approaches (Bayesian state space models) which are ideally suited to synthesize information at multiple levels and to account for error propagation.

1. Advantages of Bayesian state space models:
 - a. Provide a statistical distribution for output variables, i.e. uncertainty in abundance or probability of persistence.
 - b. Output uncertainty can easily be decomposed into contributions from modeled steps, i.e. what proportion of the uncertainty comes from climate change projections and what proportion comes from demographic or spatially-driven processes.
 - c. Observation error and process error can be modeled separately - process error reflects lack of complete knowledge of the mechanisms responsible for modeled dynamics while observation error reflects incomplete or imperfect sampling.
 - d. Can identify additional data collection needs and can determine minimum data needs for new systems.

- e. Quite complex models can be specified.
 - f. Can unify data collected at multiple levels of complexity and can combine these data sources in a statistically-defendable way to model complex processes including climate change effects.
2. As a result of the inherent flexibility in Bayesian hierarchical models, the general framework proposed here could easily be adapted to other systems. In this proposal, we will work out the kinks for a well-studied system (stream salmonids), and will incorporate what we have learned into a generic modeling framework that can be adapted to model downscaled climate change effects on other species and locations.

3.1 General solution statement

We are adopting a strategy best characterized as development of a "Modular Hierarchical State-Space Model" (MHSSM) in a Bayesian framework. "State-Space" modelling is a strategy whereby process and observation are both explicitly modelled. This is necessary for a number of reasons: for example, because natural populations can rarely be completely censused, capture histories of individuals typically reflect both capture rate (i.e., observation) and survival rate (i.e., process). By explicitly and jointly modelling survival and capture, it is possible to robustly obtain unbiased estimates of both capture and survival. "Hierarchical" modelling allows use of information from multiple locations and sources. This has several potential benefits, including the potential for data from well-studied systems to be used to inform analyses of important, but perhaps less-studied locations. Most importantly, and of particular interest here, is that the "Modular" aspect of our modelling allows us to simultaneously model different aspects of 'how populations work'. The modular approach allows us to simultaneously consider how different aspects of 'how populations work', such as how survival and movement contribute to population outcomes, and also, how population outcomes may be driven by environmental variation acting through multiple aspects of population biology. Importantly, by implementing our MHSSM in a Bayesian analysis, we can also evaluate uncertainty in all aspects of our joint models of population processes, leading to complete acknowledgement of 'what we know' and 'what we do not know'. It is also possible to integrate over uncertainty in environmental data, for example, if data are hind-casted (potentially, flow from hydrological models), or forecasted (potentially, predictions of future environmental conditions). Critically, robust assessment of uncertainty in high-level population outcomes, such as persistence probability or

population growth rate is critical to management decisions, but is remarkably rare, both in contemporary fundamental population biology and in conservation science.

A schematic of our MHSSM is presented in figure 1. We provide some mathematical description of the model below. The modularity of the MHSSM is depicted by red boxes. In the form depicted in figure 1, there are four modules. The modules pertain to growth and size, survival, movement, and reproductive success. Simultaneous evaluation of the modular model components, including inference of how the components depend on one another (i.e., size-dependence of survival, movement, and reproductive success), and how each component depends on multiple aspects of environmental variation provides information to evaluate population outcomes (blue box in figure 1). Importantly, because the modules are evaluated simultaneously, the joint uncertainty in model parameters can be carried through to inference of the high-level population outcomes. In practice, this is accomplished by running the modules in a single Bayesian analysis and obtaining multiple samples of the joint posterior distribution of all of the parameters that characterize each module. The calculations that provide inference of the population outcomes are then repeated for each of the many samples of the joint posterior distribution to provide estimates and inference of statistical uncertainty in the parameters that ultimately matter.

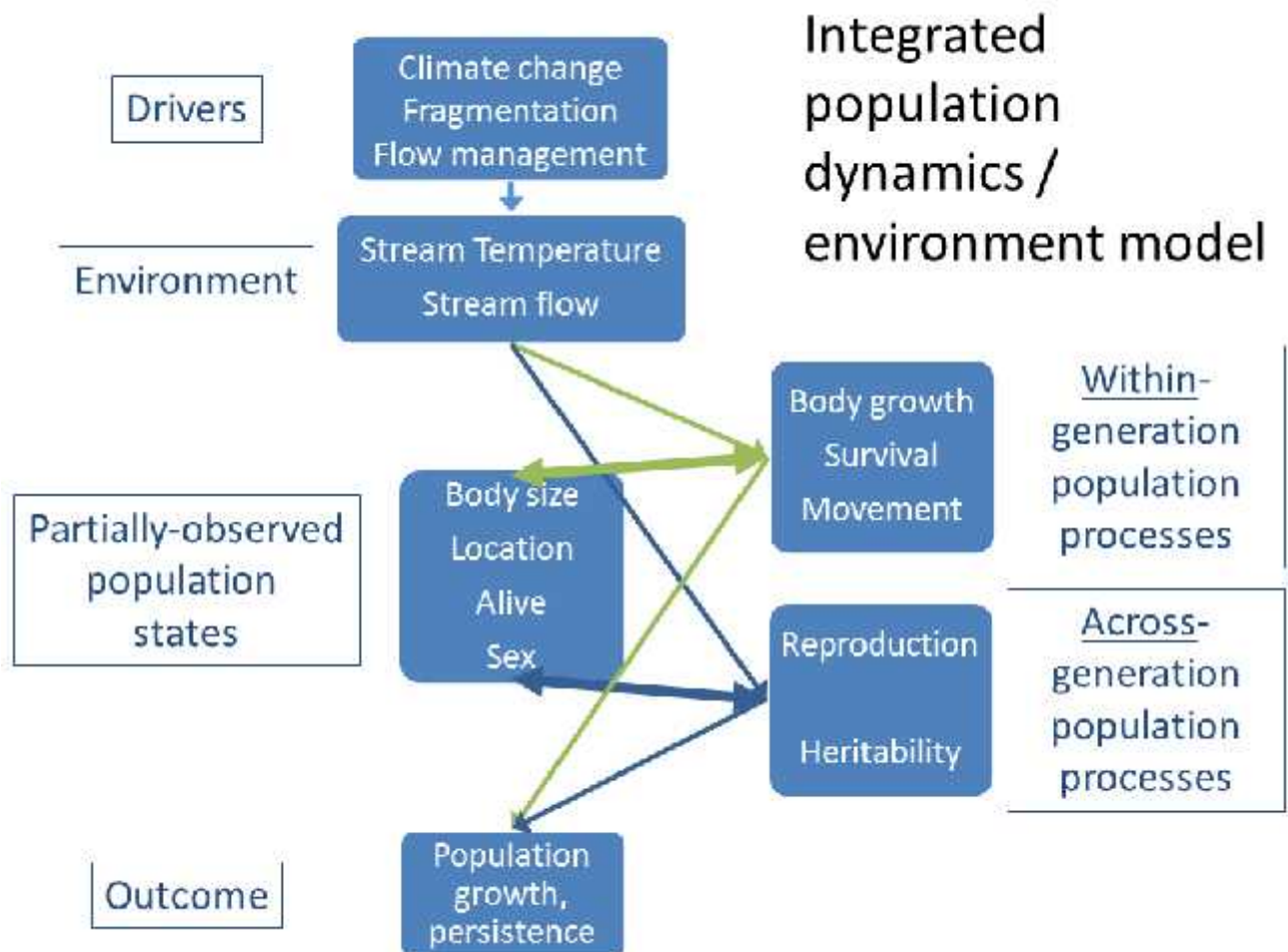


Figure 1. Depiction of major model components and their interactions.

Calculation of population outcomes from the joint posterior distribution of the parameters of multiple modular model components relies on both established and cutting-edge demographic theory. We use population projection models to take all of the population process information and synthesize it into population outcomes. Specifically, we use integral projection models (IPMs). The choice of IPMs is very important in the context of understanding fish populations, and these models are much more capable of incorporating the fact that population processes are likely to be highly influenced by size, and also that the effect of the environment on population parameters may act in large part through variation in growth and body size. Notably, alternative population projection models (size- and stage-structured models) are not specifically designed to handle fundamentally-important structuring variables that are continuous

in nature, such as body size. Established methods in demography allow explicit characterization of how population outcomes depend of potentially variable quantities, in our case multiple aspects of the environment. Specifically, determination and quantitative inference of the environmental-dependence of both population processes and population outcomes will be accomplished via sensitivity and elasticity analysis.

3.1.1 Specific research questions

The modelling framework we are implementing is very flexible, and allows us to address many important and specific questions. For the purpose of illustration, some of the most critical questions follow. See section 4.3.3 for more detail on how the specific methods we are implementing will allow estimation of specific parameters to address each question.

(3.2.i) What is the population growth rate in our primary study area?

(3.2.ii) To what extent does each vital rate (age-specific survival and reproduction) influence population growth rate?

(3.2.iii) How much does variation in environmental variables (stream temperature, flow) influence population growth rate?

(3.2.iv) Through which vital rates do environmental variables most influence population growth rate?

(3.2.v) What portion of the effect of environmental variation on vital rates and population growth rate is mediated by variation in (body) growth rate and body size?

(3.2.vi) How much does population growth rate vary in time?

Taken together, the improved understanding of population biology resulting from addressing these questions, as well as the availability of a cutting-edge mathematical apparatus for estimation of demographic parameters, will allow us to answer specific management questions. For example, when we know the answers to these questions, we will not only be able to determine which aspects of environmental changes will most affect brook trout populations, but we will also be able to put this knowledge to use to identify environmental changes that are likely to be most harmful to brook trout populations, and will also to allow us to quantitatively evaluate the relative benefits of different conservation and mitigation strategies.

[↑ Index](#)

4 Key Features

4.1 Available data

The bulk of the modeling will be based on data from our long-term study site in Whately, MA. Since 1997, we have tagged over 25,000 salmonids in the 1-km long study area. We have also sampled fish from three 300-m long tributaries since 2002. All samplings are seasonal (4 times/year). Captured fish longer than 60 mm are tagged with 12-mm PIT tags and length, body mass and location are recorded for all captured fish. Details of the study site and sampling methods can be found in (Letcher and Gries 2003, Letcher et al. 2004, 2011).

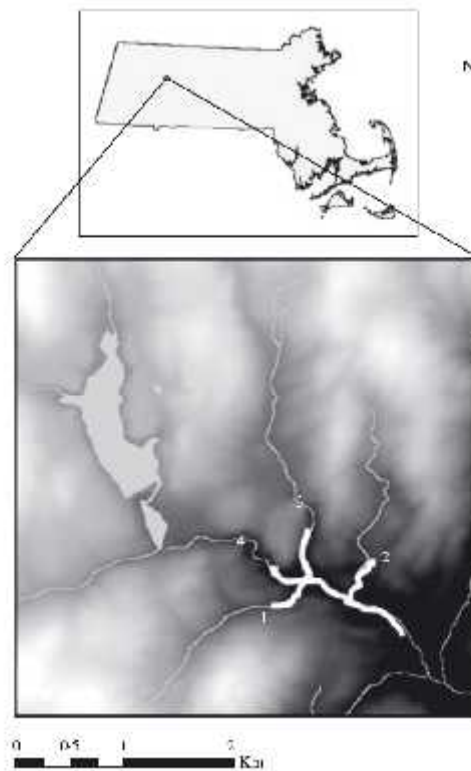


Figure 1. Study site location in western Massachusetts. Labels 1-3 refer to tributaries and 4 is the West Brook mainstem.

We have also conducted shorter-duration studies in other locations (see Fig. 2) and data from these studies will be combined with analyses of the West Brook data to allow a hierarchical analysis of brook trout population dynamics. We are also working with state and federal managers to compile a dataset of brook trout sampling sites used by managers to characterize brook trout populations in each state.

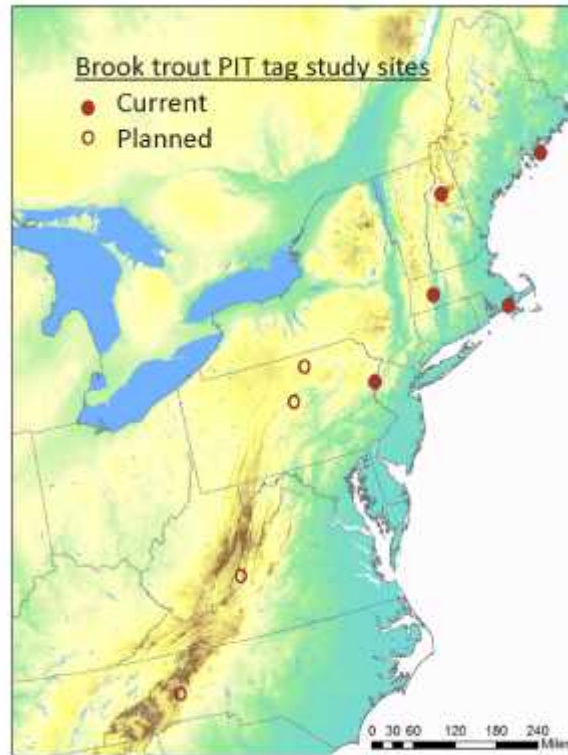


Figure 2. Study site locations of the West Brook and shorter-duration brook trout PIT tag studies.

We are also exploring statistical methods for joining our detailed models that use longitudinal data (individuals captured more than once) with cross-sectional data (many samples across space, e.g. state index site sampling efforts). Combined models of this sort may provide the best approach for scaling up our detailed models across space.

4.2 Model components

Our primary goal is to understand how multiple aspects of environmental variation influence different population processes to ultimately affect variation in population outcomes. Our goal is thus neither necessarily to uncover new aspects of 'how populations work', except in an integrative sense, nor is our goal necessarily, or in any case initially, to develop innovative models of each process (module). For the most part we propose to

initially seek simple models of each population process, how each process is influenced by the environment and how the modules interact. In some cases, these models are highly innovative and novel in and of themselves, but for the most part we seek to use simple generalized additive models for each model module. Here we propose a basic form for each model component.

One general feature of our models for brook trout population processes will be that we will use size, rather than age, as an explicit structuring variable. Brook trout typically exhibit extensive intra-population variation in growth rates (Letcher et al. 2011), and consequently length-frequency distributions of consecutive cohorts often overlap substantially.

Combined with extensive information on the size-dependence of vital rates, especially fecundity and reproductive success (Letcher et al. 2007), it is both biologically sensible and pragmatic given the available data to model most features of brook trout population dynamics as primarily size-dependent processes. For brevity, we do not develop specific models of the size dependence of each demographic process below. For the most part, penalized splines will be used to estimate functions of size dependent growth, survival, and reproductive success. Penalized splines are useful because they can allow estimation of functions without a priori assumptions about their functional form while minimizing the effective number of modelled parameters. Spline-based functions will be used from estimation of size dependence in the characterization of the functions denoted by g , ϕ , r , and possibly m (growth, survival, reproductive success, movement), below.

4.2.1 Module description: growth rate and body size

The state equation of the growth model follows the general form

$$S_{i,t+1} = S_{i,t} + g_{i,t}$$

where s represents size and g represents growth. Here and below, i indexes individuals and t indexes time (sampling occasions). The growth rate function can be arbitrarily complex, and will include size (using a spline function, see 4.2, second paragraph) and environment (temperature, flow) dependent terms.

4.2.2 Module description: Module description: survival

Our survival model takes the basic form of a Cormack-Jolly-Seber model (Lebreton et al. 1992) in a standard state space formulation (Royle and Dorazio 2008),

$$\begin{aligned}
L(a_{i,t}|t = f_i) &= 1, \\
L(\mathbf{a}_i|\phi_{i,t}, t > f_i) &= [a_{i,t}|a_{i,t-1}, \phi_{i,t}], \quad [a_{i,t}|a_{i,t-1}, \phi_{i,t}] \sim \text{bern}(a_{i,t-1}\phi_{i,t}), \\
L(\mathbf{y}|\phi, p) &= \prod_{i=1}^N \prod_{t=f_i}^{t_{max}} L(y_{i,t}|p) = [o_{i,j}|a_{i,t}, p], \quad [y_{i,t}|a_{i,t}, p] \sim \text{bern}(a_{i,t}p),
\end{aligned}$$

where a is an array of index variables denoting the state of alive vs. dead of individuals, indexed by i at sampling times, t . f_i are the times at which individuals were first cited. ϕ is the per-interval survival probability and p is the per-sample detection probability. \mathbf{y} is an array of encounter histories for all individuals at all sampling occasions, with 1 denoting that an individual was sighted and 0 otherwise. $\text{bern}(x)$ denotes a draw from a Bernoulli distribution (a binomial distributing with one trial), with probability x of the outcome being 1 rather than 0. $\phi_{i,t}$ is itself modelled as a function of individual and environmental covariates, i.e., body size, temperature, flow, via logistic generalized additive models.

4.2.3 Module description: movement

The state equation of movement model follows the general form

$$l_{i,t+1} = l_{i,t} + m_{i,t}$$

where l represents location, and m represents movements. $m_{i,t}$ are drawn from distributions describing individual movement, that can incorporate different expected values of absolute displacement due to individual covariates such as size, and environmental covariates, such as temperature. We have found that a three parameter t distribution is very robust for modelling movement, but in principle other movement kernels are possible. For example, we could use a double exponential movement kernel, which is commonly applied in the salmonid movement literature (Rodriguez 2002, Skalski 2007). This movement model augments locations for all individuals; augmented values when individuals are not captured are not expected to be precise, but this modelling framework provides high quality information on population-level parameters, such as local density and the proportion of individuals that are inside, vs. outside, of the study area, and thus available for capture.

Modelling of movement thus has two very important benefits. First, this will ultimately allow us to study the consequences of connectivity among different parts of the study system for the performance of brook trout throughout their life cycle. For brevity, we do not specifically further develop this aspect of our study in this document. Second, permanent and temporary emigration are population processes that traditionally make estimation of vital rates, particularly survival rates, very difficult. We have recently applied

models very similar to those described in the previous section (4.2.2) in conjunction with a movement model very similar to that described here. We have shown that this can yield unbiased estimates of survival rate, even in open populations, essentially, because the presence of individuals on the study site can be modelled probabilistically.

4.2.4 Module description: reproduction

The model module for reproduction takes the form of an overdispersed Poisson generalized additive model

$$\log(E[RS_{i,t}]) = s_{i,t} + f(E_{i,t}) + e_{i,t}$$

where $E[RS_{i,t}]$ is the logarithm of the expected reproductive success of individual i at time t . $f()$ represents an arbitrary function of several predictors, including for example, body size, environmental variables (E) and interactions among these, and e are latent-scale residual deviates.

4.3 Population projection

4.3.1 Further justification for model selection

The components of the model described in section 4.2 characterize all of the aspects of the biology of a stream-dwelling salmonid that are required to parameterize population projection models. Population projection models lie at the core of population demography, and among other things, can be used to characterize a population's growth rate, and the sensitivity of population growth rates to environmental variables. The best-known population projection models are age- and stage-structured population projection models (Caswell 2001). However, because so many aspects of the biology of fishes are highly influenced by body size, it is very important to acknowledge the structuring effect of size, and the influence of the environment acting through growth rate variation. Some progress has been made to include growth rate and body size variation in by discretizing populations' size distributions and using size classes as a structuring variable in stage structured models (Rivot et al. 2004). However, recent development of integral projection models (IPMs) provides a population projection approach that (i) fully incorporates the continuous nature of continuous traits, such as body size, that influence population demography, and (ii) allows construction of models that can be directly built upon empirical estimates of vital rates, and the variation therein as determined by variation in body size. This second feature of IPMs allows us (iii) to explicitly and directly merge the estimation step of population analysis (section 4.2) with the population projection model and (iv) when applied in a

Bayesian context allows us to fully consider uncertainty in all model parameters, including estimates obtained by methods described in section 4.2 and also in emergent parameters

4.3.2 Integral projection model construction

Following Coulson et al. (2010), the most general description of an IPM is

$$n(1, t + 1, z) = \sum_a \int dz' D(a, t, z | z') R(a, t, z') n(a, t, z')$$
$$n(a + 1, t + 1, z) = \int dz' G(a, t, z | z') S(a, t, z') n(a, t, z'), \quad a \geq 1$$

where $n(a, t, z)$ represents the density of individuals age a at time t with trait values z (z may be multivariate, and may be composed of continuous or discrete characters). The functions D , R , G and S represent inheritance, reproduction, ontogenetic change, and survival, respectively. Note that R , G , and S are IPM components that correspond directly to model modules described in section 4.2. Initially, we will apply simple models of inheritance, D based on Letcher et al. (2011), but ultimately, incorporation of quantitative genetic methods to estimate genetic parameters of characters such as growth rate will allow incorporation of evolutionary dynamics into our models. While not described here, extension of the basic IPM framework to include spatial structure is relatively straightforward, is one of our longer term goals, and will provide important insights into the importance of spatial connectivity for conservation of organisms such as brook trout. Age structure at time t is obtained by integrating over z for all a , and population size is obtained integrating over all z and a given t . Given constant functions D , R , G , and S , very good approximate numerical methods are available for deriving important demographic quantities such as population growth rate, age structure, and reproductive value. Similarly, metrics of the dependence of population processes, such as population growth rate, on environmental variables are obtainable using good approximate analytical solutions for deterministic cases and relatively simple computationally intensive numerical methods under environmental and demographic stochasticity. Particular examples of the uses of parameters derived in the general IPM framework are given in the next section in the context of some of our study's specific research questions.

4.3.3 Critical parameters to be obtained by population projection

Here, we repeat the specific (example) research questions from section 3.2, and now provide further description regarding what parameters obtained from our population projection models allow us to address each question.

(4.4.4.i) What is the population growth rate in our study area?

IPMs, and population projection models in general allow calculation of λ , the population growth rate.

(4.4.4.ii) To what extent does each vital rate (age-specific survival and reproduction) influence population growth rate?

Sensitivities and elasticities of λ to variation in vital rates are the absolute and relative changes in λ that are caused by changes in particular vital rates.

(4.4.4.iii) How much does variation in environmental variables (temperature, flow) influence population growth rate?

Integrated sensitivities and elasticities (van Tienderen 2000) can be calculated for IPMs and provide summaries of the total effect of given environmental variables on population growth rate, as determined by all causative effects of the environmental variables under consideration on all vital rates.

(4.4.4.iv) Through with vital rates do environmental variables most influence population growth rate?

Path analysis (van Tienderen 2000) provides a means of separating the total integrated sensitivities and elasticities of population growth rate to environmental variables into the separate causative effects that the environment has on each vital rate.

(4.4.4.v) What portion of the effect of environmental variation on vital rates and population growth rate is mediated by variation in (body) growth rate and body size?

As for 4.4.4.iv, path analysis provides a means by which to separate the downstream effects of environmental variation on growth rates, which may subsequently affect vital rates through changes in a population's size structure, from effects of the environment acting more directly on the vital rates.

(4.4.4.vi) How much does population growth rate vary in time?

Variation in demographic parameters, such as population growth rate, can come from two sources. Environmental stochasticity is variation caused by variation in environmental drivers of variation in vital rates. The estimation and population projection models that we have constructed allow us to calculate population growth rates under different environmental conditions, and also allow us to use numerical methods to obtain trajectories of population size, population growth rate, size and age structures under arbitrary (i.e., real, simulated, and/or projected) time series of environmental data. Demographic stochasticity

is variation in vital rates and population growth rate that arises from stochasticity in vital rates due to small population size, and is thus a source of variation in population performance that is of particular interest in populations of conservation concern. Given the estimation and population projection models that we describe above, numerical methods applicable to investigation of environmental stochasticity are applicable to demographic stochasticity as well.

4.4 Data processing

All data are stored in a Microsoft Access relational database. Data for analysis are extracted and further manipulated using R. we have written the code for parameter estimation in the Bugs language (Lunn et al. 2000) and run the Bugs language using JAGS (<http://mcmc-jags.sourceforge.net/>). The interface between R and JAGS is accomplished with the r package rjags.

 [Index](#)

5 Alternatives Considered and Rejected

We have used simple matrix projection models to develop models of size-dependent growth (Xu et al. 2010), size-dependent survival (Letcher and Horton 2008), and movements (Letcher et al. 2007), but these models do not capture important dynamics and their interactions fully. For all these reason explained above, we have adopted a Bayesian modular modelling approach for developing population projection models.

 [Index](#)

6 Major Implementation Constraints

These are complex models and working out all the details and interactions takes time and thorough simulation. However, once all the details are worked out, the parameter estimates extend easily to development of the IPM and then to use by managers. We are working closely with resource managers to determine the best format for use of the models to evaluate alternate management scenarios.

 [Index](#)

7 Major Risks and Dependencies

7.1 Major risks

Scaling up: the major risk of this project is the ability to effectively scale up dynamics from our detailed study sites to other watersheds. We will take two approaches to this problem. First we will combine PIT tag data from all long-term study sites into one analysis, using location as a random effect. This should account for the variation across study sites, but will not account for variance across all possible sites. To approach defining population parameters (across sites), our second approach will be to combine data from cross-sectional studies (agency index sites) into the analysis. This will help define means and variances across the range of index sites.

7.2 Dependencies

There are zero dependencies for the PIT tag datasets. We will rely heavily on external datasets for the cross-sectional data analysis, however. We will be able to account for a large portion of the variation in data quality and data collection techniques across collection sites and times, but there will inevitably be some variation that will not be included in the models.

[↑ Index](#)

8 Acknowledgments

[↑ Index](#)

9 Literature Cited

Coulson, T., S. Tuljapurkar, and D. Z. Childs. 2010. Using evolutionary demography to link life history theory, quantitative genetics and population ecology. *The Journal of animal ecology*:1-15. doi: 10.1111/j.1365-2656.2010.01734.x.

Lebreton, J. D., K. P. Burnham, J. Clobert, and D. R. Anderson. 1992. Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. *Ecological monographs* 62:67–118. Eco Soc America. Retrieved November 8, 2011, from <http://www.esajournals.org/doi/abs/10.2307/2937171>.

Letcher, B. H., J. a Coombs, and K. H. Nislow. 2011. Maintenance of phenotypic variation: repeatability, heritability and size-dependent processes in a wild brook trout population. *Evolutionary Applications* 4:602-615. doi: 10.1111/j.1752-4571.2011.00184.x.

- Letcher, B. H., T. L. Dubreuil, M. J. O'Donnell, M. Obedzinski, K. Griswold, and K. H. Nislow. 2004. Long-term consequences of variation in timing and manner of fry introduction on juvenile Atlantic salmon (*Salmo salar*) growth, survival, and life-history expression. *Canadian Journal of Fisheries and Aquatic Sciences* 61:2288-2301. doi: 10.1139/f04-214.
- Letcher, B. H., K. H. Nislow, J. Coombs, M. J. O'Donnell, and T. L. Dubreuil. 2007. Population response to habitat fragmentation in a stream-dwelling brook trout population. *PloS one* 2:e1139. doi: 10.1371/journal.pone.0001139.
- Letcher, B. H., and G. Gries. 2003. Effects of life history variation on size and growth in stream-dwelling Atlantic salmon. *Journal of Fish Biology* 62:97-114. doi: 10.1046/j.0022-1112.2003.00009.x.
- Letcher, B. H., and G. E. Horton. 2008. Seasonal variation in size-dependent survival of juvenile Atlantic salmon (*Salmo salar*): performance of multistate capture-mark-recapture models. *Canadian Journal of Fisheries and Aquatic Sciences* 65:1649-1666. doi: 10.1139/F08-083.
- Lunn, D. J., A. Thomas, N. Best, and D. Spiegelhalter. 2000. WinBUGS -- a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing* 10:325-337.
- Rivot, E., E. Prévost, E. Parent, and J.-L. Bagliniere. 2004. A Bayesian state-space modelling framework for fitting a salmon stage-structured population dynamic model to multiple time series of field data. *Ecological Modelling* 179:463-485. doi: 10.1016/j.ecolmodel.2004.05.011.
- Rodríguez, M. A. 2002. Restricted movement in stream fish: the paradigm is incomplete, not lost. *Ecology* 83:1-13.
- Royle, J. A., and R. Dorazio. 2008. Hierarchical modeling and inference in ecology. Page 444. Academic Press, San Diego.
- Skalski, G. T. 2007. Joint estimation of migration rate and effective population size using the island model. *Genetics* 177:1043-57. doi: 10.1534/genetics.107.078998.
- van Tienderen, P. H. 2000. Elasticities and the Link between Demographic and Evolutionary Dynamics. *Ecology* 81:666. doi: 10.2307/177368.
- Xu, C. L., B. H. Letcher, and K. H. Nislow. 2010. Size-dependent survival of brook trout *Salvelinus fontinalis* in summer: effects of water temperature and stream flow. *Journal of Fish Biology* 76:2342-2369. doi: 10.1111/j.1095-8649.2010.02619.x.

[↑ Index](#)

10 Appendix A – Diagrams/Figures

11 Appendix B – Model code

JAGS (BUGS) code for the age-based model

R code to call the JAGS model is available on request

```
#####  
# CJS model with emigration  
# reparameterized so that events happening during an interval are related  
# to the line in dMData pertaining to the occasion preceeding the interval  
# Michael Morrissey, Krzysztof Sakrejda-Leavitt, Ben Letcher  
#####  
  
var probEm[nAllRows], z[nAllRows], available[nAllRows], probSeenGoing[nAllRows],  
    matBetas[nAgeInSamples,nCohorts], maturing[nAllRows], gr[nAllRows], mu[nAllRows],  
    size[nAllRows], mature[nAllRows], cohortGrowthIntercept[nAgeInSamples,nCohorts],  
    phiAlphas[nAgeInSamples,nCohorts], l[nAllRows], location[nAllRows],  
    matAlphas[nAgeInSamples,nCohorts], matureS[nAllRows], matAlphaMu[14];  
  
model {  
  
#####  
# Growth model  
#####  
  
# Size at recruitment (state):  
for( i in 1: nAgeInSamplesOneRows ) {  
  size[ageInSamplesOneRows[i]] ~dnorm(  
    cohortSizeIntercept[cohortIndex[ageInSamplesOneRows[i]]],  
    1/recruitSizeSigma^2 )  
}  
  
# Size after recruitment (growth):
```

NALCC Project Component Name: *Population persistence modelling*

Document Type: Project Documentation

```
for(i in 1:nEvalJSRows){
  mu[evalJSRows[i]] <-
  cohortGrowthIntercept[ageInSamples[evalJSRows[i]], cohortIndex[evalJSRows[i]]] +
  b[ageInSamples[evalJSRows[i]]] *
  (size[evalJSRows[i]]-sizeStd[ageInSamples[evalJSRows[i]]]) +
  t[ageInSamples[evalJSRows[i]]] *
  (temperature[evalJSRows[i]] - tempStd[season[evalJSRows[i]]]) +
  f[ageInSamples[evalJSRows[i]]] *
  (flow[evalJSRows[i]] - flowStd[season[evalJSRows[i]]]) +
  m[ageInSamples[evalJSRows[i]]] *
  mature[evalJSRows[i]]

  gr[evalJSRows[i]] ~ dnorm( mu[evalJSRows[i]]*(intervalDays[evalJSRows[i]]*4/365),
    1/(growthSigma[ageInSamples[evalJSRows[i]]]^2*
      (intervalDays[evalJSRows[i]]*4/365)) )
  size[evalJSRows[i]+1] <-
  size[evalJSRows[i]] + gr[evalJSRows[i]]
}

# Observed size (measurements accurate when fish is captured/measured).
for(i in 1:nAllRows){
  obsLength[i] ~ dnorm(size[i],1)
}

#####
# Priors (growth) : ...
#####

# Size of individuals in cohort at recruitment.
sizeIntercept ~ dnorm(65,0.0001)
sizeInterceptSigma ~ dgamma(2,1/20)
recruitSizeSigma ~ dgamma(2,1/20)
for(c in 1:nCohorts){
  cohortSizeIntercept[c] ~ dnorm(sizeIntercept, 1/sizeInterceptSigma^2)
}

# Age-cohort specific growth model intercept
for(a in 1:nAgeInSamples){
  growthIntercept[a] ~ dnorm(0,0.0001)
  cohortSigma[a] ~ dgamma(2,1/10)
  growthSigma[a] ~ dgamma(2,1/20)
  for(c in 1:nCohorts){
    cohortGrowthIntercept[a,c] ~ dnorm(growthIntercept[a], 1/cohortSigma[a]^2)
```

Author: *Ben Letcher and Michael Morrissey*

Page 19 of 29

Created on 14/07/2011 14:26:00

```
}
}

# Additional linear effects for growth model.
for(a in 1:(nAgeInSamples-1)){
  b[a]~dnorm(0,0.01) # effect of size (size-dependent growth)
  f[a]~dnorm(0,0.01) # effect of flow
  t[a]~dnorm(0,0.01) # effect of temperature
  m[a]~dnorm(0,0.01) # effect of maturity
}

#####
# Survival parameters
#####

for(i in 1:nEvalRows){
  logit(phi[evalRows[i]]) <-
    phiAlphas[ ageInSamples[evalRows[i]], cohortIndex[evalRows[i]] ] +
    phiSizeBetas[ ageInSamples[evalRows[i]] ] *
      (size[evalRows[i]]-sizeMu[ageInSamples[evalRows[i]]]) /
      sizeSD[ ageInSamples[evalRows[i]] ] +
    phiTempBetas[ ageInSamples[evalRows[i]] ] *
      (temperature[evalRows[i]]-tempMu[season[evalRows[i]]]) /
      tempSD[ season[evalRows[i]] ] +
    phiFlowBetas[ ageInSamples[evalRows[i]] ] *
      (flow[evalRows[i]]-flowStd[season[evalRows[i]]]) /
      flowSD[ season[evalRows[i]] ] +
    phiMaturityBetas[ ageInSamples[evalRows[i]] ] *
      mature[evalRows[i]]
}

# Priors (survival): for intercepts and slopes, values are independent among
# ages and hierarchical within ages, and for effects of temperature and
# flow, effects are independent among ages and not variable among cohorts

# parameter expanded priors for random latent scale survival intercepts
xiPhiAlpha ~ dnorm(0,0.001)
phiAlphaTauEta ~ dgamma(0.5,0.5)
phiAlphaSigma <- abs(xiPhiAlpha)/sqrt(phiAlphaTauEta)

for( a in 1 : nAgeInSamples ){
```

NALCC Project Component Name: *Population persistence modelling*

Document Type: Project Documentation

```
for( c in 1:nCohorts ){
  phiAlphas[ a,c ] <- phiAlphaMu[ a ]+xiPhiAlpha*phiAlphaEta[ a,c ]
  phiAlphaEta[ a,c ] ~ dnorm( 0, phiAlphaTauEta )
}

phiAlphaMu[a]      ~ dt(0, 1, 1)
phiSizeBetas[a]   ~ dt(0, 1, 1)
phiTempBetas[a]   ~ dt(0, 1, 1)
phiFlowBetas[a]   ~ dt(0, 1, 1)
phiMaturityBetas[a] ~ dt(0, 1, 1) # note: expect no information when no maturity!

}

#####
# Maturation parameters
#####

for( i in 1:nEvalJSRows ) {
  logit( probMat[ evalJSRows[i]+1 ] ) <-
    matAlphas[ ageInSamples[ evalJSRows[i]+1 ], cohortIndex[ evalJSRows[i]+1 ] ] +
    matSizeBetas[ ageInSamples[evalJSRows[i]+1 ] ] *
    (size[ evalJSRows[i]+1 ] - sizeMu[ ageInSamples[ evalJSRows[i]+1 ] ] ) /
    sizeSD[ ageInSamples[ evalJSRows[i]+1 ] ]
}

# Priors (maturation): intercepts and slopes, values are independent among
# ages and hierarchical within ages

# parameter expanded priors for random latent scale maturation intercepts
xiMatAlpha ~ dnorm(0,0.001)
matAlphaTauEta ~ dgamma(0.5,0.5)
matAlphaSigma <- abs(xiMatAlpha)/sqrt(matAlphaTauEta)

for(a in 1:3 ){ # model is really only for summers (AIS 4, 8, 12)

  for( c in 1:nCohorts ){
    matAlphas[ summers[a],c ] <- ( matAlphaMu[ summers[a] ] +
      xiMatAlpha * matAlphaEta[ summers[a],c ] )
    matAlphaEta[summers[a],c]~dnorm(0, matAlphaTauEta)
  }

  matAlphaMu[summers[a]] <- matAlphaMuConstant
  matSizeBetas[summers[a]] <- matSizeBetaConstant
```

Author: *Ben Letcher and
Michael Morrissey*

Page 21 of 29

Created on 14/07/2011 14:26:00

NALCC Project Component Name: *Population persistence modelling*

Document Type: Project Documentation

```
}
matAlphaMuConstant ~ dt(0, 1, 1)
matSizeBetaConstant ~ dt(0, 1, 1)

for(a in 1 : 11 ){ # null values for non-summer AIS

  for( c in 1:nCohorts ){
    matAlphas[ nonSummers[a],c ] <- -10
    matAlphaEta[nonSummers[a],c] <- 1
  }

  matAlphaMu[nonSummers[a]] <- -10
  matSizeBetas[nonSummers[a]] <- 0
}

#####
# Location/movement module
#####

# note that this is set up to sample location from the first capture of
# individuals. To have location-specific growth, etc, it will become necessary
# to use all of the JS rows...

# location at recruitment:
for( i in 1: nFirstObsRows ) {
  loc[ firstObsRows[i] ] ~ dunif( 4000, 5500 )
}

# location after recruitment:
for(i in 1:nEvalRows){
  loc[ evalRows[i]+1 ] ~ dt( loc[evalRows[i]]
    + smoltDrift*smolting[evalRows[i]]
    + matDrift[season[evalRows[i]]]*mature[evalRows[i]],
    tauMove[ ageInSamples[ evalRows[i] ], cohortIndex[ evalRows[i] ] ],
    kMove[ ageInSamples[ evalRows[i] ], cohortIndex[ evalRows[i] ] ] )

  inside[evalRows[i]+1] <- ( loc[evalRows[i]+1] >= 4242.25 ) *
    ( loc[evalRows[i]+1] <= 5493.85 )
  inside47[evalRows[i]+1] <- ( loc[evalRows[i]+1] > 4383.25 ) *
    ( loc[evalRows[i]+1] <= 5360.4 )
}
```


NALCC Project Component Name: *Population persistence modelling*

Document Type: Project Documentation

```
# Observed location
for(i in 1: nEvalRows){
  sectionAdj[evalRows[i]] ~ dinterval( loc[evalRows[i]], sectionBreaks )
}
for(i in 1: nLastObsRows){
  sectionAdj[lastObsRows[i]] ~ dinterval( loc[lastObsRows[i]], sectionBreaks )
}

# Priors - no heterogeneity right now
for(a in 1:nAgeInSamples){
  sigmaMoveAge[a] ~ dgamma(2,1/250)
  tauMoveAge[a] <- 1/sigmaMoveAge[a]^2
  kMoveAge[a] ~ dgamma(2,1/100)
  for(c in 1:nCohorts){
    tauMove[a,c] <- tauMoveAge[a]
    kMove[a,c] <- kMoveAge[a]
  }
}
smoltDrift ~ dnorm(0,0.001)
for(s in 1:4){
  matDrift[s] ~ dnorm(0,0.001)
}

#####
# Smolting parameters:
#####

for(i in 1:nEaloRows){
  logit(probSmolt[ealoRows[i]]) <-
    smoltAlphas[ ageInSamples[ealoRows[i]], cohortIndex[ealoRows[i]] ] +
    smoltSizeBetas[ ageInSamples[ealoRows[i]] ] *
    (size[ealoRows[i]] - sizeMu[ageInSamples[ealoRows[i]]]) /
    sizeSD[ ageInSamples[ealoRows[i]] ] +
    smoltMaturityBetas[ ageInSamples[ealoRows[i]] ] *
    mature[ealoRows[i]]
}

#####
# Priors (smolting): for intercepts and slopes, values are independent among
# ages and hierarchical within ages
#####
```

NALCC Project Component Name: *Population persistence modelling*

Document Type: Project Documentation

```
# parameter expanded priors for random latent scale smolting intercepts
xiSmoltAlpha ~ dnorm(0,0.001)
smoltAlphaTauEta ~ dgamma(0.5,0.5)
smoltAlphaSigma <- abs(xiSmoltAlpha)/sqrt(smoltAlphaTauEta)

for(a in 1 : 3 ){ # model is really only for springs (AIS 3, 7, 11)

  for( c in 1:nCohorts ){
    smoltAlphas[ springs[a],c ] <- ( smoltAlphaMu[springs[a]]
      +xiSmoltAlpha*smoltAlphaEta[springs[a],c ] )
    smoltAlphaEta[springs[a],c]~dnorm(0,smoltAlphaTauEta)
  }

  smoltAlphaMu[springs[a]] <- smoltAlphaMuConstant
  smoltSizeBetas[springs[a]] <- smoltSizeBetaConstant
  smoltMaturityBetas[springs[a]] <- smoltMaturityBetaConstant
}

# this makes the reaction norm non-heterogeneous by age, which is necessary
# in WB because of little or no data at some ages
smoltAlphaMuConstant ~ dt(0, 1, 1)
smoltSizeBetaConstant ~ dnorm(0,0.01)
smoltMaturityBetaConstant ~ dnorm(0,0.01)

for(a in 1 : 11 ){ # null values for non-spring AIS

  for( c in 1:nCohorts ){
    smoltAlphas[ nonSprings[a],c ] <- -10
    smoltAlphaEta[nonSprings[a],c] <- 1
  }

  smoltAlphaMu[nonSprings[a]] <- 0
  smoltSizeBetas[nonSprings[a]] <- 0
  smoltMaturityBetas[nonSprings[a]] <- 0
}

#####
# Recapture parameters
#####

# intercepts independent by ages and drawn from separate distributions for
# different cohorts at each age, with flow effects as fixed effects and
# heterogeneous by age and common across cohorts
```

Author: *Ben Letcher and Michael Morrissey*

Page 24 of 29

Created on 14/07/2011 14:26:00

NALCC Project Component Name: *Population persistence modelling*

Document Type: Project Documentation

```
for(i in 1:nEvalRows){
  logit(p[evalRows[i]+1]) <-
    PAlphas[ ageInSamples[evalRows[i]+1], cohortIndex[evalRows[i]+1] ] +
    PFlowBetas[ ageInSamples[evalRows[i]+1] ] *
      ( flowForRecapture[evalRows[i]+1] - flowMu[season[evalRows[i]+1]] ) /
      flowSD[season[evalRows[i]+1]] +
    aboveBelow47[ ageInSamples[evalRows[i]+1] ] * (1-inside47[ evalRows[i]+1 ] ) +
    aboveBelow47[ ageInSamples[evalRows[i]+1] ] * (1-inside47[ evalRows[i]+1 ] )
}

#####
# Priors (recapture): age-specific mean intercepts with heterogeneity
# among cohorts, and age-specific regression coefficients for flow
#####

# parameter expanded priors for random latent scale capture intercepts
xiPAlpha ~ dnorm(0,0.001)
pAlphaTauEta ~ dgamma(0.5,0.5)
pAlphaSigma <- abs(xiPAlpha)/sqrt(pAlphaTauEta)

for(a in 1 : ( nAgeInSamples)){

  for( c in 1:nCohorts ){
    PAlphas[ a,c ] <- PAlphaMu[a]+xiPAlpha*pAlphaEta[a,c]
    pAlphaEta[a,c] ~ dnorm(0,pAlphaTauEta)
  }

  PAlphaMu[a] ~ dt(0, 1, 1)
  PFlowBetas[a] ~ dt(0, 1, 1)

  aboveBelow47[a] ~ dt(0, 1, 1)
}

poorWinterConditions ~ dt(0, 1, 1)

#####
# Latent sex: this really just allows re-scaling the probability of maturing
# to lie between zero and one, rather than hoping it comes out to less than 0.5.
# Note though that female-biased emmigration can make the maturation rate in any
```

Author: *Ben Letcher and Michael Morrissey*

NALCC Project Component Name: *Population persistence modelling*

Document Type: Project Documentation

```
# subsequent intervals go > 0.5. i.e.- Only _initial_ sex ratio is 50/50.
#####

for(i in 1:nInd){
  sex[i] ~ dbern(0.5)
}

#####
#####
#####
# Likelihood
#####
#####
#####

# Initial conditions:
# 1) individuals enter the sample with probability 1
# 2) individuals enter the sample alive, with probability 1
for(i in 1:nFirstObsRows){
  z[firstObsRows[i]] <- 1
  notSmolted[firstObsRows[i]] <- 1
}
# Also, fill in unused entries to allow setting the monitor
# from rjags.
for(i in 1:nBeforeFirstRows){
  z[beforeFirstRows[i]] <- 1
  notSmolted[beforeFirstRows[i]] <- 1
  going[beforeFirstRows[i]] <- 0
  smolting[beforeFirstRows[i]] <- 0
}

for(i in 1:nAgeInSamplesOneRows){
  mature[ageInSamplesOneRows[i]] <- 0
}

#####
# State-space CJS model for survival and capture. Survival is scaled to the
# actual number of days at large. Capture is conditional on being alive and
# being in the site. Capture probability is scaled based on whether a particular
# sample covered the entire study site or not. The dead/alive state is indicated
# in 'z'. 'phi' is the survival parameter (calculated above). 'p' is the
# capture probability (calculated above).
#####
```

Author: *Ben Letcher and Michael Morrissey*

NALCC Project Component Name: *Population persistence modelling*

Document Type: Project Documentation

```
for(i in 1:nEvalRows){

  # State of survival
  z[evalRows[i]+1]~dbern(survProb[evalRows[i]])
  survProb[evalRows[i]] <-
    phi[evalRows[i]]^(intervalDays[evalRows[i]]*4/365) * z[evalRows[i]]

  # note that capture probabilities have to be calculated for all occasions after
  # an individual was first captured. to do this using evalRows, which is first
  # capture to second last capture occasion or potential capture occasion, 1
  # has to be consistently added to evalRows

  # Observation of live encounters
  captured[ evalRows[i]+1 ] ~ dbern(obsProb[evalRows[i]+1])
  obsProb[ evalRows[i]+1 ] <-
    p[ evalRows[i]+1 ] * # Capture probability (calculated above).
    z[ evalRows[i]+1 ] * # Must be alive to be capturable.
    notSmolted[ evalRows[i]+1 ] * # Must not have smolted.
    inside[ evalRows[i]+1 ] * # Must be in the site to be capturable (includes above and below).
    sectionsSampledWithOutside[ year[ evalRows[i]+1 ] ,
      season[ evalRows[i]+1 ] ,
      dinterval( loc[ evalRows[i]+1 ] , sectionBreaks )+1 ]
}

#####
# Maturation model: Maturation can only occur "in" the summer (before Fall
# sample). Maturation process is set to zero for all non-summer rows.
# Maturity is a process which goes to 1 in
#####

# Observation of maturity - augmentation of NAs, no observation error assumed:
for( i in 1:nSummerObsRows ){
  seenMatureYear[summerObsRows[i]]~dbern(mature[summerObsRows[i]])
}

for( i in 1:nEvalJSRows ){

  maturing[ evalJSRows[i]+1 ]~ dbern(probMat[ evalJSRows[i]+1 ] *
    sex[integerID[ evalJSRows[i]+1 ] ] *

```

NALCC Project Component Name: *Population persistence modelling*

Document Type: Project Documentation

```
z[ evalJSRows[i]+1 ] * notSmolted[ evalJSRows[i]+1 ] )

mature[ evalJSRows[i]+1 ] <-      # maturing[evalJSRows[i]+1]
  ( summerIndicator[ evalJSRows[i]+1 ] ) * maturing[ evalJSRows[i]+1 ] +
  ( 1 - summerIndicator[ evalJSRows[i]+1 ] ) * mature[ evalJSRows[i] ]

}

#####
# State part of emigration (i.e., emigration other than smolting) and smolting
# model. State works in two parts. First probabilities of emigration and
# smolting are worked out conditionally on being available. Then realizations
# of emigration and smolting are worked out. Availability for the next
# interval is determined jointly from (not) smolting and (not) emigrating.
#####

# ealo -> evalAndLastObsRows...
for ( i in 1:nEaloRows ) {
  smoltProb[ealoRows[i]] <-
    probSmolt[ealoRows[i]] *      # - note not scaled by the number of
      # days in interval for individual.
    notSmolted[ealoRows[i]] *      # - can't smolt twice!
    z[ealoRows[i]]                # - must be alive to smolt.
}

for ( i in 1:nEvalRows ) {
  smolting[evalRows[i]] ~ dbern(smoltProb[evalRows[i]])
  notSmolted[evalRows[i]+1] <- notSmolted[evalRows[i]] * (1 - smolting[evalRows[i]])
}

for(i in 1:nLastObsRows){
  smolting[lastObsRows[i]] ~ dbern(smoltProb[lastObsRows[i]])
}

for ( i in 1:nEaloRows ) {
  smoltObsWithNAs[ealoRows[i]] ~ dbern(smolting[ealoRows[i]])
}

}
```

 [Index](#)